

# The 21-cm signal from the cosmic dawn: metallicity dependence of high-mass X-ray binaries

Harman Deep Kaur <sup>1</sup>★, Yuxiang Qin <sup>1,2,3</sup>, Andrei Mesinger <sup>1</sup>, Andrea Pallottini <sup>1</sup>, Tassos Fragos<sup>4</sup> and Antara Basu-Zych <sup>5,6,7</sup>

<sup>1</sup>*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

<sup>2</sup>*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

<sup>3</sup>*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*

<sup>4</sup>*Département d'Astronomie, Université de Genève, Chemin Pegasi 51, CH-1290 Versoix, Switzerland*

<sup>5</sup>*Department of Physics, University of Maryland Baltimore County, Baltimore, MD 21250, USA*

<sup>6</sup>*NASA Goddard Space Flight Center, Laboratory for X-ray Astrophysics, Greenbelt, MD 20771, USA*

<sup>7</sup>*Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA*

Accepted 2022 April 28. Received 2022 April 28; in original form 2022 March 21

## ABSTRACT

X-rays from high-mass X-ray binaries (HMXBs) are likely the main source of heating of the intergalactic medium (IGM) during Cosmic Dawn (CD), before the completion of reionization. This Epoch of Heating (EoH;  $z \sim 10\text{--}15$ ) should soon be detected via the redshifted 21-cm line from neutral hydrogen, allowing us to indirectly study the properties of HMXBs in the unseen, first galaxies. Low-redshift observations, as well as theoretical models, imply that the integrated X-ray luminosity to star formation rate of HMXBs ( $L_X/\text{SFR}$ ) should increase in metal-poor environments, typical of early galaxies. Here, we study the impact of the metallicity ( $Z$ ) dependence of  $L_X/\text{SFR}$  during the EoH. For our fiducial models, galaxies with star formation rates of order  $10^{-3}\text{--}10^{-1} M_\odot \text{ yr}^{-1}$  and metallicities of order  $10^{-3}\text{--}10^{-2} Z_\odot$  are the dominant contributors to the X-ray background (XRB) during this period. Different  $L_X/\text{SFR}\text{--}Z$  relations result in factors of  $\sim 3$  differences in these ranges, as well as in the mean IGM temperature and the large-scale 21-cm power, at a given redshift. We compute mock 21-cm observations adopting as a baseline a 1000 h integration with the upcoming Square Kilometer Array (SKA) for two different  $L_X/\text{SFR}\text{--}Z$  relations. We perform inference on these mock observations using the common simplification of a constant  $L_X/\text{SFR}$ , finding that constant  $L_X/\text{SFR}$  models can recover the IGM evolution of the more complicated  $L_X/\text{SFR}\text{--}Z$  simulations only during the EoH. At  $z < 10$ , where the typical galaxies are more polluted, constant  $L_X/\text{SFR}$  models overpredict the XRB and its relative contribution to the early stages of the reionization.

**Key words:** galaxies: high-redshift – (galaxies:) intergalactic medium – (cosmology:) dark ages, reionization, first stars – cosmology: theory – (cosmology:) early Universe – X-rays: binaries – X-rays: diffuse background – X-rays: galaxies.

## 1 INTRODUCTION

High-mass X-ray binaries (HMXBs) are expected to be the dominant source of heating in the early Universe, soon after the formation of the first galaxies (e.g. Furlanetto 2006; McQuinn 2012; Fragos et al. 2013; Pacucci et al. 2014). Their X-rays have long mean free paths, and are thus able to penetrate deep into the intergalactic medium (IGM), at a time when reionization was still in its infancy ( $z \sim 10\text{--}20$ ; Mirabel et al. 2011; McQuinn 2012; Madau & Fragos 2017; Eide et al. 2018). As a result, X-rays drive large-scale IGM temperature fluctuations, during this so-called Epoch of Heating (EoH; e.g. Pritchard & Furlanetto 2007; Santos et al. 2010; Mesinger, Furlanetto & Cen 2011; Visbal et al. 2012; Pacucci et al. 2014; Muñoz et al. 2022). The corresponding signal from the 21-cm line of neutral hydrogen is expected to be detectable by the ongoing Hydrogen Epoch of Reionization (HERA; e.g. DeBoer et al. 2017)<sup>1</sup> and Square

Kilometer Array (SKA)<sup>2</sup> interferometers (for a recent review, see e.g. Mesinger 2019).

Current theoretical models of the EoH are based on empirical scaling relations (e.g. Lehmer et al. (e.g. Lehmer et al. 2010; Kaaret, Schmitt & Gorski 2011; Basu-Zych et al. 2012; Mineo, Gilfanov & Sunyaev 2012; Kouroumpatzakis et al. 2020) between the (population-averaged) X-ray luminosity ( $L_X$ ) and star formation rates (SFRs) of local galaxies:  $L_X/\text{SFR}$ ). However, both theoretical models of HMXB evolution (e.g. Linden et al. 2010; Fragos et al. 2013) and recent observations (e.g. Basu-Zych et al. 2012; Prestwich et al. 2013; Douna et al. 2015; Brorby et al. 2016; Lehmer et al. 2021) suggest that this  $L_X/\text{SFR}$  relation should have a strong dependence on metallicity. Theoretically, we expect that weaker stellar winds resulting from lower metallicity environments would result in more X-ray luminous binaries, due to both a reduced expansion of the binary orbit and a reduced radial expansion of the companion that impacts when Roche-

\* E-mail: [harman.kaur@sns.it](mailto:harman.kaur@sns.it)

<sup>1</sup><https://reionization.org/>

<sup>2</sup><https://www.skatelescope.org/>

Lobe overflow occurs. This metallicity dependence can be important for the EoH, since the first galaxies are expected to form inside pristine gas with a rapidly evolving metallicity (e.g. Wise et al. 2012; Xu, Wise & Norman 2013; Pallottini et al. 2014; Jaacks, Finkelstein & Bromm 2019; Ucci et al. 2021). The cosmic 21-cm signal at those redshifts can therefore be used to study the properties of the first generations of HMXBs (e.g. Madau & Fragos 2017; Qin et al. 2020a).

In this work, we study the imprint of metallicity-dependent  $L_X/\text{SFR}$  relations in the 21-cm signal from the Cosmic Dawn.<sup>3</sup> Assuming a mass-metallicity relation, we compute the evolution of the 21-cm signal corresponding to two different scalings of  $L_X/\text{SFR}$  with metallicity.<sup>4</sup> We quantify how these relations impact the X-ray emissivity and IGM temperature evolution, and also make forecasts for the corresponding 21-cm signal.

This paper is organized as follows. In Section 2, we present the steps for computing the X-ray emissivity during the cosmic dawn. In Section 3.1, we compute the associated thermal evolution of the IGM, followed by the 21-cm signal in Section 3.2. In Section 4, we quantify if simpler, constant  $L_X/\text{SFR}$  models can recover the same IGM evolution as predicted from the fiducial  $L_X/\text{SFR}-Z$  relations. Finally, we conclude in Section 5. We assume a standard  $\Lambda$ CDM cosmology with the following cosmological parameters:  $h = 0.678$ ,  $\Omega_m = 0.308$ ,  $\Omega_b = 0.0484$ ,  $\Omega_\Lambda = 0.692$ ,  $\sigma_8 = 0.815$ ,  $n_s = 0.968$  based on result from Planck Collaboration XIII (2016a).

## 2 X-RAY EMISSION FROM EARLY GALAXIES

In this section, we present our model for computing the inhomogeneous X-ray background during the Cosmic Dawn. We begin by discussing the empirical and theoretical  $L_X/\text{SFR}-Z$  relations we use in Section 2.1. We then show the adopted mass-metallicity relation for assigning metallicities to dark matter haloes at high redshifts in Section 2.2. Finally in Section 2.3, we compute the corresponding X-ray backgrounds by integrating over the halo mass functions and SFRs of galaxies.

To compute the metallicity-dependent X-ray backgrounds and their impact on the IGM, we modify the public code, 21cmFAST.<sup>5</sup> 21cmFAST uses a combination of perturbation theory, excursion-set formalism and light-cone integration to compute 3D realizations of the IGM density, velocity, ionization, temperature, and 21-cm intensity fields. A detailed description is available in Mesinger & Furlanetto (2007), Mesinger et al. (2011), and Murray et al. (2020). Our simulations are 500 comoving Mpc per side, using a  $256^3$  grid smoothed down from  $1024^3$  initial conditions. These volumes are large enough to accurately sample cosmic variance during the EoH (Kaur, Gillet & Mesinger 2020; Balu et al. in preparation).

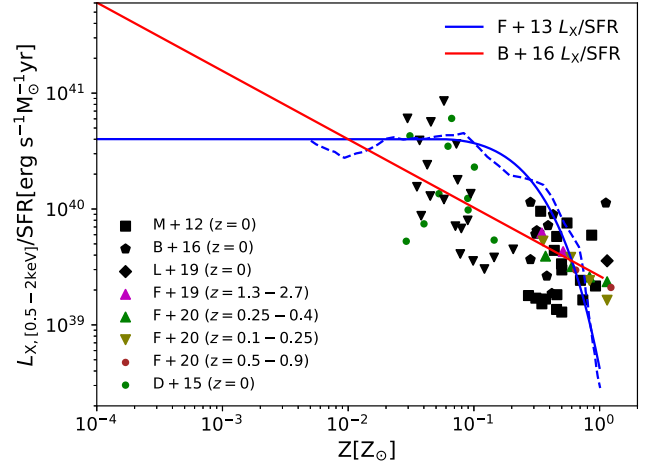
### 2.1 Metallicity-dependent X-ray luminosity to SFR relations

We assume that early galaxy populations can be characterized by an average  $L_X/\text{SFR}$  relation. This relation is often taken to be a constant when computing the EoH and associated 21-cm signal (though see

<sup>3</sup>The 21-cm signal depends on metallicity also through its impact on the stellar population (e.g. Magg et al. 2021; Muñoz et al. 2022), as well as dust attenuation of UV photons (e.g. Mirocha & Furlanetto 2019). Here we focus on its role in setting the X-ray luminosity to SFR relation, deferring more focused studies to future work.

