


**HERA bound on x-ray luminosity when accounting for population III stars**Hovav Lazare<sup>ⓧ,\*</sup>, Debanjan Sarkar<sup>ⓧ,†</sup> and Ely D. Kovetz<sup>ⓧ,‡</sup>*Department of Physics, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel* (Received 16 August 2023; accepted 16 January 2024; published 21 February 2024)

Recent upper bounds from the Hydrogen Epoch of Reionization Array on the cosmological 21-cm power spectrum at redshifts  $z \approx 8, 10$ , have been used to constrain  $L_{X<2 \text{ keV}}/\text{SFR}$ , the soft-band x-ray luminosity measured per unit star formation rate (SFR), strongly disfavoring values lower than  $\approx 10^{39.5} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ . This conclusion is derived from seminumerical models of the 21-cm signal, specifically focusing on contributions from atomic cooling galaxies that host population II stars. In this work, we first reproduce the bounds on  $L_{X<2 \text{ keV}}/\text{SFR}$  and other parameters using a pipeline that combines machine learning emulators for the power spectra and the intergalactic medium characteristics, together with a standard Markov chain Monte Carlo parameter fit. We then use this approach when including molecular cooling galaxies that host population III stars in the cosmic dawn 21-cm signal, and show that lower values of  $L_{X<2 \text{ keV}}/\text{SFR}$  are hence no longer strongly disfavored. The revised Hydrogen Epoch of Reionization Array bound does not require high-redshift x-ray sources to be significantly more luminous than high-mass x-ray binaries observed at low redshift.

DOI: [10.1103/PhysRevD.109.043523](https://doi.org/10.1103/PhysRevD.109.043523)**I. INTRODUCTION**

The 21-cm signal carries immense potential to probe the cosmic dawn and Epoch of Reionization (EoR), two pivotal phases in the early Universe [1–5]. It arises from the hyperfine transition of neutral hydrogen atoms, and its measurement enables us to peer into the Universe’s infancy [6]. By studying the fluctuations in the 21-cm signal, we can unravel the processes that drove the formation of the first stars and galaxies [7–9], as well as the subsequent ionization of the intergalactic medium (IGM) [10,11]. The 21-cm signal is a powerful tool for understanding cosmic evolution and unlocking the mysteries of our cosmic origins [2,3,12].

Several ground-breaking experiments have been designed to target the 21-cm signal [13–19], including the Hydrogen Epoch of Reionization Array (HERA)<sup>1</sup> [20]. HERA has recently released two datasets [21,22], that were used to set an upper bound on the 21-cm fluctuation power spectrum from the epoch of reionization ( $z \sim 8$  and 10). This upper limit [21,22], together with other observations, including the galaxy ultraviolet (UV) luminosity function [23–26], the reionization optical depth [27] and the IGM neutral fraction [28] was then used to place the first constraints on properties related to the evolution of the first galaxies and the epoch of reionization [29].

A major finding of Ref. [29] has to do with the parameter  $L_{X<2 \text{ keV}}/\text{SFR}$  that quantifies the x-ray luminosity sourced by the endpoints of the stars that drive cosmic dawn and is defined as the ratio of the integrated soft-band ( $<2 \text{ keV}$ ) x-ray luminosity to the star formation rate (SFR). The HERA upper limit [21,22] disfavors  $L_{X<2 \text{ keV}}/\text{SFR} \lesssim 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  values with high significance. Taken at face value, this would imply that galaxies at high redshifts were more x-ray luminous than metal-enriched galaxies in the local Universe [30,31].

Crucially, the analysis in Ref. [29] considered only population II (or PopII) stars [32] as the sources of heating and ionization. The first generation of stars (PopIII) [33,34] are believed to have formed inside metal-poor galaxies (sometimes called molecular cooling galaxies or MCGs) hosted by mini halos where gas is cooled by molecular line transitions [34–41]. While we have very limited knowledge about the formation and evolution of MCGs, a number of simulations that have modeled them have shown that they can significantly prepone the onset of cosmic dawn [42–47]. In addition, MCGs contribute to the total x-ray and ionizing photon budget of the Universe [48,49]. Ignoring MCGs in any analysis could therefore lead to an overestimation of the x-ray and ionizing efficiencies of the atomic cooling galaxies (ACGs) that are believed to host mostly PopII stars [46]. This concern has also been raised in Ref. [22].

In this paper, we reanalyze the HERA data including MCGs, using machine learning to emulate the 21-cm signal [50–56] and other global quantities. The implication for the  $L_{X<2 \text{ keV}}/\text{SFR}$  constraint is summarized in Fig. 1.

\*hovav1@post.bgu.ac.il

†debanjan@post.bgu.ac.il

‡kovetz@bgu.ac.il

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

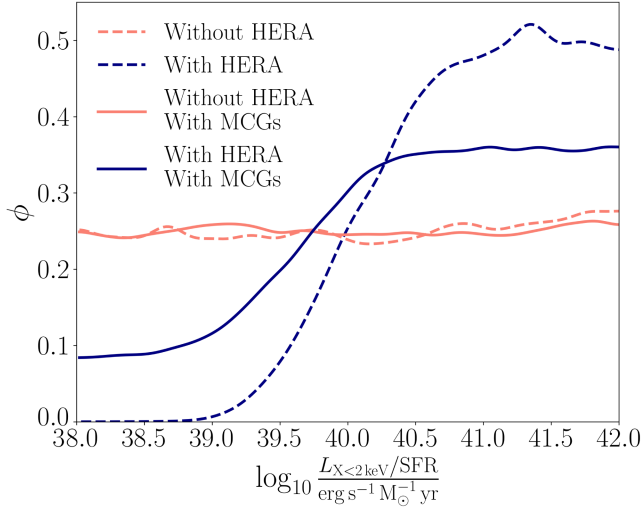


FIG. 1. Marginalized posteriors for  $L_{X<2 \text{ keV}}/\text{SFR}$ , with (blue) and without (orange) HERA and with (solid) and without (dashed) molecular cooling halos. With the inclusion of MCGs, the preference for strong x-ray heating weakens.

Simulations play a crucial role in modeling the 21-cm signal and interpreting the observational data [57–74]. Due to the complexity of the astrophysical processes involved, simulations help us understand the underlying physics and generate theoretical predictions for the 21-cm signal. However, simulations can be computationally expensive, if we need to simulate these complex processes accurately. Furthermore, in the traditional Bayesian parameter-inference pipeline, where we use simulations in each step to compare the outcomes against observations, we require a significant amount of time before the process converges. Even with faster seminumerical codes such as 21cmFAST<sup>2</sup> [75], the Markov chain Monte Carlo (MCMC) process with a minimal number of parameters can take up to several weeks to converge. We, therefore, need alternative methods.

Machine learning (ML) techniques have emerged as powerful tools for emulating the 21-cm signal and circumventing the time-consuming nature of traditional simulators [50–56]. The importance of machine learning in this context lies in its ability to significantly accelerate the parameter inference process. By training a machine learning model on a large set of precomputed simulations, it becomes possible to generate fast and accurate emulators that can predict the 21-cm signal for a given set of cosmological and astrophysical parameters. Moreover, machine learning emulators can be trained on a diverse range of simulations, enabling the exploration of different cosmological scenarios and astrophysical processes.

With the goal of revisiting the analysis in Ref. [29] to check the robustness of the  $L_{X<2 \text{ keV}}/\text{SFR}$  constraint to the inclusion of the PopIII stars or MCGs contribution, we have

built an ML-based emulator, that emulates the 21-cm power spectrum  $\Delta_{21}^2(k)$  as a function of wave vector  $k$  at different redshifts ranging from cosmic dawn to the epoch of reionization. We also built separate emulators for the Cosmic Microwave Background (CMB) optical depth and IGM neutral fraction.

Our paper is organized as follows. In Sec. II we briefly describe the models used in the HERA analysis [29] and the results of their analysis. In Sec. III we describe the HERA phase-I data and the likelihood we use for our analysis, including the additional observables. In Sec. IV we describe the artificial neural network architecture our emulator is based on, how we build our training, validation and test sets, and demonstrate the performance of emulator. In Sec. V we explain how we include the PopIII contribution in the modeling of the 21-cm signal. We then present our results in Sec. VI. After reproducing the HERA parameter constraints with our pipeline when considering only PopII stars, we include the MCGs in our simulations and train corresponding emulators. We then perform the MCMC analysis by considering both the ACG and MCG parameters. In particular, we present the marginalized posterior of  $L_{X<2 \text{ keV}}/\text{SFR}$  from this analysis (summarized in Fig. 1). We find that with the mini halos, we can no longer discard the  $L_{X<2 \text{ keV}}/\text{SFR} \lesssim 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  x-ray luminosity values easily. Although there is a small decline in posterior probability at  $L_{X<2 \text{ keV}}/\text{SFR} < 10^{39.5} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ , the  $L_{X}/\text{SFR} \sim 10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  values are still allowed. This suggests that the relationship between star formation and soft x-ray luminosity for high redshift galaxies may not be very different from the local galaxies [30,31]. We discuss and elaborate on this conclusion in Sec. VII.

## II. SUMMARY OF THE HERA ANALYSIS

Reference [29] have considered four different models of reionization to interpret the HERA observation. We briefly discuss the models below, and the inferences based on those.

- (1) Density-driven linear bias model: In this model, the 21-cm fluctuations are assumed to follow the density fluctuations and the 21-cm power spectrum is proportional to the matter power spectrum. A bias parameter, which is a function of quantities like  $x_{\text{HI}}$ ,  $T_{\text{S}}$  and  $T_{\text{CMB}}$ , relates the two power spectra.
- (2) Phenomenological reionization-driven model: In this model, IGM is considered as a two phased system consisting of fully ionized bubbles and medium with uniform temperature outside of the bubbles [76].
- (3) Seminumerical model I (21cmFAST): This model assumes that the star-forming galaxies reside inside dark matter haloes and these galaxies are responsible for the heating and ionization of the IGM. This utilizes empirical scaling relations to relate galaxy properties to their host dark matter haloes. The galaxy properties include stellar to halo mass ratio,

<sup>2</sup>[github.com/21cmfast/21cmFAST](https://github.com/21cmfast/21cmFAST).

escape fraction of ionizing UV radiation, the star formation rate and x-ray luminosity [75].

- (4) Seminumerical model II: This is an alternate seminumerical model [77], built on similar assumptions as 21cmFAST. On top of that, this model allows for an excess radio background. Excess radio background, on top of CMB, produces an enhanced 21-cm absorption signal and can also increase the amplitude of the 21-cm power spectrum.

The first two models relate the 21-cm signal directly to the IGM properties, without trying to model the source properties. On the other hand, the seminumerical models try to model the properties of the radiation sources and their evolution in redshift.

The analysis of Ref. [29] based on all of the above models suggest that the IGM at  $z = 10.4$  is heated over the adiabatic cooling limit at  $>95\%$  confidence, and this rule out the “cold reionization” scenarios in which the IGM continues to adiabatically cool until it reionizes.

Considering the seminumerical model I, Ref. [29] quotes a lower limit on the parameter  $L_{X<2 \text{ keV}}/\text{SFR}$  which describes the heating ability of the EoR galaxies per unit of star formation. We discuss the significance of  $L_{X<2 \text{ keV}}/\text{SFR}$  later in this paper.

HERA observations, Based on the seminumerical model II, and combined with the Chandra x-ray background constraints, rule out most of the models that explain the radio background observed by LWA [78] and ARCADE-2 [79] as originating at  $z \gtrsim 8$ .

Note that the inferences in Ref. [29] are very much dependent on the choice of physical models and the interpretations of different quantities change with models. In our current analysis, we mainly focus on seminumerical model I (21cmFAST), and we discuss the findings based on this model in detail.

### III. DATA AND LIKELIHOOD

#### A. The 21-cm observables

The emission or absorption of the 21-cm line is characterized by the spin temperature,  $T_S$ , and is usually measured as the differential brightness temperature,  $\delta T_{21}$ , with respect to the brightness temperature of the low-frequency radio background,  $T_{\text{rad}}$ . In the usual scenario,  $T_{\text{rad}}$  is taken to be the CMB temperature  $T_{\text{cmb}} = 2.7254 \times (1+z)$  K where  $z$  refers to the cosmological redshift. The differential brightness temperature,  $\delta T_{21}$ , can be expressed as [80–82]

$$\delta T_{21}(\nu) = \frac{T_S - T_{\text{rad}}}{1+z} (1 - e^{-\tau_{\nu_0}}). \quad (1)$$

Here, the factor  $(1 - e^{-\tau_{\nu_0}})$  exhibits the effect of propagation through a medium, where  $\tau_{\nu_0}$  is the optical depth. Note that, in this equation, the astrophysical information is mostly encoded in  $T_S$ , whereas the state of the propagating medium is captured by the optical depth  $\tau_{\nu_0}$ .

The measurement of  $\delta T_{21}$  is done in two different ways:

- (i) Global signal: Here,  $\delta T_{21}(\mathbf{n}, z)$  is first measured along different directions  $\mathbf{n}$  of the sky, keeping the redshift (or frequency) fixed. The measurements are finally averaged over the directions to convert into an average quantity  $\langle \delta T_{21}(z) \rangle$ , which is known as the global signal [83,84].
- (ii) Fluctuations: Here, telescopes measure the spatial fluctuations of the  $\delta T_{21}(\mathbf{x}, z)$  field. The fluctuations are then interpreted in terms of statistics like the power spectrum [76,85–90], bispectrum [91–97], etc. The 21-cm power spectrum is the most relevant quantity for this work, and is defined as

$$\langle \tilde{\delta T}_{21}(\mathbf{k}_1) \tilde{\delta T}_{21}(\mathbf{k}_2) \rangle = (2\pi)^3 \delta^D(\mathbf{k}_1 - \mathbf{k}_2) P_{21}(\mathbf{k}_1), \quad (2)$$

where  $\langle \dots \rangle$  denotes an ensemble average,  $\tilde{\delta T}_{21}(\mathbf{k})$  is the Fourier transform of  $\delta T_{21}(\mathbf{x})$  and  $\delta^D$  is the Dirac delta function. Specifically, the quantity of interest for us is  $\Delta_{21}^2(\mathbf{k}) \equiv k^3 P_{21}(\mathbf{k}) / (2\pi^2)$ , which will be the output of HERA measurements, and of our simulations and emulators.

#### B. HERA phase-I data

The upper limits on the 21-cm signal from HERA were first published in Ref. [21], and improved limits were published in Ref. [22]. These limits are based on 94 nights of observation using the HERA phase I experimental configuration. For detailed information on the configuration, the reader is referred to Ref. [22]. Out of all the different bands of observation, in this work we concentrate on two bands: band 2, centered at  $z = 7.9$ ; and band 1, centered at  $z = 10.4$ . These bands are largely free of radio frequency interference. Following Ref. [22], we use the observed power spectrum computed using all fields for each redshift band. For a comprehensive reading about the data, the reader is referred to Refs. [21,22].

#### C. The data likelihood

The form of the likelihood function is important for any Bayesian analysis. The 21-cm power spectrum measurements from HERA suffer from (largely unknown) systematics which need to be taken into account in the likelihood analysis. Below, we briefly explain how to derive the likelihood in the presence of systematics.

First, we divide our data into two vectors, one for every redshift bin,

$$\mathbf{d}_1 = \begin{pmatrix} c\Delta_{21}^2(k_1, z_1) \\ \Delta_{21}^2(k_2, z_1) \\ \vdots \end{pmatrix}, \quad \mathbf{d}_2 = \begin{pmatrix} c\Delta_{21}^2(k'_1, z_2) \\ \Delta_{21}^2(k'_2, z_2) \\ \vdots \end{pmatrix}. \quad (3)$$

In general  $k_i \neq k'_i$ . Considering a single data vector, in order to reduce the covariance between the neighboring  $k$  bins, we follow the method adopted in Ref. [29]. Starting from a minimum  $k$  bin (which is decided on the basis of the signal-to-noise ratio), we include the alternate  $k$  bins in our analysis such that the bins do not have significant overlap. In the following derivation, we show the form of the likelihood function for a single data vector, and the composite likelihood for the two data vectors may be calculated as  $\mathcal{L} \propto \mathcal{L}_1(\mathbf{d}_1) \cdot \mathcal{L}_2(\mathbf{d}_2)$ .

Following Ref. [29], we assume a Gaussian likelihood for the data, and denote the unknown systematics as  $\mathbf{u}$ , so that one can express the likelihood function for a data vector  $\mathbf{d}$  as

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}, \mathbf{u}) \propto \exp\left(-\frac{1}{2}r(\theta, \mathbf{u})^T \Gamma r(\theta, \mathbf{u})\right), \quad (4)$$

where  $r(\theta, \mathbf{u}) \equiv \mathbf{d} - \mathbf{W} \cdot \mathbf{m}(\theta) - \mathbf{u}$ ,  $\mathbf{m}(\theta)$  is the emulator prediction for the data vector  $\mathbf{d}$ ,  $\mathbf{W}$  is the Window function (which accounts for the point spreading in Fourier  $k$  space [98–102]), and  $\Gamma \equiv \Sigma^{-1}$  where  $\Sigma$  is the covariance matrix. In the parameter estimation problem, we are interested in finding the posterior probability for obtaining a parameter set given the observation data. This can be expressed in terms of the likelihood, based on Bayes' theorem, as

$$p(\theta|\mathbf{d}, \mathcal{M}, \mathbf{u}) \propto \mathcal{L}(\mathbf{d}|\theta, \mathcal{M}, \mathbf{u})p(\theta|\mathcal{M})p(\mathbf{u}), \quad (5)$$

where  $p(\theta|\mathcal{M})$  is the prior function of the parameter set  $\theta$  given the model  $\mathcal{M}$ , and  $p(\mathbf{u})$  denotes the priors on the systematics. In general, priors on systematics should come from the experiments. However, in case of HERA, the systematics are largely unknown. Therefore, we follow Ref. [29] and marginalize over the systematics prior range in order to obtain the likelihood. We assume a multivariate uniform distribution for  $\mathbf{u}$ , where the systematic noise in each  $k$  and  $z$  bins are independent. The marginalization can be done using,

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}, \mathbf{u}) \propto \int_{\mathbf{u}_{\min}}^{\mathbf{u}_{\max}} \exp\left(-\frac{1}{2}\mathbf{r}(\theta, \mathbf{u})^T \Gamma \mathbf{r}(\theta, \mathbf{u})\right) d\mathbf{u}. \quad (6)$$

Assuming  $\Gamma$  to be diagonal, the integral takes a simple form

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}, \mathbf{u}) \propto \prod_i^{N_d} \int_{u_{i,\min}}^{u_{i,\max}} \exp\left(-\frac{[y_i - u_i]^2}{2\sigma_i^2}\right) du_i, \quad (7)$$

where the index “ $i$ ” denotes the  $k$  bin number,  $y_i = d_i - [Wm(\theta)]_i$  is the difference between the model and the data at each bin,  $N_d$  is the number of  $k$  bins, and  $\sigma_i$  are the diagonal elements of the covariance matrix. In order to calculate the integral, we set  $\mathbf{u} \in [0, \infty]$ , assuming

systematics to be positive (although negative systematics can occur in some cases [103] and those are subject to interpretation). Marginalization over the systematics using the assumed bounds yields

$$\mathcal{L}(\mathbf{d}|\theta, \mathcal{M}) \propto \prod_i^{N_d} \frac{1}{2} \left(1 + \operatorname{erf}\left[\frac{y_i}{\sqrt{2}\sigma_i}\right]\right), \quad (8)$$

where “erf” is the error function. We note that the marginalized likelihood function is sensitive to the upper bounds of the data points, which is particularly useful for the HERA power spectrum data.

### 1. Additional observables

Note again that HERA measurements place only an upper bound on the 21-cm power spectrum. Therefore, the Likelihood in Eq. (8) does not impose any tight bounds on the model parameters. Following Ref. [29], we use additional observables from other experiments to complement the HERA power spectrum data. These additional observables are (i) the upper bound on the neutral hydrogen fraction  $x_{\text{HI}} < 0.06 + 0.05(1\sigma)$  at  $z \sim 5.9$ , measured by the dark fraction on high redshift quasar spectra [28], (ii) the faint galaxy UV luminosity functions at  $z = 6, 7, 8, 10$  [23,24,26],<sup>3</sup> and (iii) the Thomson scattering optical depth of CMB photons  $\tau_e = 0.0569_{-0.0066}^{+0.0081}$ , using Planck [27] data from Ref. [104]. We shall see later that these additional observables help to place tight bounds on some of the model parameters, while HERA helps constrain some of the remaining parameters.

## IV. ARTIFICIAL NEURAL NETWORK

In order to perform a MCMC analysis, we need to be able to produce realizations of the 21-cm power spectra and other global quantities, given a set of input parameters. However, to simulate the 21-cm signal for a large cosmological volume with the required resolution using the latest version of 21cmFAST [47] takes  $\mathcal{O}(1 \text{ hour})$ . The traditional MCMC pipeline requires to generate a large number of simulations in a sequential manner to be able to construct the final posterior, which takes a lot of time and can be prohibitive. In order to address this, we have developed an artificial neural network (ANN)-based 21-cm signal emulator, which we describe below.

### A. ANN architecture

#### 1. Simulations

It is relatively well accepted that the galaxies form inside the dark matter halos, with different feedback

<sup>3</sup><https://github.com/21cmfast/21CMMC/tree/master/src/py21cmmc/data>.

processes shape their evolution [105–110]. These galaxies are the main sources of various radiation fields that affect the 21-cm signal from the beginning of cosmic dawn to the epochs that follow. We use the seminumerical code  $21\text{cmFAST}$  that simulates various galaxy properties, depending on the host halo mass, and produces these radiation fields [45–47,75,111–113]. The code then uses the same radiation fields to deduce the 21-cm fluctuations at different redshifts. Note that the seminumerical calculations of  $21\text{cmFAST}$  are tuned to reproduce the galaxy UV luminosity functions [114–116] observed at reionization as well as the observed opacity of IGM.

Given a set of cosmological and astrophysical parameters,  $21\text{cmFAST}$  generates realizations of the 21-cm fluctuations in coeval and/or light cone boxes. Those boxes are then used to compute the global quantities like the mean brightness temperature of the 21-cm signal  $\langle\delta T_{21}\rangle$ , optical depth of reionization  $\tau_{\text{reio}}$ , etc., and 21-cm fluctuation power spectrum  $\Delta_{21}^2(\mathbf{k})$ . In addition to this,  $21\text{cmFAST}$  is also able to output the UV luminosity function of galaxies at various redshifts. We create an ensemble of the observables (which we call datasets) by changing the model parameters, and this ensemble will be used to train and validate our emulator, as we describe later in this section. For the 21-cm power spectrum, in this work we are mainly concerned about two redshift bins  $z = 7.9$  and  $10.4$  where HERA observations have the best detection [29].

The definitions of the important astrophysical parameters used in this work are given in Table I. The  $21\text{cmFAST}$  code models the effects of both the PopIII and PopII stars. PopIII [33,34] refers to the first generation of stars believed to reside in MCGs or minihalos, while PopII [32] refers to the second generation of stars which predominantly reside in atomically cooled galaxies or ACGs. We shall discuss these in more detail in Sec. V. The parameters related to the ACGs (MCGs) are represented by a “10” (“7” or “mini”) in the subscript (referring to the typical halo mass). In the next few sections, we shall assume that the cosmic dawn and reionization processes are driven solely by the PopII stars. We build the dataset in the following section using

$21\text{cmFAST}$  where we consider only PopII stars. We note that the same modeling has been considered in Ref. [29].

## 2. Building the dataset

In this section, we discuss the method of composing the dataset which will then be divided into training, validation and testing sets. We start by determining the desired range over which we want to vary each parameter. Having its range fixed, we sample the whole parameter space using a latin hypercube (LH) sampler which seeks to produce uniform sampling densities when all points are marginalized onto any one dimension [117]. After getting the sampled parameter sets, we run the  $21\text{cmFAST}$  code to obtain the 21-cm power spectrum and global quantities like  $\tau_{\nu}$ ,  $x_{\text{HI}}$  for each combination. We then save these quantities along with the parameter values.

We now discuss the dataset which consists of the 21-cm power spectrum as a function of  $k$ . Note that, we have a limited number of  $k$  bins from the simulations to begin with. Before the training process, we use interpolation to increase the number of power spectrum  $k$  bins to 70, ranging from  $k = 0.1$ – $1.73 \text{ Mpc}^{-1}$ , so that the variation with the wave number is smoother. The dataset is then classified into six different classes (denoted by index “ $n$ ”) according to the amplitude difference between  $k$  bins. The first four ( $n = 0$ – $3$ ) classes are for signals where the difference between the maximal amplitude and the minimal amplitude of the power spectrum, is smaller than  $10^n \text{ mK}^2$  but greater than  $10^{(n-1)} \text{ mK}^2$ . Class 5 is for signals where the difference between the maximum and minimum power spectrum values in different bins, i.e.  $[\max(\text{power spectrum}) - \min(\text{power spectrum})]$  is  $> 10^3 \text{ mK}^2$ . Finally, class 6 is reserved for signals where the minimal amplitude is not in the lowest  $k$  bin (these are exotic-shaped signals). This classification is done so that when the emulator is trained, we can take approximately the same number of samples from each class, so that the trained emulator will not be biased towards any of the classes. Without doing so, emulating exotic signals tends to be more difficult, since

TABLE I. The main astrophysics parameters.

Parameters	Description
$f_{\star}$ .....	Stellar to halo mass ratio for haloes
$\alpha_{\star}$ .....	Stellar-to-halo mass power-law index
$f_{\text{esc}}$ .....	Escape fraction of ionizing photons from the haloes
$\alpha_{\text{esc}}$ .....	Escape fraction of ionizing photons to halo mass power-law indices
$M_{\text{turn}}$ .....	Mass scale below which inefficient cooling and/or feedback suppresses efficient star formation
$t_{\star}$ .....	Star formation timescale in units of $H^{-1}(z)$
$L_{X<2 \text{ keV}}/\text{SFR}$	Soft-band x-ray luminosity per SFR in units of $\text{erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ for ACGs
$E_0$ .....	Minimum x-ray energy (in eV) escaping the galaxies into the IGM
$\alpha_X$ .....	Spectral index of x-ray sources

they are rare in the unmanipulated dataset. Note also that the “class 0” signals are also very hard to emulate even after this process because most of them are actually “dead” or flat signals having no features that the emulator can learn. This happens when for some parameter combination, the fluctuations at all scales cease at a specific redshift. For example, when the IGM becomes sufficiently ionized at low redshifts that the fluctuations in the electron density field become more or less uniform, and the power spectrum is flat (and even becomes zero in some cases). Emulating the flat signals is extremely difficult, and we have decided not to include these at all in the emulating process. Including flat signals in the emulating process leads to unreliable predictions. A method for predicting these flat signals (for use in the MCMC pipeline) will be explained later. The bottom line is that we have discarded the class 0 signals when training the emulator.

For the training process, we take  $\sim 2500$  randomly sampled signals from each class for  $z = 7.9$ , and  $\sim 3000$  signals from each class for the  $z = 10.4$  band. We choose the numbers such that we have a similar number of samples in each class, and such that no class is over/under-represented. Finally, the dataset is divided into training (85%), validation (10%) and testing (5%) sets. Before the training process, the astrophysical parameters are normalized to be in the range  $[-1, 1]$ . We do not normalize the power spectrum as our initial tests suggested a better performance with the original power spectrum.