<sup>4</sup>Throughout we use ‘metallicity’ to refer to the gas-phase metallicity,  $12 + \log_{10}(\text{O}/\text{H})$ . Solar gas-phase metallicity ( $12 + \log_{10}(\text{O}/\text{H})_\odot$ ) is taken to be 8.69 (Asplund et al. 2004). Absolute metallicity,  $Z_\odot$  is taken to be 0.02.

<sup>5</sup><https://github.com/21cmfast/21cmFAST>



**Figure 1.** Soft-band  $L_X/\text{SFR}$  versus  $Z$ . The F + 13 curve (solid blue) is constructed by fitting the XRB synthesis model 265 (dashed blue) of Fragos et al. 2013. The red line is the empirical fit of observation data of local Lyman break analogs (black pentagons) from Brorby et al. 2016. To compare these models with observations, we also plot some observational estimates from the literature. Black squares are  $L_X/\text{SFR}$  estimates of  $z = 0$  HMXB dominant star-forming galaxies from Mineo et al. (2012), constructed using X-ray, infrared, and UV data from *Chandra*, *SPITZER*, *GALEX*, and 2MASS archives. Their corresponding metallicities were taken from Douna et al. (2015). Black diamonds are resolved samples of galaxies observed in both infrared (SINGS survey) and *Chandra* X-ray from Lehmer et al. (2019). Green circles correspond are from Douna et al. (2015), estimated primarily from blue compact dwarf galaxies, with upper limits denoted by upside-down triangles in black. Stacked *Chandra* data from the COSMOS Legacy survey (Fornasini et al. 2019; Fornasini, Civano & Suh 2020) in the redshift range  $z = 0.1-2.7$  is shown with corresponding triangles. All data points are converted into the soft-band.

e.g. Madau & Fragos 2017; Eide et al. 2018). Here, we explore how the 21-cm signal is impacted by the metallicity dependence of this relation, assuming two very different  $L_X/\text{SFR}-Z$  scalings.

The first is taken from Brorby et al. (2016) (hereafter, B + 16), who studied local, metal-poor, star-forming galaxies obtaining:

$$\log_{10} \left( \frac{L_{X,[0.5-8 \text{ keV}]}}{\text{erg s}^{-1}} \right) = a \log_{10} \left( \frac{\text{SFR}}{M_\odot \text{ yr}^{-1}} \right) + b \log_{10} \left( \frac{(\text{O}/\text{H})}{(\text{O}/\text{H})_\odot} \right) + c, \quad (1)$$

where  $L_{X,[0.5-8 \text{ keV}]}$  is the X-ray luminosity in 0.5–8 keV range,  $a = 1.03 \pm 0.06$ ,  $b = -0.64 \pm 0.17$ ,  $c = 39.46 \pm 0.11$ .

The second  $L_X/\text{SFR}-Z$  relation is taken from stellar evolution models of HMXBs (Fragos et al. 2013). Specifically, we use model 265 (hereafter, F + 13), which is also the maximum likelihood model in Madau & Fragos (2017). We find that the following Schechter function provides a good fit for its  $L_X/\text{SFR}-Z$  relation:

$$\log_{10} \left( \frac{L_{\text{bol}}/\text{SFR}}{\text{erg s}^{-1} M_\odot^{-1} \text{ yr}} \right) = A + \alpha \log_{10} (Z/Z_{\text{turn}}) - Z/Z_{\text{turn}}. \quad (2)$$

Here,  $L_{\text{bol}}/\text{SFR}$  is the bolometric X-ray luminosity per SFR in  $\text{erg s}^{-1} M_\odot^{-1} \text{ yr}$ ,  $A = 41.3$ ,  $\alpha = 0.3$ , and  $Z_{\text{turn}} = 8 \times 10^{-3}$ . We also set the X-ray luminosity to constant below the peak value at  $Z/Z_{\text{turn}} \leq \alpha/\ln 10$ , to prevent the empirical fit from decreasing towards the lowest metallicities, in agreement with the constant  $L_X/\text{SFR}$  predicted by the F + 13 model at these metallicities.

We show the B+16 and F + 13 relations in Fig. 1, together with some observational data (see caption for details). Since only soft X-

rays are efficient in heating the IGM, we convert all luminosities into the soft-band (0.5–2 keV), using the two component plus a power-law model from Mineo et al. (2012). We note that observations based on individual galaxies show significant scatter, especially at low SFRs (roughly corresponding to low metallicities). This is due to sparse sampling of the relatively broad luminosity functions (LFs) of X-ray binaries, which require  $\sim$  tens of galaxies to get unbiased estimates of the population-average values of  $L_X/\text{SFR}$  (e.g. Lehmer et al. 2021).

From the figure, we see that both B+16 and F + 13  $L_X/\text{SFR}-Z$  relations pass through the observational data; however, they show dramatically different trends towards very low metallicities. As we shall see further below, galaxies with  $Z \lesssim 10^{-2}Z_\odot$  likely dominate the X-ray background during the EoH. In this regime, the model from F + 13 asymptotes to a much lower value of  $L_X/\text{SFR}$  than the empirical relation from B + 16. The theoretical explanation for this is that once metallicity goes below a certain value (roughly  $0.05 Z_\odot$ ), stellar winds become inefficient and only have a marginal impact on the stellar evolution (e.g. Fragos et al. 2013). However, these models do not vary the initial mass function (IMF), which could have an additional impact on the HMXB luminosity at very low metallicities (c.f. fig. 7 in F + 13). Moreover, we note that the B + 16  $L_X/\text{SFR}-Z$  relation is derived empirically, and was not intended to be extrapolated to the lowest metallicities we show in the figure.

Therefore, the  $L_X/\text{SFR}-Z$  relation is highly uncertain at the lowest metallicities. It is important to note, however, that here we are not arguing for the validity of any specific  $L_X/\text{SFR}-Z$  relation. Instead, we are exploring the observational impact of  $L_X/\text{SFR}-Z$  relations during the Cosmic Dawn. Therefore, the B + 16 and F + 13 scalings are useful in that they assign very different properties to the first generations of HMXBs. Indeed, we would want to eventually *infer these relations directly from future 21-cm data*.

## 2.2 Mass–metallicity relation

In the previous section, we presented the two different  $L_X/\text{SFR}-Z$  relations we use in this study. We now discuss our adopted prescription for assigning  $Z$  to high-redshift galaxies.

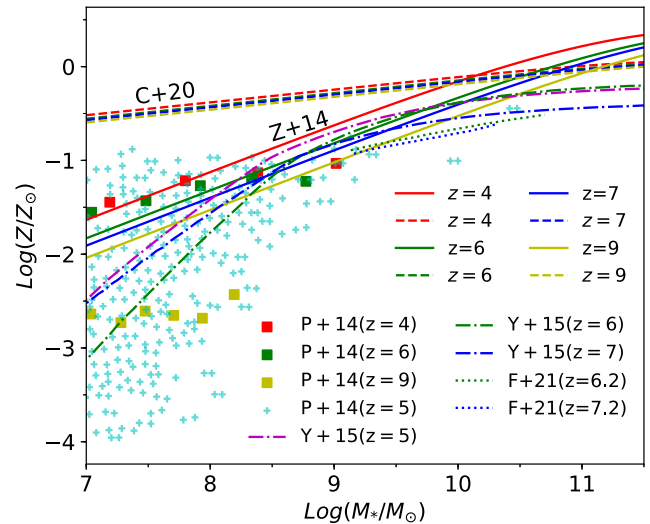
The galaxy-averaged gas-phase metallicity shows a strong correlation with the galaxy’s stellar mass ( $M_*$ ), commonly referred to as the mass-metallicity relation (MZR; Lara-López et al. 2010; Mannucci et al. 2010; Yates, Kauffmann & Guo 2012; Zahid et al. 2014; Cresci, Mannucci & Curti 2019; Curti et al. 2020).<sup>6</sup> It has been observed in the local Universe and even at redshifts as high as  $z = 3.5$  (e.g. Tremonti et al. 2004; Erb et al. 2006; Ellison et al. 2008; Maiolino et al. 2008; Mannucci et al. 2010; Zahid et al. 2014; Sanders et al. 2018; Curti et al. 2020; Sanders et al. 2021), and has also been found in theoretical works (e.g. Yates et al. 2012; Ucci et al. 2021).

Here we adopt the empirical MZR from Zahid et al. (2014):

$$Z = Z_0 + \log_{10}(1 - e^{-(M_*/M_0)^\gamma}), \quad (3)$$

where  $Z_0$  is the saturation metallicity,  $M_0$  is a characteristic stellar mass scale above which the metallicity asymptotically approaches  $Z_0$ . The relation reduces to a power law with index  $\gamma$  at  $M_* < M_0$ , and flattens at higher masses. The redshift dependence is

<sup>6</sup>The mass–metallicity relation can also be understood as the 2D projection of the more general 3D fundamental metallicity relation (FMR), relating metallicity to the stellar mass and SFR (e.g. Mannucci et al. 2010; Hunt et al. 2012). In our analytic model, the stellar mass and SFR are themselves related deterministically (i.e. through the main sequence of star-forming galaxies), thus it suffices to only specify one or the other.