### 3. NN architecture

In this section, we discuss the architecture of our NN emulators. All of our NN emulators and classifiers are built using the TensorFlow [118] and Keras [119] libraries defined in PYTHON3. Here we consider a nine-parameter emulator model based on 21cmFAST. According to this model, only the halos which are cooled by atomic line transitions host sources that start cosmic dawn and cosmic reionization. The choice of the parameters is based on Ref. [29]. As explained above, there is another possible cooling mechanism, which makes it possible to form sources inside minihalos. Using 21 cmFAST simulations, it is possible to model these mini halos, and we shall discuss this in more detail in the later sections.

Therefore, the emulator has nine input parameters, 70 output  $k$  bins, and six fully connected hidden layers—of 288, 512, 512, 288, 512, 1024 dimensions, respectively. The number of layers and the number of neurons in each was chosen after an optimization process. We have tested the emulator for a number of conventional activation functions, namely ReLU [120] and tanh, etc. We finally settled with the following function (taken from Ref. [121])

$$f(\mathbf{x}, \boldsymbol{\alpha}) = \mathbf{x} \odot (\boldsymbol{\alpha} + (1 - \boldsymbol{\alpha}) \odot (1 + e^{-\mathbf{x}})^{-1}), \quad (9)$$

where  $\mathbf{x}$  are the layer inputs,  $\boldsymbol{\alpha}$  is a vector of trainable parameters, and  $\odot$  denotes the element-wise product. The activation is followed by a batch normalization layer [122], which standardizes the input for the next layer. After the forward propagation, the loss is calculated using a relative percentage error

$$\mathcal{L}(y_{\text{true}}, y_{\text{pred}}(\theta)) = \frac{|y_{\text{true}} - y_{\text{pred}}(\theta)|}{y_{\text{true}}} \times 100. \quad (10)$$

The training process is performed using mini batches of size 128 samples, so that the trainable parameters are updated using the gradients of the loss function, after forward propagating each batch. The training process was done using the Adam [123] optimizer, where the initial learning rate is 0.01, and is reduced when the prediction on the validation set does not improve over more than 10 epochs. The training phase ends when there is no improvement over more than 30 epochs, and the best weights are stored at that moment. Finally, we examine our emulator performance on the test set, and the results are presented in Fig. 2.

### 4. Retraining

After training the emulator, we review its performance very carefully on the test set. In the context of the power spectrum, we find that on rare occasions (less than 5%), the emulator prediction differs from the real value by almost 100%. This may bias the MCMC results, particularly the median value. We do the following to make sure that this error does not affect our results.

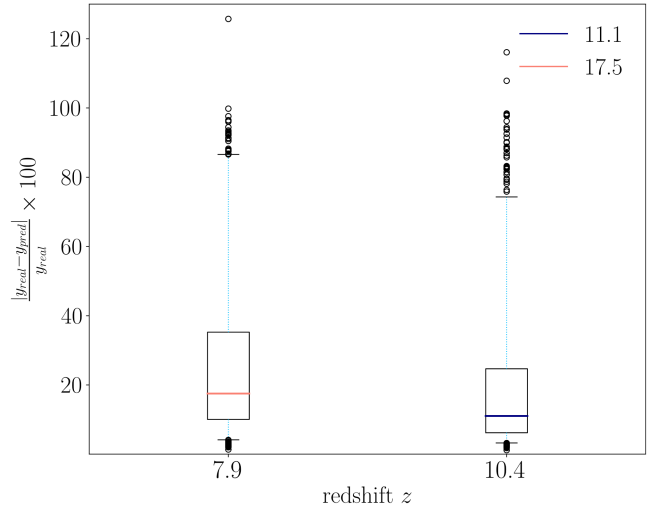


FIG. 2. Emulator loss statistics on the testing sets, of size  $\sim 650$  signals for each band. The box represents all the samples that fall between the 25th and the 75th percentile, and the line in the middle of it is the median. The black circles are outliers that extend beyond the 95th percentile.

After the first training, we determine a rough  $\sim 1\sigma$  region in the parameter space by running the MCMC analysis using the emulator. We then generate  $\sim 5000$  more signals by sampling the rough  $1\sigma$  parameter space using the LH method, where the  $1\sigma$  is considered for all the parameters except  $L_{X<2\text{ keV}}/\text{SFR}$ , for which the emulator has to be accurate over the full prior range. We then retrain our emulator on this resampled dataset.

Note that, for some astrophysical parameters such as  $\alpha_*$ , the  $1\sigma$  constraint range is very small as we shall show later in Fig. 7. For those parameters, we choose a wider range when sampling. We show the performance results after the retraining process in Figs. 3 and 4. We see that the retrained emulator gives a median error of 2–3% in the  $1\sigma$  parameter space. Now considering the whole range of parameters that we used for initial training, we see that our retrained emulator exhibits  $<20\%$  median error outside of the  $1\sigma$  parameter space. This suggests that the emulator will have better precision, at least within the  $1\sigma$  parameter space, such that it will not bias the MCMC analysis by much and will provide better  $1\sigma$  constraints.

We shall see later (see also Ref. [29]) that most of the astrophysical parameters, courtesy of the additional observables used in addition to the HERA power spectrum data, are very well constrained. Other parameters, like  $L_{X<2\text{ keV}}/\text{SFR}$  (which is meaningfully constrained only at the lower side) or  $E_0$ , have almost no constraints from the observations. Now, due to the property of the Markov chains, the MCMC algorithm will mostly sample the region of the parameter space where our retrained emulator performs extremely well. This means we shall have many accurate MCMC samples in the vicinity of the  $1\sigma$  range, and very few less-accurate samples outside this region. This

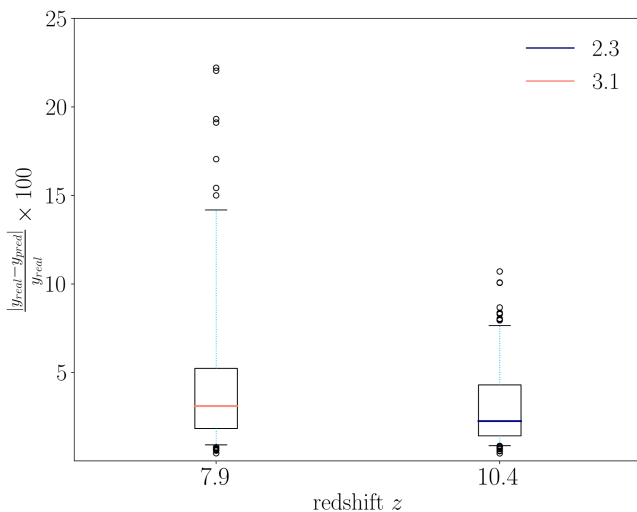


FIG. 3. Retrained emulator loss statistics on our testing sets, of size  $\sim 250$  signals for each band. The box represents all the samples that are between the 25th and the 75th percentiles, and the line in the middle of it is the median. The black circles are outliers that extend beyond the 95th percentile.

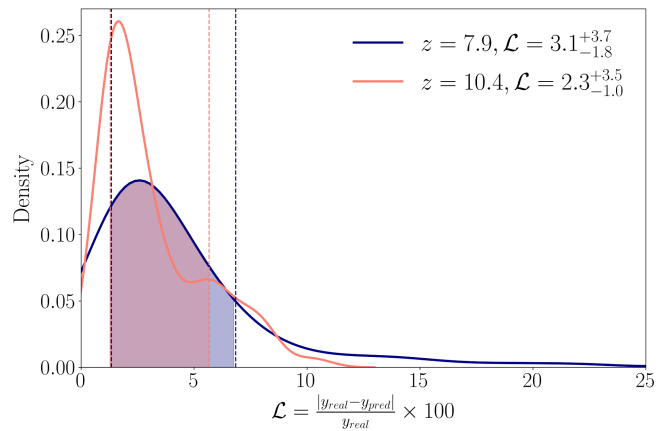


FIG. 4. Different view of the loss statistics for the retrained emulator. The dashed lines represent the 16th and 84th percentiles respectively, and their exact value is written along with the median value in the figure legend.

weighs down any possible bias and ensures our predictions are robust.

## 5. Classification

We have already mentioned in Sec. IV A 2 that the class 0 signals or flat power spectra are very hard to emulate. The power spectrum amplitude in this class is close to zero,  $\lesssim 3\text{ mK}^2$  for all the  $k$  bins. When we run the MCMC process, for each set of parameters we need to know whether the parameters will produce flat signals even before we feed those to the ANN. We, therefore, need machinery that predicts the occurrence of the flat signal for a given parameter set.

We have built a neural network classifier for the above purpose. This classifier is trained at each frequency band to predict the occurrence of the flat 21-cm power spectrum. We discuss the architecture of the classifier in Appendix C. We find that the classifier’s prediction accuracy is  $\gtrsim 97\%$ . This indicates that, on average, we shall have a misclassification less than once in every 30 MCMC iterations. This does not boast a serious error in the MCMC analysis.

Considering the flat signal, we choose the following path. When the classifier predicts a flat signal for some parameter combination, we generate a random power spectrum at each  $k$  bin where the amplitude is drawn from a Gaussian distribution having a mean of  $2\text{ mK}^2$  and a standard deviation of  $1\text{ mK}^2$ . This we do irrespective of the parameter values. Our goal here is not to predict the flat signal accurately as these signals are still somewhat rare. Also, considering the upper bound on the 21-cm power spectrum from HERA [and our chosen likelihood function in Eq. (8)] it is clear that any flat signal below the upper limit is allowed. Therefore, HERA data do not really discard these parameter values, rather the constraints come from the additional likelihoods. Therefore, we surmise that

our solution for tackling the flat power spectra does not pose any serious issues.

### B. Emulating the other observables

We extend our NN for emulating  $x_{HI}$  and  $\tau_e$ . These are easy to emulate as for a fixed redshift both of them are just single numbers. We have therefore used a relatively simple neural network architecture for emulating both  $x_{HI}$  and  $\tau_e$ . As we discuss in Appendix A, we find that the emulators performs extremely well and the emulation error is negligible. As for the simulated UV luminosity functions, we do not have any emulators for these. UV luminosity functions can be calculated analytically very quickly via 21cmFAST, and in the MCMC analysis we generate these directly by plugging in the parameters.

## V. THE ROLE OF MINIHALOS

The current understanding of the stellar composition in galaxies at low redshifts is quite comprehensive. However, our knowledge about the initial galaxies that sparked cosmic reionization is limited. According to the hierarchical model of structure formation, the first generation of stars, known as PopIII stars, likely formed inside small molecular cooling galaxies or MCGs around  $z \sim 20\text{--}30$  [34–41]. Inside MCGs, the gas cools mainly through the  $H_2$  rotational–vibrational transitions efficient at  $T_{\text{vir}} \sim 10^3\text{--}10^4$  K. These early galaxies were believed to be hosted by halos with virial temperatures  $\lesssim 10^4$  K, corresponding to total halo masses of  $\lesssim 10^8 M_\odot$  during the epochs of reionization and cosmic dawn [41,44,45,47,124]. These small halos are often called minihalos. Over time, feedback mechanisms suppressed star formation in these MCGs, leading to the emergence of heavier ACGs with halo masses  $\gtrsim 10^8 M_\odot$ . Therefore, most of the ACGs at high redshifts are “second-generation” galaxies, forming out of MCGs. It is believed that the second generation of stars, known as PopII stars, predominantly reside in ACGs [46].

PopII stars likely drove most of the heating and ionization process at  $z \sim 5\text{--}10$  [125–127]. However, there should be a small but significant contribution from the PopIII stars or MCGs as well. The star formation in MCGs is subject to a number of feedback processes like the Lyman-Werner radiation feedback [36,74,128–132], baryon supersonic streaming velocities [133–139], etc. Depending on these, the radiation contribution from the MCGs varies. MCGs spark the cosmic dawn epoch earlier compared to the scenarios where only ACGs exist [74]. X-ray photons from MCGs also pre-empt the epoch of heating [74], although MCGs are believed to have a negligible contribution to reionization. As emphasized above, in Ref. [29], the 21-cm signal was modeled without taking the contribution of the PopIII stars or MCGs into account. One of the main conclusions of Ref. [29] is that the early galaxies were more x-ray luminous than their present-day counterparts.

Our main goal in this paper is to check the validity of this result in the presence of a contribution to the 21-cm signal from PopIII stars.

One of the most salient features in the latest version of 21cmFAST code [47] is that it now includes the contributions of both the MCGs and ACGs. The code assumes MCGs host PopIII stars with a simple stellar-to-halo mass relation, distinct from that of ACGs hosting PopII stars. For a comprehensive reading about how 21cmFAST incorporates PopII and PopIII stars in the calculations, the reader is referred to Refs. [45,47].

As mentioned earlier, star formation in minihaloes depends on feedback processes which further decide the molecular-cooling turnover halo mass. Haloes below this mass are not capable of forming stars. In 21cmFAST, this turnover mass is a product of three factors [47],

$$M_{\text{mol}} = M_0(z) f_{v_{\text{cb}}}(v_{\text{cb}}) f_{\text{LW}}(J_{21}), \quad (11)$$

where  $M_0(z)$  is the molecular cooling threshold in the absence of feedback with  $M_0(z) = \tilde{M}_0 \times (1+z)^{-3/2}$  where  $\tilde{M}_0 = 3.3 \times 10^7 M_\odot$  (corresponding to a virial temperature of  $10^3$  K [37]),  $f_{v_{\text{cb}}}(v_{\text{cb}})$  and  $f_{\text{LW}}(J_{21})$  encodes the effects of respectively the baryon streaming velocity and Lyman-Werner background radiation.  $f_{\text{LW}}(J_{21})$  is parametrized as

$$f_{\text{LW}}(J_{21}) = 1 + A_{\text{LW}}(J_{21})^{\beta_{\text{LW}}}, \quad (12)$$

with  $A_{\text{LW}}$  and  $\beta_{\text{LW}}$  are two free parameters that quantifies the strength of the feedback, and  $J_{21}$  being the Lyman-Werner radiation intensity. Similarly,  $f_{v_{\text{cb}}}(v_{\text{cb}})$  is parametrized as

$$f_{v_{\text{cb}}}(v_{\text{cb}}) = \left( 1 + A_{v_{\text{cb}}} \frac{v_{\text{cb}}}{v_{\text{rms}}} \right)^{\beta_{v_{\text{cb}}}}, \quad (13)$$

where  $v_{\text{cb}}$  is the local value of the streaming velocity,  $v_{\text{rms}}$  is the rms of the streaming velocity,  $A_{v_{\text{cb}}}$  and  $\beta_{v_{\text{cb}}}$  are two free parameters that determine the strength of the feedback. We have kept the feedback parameter values fixed to  $A_{\text{LW}} = 2.0$ ,  $\beta_{\text{LW}} = 0.6$ ,  $A_{v_{\text{cb}}} = 1$  and  $\beta_{v_{\text{cb}}} = 1.8$  throughout all of the simulations that include MCGs. Note that the feedback, or in other words, the parameters, are largely unknown at the redshifts that we are interested in. The contribution of MCGs in cosmic dawn and reionization vary with the variation in these parameters. We do not vary these in our analysis as these would increase the number of parameters. However, we expect that our results will still be valid even when these parameters are varied in the vicinity of the fiducial values.