**Figure 2.** Gas-phase metallicity versus stellar mass relation (MZR) from Zahid et al. (2014) (equation 3) is shown with solid curves ( $Z + 14$ ) and from Curti et al. (2020) with dashed curves ( $C + 14$ ). The curves are colour-coded according to redshift. We also plot mean relations from the hydrodynamic simulations of Pallottini et al. (2014) ( $P + 14$ , squares), Yue et al. (2015) ( $Y + 15$ , dot-dashed curves), and semi-analytic models of Fontanot et al. (2021) ( $F + 21$ , dotted curves). Cyan crosses denote values from individual galaxies in Pallottini et al. (2014) at  $z = 5$ , illustrating the galaxy-to-galaxy scatter.

incorporated in  $M_0 \equiv 10^b$ . By fitting to data from the Sloan Digital Sky Survey (SDSS), the Deep Extragalactic Evolutionary Probe 2 (DEEP2), FMOS-COSMOS, and Smithsonian Hectospec Lensing Survey (SHELS), that span a redshift range up to 1.6, Zahid et al. (2014) find a set of best-fitting parameters as follows:  $Z_0 = 9.100$ ,  $b = 9.135 + 2.64 \log_{10}(1 + z)$ , and  $\gamma = 0.522$ .

We illustrate this relation in Fig. 2 at  $z = 4-9$  using solid curves. For reference, we also show data from the cosmological, hydrodynamic simulations of Pallottini et al. (2014), in which they follow interstellar medium (ISM) and IGM metal enrichment.  $Y + 15$  curves are from Yue et al. (2015) where they use the stellar mass to UV magnitude relation from Duncan et al. (2014), combined with the FMR from Mannucci et al. (2010). We also show  $Z-M_*$  predictions from semi-analytical GAEA models of Fontanot et al. (2021), tested using the VANDELS survey (McLure et al. 2018; Pentericci et al. 2018). Additionally, we show the empirical MZR from Curti et al. (2020), which depends on both stellar mass and SFR. This relation is more similar to the FMR (Mannucci et al. 2010), which is usually regarded as redshift-independent. In making this comparison figure, we used the stellar – halo mass and the SFR – stellar mass relations from Park et al. (2019) when required; these were calibrated to reproduce high-redshift UV LFs and other EoR data (see Section 2.3 for more details).

From Fig. 2 we see that the high-redshift simulations and semi-analytic models are in reasonable agreement with the MZR from Zahid et al. (2014), though the evolution towards the highest redshifts and smallest masses is highly uncertain. Therefore, the 21-cm forecasts we present in Section 4 should be taken with caution.

## 2.3 X-ray emissivity during the Cosmic Dawn

The  $L_X/\text{SFR}-Z$  relations and MZR discussed in the previous sections allow us to compute an average X-ray luminosity for a galaxy, given its SFR and stellar mass. To compute the X-ray background,

we then need to relate the SFR and stellar mass to the typical mass of the host halo, since then we can integrate over well-established halo mass functions (HMFs). Here, we use the simple and flexible power-law scaling relations from Park et al. (2019), shown to recover various reionization era observables. Specifically, we assume a power law for the stellar-to-halo mass relation:

$$\frac{M_*}{M_h} = f_{*,10} \left( \frac{M_h}{10^{10} M_\odot} \right)^{\alpha_*} \left( \frac{\Omega_b}{\Omega_m} \right), \quad (4)$$

where  $f_*(M_h) = f_{*,10}(M_h/10^{10}M_\odot)^{\alpha_*}$  is the fraction of the galactic gas in stars and  $\alpha_*$  is the power-law index. Similarly, we relate the fraction of UV ionizing photons which escape into the IGM,  $f_{\text{esc}}$  to halo mass with  $f_{\text{esc}}(M_h) = f_{\text{esc},10}(M_h/10^{10}M_\odot)^{\alpha_{\text{esc}}}$ . The SFR can be expressed as

$$\text{SFR} = \dot{M}_* = \frac{M_*}{t_* H(z)^{-1}}, \quad (5)$$

where  $t_*$  is a typical star formation time-scale as a fraction of the Hubble time,  $H(z)^{-1}$ , with  $H(z)$  being the Hubble parameter at  $z$ . These simple scaling relations seem sufficient to characterize the population-averaged properties of high- $z$  galaxies. They are consistent with UV LFs at  $z = 6\text{--}10$  and  $M_{\text{UV}} > -20$  (e.g. Bouwens et al. 2014, 2016; Finkelstein 2016; McLeod, McLure & Dunlop 2016; Livermore, Finkelstein & Lotz 2017; Atek et al. 2018; Ishigaki et al. 2018; Oesch et al. 2018; Gillet, Mesinger & Park 2020; Bouwens et al. 2021).<sup>7</sup> Moreover, both hydrodynamical simulations (e.g. Xu et al. 2016; Ma et al. 2020) and semi-analytic models (e.g. Mutch et al. 2016; Sun & Furlanetto 2016; Behroozi et al. 2019) also imply power laws for the stellar-to-halo mass relation.

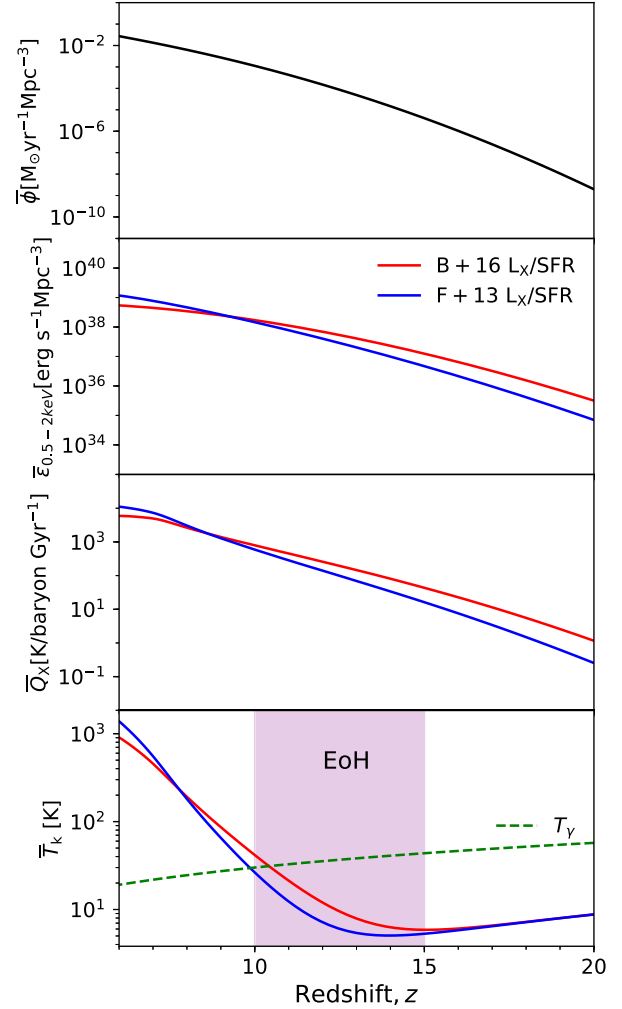
With the above, we can express the comoving, soft-band X-ray emissivity (in  $\text{erg s}^{-1} \text{Mpc}^{-3}$ ) at position  $\mathbf{x}$  and redshift  $z$  as

$$\epsilon_{0.5\text{--}2\text{keV}}(\mathbf{x}, z) = \int_0^\infty dM_h \frac{dn}{dM_h} e^{-\left(\frac{M_{\text{turn}}}{M_h}\right)} \dot{M}_* \frac{L_X}{\text{SFR}}. \quad (6)$$

Here,  $dn/dM_h$  is the local (conditional) halo mass function,  $e^{-(M_{\text{turn}}/M_h)}$  accounts for the suppression in star formation in haloes below a characteristic mass scale (i.e.  $M_{\text{turn}}$ ) due to inefficient cooling and/or feedback (e.g. Hui & Gnedin 1997; Barkana & Loeb 2001; Springel & Hernquist 2003; Mesinger & Dijkstra 2008), and  $L_X/\text{SFR}(Z, M_*, z)$  is related to the halo mass through equations (1)–(5). When computing mock observations below, we use these fiducial values:  $f_{*,10} = 0.05$ ,  $\alpha_* = 0.5$ ,  $f_{\text{esc},10} = 0.1$ ,  $\alpha_{\text{esc}} = -0.5$ ,  $M_{\text{turn}} = 5 \times 10^8 M_\odot$ , and  $t_* = 0.5$ . These values correspond to the maximum a posteriori model in Park et al. (2019), which used UV LFs, the electron scattering optical depth to CMB, and the QSO forest dark fraction in the likelihood.

In the top panel of Fig. 3, we show our fiducial star formation rate density (SFRD) evolution, with the corresponding X-ray emissivity shown in the second panel (for both B + 16 and F + 13  $L_X/\text{SFR}-Z$  relations). Because structure formation is hierarchical, the mass and metallicity of the typical galaxy population increase with time. At early times, most galaxies were extremely metal-poor and so the B + 16  $L_X/\text{SFR}-Z$  relation implies an X-ray emissivity that is a factor of  $\sim 5$  larger than the one from F + 13 at  $z \sim 20$ . When

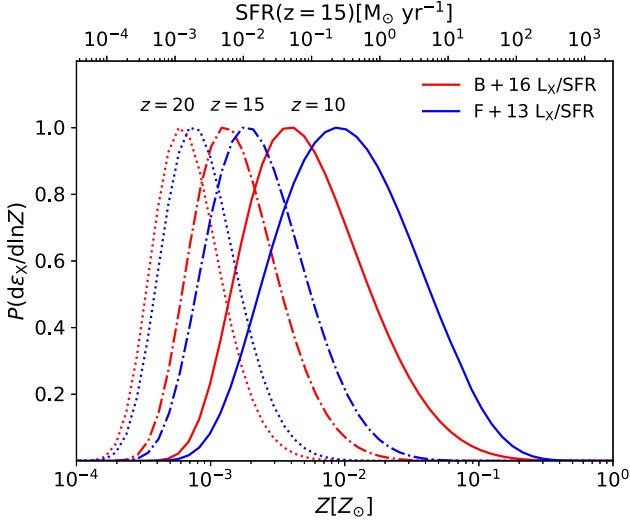
<sup>7</sup>Here, we are only interested in the faint-end galaxies that dominate the photon budget (e.g. Qin et al. 2021). We therefore do not include a separate power law at the bright end, typically associated with AGN feedback (e.g. Furlanetto 2006; Behroozi & Silk 2015; Mirocha, Furlanetto & Sun 2017; Rudakovskiy et al. 2021; Sabti, Muñoz & Blas 2022). Such bright galaxies sit on the exponential tail of the HMF, and are thus too rare to be important for determining cosmic radiation fields.