The 21cmFAST code has a number of astrophysical parameters to characterize the galaxy properties. In Table I, we mention some of the important astrophysical parameters used in this work. We follow the same method to build the dataset and utilize the same pipeline developed in Sec. IV to emulate the 21-cm power spectrum and other



observables in the presence of MCGs. We discuss the performance of this emulator in Appendix B. Overall, we find that we get similar results even after introducing the additional parameters for MCGs.

In Fig. 5 we show the importance of accounting for the MCGs when interpreting the HERA data, plotting the power spectrum for  $L_{X<2\text{ keV}}/\text{SFR} = 10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ , where we use only ACGs in the simulation, we see that this value of  $L_{X<2\text{ keV}}/\text{SFR}$  is clearly disfavored by the HERA observations. This is consistent with the findings in Ref. [29]. However, when adding MCGs to our simulations, we find that the same value of  $L_{X<2\text{ keV}}/\text{SFR}$  is consistent with the HERA data. A prediction for this behavior, was already mentioned in Ref. [22]. The bottom line is that with MCGs, even the smallest values of  $L_{X<2\text{ keV}}/\text{SFR}$  in our prior range cannot be ruled out by current HERA data. Later, we shall see the same in the posterior distribution of  $L_{X<2\text{ keV}}/\text{SFR}$  from our full MCMC analysis.

## VI. RESULTS

### A. Bayesian inference

We now perform the MCMC analysis based on our pipeline that combines the emulator discussed in Sec. IV

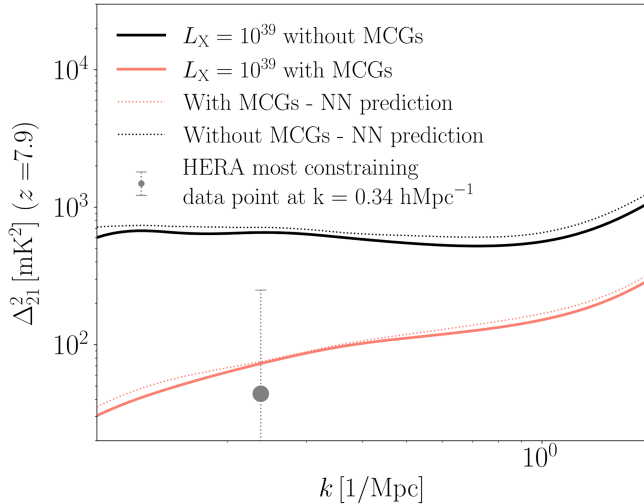


FIG. 5. Comparison between simulated power spectra with and without MCGs, and the emulator performance for each signal. The parameters taken for the simulation are the medians of the marginalized posterior distributions depicted in Fig. 8, for all the parameters (MCG parameters were not included in the “without MCGs” simulation) except  $L_{X<2\text{ keV}}/\text{SFR}$ , for which we have taken  $10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ . We quote some of the other parameters used in this figure for convenience:  $\log_{10} f_{*,10} = -1.25$ ,  $\log_{10} f_* = -2.65$ ,  $\alpha_* = 0.52$ ,  $\alpha_{*,\text{mini}} = -0.03$ ,  $\log_{10} f_{\text{esc},10} = -1.45$ ,  $\log_{10} f_{\text{esc},7} = -2.31$ ,  $\alpha_{\text{esc}} = 0.55$ . Note that, for the “without MCGs” scenario (red curves), we have considered only ACGs in the simulations (same as the scenario assumed in Fig. 7) with the appropriate parameter values as mentioned above, ignoring any contribution from the MCGs.

and the likelihood mentioned in Sec. III. In the MCMC analysis, we have two main goals: (I) To reproduce the results of Ref. [29] which includes only ACGs in the analysis, and (II) to study the impact of MCGs on the x-ray heating constraints. Considering the two goals, we have two scenarios for the likelihood: (A) Use a likelihood for only the additional observables mentioned in Sec. III C 1, which we denote as “External likelihood,” and (B) consider a “Full likelihood” that consists of the external likelihood and the likelihood for the HERA power spectrum upper bound. We, therefore, have four different MCMC inferences and these are (I-A): ACGs + External likelihood, (I-B): ACGs + Full likelihood, (II-A): MCGs + External likelihood, and (II-B): MCGs + Full likelihood. We discuss the results below. For the MCMC runs, we use the EMCEE [140] sampler. In Table II, we show the prior ranges used for this work and we use uniform flat priors for all the variables.

For inferences of type I-A and I-B, we have used the same prior range as in Ref. [29]. For inferences with minihalos, we drop some parameters in order to reduce the parameter space. We find that  $t_*$  (0.5) and  $\alpha_X$  (1.0) parameters are not very well constrained. Therefore, even if we drop them, there should not be any major effect on the inference. We keep them fixed at the values mentioned within the brackets. Furthermore, in principle, PopIII stars should have different x-ray luminosity (denoted by  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR}$ ) than PopII stars (denoted by  $L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$ ). However, we have no prior knowledge of  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR}$  of the PopIII stars, which reside mainly in MCGs. Considering this uncertainty, and, in order to reduce the number of parameters, we assume  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} = L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$  and call this the effective  $L_{X<2\text{ keV}}/\text{SFR}$ .

TABLE II. Prior range summary for all the astrophysical parameters used in the different MCMC runs. Note that only a subset of them is used in each run, as explained in the text.

Parameter name	Priors	
	Lower bound	Upper bound
$\log_{10} f_{*,10}$	-3.0	0.0
$\log_{10} f_{*,7}$	-3.5	-1.0
$\alpha_*$	-0.5	1.0
$\alpha_{*,\text{mini}}$	-0.5	0.5
$\log_{10} f_{\text{esc},10}$	-3.0	0.0
$\log_{10} f_{\text{esc},7}$	-3.0	0.0
$\alpha_{\text{esc}}$	-1.0	1.0
$\log_{10} [M_{\text{turn}}/M_{\odot}]$	8.0	10.0
$t_*$	0.0	1.0
$\log_{10} \frac{L_{X<2\text{ keV}}/\text{SFR}}{\text{erg s}^{-1} M_{\odot}^{-1} \text{ yr}}$	38.0	42.0
$E_0/\text{keV}$	0.1	1.5
$\alpha_X$	-1.0	3.0

The prior range of the MCG parameters is taken to be the same as the ACG parameters.

Note that the above choice may affect the prior volume nontrivially. Making  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} = L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$  automatically reduces the effective value of  $L_{X<2\text{ keV}}/\text{SFR}$ . This would not be the case if  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR}$  is made to vary independently. However, as the PopIII stars are metal poor, our naive expectation is that  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} \gtrsim L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$ . Therefore, we expect our results to be consistent and even somewhat conservative under the assumption we made.

### B. Weaker constraints on $L_X$

In the top panel of Fig. 6, we show the marginalized posteriors for  $L_{X<2\text{ keV}}/\text{SFR}$ , the parameter most constrained by HERA, for inferences I-A and I-B. This figure is directly comparable to Fig. 28 of Ref. [22] (see also Fig. 7 in Ref. [29]) and our result is largely in agreement with theirs. The external likelihoods do not impose any constraint on  $L_{X<2\text{ keV}}/\text{SFR}$ . The HERA data helps to place bounds at low  $L_{X<2\text{ keV}}/\text{SFR}$ , and we can see that values below  $L_{X<2\text{ keV}}/\text{SFR} \sim 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  are heavily disfavored, while higher values are preferred.

The amplitude of the power spectrum increases as  $L_{X<2\text{ keV}}/\text{SFR}$  is decreased, and below  $L_{X<2\text{ keV}}/\text{SFR} \sim 10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  (as can be seen from Fig. 5, where we change the  $L_{X<2\text{ keV}}/\text{SFR}$  values while keeping the other parameters at the median values as mentioned in Fig. 7), the power spectrum amplitude exceeds the HERA  $1\sigma$  upper bound. On the other hand, we find an almost flat posterior towards the large  $L_{X<2\text{ keV}}/\text{SFR}$ . This result somewhat disagrees with Ref. [22] where the  $L_{X<2\text{ keV}}/\text{SFR}$  posterior shows a steady decline at  $L_{X<2\text{ keV}}/\text{SFR} > 10^{40.8} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ . This is possibly due to the unequally weighted sample points of the MultiNest sampler [141] used by 21cmMC [142]. The key point in Fig. 6 is that for models where only ACGs (or PopII stars) are present, the HERA upper limit suggests more x-ray luminosity for the high-redshift ( $z > 6$ ) sources.

The bottom panel of Fig. 6 shows the marginalized posteriors for  $L_{X<2\text{ keV}}/\text{SFR}$ , for inferences II-A and II-B. As discussed earlier, both these inferences are based on models where we include MCGs (or PopIII stars) and ACGs (or PopII stars), and  $L_{X<2\text{ keV}}/\text{SFR}$  here represents an effective value for the x-ray luminosity of these galaxies. The External likelihoods do not place any bounds on  $L_{X<2\text{ keV}}/\text{SFR}$  within the prior range we used, which is expected. However, with HERA data, we find a striking difference when compared with the top panel. We see that the strong preference for the high  $L_{X<2\text{ keV}}/\text{SFR}$  ( $L_{X<2\text{ keV}}/\text{SFR} > 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ ) values is severely relaxed as soon as we include MCGs. This is also evident from Fig. 5. Although there is a decline in posterior

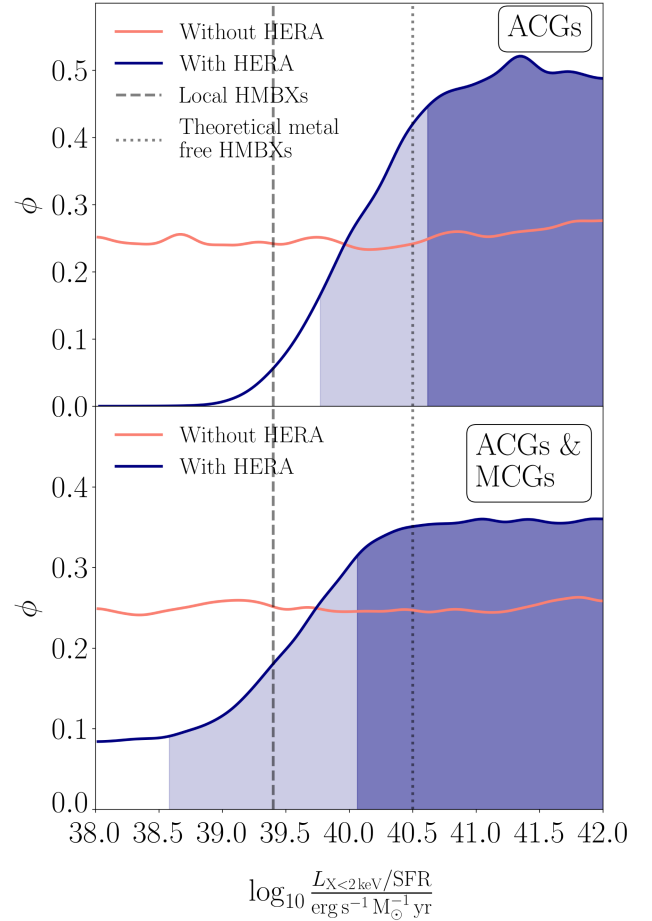


FIG. 6.  $L_{X<2\text{ keV}}/\text{SFR}$  marginalized posterior distribution. The dark shaded area denotes 68% highest posterior density (HPD), and the light shaded area denotes 95% HPD. The dashed black line represents the average  $L_{X<2\text{ keV}}/\text{SFR}$  as inferred from observations of high-mass x-ray binaries (HMXBs) in local, star-forming galaxies [30]. The dotted black line denotes the theoretical prediction for metal-free HMXBs presented in Ref. [147] (such galaxies are expected to be common in the early Universe). The average observed value is outside the  $2\sigma$  region of the posterior distribution in the top plot, suggesting a possible tension between the observations. However, the bottom plot implies that this tension is relaxed when including MCGs in the 21-cm simulations.

probability below  $L_{X<2\text{ keV}}/\text{SFR} \sim 10^{39.5} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ , this is not as significant as in the top panel where the decline is very sharp at  $L_{X<2\text{ keV}}/\text{SFR} < 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ . Overall, Fig. 6 clearly indicates that the lower  $L_{X<2\text{ keV}}/\text{SFR}$  values are still possible when we add MCGs to the analysis.

### C. Posteriors for other astrophysical parameters

In Fig. 7, we show 1D and 2D marginalized posterior distributions for inferences I-A and I-B. This figure is directly comparable to Fig. 6 of Ref. [29]. We find that our results resemble the results of Ref. [29] except for one

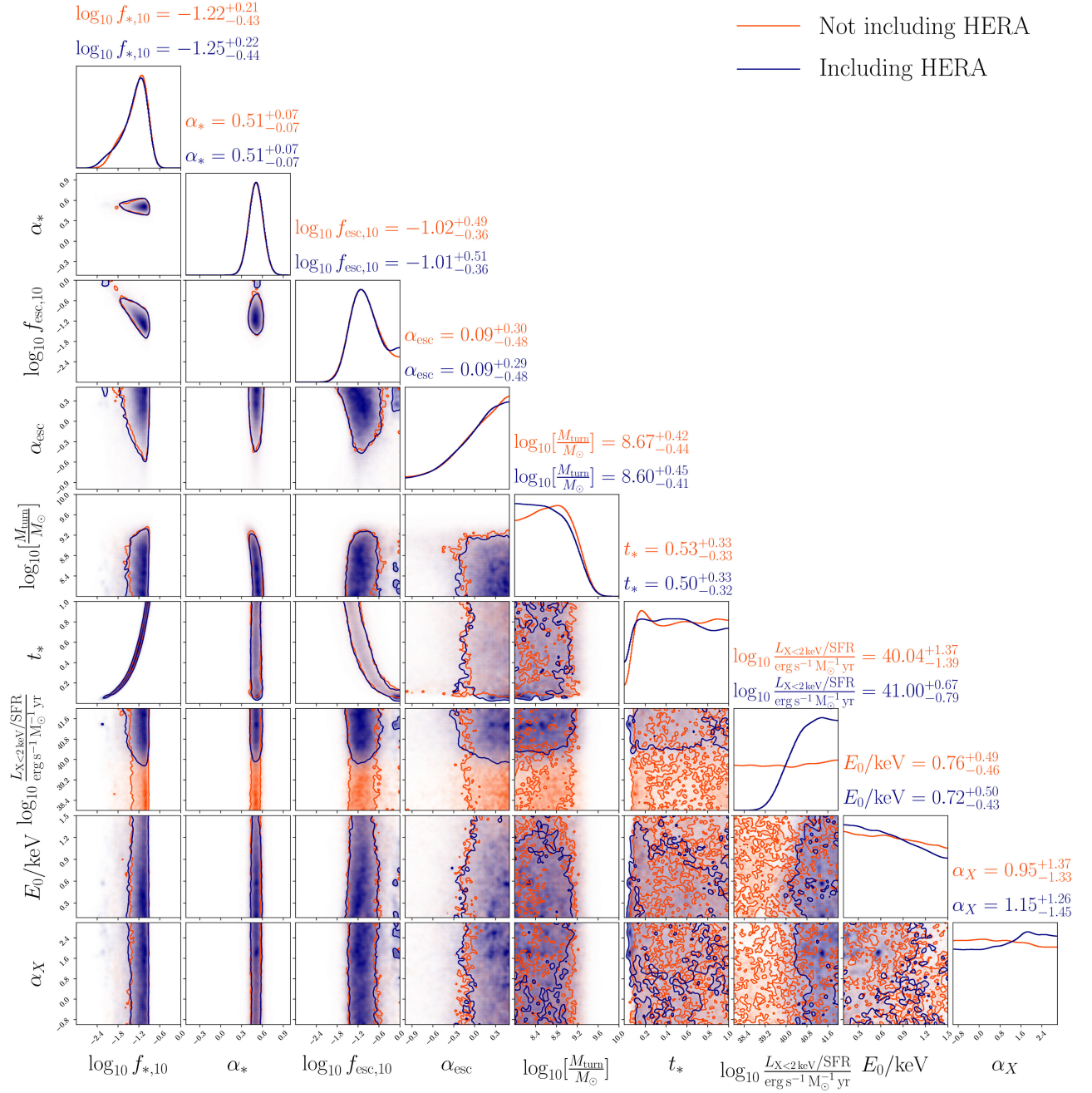


FIG. 7. Posterior distribution with and without HERA data. Posteriors without HERA data are produced using upper bound on the neutral fraction  $\bar{x}_{HI}$  at  $z = 5.9$ ,  $\tau_e$ —CMB optical depth, and galaxy UV luminosity function at redshift  $z = 6-10$  ( $M_{UV} > -20$ ). The most significant effect of adding the HERA data, is ruling out the possibility of low values of  $L_{X<2\text{ keV}}/\text{SFR}$  and setting the 68% HPD on  $L_{X<2\text{ keV}}/\text{SFR} = [10^{40.6}, 10^{42}] \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ , which is in good agreement with the results presented at [22],  $L_{X<2\text{ keV}}/\text{SFR} = [10^{40.4}, 10^{41.7}] \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ .

detail. In Ref. [29], we see that there is a sudden decline in the 1D posterior at the two ends for parameters  $\alpha_X$  and  $E_0$ , and at the highest end for  $t_*$  and  $L_{X<2\text{ keV}}/\text{SFR}$ . The reason for this in the case of  $L_{X<2\text{ keV}}/\text{SFR}$  was discussed previously, and it is likely due to the unequally weighted sample points in 21cmMC, and we suspect the same is happening here for the other parameters.

Figure 8 is the same as Fig. 7 but for inferences II-A and II-B. Here also most of the parameters are largely constrained by the external data, and HERA data only helps to put a lower bound on  $L_{X<2\text{ keV}}/\text{SFR}$ . Median values of most of the ACG parameters are almost similar to Fig. 7 except for  $f_{\text{esc},10}$ . With MCGs in the scenario, we find that a smaller median value of  $f_{\text{esc},10}$  is now preferred. This is

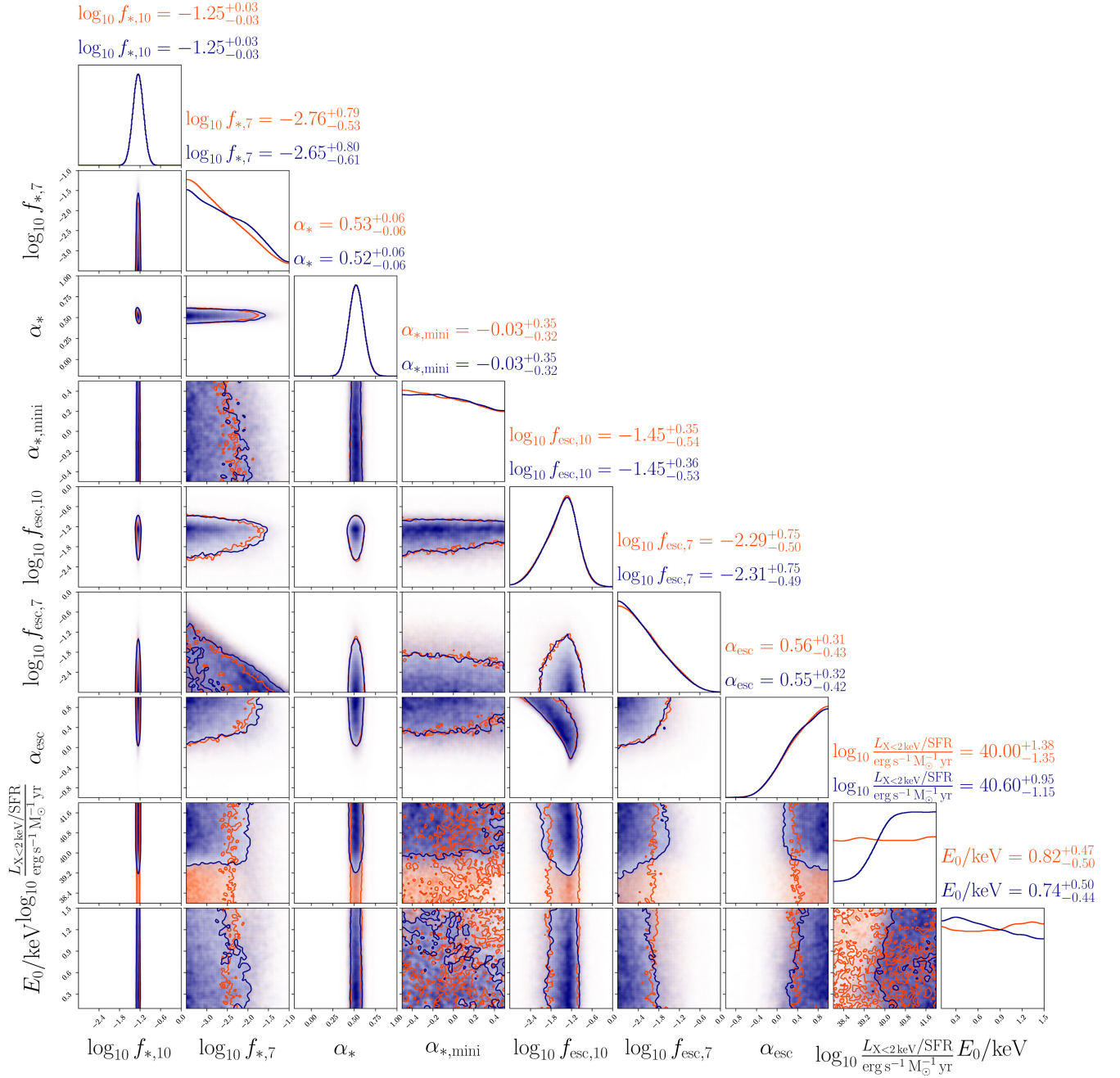


FIG. 8. Posterior distribution with and without HERA data, including MCGs. The most significant difference from what presented at Fig. 7, are the relaxed constraints on  $L_{X<2\text{keV}}/\text{SFR}$ , setting the 68% HPD on  $L_{X<2\text{keV}}/\text{SFR} = [10^{40}, 10^{42}] \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ , and the 95% HPD on  $[10^{38.6}, 10^{42}] \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ .

expected as less radiation is required to escape from the ACGs since MCGs also contribute to the ionizing photon budget. We also see that  $\log_{10} f_{*,10}$  here is more tightly constrained than in Fig. 7, although the median values are similar in both figures. Considering the MCG parameters, we find that the total likelihood places only upper bounds on  $\log_{10} f_{*,7}$  and  $\log_{10} f_{\text{esc},7}$ , and almost no statistically significant bound on  $\alpha_{*,\text{mini}}$ .

#### D. Considering different priors on MCGs parameters

As discussed in [45], it is important to note that the constraints over  $L_{X<2\text{keV}}/\text{SFR}$ , achieved when introducing MCGs, are highly dependent on the allowed values of  $\log_{10} f_{*,7}$  and  $\log_{10} f_{\text{esc},7}$ , i.e. on the choice of priors. In the above analysis, we did not allow  $\log_{10} f_{*,7}$  to be smaller than  $-3.5$ . When lower values are permitted, the external

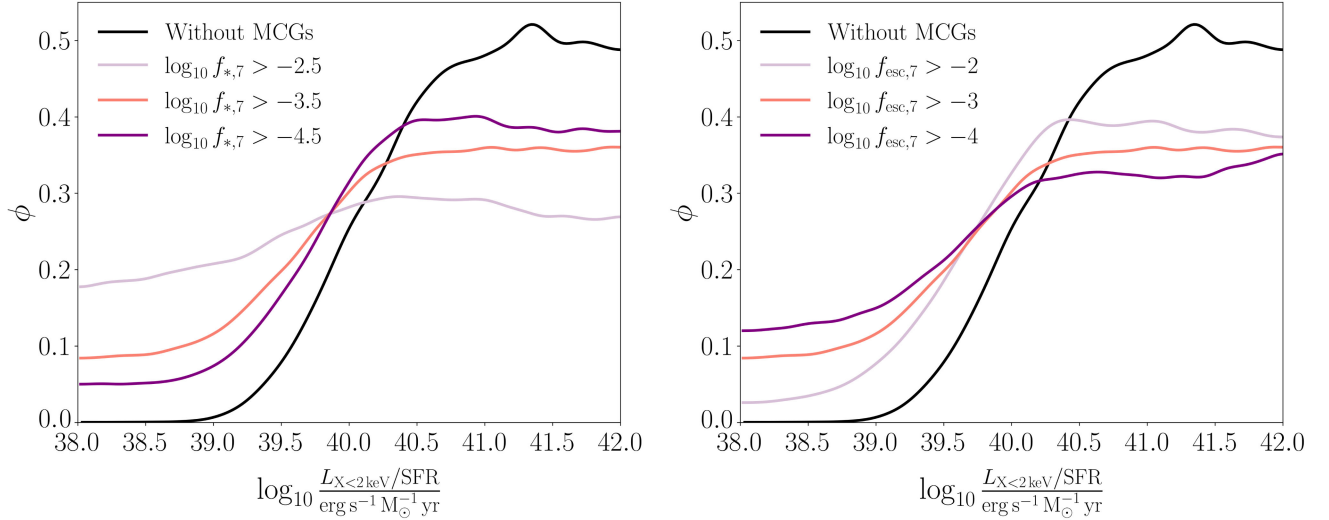


FIG. 9. Marginalized posterior distribution of  $L_{X<2 \text{ keV}}/\text{SFR}$  obtained when including MCGs, while varying the lower bound of  $\log_{10} f_{*,7}$  (left) or the lower bound of  $\log_{10} f_{\text{esc},7}$  (right). When lower values of  $\log_{10} f_{*,7}$  are enabled, the posterior distribution tends towards the results of I-B (ACGs only). In case we restrict  $\log_{10} f_{*,7}$  to a higher range, the  $L_{X<2 \text{ keV}}/\text{SFR}$  constraints tend to vanish, as less s-ray luminosity is now required to lower the amplitude of the signal at the redshifts measured by HERA. Similarly, when higher values of  $\log_{10} f_{\text{esc},7}$  are forced, the posterior distribution tends towards the results of I-B (ACGs only). If we allow  $\log_{10} f_{\text{esc},7}$  to be smaller, the  $L_{X<2 \text{ keV}}/\text{SFR}$  constraints tend to relax.

likelihoods push  $\log_{10} f_{*,7}$  towards them. The star-formation-rate density of MCGs depends linearly on  $f_{*,7}$  [47], thus it decreases as well. In turn, that decreases the number of x-ray producing remnants in MCGs, which increases the overall  $L_{X<2 \text{ keV}}/\text{SFR}$  required so that the power spectrum amplitude remains below the HERA upper bounds. The resulting constraints are very similar to the ones produced when taking only ACGs into account, as shown in Fig. 9 (left panel). The opposite effect occurs when increasing the lower bound on  $\log_{10} f_{*,7}$ . Higher values of  $\log_{10} f_{*,7}$ , i.e. higher MCGs star-formation-rate density, yield more x-ray producing sources, thus decreasing the required  $L_{X<2 \text{ keV}}/\text{SFR}$  for the signal to be consistent with the current HERA data. This behavior is illustrated as well in Fig. 9 (left panel).