**Figure 3.** From top to bottom: Redshift evolution of the (volume-averaged) star formation rate density (SFRD), soft-band X-ray emissivity (equation 6), X-ray heating rate per baryon,  $\bar{Q}_X$  (equation 10), and kinetic temperature of the neutral IGM,  $\bar{T}_k$  (equation 9) computed using 21cmFAST. Red and blue curves correspond to the  $L_X/\text{SFR}-Z$  relations from B + 16 and F + 13, respectively. In the bottom panel, we also show the cosmic microwave background (CMB) temperature ( $T_\gamma$ ) evolution with the dashed line, and approximately demarcate the Epoch of Heating (EoH).

the characteristic metallicity of star-forming galaxies surpasses  $Z \gtrsim 0.01Z_\odot$ , the F+13 relation implies higher X-ray luminosities than B + 16 (see Fig. 1). From the second panel of Fig. 3, we see this transition happening at  $z \sim 10$ , with F + 13 implying a higher emissivity over the range  $6 < z < 10$ . As can be seen from Fig. 1, we expect this trend to reverse again, with B + 16 having higher luminosities at  $Z \gtrsim Z_\odot$ . However, such highly metal-enriched galaxies are too massive and rare to be relevant during the Cosmic Dawn.

We further quantify properties of the relevant galaxies in Fig. 4, where we plot the fractional contribution to the X-ray emissivity of galaxies within a given logarithmic metallicity bin. The probability distribution functions (PDFs) shift to higher metallicities and become wider with decreasing redshift. This is driven by the evolution of the HMF which shifts to larger masses and flattens with time. We see explicitly that the transition when the F+13 emissivity surpasses that of B + 16 at  $z \lesssim 10$  corresponds to when the mean of the PDFs goes above  $Z \sim 0.01Z_\odot$ . As we discuss further below, the EoH for these



**Figure 4.** Fractional contribution to the total X-ray emissivity sourced by galaxies per logarithmic metallicity bin. B + 16 and F + 13  $L_X/\text{SFR}$ – $Z$  relations are shown with red and blue curves, respectively. Here, different line-styles represent different redshifts. On the top  $x$ -axis, we denote the corresponding SFR at  $z = 15$ . For viewing purposes, we normalize all PDFs to peak at unity.

models corresponds to  $10 \lesssim z \lesssim 15$ . As a result, the EoH is driven by galaxies with metallicities  $10^{-3} \lesssim Z/Z_\odot \lesssim 10^{-2}$  and SFRs roughly in the range of  $10^{-3} \lesssim \dot{M}_*/M_\odot \text{yr}^{-1} \lesssim 10^{-1}$  (c.f. the top axis shows the corresponding SFRs implied by our model at  $z = 15$ ).

### 3 EVOLUTION OF IGM PROPERTIES

#### 3.1 Temperature

We now compute the global evolution of the thermal and ionization state of the IGM, for the two X-ray emissivities shown in the previous section. We can write the X-ray specific intensity,  $J$  (in units of  $\text{erg s}^{-1} \text{keV}^{-1} \text{cm}^{-2} \text{sr}^{-1}$ ) for photons with energy  $E_X$ , as seen by a gas element at position  $\mathbf{x}$  and redshift  $z$ , as

$$J(\mathbf{x}, z, E_X) = \frac{(1+z)^3}{4\pi} \int_z^\infty dz' \frac{cdt}{dz'} \epsilon_\nu e^{-\tau}. \quad (7)$$

Here  $\epsilon_\nu(\mathbf{x}, z')$  is the *specific* X-ray emissivity, evaluated in the rest frame  $E_0 = E_X(1+z)/(1+z)$ , assuming a power-law spectral energy distribution (SED) with energy index  $\alpha_X = 1$  and integrated soft-band value according to equation (6). The term  $\exp[-\tau(z, z', E_X)]$  accounts for attenuation by hydrogen and helium through a two-phased IGM between  $z$  and  $z'$  [see Mesinger et al. (2011) for more details].

This X-ray intensity drives the evolution of the ionized fraction ( $x_e$ ) and kinetic temperature ( $T_k$ ) of the (mostly neutral) IGM gas outside the H II regions proximate to galaxies according to

$$\frac{dx_e(\mathbf{x}, z)}{dz} = \frac{dt}{dz} (\Gamma_X - \alpha_A C x_e^2 n_b f_H) \quad (8)$$

and

$$\frac{dT_k(\mathbf{x}, z)}{dz} = \frac{2}{3k_b(1+x_e)} \frac{dt}{dz} \sum Q_p + \frac{2T_k}{3n_b} \frac{dn_b}{dz} - \frac{T_k}{1+x_e} \frac{dx_e}{dz}, \quad (9)$$

where  $n_b$  is the baryon number density,  $\Gamma_X$  is the ionization rate per baryon from X-rays,  $\alpha_A$  is the case-A recombination coefficient,  $C = \langle n_H^2 \rangle / \langle n_H \rangle^2$  is the clumping factor on the scale of the simulation cell,  $n_H$  is the hydrogen number density,  $f_H$  is the hydrogen number fraction,  $k_b$  is the Boltzmann constant, and the radiative heating rate per baryon,  $Q_p$ , includes both Compton heating and X-ray heating. The X-ray heating and ionization rates can be expressed as

$$Q_X(\mathbf{x}, z) = \int d\nu \frac{4\pi J}{h\nu} \sum_i (h\nu - E_i^{\text{th}}) f_{\text{heat}} f_i x_i \sigma_i, \quad (10)$$

$$\Gamma_X(\mathbf{x}, z) = \int d\nu \frac{4\pi J}{h\nu} \sum_i f_i x_i \sigma_i F_i, \quad (11)$$

where

$$F_i = (h\nu - E_i^{\text{th}}) \left( \frac{f_{\text{ion,H I}}}{E_{\text{H I}}^{\text{th}}} + \frac{f_{\text{ion,He I}}}{E_{\text{He I}}^{\text{th}}} + \frac{f_{\text{ion,He II}}}{E_{\text{He II}}^{\text{th}}} \right) + 1. \quad (12)$$

Here,  $i$  stands for the atomic species: H, He I, and He II;  $E_i^{\text{th}}$  is their corresponding ionization threshold;  $f_i$  is their number fraction;  $x_i$  is the ionization fraction (which for H I and He I we take to be  $x_i = (1 - x_e)$  and for He II  $x_i = x_e$ );  $\sigma_i$  is the photoionization cross-section;  $f_{\text{heat}}$  is the fraction of the primary ionized electron's energy dissipating as heat; and  $f_{\text{ion},j}$  is its energy contributing to secondary ionization of the species  $j$ , taken from Furlanetto & Stoever (2010). Following the ISM simulations discussed in Das et al. (2017), we assume photons with energies below  $E_0 = 0.5 \text{ keV}$  are absorbed by the host galaxies and unable to escape into the IGM.

In the bottom two panels of Fig. 3, we show the volume-averaged  $Q_X$  and  $T_k$  corresponding to our two fiducial  $L_X/\text{SFR}$ – $Z$  relations. In the temperature panel, we also show the evolution of the CMB temperature that provides the radio background in standard 21-cm models. We roughly demarcate the EoH as shown in the figure.

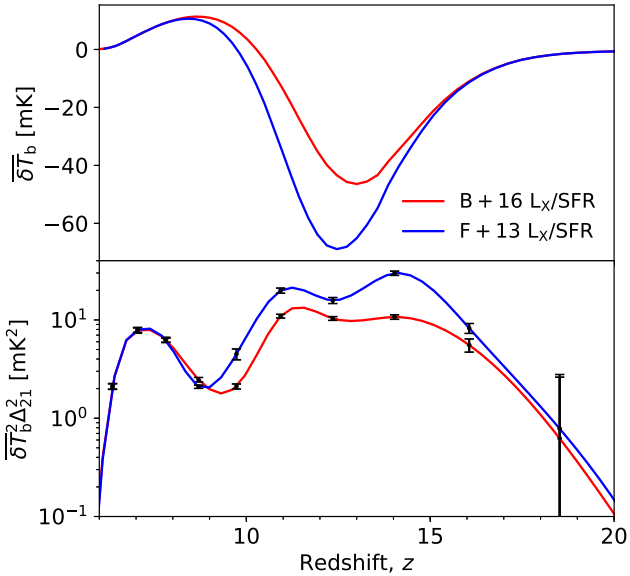
The heating rate and temperature follow the same qualitative trend seen in the emissivity panel, with B + 16 resulting in a higher temperature at redshifts  $z \gtrsim 8$ . The transition redshift at which F + 13 surpasses B + 16 is somewhat lower for temperature (bottom panel) compared to the emissivity (second panel). This is because the temperature depends on the light-cone integral over the emissivity, and not just its instantaneous value. The flattening of the heating rate seen at  $z \lesssim 8$  is due to reionization by UV photons, which decreases the neutral fractions of hydrogen and helium.

#### 3.2 21-cm signal

As mentioned in the Introduction section, we use the cosmic 21-cm signal as our fiducial observational data set, as it has the largest potential of constraining the thermal state of the gas in those early epochs. Indeed, the recent HERA season 1 observations (The HERA Collaboration et al. 2022a) already constrain the  $L_X/\text{SFR}$  of high- $z$  galaxies, disfavouring the local  $L_X/\text{SFR}$  relation by  $> 1\sigma$  (The HERA Collaboration et al. 2022b). As discussed above, this is consistent with expectations that the first sources were more X-ray efficient due to their lower metallicities.