Similar behavior is seen when the lower bounds on  $\log_{10} f_{\text{esc},7}$  are varied. Our fiducial value is  $\log_{10} f_{\text{esc},7} > -3$ . When  $\log_{10} f_{\text{esc},7}$  is forced to be larger, ionization occurs earlier, thus increasing the optical depth ( $\tau$  is calculated as an integral, where the integrand is proportional to  $1 - x_{\text{HI}}$ ). In order to fit for the optical depth constraint from Ref. [27],  $\log_{10} f_{*,7}$  must be smaller, and so the  $L_{X<2 \text{ keV}}/\text{SFR}$  posterior tends towards the one achieved in analysis I-B, as discussed above. In the opposite scenario, we enable lower values of  $\log_{10} f_{\text{esc},7}$ , and two related processes take place. Since  $\log_{10} f_{\text{esc},7}$  is smaller, ionization occurs later, such that the optical depth is smaller as well. In addition, the neutral fraction at  $z = 5.9$ , also increases when  $\log_{10} f_{\text{esc},7}$  is lowered. The constraints on these two observables force a higher value of  $\log_{10} f_{*,7}$ , and as discussed above this enhances the effect of including

MCGs in the analysis. The  $L_{X<2 \text{ keV}}/\text{SFR}$  posteriors in the different cases are shown in Fig. 9 (right panel). The bottom line is that the above analysis is prior dependent, and caution is warranted when reaching conclusions for models with MCGs.

## VII. CONCLUSIONS

In this work, we have introduced an ANN-based emulator of the 21-cm signal which was trained on data produced by the 21cmFAST seminumerical code. We found this emulator to be accurate enough to predict the 21-cm signal over a range of redshifts  $z$  and wave numbers  $k$ .

We then used this emulator in an MCMC pipeline to reproduce the parameter constraints presented in Refs. [22,29], based on the HERA phase-I upper limits on the 21-cm power spectrum and the external likelihoods containing information on the high-redshift UV luminosity function, the IGM neutral fraction and the reionization optical depth. Here, the underlying astrophysical model assumption is that the ACGs (which host the PopII stars) sparked the cosmic dawn and are responsible for the subsequent heating and reionization of the IGM. Our emulator-based MCMC pipeline seems to produce similar results as presented in Ref. [29], which validates our pipeline. One of the most important results of this analysis, based on the posterior of the parameter  $L_{X<2 \text{ keV}}/\text{SFR}$ , can be stated as follows. If PopII stars (or ACGs) dominate the x-ray heating and reionization of the IGM, then we expect the high redshift ( $z > 6$ ) galaxies to be more x-ray luminous (with  $L_{X<2 \text{ keV}}/\text{SFR} \gtrsim 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ ) than

their present-day counterparts (for which  $L_{X<2\text{ keV}}/\text{SFR} \approx 10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ ) [30,31].

Our primary goal in this work was to check the validity of the above result when we include MCGs in the analysis. A number of recent simulations show that MCGs, which predominantly host PopIII stars, also contribute to the total photon budget during cosmic dawn and reionization. Ignoring MCGs in any analysis could therefore lead to an overestimation of the x-ray and ionizing efficiencies of the ACGs. The same concern has also been mentioned in Ref. [22].

We therefore trained our emulator with the results obtained after including PopIII stars in the simulations, and finally ran an MCMC analysis that fits for additional parameters corresponding to the PopIII stars. The most interesting result that emerges out of the final posterior distributions we obtained is that including both ACGs and MCGs in the simulations relaxes the preference for the high x-ray luminosity of the high redshift sources as we discussed in the previous paragraph, and now values as low as  $L_{X<2\text{ keV}}/\text{SFR} \lesssim 10^{39} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  are still allowed by the HERA power spectrum data. This is due to the fact that MCGs contribute to the x-ray heating and therefore the ACGs do not need to be as x-ray efficient. This indicates that the x-ray luminosity of the high redshift sources may not be very different from their low redshift counterparts. It is important to note, that although we do see a small decline in the posterior distribution at the lower  $L_{X<2\text{ keV}}/\text{SFR}$  regime when including MCGs (as in Figs. 1 and 6) we cannot determine confidently that higher values are preferred. This is because the chosen likelihood function [see Eq. (8)] gives higher probability for power spectra well below the HERA upper bounds. In fact, as we see in Fig. 5, the spectra of this low x-ray signal is just below the HERA bound, making it a valid scenario. Contrary to that, the same figure shows that when MCGs are not included, the spectra obtained with large  $L_{X<2\text{ keV}}/\text{SFR}$  is well above the upper bound. As mentioned above, this result depends on the choice of priors, and specifically on the allowed ranges of  $\log_{10} f_{*,7}$  and  $\log_{10} f_{*,7}$ , and caution has to be taken when applying this conclusions.

Before closing, we check the validity of our results by comparing them with the 21-cm global signal  $\langle \delta T_{21} \rangle$  measured by the Shaped Antenna Measurement of the Background Radio Spectrum 3 (SARAS 3) [143] experiment in the frequency band 55 to 85 MHz. Note that SARAS 3 did not yield a statistically significant detection of  $\langle \delta T_{21} \rangle$ , rather they provide rms residuals of amplitude 213 mK which can be considered as upper and lower bounds of the signal. Figure 10 shows that our results do not contradict the SARAS 3 results. Note that existing global signal [83,143–146] measurements do not have much constraining power and hence are not included in our analysis. However, with more precise experiments in the future, global signal measurements will achieve more

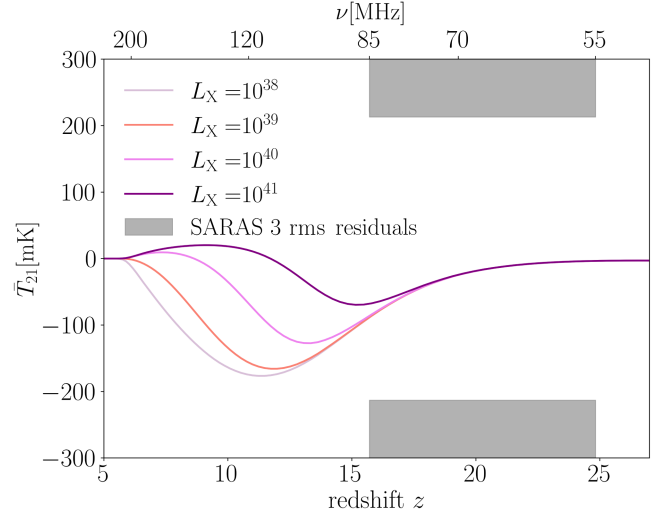


FIG. 10. Comparison of the 21-cm global signal obtained with median values from the posterior distribution from the inference which accounts for MCGs, and different  $L_{X<2\text{ keV}}/\text{SFR}$  values. These values are compared to the SARAS 3 brightness temperature bounds, demonstrating that introducing MCGs while enabling lower values of  $L_{X<2\text{ keV}}/\text{SFR}$  does not create a conflict with earlier cosmological measurements.

constraining power. In that case, our pipeline can be easily extended to emulate the global signals [52–54]. We shall explore this possibility in future work.

Note that in our analysis, we have assumed that the x-ray luminosities of both the PopII and PopIII stars are the same, i.e.  $L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR} = L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} = L_{X<2\text{ keV}}/\text{SFR}$ . We call  $L_{X<2\text{ keV}}/\text{SFR}$  as the effective x-ray luminosity. This choice is made just to reduce the total number of parameters. However, this choice affects the prior volume in a non-trivial manner. Making  $L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR} = L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR}$  automatically reduces the effective value of  $L_{X<2\text{ keV}}/\text{SFR}$ . This would not be the case if  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR}$  is made to vary independently. In this case, if the likelihood had preferred a lower value  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} \ll L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$ , we would have obtained the same conclusion as in Ref. [29]. However, this is unlikely as the theoretical modeling indicates that the low-metallicity stars, like PopIII, are expected to exhibit high  $L_{X<2\text{ keV}}/\text{SFR}$  [147], and we anticipate  $L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} \gtrsim L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR}$ . Therefore, we expect that our conclusions will still be valid under the assumption  $L_{X<2\text{ keV}}^{(\text{II})}/\text{SFR} = L_{X<2\text{ keV}}^{(\text{III})}/\text{SFR} = L_{X<2\text{ keV}}/\text{SFR}$ .

Future study will also try to address the limitations of the emulator used in this work, most significant of which is the inability to predict the flat or zero signals reliably. We have proposed a workaround considering the present scenario. However, with a more precise measurement of the power spectrum, we will need to predict signals of all kinds with better precision. Below, we discuss a few possible appro-

aches to improving our emulation pipeline: (i) Using generative adversarial networks [148], which implement some unsupervised learning methods to train models. (ii) Training the ANN on the likelihood function instead of the power spectrum. In that case, the ANN does not have to learn the complicated features of the signal. Rather, it learns to make decisions as to how acceptable the signal is based on the likelihood. (iii) Since in future experiments we expect to have more redshift bands, we can treat the power-spectra observations as two-dimensional data, where one dimension is the redshift, and the other is the wave number. We therefore can exploit the developing research in the field of convolutional neural networks [149,150], which uses for machine learning algorithms where the data has more than one dimension, for our goal. A different approach to the computational problem of producing samples of our model in a reasonable time, is using analytic codes like ZEUS21 [71]. This code provides good approximations for all of 21cmFAST (ACG-only) outputs, in  $\mathcal{O}(1-10)$  sec.

The conclusions of our work may have a number of consequences which we note below: (i) MCGs spark cosmic dawn and consequently, the Ly- $\alpha$  coupling starts early. This allows the spin temperature to catch up with the kinetic temperature at a higher redshift compared to the scenario where only ACGs exist. Now, a combination of less x-ray luminous ( $L_{X<2 \text{ keV}}/\text{SFR} \lesssim 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$ ) MCG and ACGs would make the absorption signal deeper and extended over several redshifts (compared to more luminous  $L_{X<2 \text{ keV}}/\text{SFR} > 10^{40} \text{ erg s}^{-1} M_{\odot}^{-1} \text{ yr}$  sources). This bodes well for global signal experiments as deeper and more extended absorption signals are easier to detect. (ii) Since the amplitude of the 21-cm power spectrum increases as we decrease  $L_{X<2 \text{ keV}}/\text{SFR}$ , this could also be fortunate for experiments like HERA and SKA that target 21-cm fluctuations. (iii) Low x-ray efficiency of the early sources provides room for other heating mechanisms to be effective, such as Ly- $\alpha$  heating due to resonant scattering of the Ly- $\alpha$  photons with the atoms in IGM [77,151].

It is crucial to take caution when interpreting results related to the modeling of the 21-cm signal at high redshifts, as these outcomes heavily hinge on the intricate modeling of various processes. Notably, different codes employing similar source populations but employing distinct algorithms may yield disparate results. 21cmFAST is one such model that uses a specific algorithm to model the cosmic dawn and reionization. It would be worth checking the consistency of the results presented here with other similar codes.

Finally, we emphasize that the emulator-based analysis pipeline developed in this work is flexible, fast and does not require a prohibitive number of simulations. We therefore plan to employ it to place 21-cm bounds on models of dark matter [152–154], primordial magnetic fields [155] and various other extensions to  $\Lambda$ CDM.

## ACKNOWLEDGMENTS

We thank Rennan Barkana, Daniela Breitman, Anastasia Fialkov, Jordan Flitter and Caner Unal for useful discussions. We are grateful to Andrei Mesinger and Julian Muñoz for insightful comments on the manuscript. We also acknowledge the efforts of the entire 21cmFAST team in producing a state-of-the-art 21-cm public simulation code. H. L. is supported by a Dkalim excellence M.Sc. scholarship from Ben-Gurion University. E. D. K. is supported by an Azrieli Foundation faculty fellowship. E. D. K. also acknowledges joint support from the U.S.-Israel Binational Science Foundation (BSF, Grant No. 2022743) and the U.S. National Science Foundation (NSF, Grant No. 2307354), and support from the ISF-NSFC joint research program (Grant No. 3156/23).

## APPENDIX A: EMULATORS FOR $\tau_e$ AND $x_{\text{HI}}$

The emulators for  $\tau_e$  and  $x_{\text{HI}}$  are built on the same structure described in Sec. IV. The dataset for each observable contain  $\sim 20000$  simulations where the parameters are sampled using the LH [117] sampler. The whole data are randomly divided into training (85%), validation (10%) and testing (5%) sets. The astrophysical parameters are normalized to fit in the range  $[-1, 1]$  for convenience. The NN here is composed of three layers of sizes 256, 512, 512, respectively. We have used the ReLU [120] activation function, and we have checked that it performs really well. As done for the emulators described in Sec. IV, we added a batch normalization layer [122] after each activation. For the loss function, we rely on the standard L2 norm, defined as

$$\mathcal{L}(y_{\text{true}}, y_{\text{pred}}(\theta)) = \frac{1}{m} \sum_{i=1}^m (y_{\text{true}} - y_{\text{pred}}(\theta))^2, \quad (\text{A1})$$

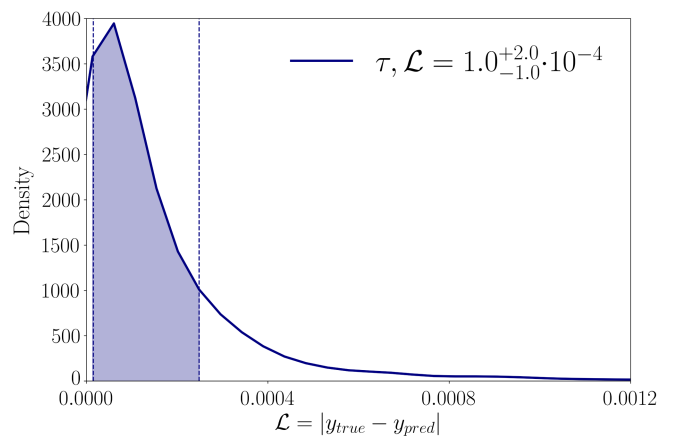


FIG. 11. Loss statistics for the  $\tau$  emulator based on testing sets having  $\sim 1000$  samples each. The dashed lines represent the 16th and 84th percentiles, respectively, and their exact value is written along with the median value in the figure legend. Note that the x-axis here is not normalized with the true values.

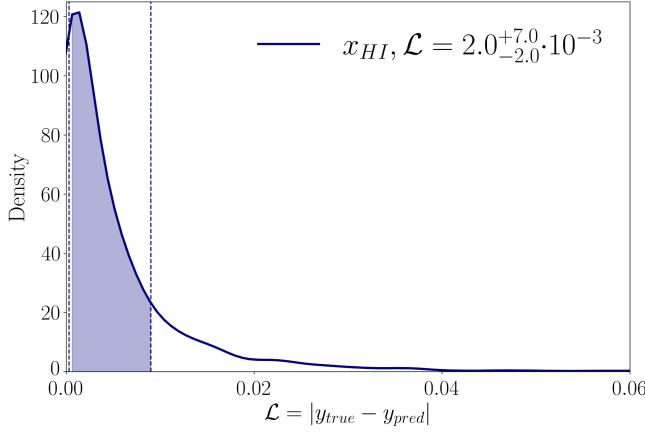


FIG. 12. Loss statistics for the  $x_{\text{HI}}$  emulator based on testing sets having  $\sim 1000$  samples each. The dashed lines represent the 16th and 84th percentiles, respectively, and their exact value is written along with the median value in the figure legend. Note that the x-axis here is not normalized with the true values.

where  $m$  is the number of samples in a batch which is taken to be 128. Here, the Adam [123] optimizer is used in the training with an initial learning rate of 0.01. Depending on the performance of the ANN on the validation set, we even reduce the learning rate by a factor of 2. In the end, we restore the best weights.

As can be seen from Figs. 11 and 12, the emulators perform really well. Even the rare errors are not significant. This convinces us that the constraints obtained from these observables will not suffer from emulation errors.

## APPENDIX B: EMULATOR WITH MCGs

Here we present the performance of our emulator where the simulations used to train it include both ACGs and MCGs. The pipeline and procedures used here are the same as described in Sec. IV, and we do not repeat these. We present the emulation results in Fig. 13. We find that the results are almost similar to what we found for the emulator where only ACGs are present.

## APPENDIX C: CLASSIFIERS

In this work, we are using a total of four different signal classifiers for the power spectrum. One for each redshift (7.9, 10.4), and for each redshift, there are two subclassifiers for emulation with (i) only ACGs and (ii) ACGs+MCGs. For each classification, we have used  $\sim 10000$  signals and divided those into training (85%) validation (10%) and testing (5%) sets. We have also normalized the

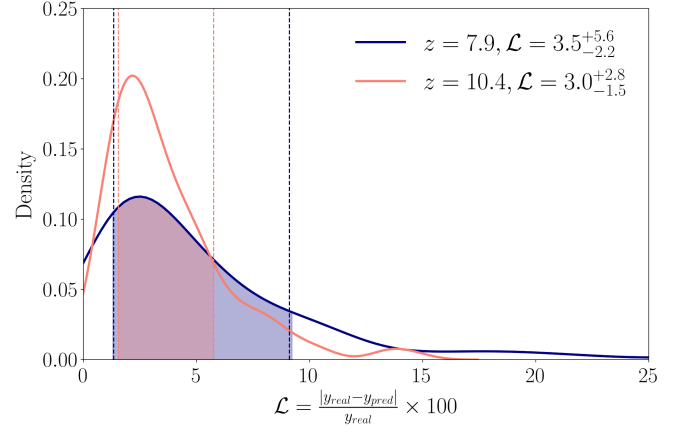


FIG. 13. Loss statistics for the power spectrum emulator, where MCGs are included in the simulations. Results are based on testing sets having  $\sim 250$  samples in each. The dashed lines represent the 16th and 84th percentiles, respectively, and their exact value is written along with the median value in the figure legend.

astrophysical parameters to fit the range  $[-1, 1]$  for faster and more efficient learning.

The NN for all the classifiers is composed of three layers of sizes 128, 512 and 256, respectively. For activation, we have used ReLU [120] followed by a batch normalization layer [122] which standardizes the input for the next layer, and a dropout [156] layer, which randomly turns off 10% of the neurons in the previous layer in order to prevent overfitting. The end layer contains one neuron that uses a “sigmoid” activation function. This is chosen in order to have the classifier output in the range  $\in [0, 1]$ . Here, an output smaller than 0.5 is classified as class 0 and  $\geq 0.5$  is classified as class 1 (as defined in Sec. IVA 5). We use the binary cross entropy loss function, which is the standard choice for classifying a dataset into two classes, defined as

$$\mathcal{L} = -\frac{1}{m} \sum_{i=1}^m y_i \log(P(y_i)) + (1 - y_i) \log(1 - P(y_i)), \quad (\text{C1})$$

where  $y_i$  are the labels,  $P(y_i)$  is the classifier probability prediction for this label, and  $m$  is the number of samples in a batch which we have chosen to be 512. We have used the Adam [123] optimizer with an initial learning rate of 0.01, which is reduced by a factor of 2 based on the performance. We store the best weights at the end. We found that the accuracy is  $\gtrsim 97\%$  for all the classifiers, and we can confidently use this in the MCMC process.



- [1] P. Madau, A. Meiksin, and M. J. Rees, 21-CM tomography of the intergalactic medium at high redshift, *Astrophys. J.* **475**, 429 (1997).
- [2] A. Loeb and M. Zaldarriaga, Measuring the small-scale power spectrum of cosmic density fluctuations through 21 cm tomography prior to the epoch of structure formation, *Phys. Rev. Lett.* **92**, 211301 (2004).
- [3] J. R. Pritchard and A. Loeb, 21-cm cosmology, *Rep. Prog. Phys.* **75**, 086901 (2012).
- [4] J. C. Pober, A. Liu, J. S. Dillon, J. E. Aguirre, J. D. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, J. N. Hewitt, D. C. Jacobs *et al.*, What next-generation 21 cm power spectrum measurements can teach us about the epoch of reionization, *Astrophys. J.* **782**, 66 (2014).
- [5] A. Liu *et al.* (Snowmass 2021 Cosmic Frontier 5 Topical Group), Snowmass2021 cosmic frontier white paper: 21 cm radiation as a probe of physics across cosmic ages, [arXiv:2203.07864](https://arxiv.org/abs/2203.07864).
- [6] S. Furlanetto, S. P. Oh, and F. Briggs, Cosmology at low frequencies: The 21 cm transition and the high-redshift universe, *Phys. Rep.* **433**, 181 (2006).
- [7] R. Barkana and A. Loeb, In the beginning: The first sources of light and the reionization of the Universe, *Phys. Rep.* **349**, 125 (2001).
- [8] R. Barkana and A. Loeb, Detecting the earliest galaxies through two new sources of 21 cm fluctuations, *Astrophys. J.* **626**, 1 (2005).
- [9] X. L. Chen and J. Miralda-Escude, The 21 cm signature of the first stars, *Astrophys. J.* **684**, 18 (2008).
- [10] J. Miralda-Escude, M. Haehnelt, and M. J. Rees, Reionization of the inhomogeneous universe, *Astrophys. J.* **530**, 1 (2000).
- [11] S. Furlanetto, M. Zaldarriaga, and L. Hernquist, The growth of HII regions during reionization, *Astrophys. J.* **613**, 1 (2004).
- [12] X. Chen, P. D. Meerburg, and M. Münchmeyer, The future of primordial features with 21 cm tomography, *J. Cosmol. Astropart. Phys.* **09** (2016) 023.
- [13] G. Mellema, L. V. E. Koopmans, F. A. Abdalla, G. Bernardi, B. Ciardi, S. Daiboo, A. G. de Bruyn, K. K. Datta, H. Falcke, A. Ferrara *et al.*, Reionization and the cosmic dawn with the square kilometre array, *Exp. Astron.* **36**, 235 (2013).
- [14] A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, P. Carroll, N. Barry, M. Rahimi, B. Pindor, C. M. Trott, J. Line, D. C. Jacobs *et al.*, First season MWA EoR power spectrum results at redshift 7, *Astrophys. J.* **833**, 102 (2016).
- [15] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh, An absorption profile centred at 78 megahertz in the sky-averaged spectrum, *Nature (London)* **555**, 67 (2018).
- [16] C. M. Trott and J. C. Pober, The status of 21 cm interferometric experiments, [arXiv:1909.12491](https://arxiv.org/abs/1909.12491).
- [17] B. Greig, C. M. Trott, N. Barry, S. J. Mutch, B. Pindor, R. L. Webster, and J. S. B. Wyithe, Exploring reionization and high- $z$  galaxy observables with recent multiredshift MWA upper limits on the 21-cm signal, *Mon. Not. R. Astron. Soc.* **500**, 5322 (2020).
- [18] F. G. Mertens, M. Mevius, L. V. E. Koopmans, A. R. Offringa, G. Mellema, S. Zaroubi, M. A. Brentjens, H. Gan, B. K. Gehlot, V. N. Pandey *et al.*, Improved upper limits on the 21-cm signal power spectrum of neutral hydrogen at  $z \approx 9.1$  from LOFAR, *Mon. Not. R. Astron. Soc.* **493**, 1662 (2020).
- [19] R. Ansari, Current status and future of cosmology with 21 cm intensity mapping, [arXiv:2203.17039](https://arxiv.org/abs/2203.17039).
- [20] D. R. DeBoer, A. R. Parsons, J. E. Aguirre, P. Alexander, Z. S. Ali, A. P. Beardsley, G. Bernardi, J. D. Bowman, R. F. Bradley, C. L. Carilli *et al.*, Hydrogen epoch of reionization array (HERA), *Publ. Astron. Soc. Pac.* **129**, 045001 (2017).
- [21] Z. Abdurashidova *et al.* (HERA Collaboration), First results from HERA Phase I: Upper limits on the epoch of reionization 21 cm power spectrum, *Astrophys. J.* **925**, 221 (2022).
- [22] Z. Abdurashidova *et al.* (HERA Collaboration), Improved constraints on the 21 cm EoR power spectrum and the x-ray heating of the IGM with HERA Phase I observations, *Astrophys. J.* **945**, 124 (2023).
- [23] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Trenti, I. Labbé, L. Bradley, M. Carollo, P. G. van Dokkum, V. Gonzalez, B. Holwerda *et al.*, UV luminosity functions at redshifts  $z \sim 4$  to  $z \sim 10$ : 10000 galaxies from HST legacy fields, *Astrophys. J.* **803**, 34 (2015).
- [24] R. J. Bouwens, P. A. Oesch, I. Labbé, G. D. Illingworth, G. G. Fazio, D. Coe, B. Holwerda, R. Smit, M. Stefanon, P. G. van Dokkum, M. Trenti, M. L. N. Ashby, J. S. Huang, L. Spitler, C. Straatman, L. Bradley, and D. Magee, The bright end of the  $z \sim 9$  and  $z \sim 10$  UV luminosity functions using all five CANDELS fields, *Astrophys. J.* **830**, 67 (2016).
- [25] P. A. Oesch, R. J. Bouwens, G. D. Illingworth, I. Labbé, and M. Stefanon, The dearth of  $z \sim 10$  galaxies in all HST legacy fields—The rapid evolution of the galaxy population in the first 500 Myr, *Astrophys. J.* **855**, 105 (2017).
- [26] N. J. F. Gillet, A. Mesinger, and J. Park, Combining high- $z$  galaxy luminosity functions with Bayesian evidence, *Mon. Not. R. Astron. Soc.* **491**, 1980 (2020).
- [27] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641**, A6 (2020); *Astron. Astrophys.* **652**, C4(E) (2021).
- [28] I. McGreer, A. Mesinger, and V. D’Odorico, Model-independent evidence in favour of an end to reionization by  $z \approx 6$ , *Mon. Not. R. Astron. Soc.* **447**, 499 (2015).
- [29] Z. Abdurashidova *et al.* (HERA Collaboration), HERA Phase I limits on the Cosmic 21 cm signal: Constraints on astrophysics and cosmology during the epoch of reionization, *Astrophys. J.* **924**, 51 (2022).
- [30] S. Mineo, M. Gilfanov, and R. Sunyaev, X-ray emission from star-forming galaxies—I. High-mass x-ray binaries, *Mon. Not. R. Astron. Soc.* **419**, 2095 (2012).
- [31] B. D. Lehmer, D. M. Alexander, F. E. Bauer, W. N. Brandt, A. D. Goulding, L. P. Jenkins, A. Ptak, and T. P. Roberts, A chandra perspective on galaxy-wide x-ray binary emission and its correlation with star formation rate and stellar mass: New results from luminous infrared galaxies, *Astrophys. J.* **724**, 559 (2010).
- [32] B. Smith, J. Wise, B. O’Shea, M. Norman, and S. Khochfar, The first Population II stars formed in externally enriched mini-haloes, *Mon. Not. R. Astron. Soc.* **452**, 2822 (2015).