The 21-cm signal can be expressed as the offset of the 21-cm brightness temperature ( $T_b$ ) from the CMB temperature ( $T_\gamma$ ) at the observed frequency  $\nu$  (Furlanetto 2006):

$$\begin{aligned} \delta T_b(\nu) &= \frac{T_s - T_\gamma}{1+z} (1 - e^{-\tau_{21}}) \\ &\approx 27 x_{\text{H I}} (1 + \delta) \left( \frac{H}{dv_r/dr + H} \right) \left( 1 - \frac{T_\gamma}{T_s} \right) \\ &\quad \times \left( \frac{1+z}{10} \frac{0.15}{\Omega_m h^2} \right)^{1/2} \left( \frac{\Omega_b h^2}{0.023} \right) \text{ mK}. \end{aligned} \quad (13)$$



**Figure 5.** Top panel: Redshift evolution of the global signal,  $\delta\bar{T}_b$  for B + 16 and F + 13  $L_X/\text{SFR}-Z$  relations shown as red and blue curves, respectively. Bottom panel: Corresponding redshift evolution of the 21-cm PS evaluated at  $k = 0.1 \text{ Mpc}^{-1}$ . Error bars denote the  $1\sigma$  uncertainty expected from a 1000 h observation assuming optimistic foregrounds with the Square Kilometer Array (see Section 4 for details).

Here,  $\tau_{21}$  is the optical depth at the 21-cm frequency,  $\delta$  is the gas over-density ( $\delta = \rho/\bar{\rho} - 1$ ),  $dv_r/dr$  is the peculiar velocity gradient along the line of sight, and  $T_S$  is the spin temperature defined by the relative abundances of the excited and ground states of the hyperfine level of neutral hydrogen atom. Although 21cmFAST solves for the exact expression in the first row of equation (13), we show the  $\tau_{21} \ll 1$  approximation in the second row for physical intuition.

In the top panel of Fig. 5, we plot the redshift evolution of the global 21-cm signal,  $\delta\bar{T}_b$ , corresponding to the B+16 and F + 13  $L_X/\text{SFR}-Z$  relations with solid red and blue lines, respectively.

Since the B + 16 scaling implies stronger X-ray heating during the EoH, its  $\delta\bar{T}_b$  has an earlier and shallower absorption trough as compared with F + 13. The differences between the global signal evolution implied by the two  $L_X/\text{SFR}-Z$  relations peak during the epoch of heating with a 25 mK difference in the depths of the absorption troughs.

In the bottom panel of Fig. 5, we plot the redshift evolution of the 21-cm power spectrum (PS) amplitude, evaluated at  $k = 0.1 \text{ Mpc}^{-1}$ . We focus on this wavemode, as it corresponds to a ‘sweet spot’ for 21-cm interferometers: large enough to mitigate foreground contamination yet small enough to have low thermal noise (e.g. Mertens et al. 2020; Trott et al. 2020; The HERA Collaboration et al. 2022a). The redshift evolution of the large-scale 21-cm power shows the characteristic three peaks, corresponding to the major astrophysical epochs: Wouthuysen-Field (Wouthuysen 1952; Field 1958; WF) coupling, EoH, and Epoch of Reionization (EoR). We also show the expected thermal and cosmic variance noise for a 1000h observation, assuming optimistic foregrounds from Pober et al. (2014), with the Square Kilometer Array (SKA; see the following section for details on how these uncertainties are calculated).

Comparing the two solid curves, we see that the B + 16 scaling results in an earlier EoH. This increases the overlap between the EoH and the epoch of WF coupling: IGM in regions with an underabundance of galaxies still has different spin and kinetic

temperatures at the same time that the IGM in regions with an overabundance of galaxies becomes hot (e.g. Mesinger, Ferrara & Spiegel 2013). This decreases the temperature contrast during the Cosmic Dawn, resulting in lower PS amplitudes by up to a factor of  $\sim 3$  for the B + 16 scaling, compared with F + 13. We note that these two scenarios can be distinguished with a 1000 h observation with SKA1-low.

#### 4 CAN CONSTANT X-RAY LUMINOSITY – SFR MODELS REPRODUCE THE METALLICITY-DEPENDENT SIGNAL?

Upcoming 21-cm observations will provide a physics-rich data set, allowing us to *infer* galaxy properties from the data directly. In particular, the signal is very sensitive to the X-ray properties of the first galaxies (e.g. Kern et al. 2017; Park et al. 2019), suggesting that the data could tell us the correct  $L_X/\text{SFR}-Z$  and MZR scalings.

However, performing inference requires parametrizing these relations and adopting physically motivated priors on the corresponding parameters. This is challenging, as our current understanding of stellar evolution and IMFs is insufficient to motivate such parametrizations and priors. In principle, these difficulties can be mitigated with Bayesian model selection (e.g. Binne & Pritchard 2019; Qin et al. 2020b), which we will explore in a follow-up work. Nevertheless, having simpler models of  $L_X/\text{SFR}$  would make inference easier and quicker.

Indeed, most current models of the X-ray background during the Cosmic Dawn assume a constant  $L_X/\text{SFR}$  (e.g. Ghara et al. 2020; Mondal et al. 2020; Greig et al. 2021a,b; The HERA Collaboration et al. 2022b; though see Madau & Fragos 2017; Eide et al. 2018), motivated by the theoretical argument that the impact of metallicity-driven winds becomes negligible below  $Z \lesssim 0.05Z_\odot$  (e.g. Fragos et al. 2013). In this section, we quantify whether this simplification can recover the evolution of IGM properties of our metallicity-dependent models presented in the previous section. We make mock 21-cm PS observations using both B + 16 and F + 13  $L_X/\text{SFR}-Z$  relations, and then perform inference assuming a constant  $L_X/\text{SFR}$ . As the figure of merit, we compare the recovered 21-cm power spectra and the X-ray heating rates to the ‘truth’ from the mock observations.

Our two mock observations correspond to the fiducial B + 16 and F + 13 models discussed in the previous section. To compute the thermal noise, we use the PYTHON module 21cmSENSE<sup>8</sup> (Pober et al. 2013, 2014) and assume a 1000 h integration (6 h per night) with the SKA1 low.<sup>9</sup> In order to tightly constrain X-ray parameters and *maximize the importance of the metallicity dependence*, we use the ‘Optimistic foreground’ flag in 21cmSENSE. This setting assumes that the foreground wedge extends only up to the full-width half-max of the primary beam; modes outside of this contaminated wedge are assumed to be free of systematics. Our mock observations span the redshift range  $z = 5.8-21.6$  and wavemode range  $k = 0.1-1 \text{ Mpc}^{-1}$ .

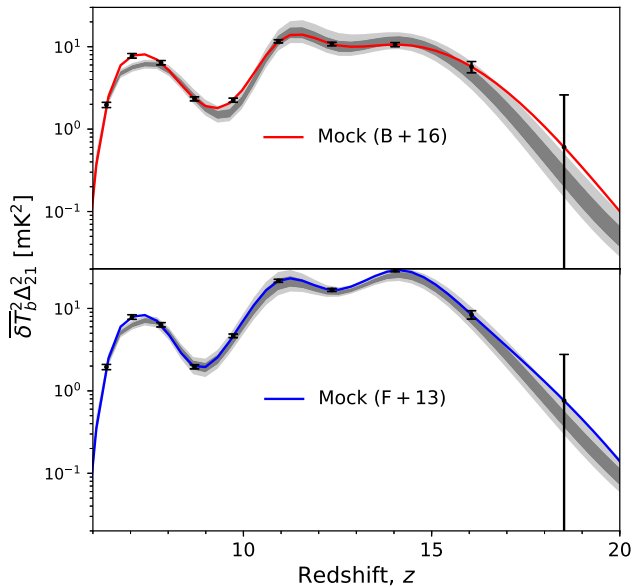
For performing inference using these mock observations, we use the public 21cmMC module<sup>10</sup> (Greig & Mesinger 2015, 2017, 2018). 21cmMC is a Bayesian sampler of 21cmFAST, which forward-models 21-cm light-cones. Here, we use the Multinest sampler<sup>11</sup> (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009), included in 21cmMC by Qin et al. (2020a).

<sup>8</sup><https://github.com/jpober/21cmSense>

<sup>9</sup><https://astronomers.skatelescope.org/>

<sup>10</sup><https://github.com/21cmfast/21cmMC>

<sup>11</sup><https://github.com/rjw57/MultiNest>



**Figure 6.** Redshift evolution of 21-cm power spectra evaluated at  $k = 0.1 \text{ Mpc}^{-1}$ . The two mock observations are shown in red (top panel; B + 16) and blue (bottom panel; F + 13), with  $1\sigma$  noise (thermal + cosmic variance) denoted with error bars (see the text for details). The dark and light shaded regions correspond to 16–84 per cent and 2–98 per cent credible intervals (C.I.), respectively, obtained assuming a model with a constant (i.e. metallicity independent)  $L_X/\text{SFR}$ .

In addition to the mock 21-cm signal, we also use the following current observations in the likelihood: (i)  $z \geq 6$  UV luminosity functions (Finkelstein 2016; Ishigaki et al. 2018; Bouwens et al. 2021); (ii) the electron scattering optical depth to CMB; Planck Collaboration XLVII et al. 2016b); and (iii) the dark fraction in the spectra of high- $z$  quasars (McGreer, Mesinger & D’Odorico 2015). We sample the following astrophysical parameters, adopting flat priors over the quoted ranges:  $\log_{10} f_{*10} \in [-3, 0]$ ,  $\log_{10} f_{\text{esc},10} \in [-3, 0]$ ,  $\alpha_* \in [-0.5, 1]$ ,  $\alpha_{\text{esc}} \in [-1, 0.5]$ ,  $\log_{10}(M_{\text{turn}}/M_\odot) \in [8, 10]$ ,  $\log_{10}(L_X/\text{SFR}/\text{erg s}^{-1}M_\odot^{-1}\text{yr}) \in [38, 44]$ , and  $E_0 \in [0.1-1.5]$  keV.

In Fig. 6, we show the recovered posteriors on the evolution of large-scale ( $k = 0.1 \text{ Mpc}^{-1}$ ) 21-cm PS. Points with error bars correspond to the mock observations assuming B + 16 (top) and F + 13 (bottom) scalings, while the shaded regions denote the posteriors obtained assuming a constant  $L_X/\text{SFR}$ . In the Appendix, we also show the PS posteriors as functions of wavenodes for all of the redshift samples we use in the mock observations.

The simple, constant  $L_X/\text{SFR}$  models are able to recover the PS evolution of the more sophisticated B + 16 and F + 13 models quite well. The posteriors are consistent with the mock data at  $1\sigma$  throughout the EoH.<sup>12</sup>

Interestingly, the largest discrepancies are found during the EoR, where the constant  $L_X/\text{SFR}$  models underestimate the B + 16 21-cm PS by up to tens of per cent. This is because most of the  $L_X/\text{SFR}$  constraining power comes from the EoH. Since the  $L_X/\text{SFR}-Z$  scaling is very steep, the EoH galaxies at  $10 \lesssim z \lesssim 15$  are considerably more efficient at emitting X-rays compared to EoR galaxies at  $5 \lesssim z \lesssim$

10. By fitting to the EoH, the constant  $L_X/\text{SFR}$  models thus end up overpredicting the X-ray background during the EoR. In the case of B + 16, this results in the neutral IGM patches being partially ionized by X-rays with long mean free paths. This decreases the 21-cm contrast between ionized and neutral regions during the patchy EoR, resulting in a smaller PS amplitude.

This is further illustrated in Fig. 7, where we plot the analogous recovery of the X-ray heating rate (left panels) and the EoR history (right-hand panels). Constant  $L_X/\text{SFR}$  models have a steeper redshift evolution of the heating rate. The inferred value of  $L_X/\text{SFR}$  corresponds to that of galaxies during the EoH at  $10 \lesssim z \lesssim 15$ , when the 21-cm signal is most sensitive to the X-ray background. However, towards the end of the EoR at  $z \sim 6$ , the inferred heating rate from constant  $L_X/\text{SFR}$  models can overestimate that from the mock observations by a factors of  $\sim \text{few} - 10$ .

This overestimate of the X-ray background also means that the EoR begins earlier in constant  $L_X/\text{SFR}$  models (see the right-hand panels of Fig. 7), driven by a larger contribution of X-rays to reionization. The earlier onset of the EoR in constant  $L_X/\text{SFR}$  models can be partially compensated by increasing the UV ionizing emission of smaller haloes. In our model, this is mostly achieved through the  $\alpha_{\text{esc}}$  parameter, which we find is indeed the only UV parameter whose marginal PDFs are not consistent with the value of the mock observation at more than  $2\sigma$  (see the Appendix). We will further quantify the recovery of galaxy properties using different parametric relations for both  $L_X/\text{SFR}$  and MZR in future work.

Our results imply that constant  $L_X/\text{SFR}$  models provide a reasonable simplification for modelling the 21-cm signal during the EoH. However, they overpredict the X-ray background at lower redshifts ( $z < 10$ ). Tying observations over large redshift intervals without properly accounting for population-evolution (e.g. metallicity) could thus introduce significant errors.

## 5 CONCLUSIONS

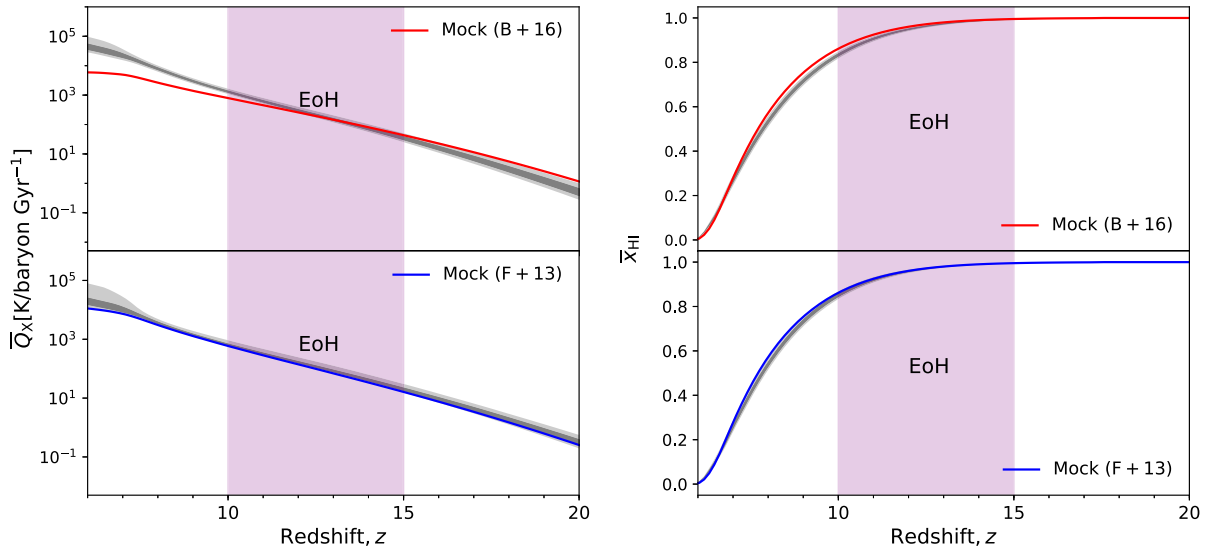
The redshifted 21-cm signal is a promising probe of heating and ionization processes in the early Universe. In particular, we expect the heating of the IGM prior to reionization to be dominated by HMXBs, hosted by galaxies too faint to be observed directly. We expect these early galaxies to be metal poor, likely implying a more X-ray luminous HMXB population than observed in local galaxies (e.g. Linden et al. 2010; Fragos et al. 2013; Basu-Zych et al. 2012; Brorby et al. 2016; Lehmer et al. 2021).

We adopt two very different  $L_X/\text{SFR}-Z$  relations: (i) an empirical power law from Brorby et al. (2016); and (ii) a theoretical stellar evolution model from Fragos et al. (2013). For the same galaxy evolution model, (i) and (ii) result in differences of up to a factor of  $\sim 2-3$  in the IGM temperature and 21-cm signal during the Cosmic Dawn.

HMXBs hosted by galaxies with SFRs of order  $10^{-3}-10^{-1} M_\odot \text{ yr}^{-1}$  and metallicities of order  $0.001-0.01 Z_\odot$  dominate the IGM heating. The assumed  $L_X/\text{SFR}-Z$  relation can shift these ranges by factors of  $\sim \text{few}$ .

We also use the two  $L_X/\text{SFR}-Z$  relations to compute mock 21-cm PS observations, assuming optimistic foreground removal and 1000 h integration with SKA. We performed inference using the common simplification of a constant  $L_X/\text{SFR}$ . The constant  $L_X/\text{SFR}$  models reproduce the IGM properties from the metallicity-dependent  $L_X/\text{SFR}$  simulations quite well. However, since the inferred value of  $L_X/\text{SFR}$  corresponds to the dominant population during the EoH ( $z \sim 10-15$ ; when the 21-cm signal is most sensitive to the IGM temperature), they overpredict the XRB at lower redshifts ( $z \lesssim$

<sup>12</sup>We remind the reader that our choice of ‘Optimistic Foregrounds’ results in the small noise errors seen in this figure and in the Appendix. As a result the recovered posteriors are very narrow. We used the ‘Optimistic Foregrounds’ setting of 21cmSense precisely to show the maximum expected bias from ignoring the metallicity dependence of  $L_X/\text{SFR}$ .



**Figure 7.** Same as Fig. 6, but here showing the volume-weighted X-ray heating rate per baryon (left-hand panels) and neutral hydrogen fraction (right-hand panels).

10). Thus, accurate inference over a broad range of redshifts should account for metallicity evolution and the  $L_X/\text{SFR}-Z$  relation.

## ACKNOWLEDGEMENTS

We gratefully acknowledge computational resources of the Center for High Performance Computing (CHPC) at Scuola Normale Superiore (SNS). YQ acknowledges that part of this work was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project #CE170100013; and some of the simulations presented in this work were run on the OzSTAR national facility at Swinburne University of Technology. AP acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. TF acknowledges support from the Swiss National Science Foundation Professorship grant (project number PP00P2\_176868). AB acknowledges support by NASA under award number 80GSFC21M0002.

## 6 DATA AVAILABILITY

The data generated during this work is available from the corresponding author on request.

## REFERENCES

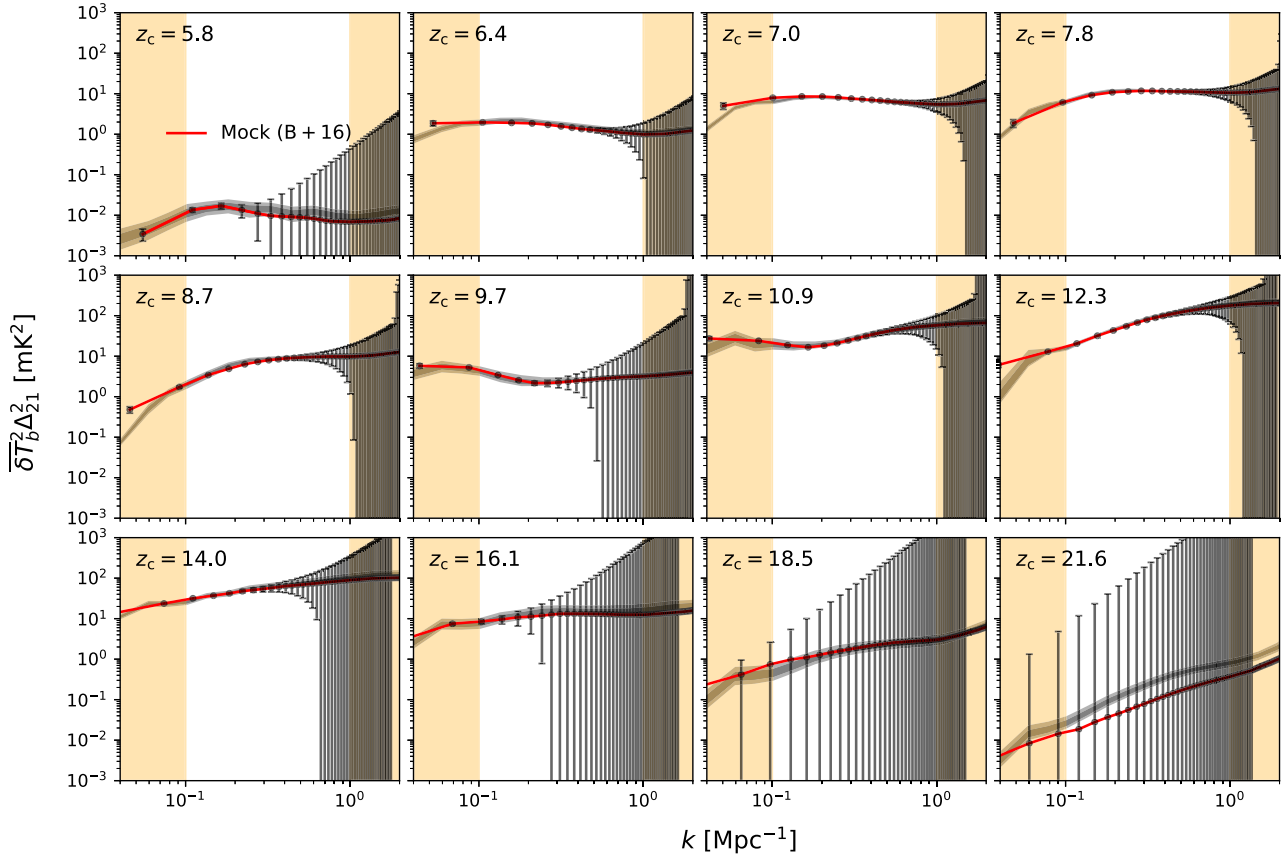
Asplund M., Grevesse N., Sauval A. J., Allende Prieto C., Kiselman D., 2004, *A&A*, 417, 751  
 Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, *MNRAS*, 479, 5184  
 Barkana R., Loeb A., 2001, *Phys. Rep.*, 349, 125  
 Basu-Zych A. R. et al., 2012, *ApJ*, 762, 45  
 Behroozi P. S., Silk J., 2015, *ApJ*, 799, 32  
 Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143  
 Binnie T., Pritchard J. R., 2019, *MNRAS*, 487, 1160  
 Bouwens R. J. et al., 2014, *ApJ*, 793, 115  
 Bouwens R. J. et al., 2016, *ApJ*, 833, 72  
 Bouwens R. J., Illingworth G. D., van Dokkum P. G., Ribeiro B., Oesch P. A., Stefanon M., 2021, *AJ*, 162, 255  
 Brorby M., Kaaret P., Prestwich A., Mirabel I. F., 2016, *MNRAS*, 457, 4081  
 Cresci G., Mannucci F., Curti M., 2019, *A&A*, 627, A42

Curti M., Mannucci F., Cresci G., Maiolino R., 2020, *MNRAS*, 491, 944  
 Das A., Mesinger A., Pallottini A., Ferrara A., Wise J. H., 2017, *MNRAS*, 469, 1166  
 DeBoer D. R. et al., 2017, *PASP*, 129, 045001  
 Douna V. M., Pellizza L. J., Mirabel I. F., Pedrosa S. E., 2015, *A&A*, 579, A44  
 Duncan K. et al., 2014, *MNRAS*, 444, 2960  
 Eide M. B., Graziani L., Ciardi B., Feng Y., Kakiichi K., Di Matteo T., 2018, *MNRAS*, 476, 1174  
 Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, *ApJ*, 672, L107  
 Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 646, 107  
 Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449–463  
 Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601  
 Field G. B., 1958, *Proc. IRE*, 46, 240  
 Finkelstein S. L., 2016, *PASA*, 33, e037  
 Fontanot F. et al., 2021, *MNRAS*, 504, 4481  
 Fornasini F. M. et al., 2019, *ApJ*, 885, 65  
 Fornasini F. M., Civano F., Suh H., 2020, *MNRAS*, 495, 771  
 Fragas T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013, *ApJ*, 776, L31  
 Furlanetto S. R., 2006, *MNRAS*, 371, 867  
 Furlanetto S. R., Stoever S. J., 2010, *MNRAS*, 404, 1869  
 Ghara R. et al., 2020, *MNRAS*, 493, 4728  
 Gillet N. J. F., Mesinger A., Park J., 2020, *MNRAS*, 491, 1980  
 Greig B., Mesinger A., 2015, *MNRAS*, 449, 4246  
 Greig B., Mesinger A., 2017, *MNRAS*, 472, 2651  
 Greig B., Mesinger A., 2018, *MNRAS*, 477, 3217  
 Greig B., Trott C. M., Barry N., Mutch S. J., Pindor B., Webster R. L., Wyithe J. S. B., 2021a, *MNRAS*, 500, 5322  
 Greig B. et al., 2021b, *MNRAS*, 501, 1  
 Hui L., Gnedin N. Y., 1997, *MNRAS*, 292, 27  
 Hunt L. et al., 2012, *MNRAS*, 427, 906  
 Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, *ApJ*, 854, 73  
 Jaacks J., Finkelstein S. L., Bromm V., 2019, *MNRAS*, 488, 2202  
 Kaaret P., Schmitt J., Gorski M., 2011, *ApJ*, 741, 10  
 Kaur H. D., Gillet N., Mesinger A., 2020, *MNRAS*, 495, 2354  
 Kern N. S., Liu A., Parsons A. R., Mesinger A., Greig B., 2017, *ApJ*, 848, 23  
 Kouroumpatzakis K. et al., 2020, *MNRAS*, 494, 5967  
 Lara-López M. A. et al., 2010, *A&A*, 521, L53  
 Lehmer B. D., Alexander D. M., Bauer F. E., Brandt W. N., Goulding A. D., Jenkins L. P., Ptak A., Roberts T. P., 2010, *ApJ*, 724, 559

- Lehmer B. D. et al., 2019, *ApJS*, 243, 3
- Lehmer B. D. et al., 2021, *ApJ*, 907, 17
- Linden T., Kalogera V., Sepinsky J. F., Prestwich A., Zezas A., Gallagher J. S., 2010, *ApJ*, 725, 1984
- Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, *ApJ*, 835, 113
- Ma X., Quataert E., Wetzel A., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2020, *MNRAS*, 498, 2001
- Madau P., Fragos T., 2017, *ApJ*, 840, 39
- Magg M. et al., 2021, preprint ([arXiv:2110.15948](https://arxiv.org/abs/2110.15948))
- Maiolino R. et al., 2008, *A&A*, 488, 463
- Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115–2127
- McGreer I. D., Mesinger A., D’Odorico V., 2015, *MNRAS*, 447, 499
- McLeod D. J., McLure R. J., Dunlop J. S., 2016, *MNRAS*, 459, 3812
- McLure R. J. et al., 2018, *MNRAS*, 479, 25
- McQuinn M., 2012, *MNRAS*, 426, 1349
- Mertens F. G. et al., 2020, *MNRAS*, 493, 1662
- Mesinger A., 2019, *The Cosmic 21-cm Revolution; Charting the First Billion Years of Our Universe*. IOP Publishing, Bristol, UK
- Mesinger A., Dijkstra M., 2008, *MNRAS*, 390, 1071
- Mesinger A., Furlanetto S., 2007, *ApJ*, 669, 663
- Mesinger A., Furlanetto S., Cen R., 2011, *MNRAS*, 411, 955
- Mesinger A., Ferrara A., Spiegel D. S., 2013, *MNRAS*, 431, 621
- Mineo S., Gilfanov M., Sunyaev R., 2012, *MNRAS*, 419, 2095
- Mirabel I. F., Dijkstra M., Laurent P., Loeb A., Pritchard J. R., 2011, *A&A*, 528, A149
- Mirocha J., Furlanetto S. R., 2019, *MNRAS*, 483, 1980
- Mirocha J., Furlanetto S. R., Sun G., 2017, *MNRAS*, 464, 1365
- Mondal R. et al., 2020, *MNRAS*, 498, 4178
- Muñoz J. B., Qin Y., Mesinger A., Murray S. G., Greig B., Mason C., 2022, *MNRAS*, 511, 3657
- Murray S., Greig B., Mesinger A., Muñoz J., Qin Y., Park J., Watkinson C., 2020, *J. Open Source Softw.*, 5, 2582
- Mutch S. J. et al., 2016, *MNRAS*, 463, 3556
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, 855, 105
- Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, *MNRAS*, 443, 678
- Pallottini A., Ferrara A., Gallerani S., Salvadori S., D’Odorico V., 2014, *MNRAS*, 440, 2498
- Park J., Mesinger A., Greig B., Gillet N., 2019, *MNRAS*, 484, 933
- Pentericci L. et al., 2018, *A&A*, 616, A174
- Planck Collaboration et al., 2016a, *A&A*, 594, A13
- Planck Collaboration et al., 2016b, *A&A*, 596, A108
- Pober J. C. et al., 2013, *AJ*, 145, 65
- Pober J. C. et al., 2014, *ApJ*, 782, 66
- Prestwich A. H., Tsantaki M., Zezas A., Jackson F., Roberts T. P., Foltz R., Linden T., Kalogera V., 2013, *ApJ*, 769, 92
- Pritchard J. R., Furlanetto S. R., 2007, *MNRAS*, 376, 1680
- Qin Y., Mesinger A., Park J., Greig B., Muñoz J. B., 2020a, *MNRAS*, 495, 123
- Qin Y., Mesinger A., Greig B., Park J., 2020b, *MNRAS*, 501, 4748
- Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, *MNRAS*, 506, 2390
- Rudakovskiy A., Mesinger A., Savchenko D., Gillet N., 2021, *MNRAS*, 507, 3046
- Sabti N., Muñoz J. B., Blas D., PRD, 2022, 928, L20
- Sanders R. L. et al., 2018, *ApJ*, 858, 99
- Sanders R. L. et al., 2021, *ApJ*, 914, 19
- Santos M. G., Ferramacho L., Silva M. B., Amblard A., Cooray A., 2010, *MNRAS*, 406, 2421
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 312
- Sun G., Furlanetto S. R., 2016, *MNRAS*, 460, 417
- The HERA Collaboration et al., 2022a, 925, 221
- The HERA Collaboration et al., 2022b, 924, 51
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Trott C. M. et al., 2020, *MNRAS*, 493, 4711
- Ucci G. et al., 2021, preprint ([arXiv:2112.02115](https://arxiv.org/abs/2112.02115))
- Visbal E., Barkana R., Fialkov A., Tseliakhovich D., Hirata C. M., 2012, *Nature*, 487, 70
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, *ApJ*, 745, 50
- Wouthuysen S. A., 1952, *AJ*, 57, 31
- Xu H., Wise J. H., Norman M. L., 2013, *ApJ*, 773, 83
- Xu H., Wise J. H., Norman M. L., Ahn K., O’Shea B. W., 2016, *ApJ*, 833, 84
- Yates R. M., Kauffmann G., Guo Q., 2012, *MNRAS*, 422, 215
- Yue B., Ferrara A., Pallottini A., Gallerani S., Vallini L., 2015, *MNRAS*, 450, 3829
- Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, *ApJ*, 791, 130

## APPENDIX: 21-CM POWER SPECTRA

Here, we show the 21-cm power spectra used in the two inferences from this paper. In Fig. A1, the black curves correspond to the B + 16  $L_X/\text{SFR}-Z$  relation, with error bars marking  $1\sigma$  noise. The recovered posterior assuming a metallicity-independent  $L_X/\text{SFR}$  is shown in red. Fig. A2 is analogous to Fig. A1, but using the F + 13 relation for the mock observation. We see that the constant  $L_X/\text{SFR}$  models reproduce the 21-cm PS quite well over the EoH. However, the steeper redshift evolution implied by the constant  $L_X/\text{SFR}$  models over-estimates the contribution of X-rays to the early stages of the EoR (see Fig. 7), which translates to a  $\sim$  tens of percent underprediction of large scale 21-cm power (c.f. Fig. 6) and a biased recovery of the ionizing escape fraction scaling with halo mass,  $\alpha_{\text{esc}}$ , shown in Fig. A3. We explore the recovery of galaxy parameters using different parametric relations for  $L_X/\text{SFR}$  and MZR in a follow-up work.



**Figure A1.** 21-cm power spectra. The mock observation corresponding to the B + 16  $L_X/\text{SFR}-Z$  scaling is shown in red with error bars representing  $1\sigma$  noise assuming a 1000 h observation with SKA1 low. The dark and light shaded regions denote 16–84 per cent and 2–98 per cent confidence intervals, respectively, obtained assuming a constant  $L_X/\text{SFR}$ .  $z_c$  denotes the central redshift of the light-cone chunk used to compute the power spectra. We demarcate with yellow stripes the  $k$  modes outside of the  $k = 0.1\text{--}1 \text{ Mpc}^{-1}$  range used to compute the likelihood.

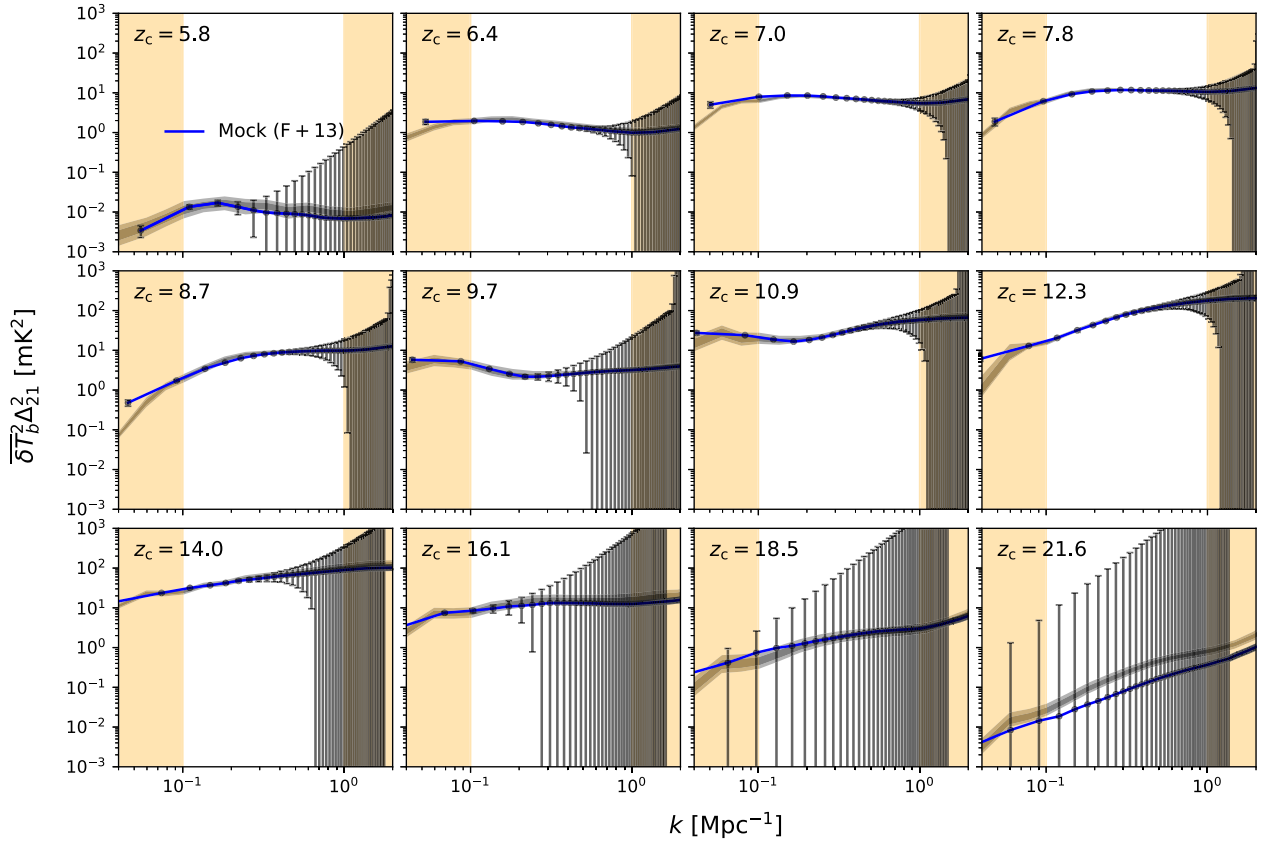
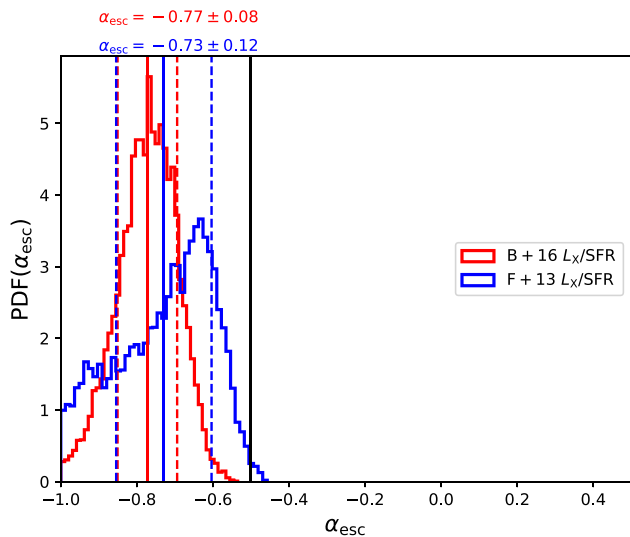


Figure A2. Same as Fig. A1 but using F + 13  $L_X$ /SFR– $Z$  scaling as our mock observation (blue curve).



**Figure A3.** Marginalized 1D PDFs of the power law scaling index of the ionizing escape fraction with halo mass,  $\alpha_{\text{esc}}$ , for constant  $L_X/\text{SFR}$  models. The PDFs inferred from mock data created using B+16 / F + 13 relations are denoted in red / blue. The mean and  $1\sigma$  rms width are denoted with solid and dashed vertical lines, respectively. The vertical black denotes the ‘true’ value used in making the mock simulations. The horizontal range shows the extent of our flat prior on  $\alpha_{\text{esc}}$ . Due to our choice of ‘Optimistic foregrounds’,  $\alpha_{\text{esc}}$  is very tightly constrained from the mock data. However, the recovered values for the constant  $L_X/\text{SFR}$  models are biased, in order to compensate for their implied additional contribution of X-rays to the very early stages of reionization, as discussed in the text. We will return to the recovery of galaxy properties in a follow-up work, including parametric models for the X-ray properties of the first galaxies.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.