- [33] V. Bromm, Formation of the first stars, *IAU Symp.* **228**, 121 (2005).
- [34] V. Bromm and R. B. Larson, The first stars, *Annu. Rev. Astron. Astrophys.* **42**, 79 (2004).
- [35] Z. Haiman, M. J. Rees, and A. Loeb, H(2) cooling of primordial gas triggered by UV irradiation, *Astrophys. J.* **467**, 522 (1996).
- [36] Z. Haiman, M. J. Rees, and A. Loeb, Destruction of molecular hydrogen during cosmological reionization, *Astrophys. J.* **476**, 458 (1997).
- [37] M. Tegmark, J. Silk, M. J. Rees, A. Blanchard, T. Abel, and F. Palla, How small were the first cosmological objects?, *Astrophys. J.* **474**, 1 (1997).
- [38] T. Abel, G. L. Bryan, and M. L. Norman, The formation of the first star in the Universe, *Science* **295**, 93 (2002).
- [39] N. Yoshida, K. Omukai, L. Hernquist, and T. Abel, Formation of primordial stars in a  $\Lambda$ CDM universe, *Astrophys. J.* **652**, 6 (2006).
- [40] Z. Haiman and G. L. Bryan, Was star-formation suppressed in high-redshift minihalos?, *Astrophys. J.* **650**, 7 (2006).
- [41] M. Trenti, Population III star formation during and after the reionization epoch, *AIP Conf. Proc.* **1294**, 134 (2010).
- [42] M. Kulkarni, E. Visbal, and G. L. Bryan, The critical dark matter halo mass for Population III star formation: Dependence on Lyman–Werner radiation, baryon-dark matter streaming velocity, and redshift, *Astrophys. J.* **917**, 40 (2021).
- [43] R. H. Mebane, J. Mirocha, and S. R. Furlanetto, The effects of Population III radiation backgrounds on the cosmological 21-cm signal, *Mon. Not. R. Astron. Soc.* **493**, 1217 (2020).
- [44] J. Mirocha, R. H. Mebane, S. R. Furlanetto, K. Singal, and D. Trinh, Unique signatures of Population III stars in the global 21-cm signal, *Mon. Not. R. Astron. Soc.* **478**, 5591 (2018).
- [45] Y. Qin, A. Mesinger, J. Park, B. Greig, and J. B. Muñoz, A tale of two sites—I. Inferring the properties of minihalo-hosted galaxies from current observations, *Mon. Not. R. Astron. Soc.* **495**, 123 (2020).
- [46] Y. Qin, A. Mesinger, B. Greig, and J. Park, A tale of two sites—II. Inferring the properties of minihalo-hosted galaxies with upcoming 21-cm interferometers, *Mon. Not. R. Astron. Soc.* **501**, 4748 (2021).
- [47] J. B. Muñoz, Y. Qin, A. Mesinger, S. G. Murray, B. Greig, and C. Mason, The impact of the first galaxies on cosmic dawn and reionization, *Mon. Not. R. Astron. Soc.* **511**, 3657 (2022).
- [48] H. Xu, K. Ahn, J. H. Wise, M. L. Norman, and B. W. O’Shea, Heating the intergalactic medium by x-rays from Population III binaries in high redshift galaxies, *Astrophys. J.* **791**, 110 (2014).
- [49] H. Xu, K. Ahn, M. L. Norman, J. H. Wise, and B. W. O’Shea, X-ray background at high redshifts from Pop III remnants: Results from Pop III star formation rates in the renaissance simulations, *Astrophys. J. Lett.* **832**, L5 (2016).
- [50] N. S. Kern, A. Liu, A. R. Parsons, A. Mesinger, and B. Greig, Emulating simulations of cosmic dawn for 21 cm power spectrum constraints on cosmology, reionization, and x-ray heating, *Astrophys. J.* **848**, 23 (2017).
- [51] W. D. Jennings, C. A. Watkinson, F. B. Abdalla, and J. D. McEwen, Evaluating machine learning techniques for predicting power spectra from reionization simulations, *Mon. Not. R. Astron. Soc.* **483**, 2907 (2019).
- [52] A. Cohen, A. Fialkov, R. Barkana, and R. Monsalve, Emulating the global 21-cm signal from cosmic dawn and reionization, *Mon. Not. R. Astron. Soc.* **495**, 4845 (2020).
- [53] C. H. Bye, S. K. N. Portillo, and A. Fialkov, 21 cmVAE: A very accurate emulator of the 21 cm global signal, *Astrophys. J.* **930**, 79 (2022).
- [54] H. T. J. Bevens, W. J. Handley, A. Fialkov, E. d. Acedo, and K. Javid, globalemu: A novel and robust approach for emulating the sky-averaged 21-cm signal from the cosmic dawn and epoch of reionization, *Mon. Not. R. Astron. Soc.* **508**, 2923 (2021).
- [55] M. Choudhury, A. Datta, and S. Majumdar, Extracting the 21-cm power spectrum and the reionization parameters from mock data sets using artificial neural networks, *Mon. Not. R. Astron. Soc.* **512**, 5010 (2022).
- [56] S. Sikder, R. Barkana, I. Reis, and A. Fialkov, Emulation of the cosmic dawn 21-cm power spectrum and classification of excess radio models using an artificial neural network, *Mon. Not. R. Astron. Soc.* **527**, 9977 (2023).
- [57] B. Ciardi, A. Ferrara, S. Marri, and G. Raimondo, Cosmological reionization around the first stars: Monte Carlo radiative transfer, *Mon. Not. R. Astron. Soc.* **324**, 381 (2001).
- [58] M. G. Santos, L. Ferramacho, M. B. Silva, A. Amblard, and A. Cooray, Fast and large volume simulations of the 21 cm signal from the reionization and pre-reionization epochs, *Mon. Not. R. Astron. Soc.* **406**, 2421 (2010).
- [59] P. R. Shapiro, I. T. Iliev, G. Mellema, K. Ahn, Y. Mao, M. Friedrich, K. Datta, H. Park, E. Komatsu, E. Fernandez *et al.*, Simulating cosmic reionization and the radiation backgrounds from the epoch of reionization, *AIP Conf. Proc.* **1480**, 248 (2012).
- [60] S. Majumdar, G. Mellema, K. K. Datta, H. Jensen, T. R. Choudhury, S. Bharadwaj, and M. M. Friedrich, On the use of seminumerical simulations in predicting the 21-cm signal from the epoch of reionization, *Mon. Not. R. Astron. Soc.* **443**, 2843 (2014).
- [61] S. Hassan, R. Davé, K. Finlator, and M. G. Santos, Simulating the 21 cm signal from reionization including non-linear ionizations and inhomogeneous recombinations, *Mon. Not. R. Astron. Soc.* **457**, 1550 (2016).
- [62] D. Sarkar, S. Bharadwaj, and S. Ananthpindika, Modelling the post-reionization neutral hydrogen (H i) bias, *Mon. Not. R. Astron. Soc.* **460**, 4310 (2016).
- [63] R. Ghara, G. Mellema, S. K. Giri, T. R. Choudhury, K. K. Datta, and S. Majumdar, Prediction of the 21-cm signal from reionization: comparison between 3D and 1D radiative transfer schemes, *Mon. Not. R. Astron. Soc.* **476**, 1741 (2018).
- [64] A. Mesinger, Reionization and cosmic dawn: Theory and simulations, *IAU Symp.* **333**, 3 (2017).
- [65] D. Sarkar and S. Bharadwaj, Modelling redshift space distortion in the post-reionization H i 21-cm power spectrum, *Mon. Not. R. Astron. Soc.* **476**, 96 (2018).

- [66] P. Ocvirk, D. Aubert, J. G. Sorce, P. R. Shapiro, N. Deparis, T. Dawoodbhoy, J. Lewis, R. Teyssier, G. Yepes, S. Gottlöber *et al.*, Cosmic Dawn II (CoDa II): A new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization, *Mon. Not. R. Astron. Soc.* **496**, 4087 (2020).
- [67] D. Sarkar and S. Bharadwaj, Redshift-space distortions of the H I 21-cm intensity mapping signal due to the internal motions within galaxies, *Mon. Not. R. Astron. Soc.* **487**, 5666 (2019).
- [68] M. Molaro, R. Davé, S. Hassan, M. G. Santos, and K. Finlator, ARTIST: Fast radiative transfer for large-scale simulations of the epoch of reionisation, *Mon. Not. R. Astron. Soc.* **489**, 5594 (2019).
- [69] R. Kannan, E. Garaldi, A. Smith, R. Pakmor, V. Springel, M. Vogelsberger, and L. Hernquist, Introducing the thesan project: Radiation-magnetohydrodynamic simulations of the epoch of reionization, *Mon. Not. R. Astron. Soc.* **511**, 4005 (2022).
- [70] A. K. Shaw, A. Chakraborty, M. Kamran, R. Ghara, S. Choudhuri, S. S. Ali, S. Pal, A. Ghosh, J. Kumar, P. Dutta *et al.*, Probing early Universe through redshifted 21-cm signal: Modeling and observational challenges, *J. Astrophys. Astron.* **44**, 4 (2023).
- [71] J. B. Muñoz, An effective model for the cosmic-dawn 21-cm signal, *Mon. Not. R. Astron. Soc.* **523**, 2587 (2023).
- [72] B. Maity and T. R. Choudhury, Efficient exploration of reionization parameters for the upcoming 21 cm observations using a photon-conserving seminumerical model SCRIPT, *Mon. Not. R. Astron. Soc.* **521**, 4140 (2023).
- [73] J. B. Muñoz, Standard ruler at cosmic dawn, *Phys. Rev. Lett.* **123**, 131301 (2019).
- [74] J. B. Muñoz, Robust velocity-induced acoustic oscillations at cosmic dawn, *Phys. Rev. D* **100**, 063538 (2019).
- [75] A. Mesinger, S. Furlanetto, and R. Cen, 21cmFAST: A fast, semi-numerical simulation of the high-redshift 21-cm signal, *Mon. Not. R. Astron. Soc.* **411**, 955 (2011).
- [76] J. Mirocha, J. B. Muñoz, S. R. Furlanetto, A. Liu, and A. Mesinger, A galaxy-free phenomenological model for the 21-cm power spectrum during reionization, *Mon. Not. R. Astron. Soc.* **514**, 2010 (2022).
- [77] I. Reis, A. Fialkov, and R. Barkana, The subtlety of Ly  $\alpha$  photons: Changing the expected range of the 21-cm signal, *Mon. Not. R. Astron. Soc.* **506**, 5479 (2021).
- [78] J. Dowell and G. B. Taylor, The radio background below 100 MHz, *Astrophys. J. Lett.* **858**, L9 (2018).
- [79] D. J. Fixsen, A. Kogut, S. Levin, M. Limon, P. Lubin, P. Mirel, M. Seiffert, J. Singal, E. Wollack, T. Villela *et al.*, ARCADE 2 measurement of the extra-galactic sky temperature at 3–90 GHz, *Astrophys. J.* **734**, 5 (2011).
- [80] S. Bharadwaj and S. S. Ali, The CMBR fluctuations from HI perturbations prior to reionization, *Mon. Not. R. Astron. Soc.* **352**, 142 (2004).
- [81] S. Furlanetto, A. Lidz, A. Loeb, M. McQuinn, J. Pritchard, J. Aguirre, M. Alvarez, D. Backer, J. Bowman, J. Burns *et al.*, Astrophysics from the highly-redshifted 21 cm line, *arXiv:0902.3011*.
- [82] P. Madau and F. Haardt, Cosmic reionization after Planck: Could quasars do it all?, *Astrophys. J. Lett.* **813**, L8 (2015).
- [83] R. A. Monsalve, A. E. E. Rogers, J. D. Bowman, and T. J. Mozdzen, Results from EDGES high-band: I. Constraints on phenomenological models for the global 21 cm signal, *Astrophys. J.* **847**, 64 (2017).
- [84] H. T. J. Bevins, A. Fialkov, E. d. Acedo, W. J. Handley, S. Singh, R. Subrahmanyam, and R. Barkana, Astrophysical constraints from the SARAS 3 non-detection of the cosmic dawn sky-averaged 21-cm signal, *Nat. Astron.* **6**, 1473 (2022).
- [85] D. Vrbanec, B. Ciardi, V. Jelic, H. Jensen, I. T. Iliev, G. Mellema, and S. Zaroubi, Predictions for the 21 cm–galaxy cross-power spectrum observable with SKA and future galaxy surveys, *Mon. Not. R. Astron. Soc.* **492**, 4952 (2020).
- [86] D. Karagiannis, J. Fonseca, R. Maartens, and S. Camera, Probing primordial non-Gaussianity with the power spectrum and bispectrum of future 21 cm intensity maps, *Phys. Dark Universe* **32**, 100821 (2021).
- [87] M. Rahimi, B. Pindor, J. L. B. Line, N. Barry, C. M. Trott, R. L. Webster, C. H. Jordan, M. Wilensky, S. Yoshiura, A. Beardsley *et al.*, Epoch of reionization power spectrum limits from Murchison widefield array data targeted at EoR1 field, *Mon. Not. R. Astron. Soc.* **508**, 5954 (2021).
- [88] J. H. Cook, C. M. Trott, and J. L. B. Line, Investigating the contribution of extended radio sources to the Epoch of reionization power spectrum, *Mon. Not. R. Astron. Soc.* **514**, 790 (2022).
- [89] M. Berti, M. Spinelli, and M. Viel, Multipole expansion for 21 cm intensity mapping power spectrum: Forecasted cosmological parameters estimation for the SKA observatory, *Mon. Not. R. Astron. Soc.* **521**, 3221 (2023).
- [90] J. S. W. Lewis, A. Pillepich, D. Nelson, R. S. Klessen, and S. C. O. Glover, The signature of galaxy formation models in the power spectrum of the hydrogen 21 cm line during reionization, *Mon. Not. R. Astron. Soc.*, [10.1093/mnras/stad3903](https://doi.org/10.1093/mnras/stad3903) (2023).
- [91] S. Yoshiura, H. Shimabukuro, K. Takahashi, R. Momose, H. Nakanishi, and H. Imai, Sensitivity for 21 cm bispectrum from Epoch of reionization, *Mon. Not. R. Astron. Soc.* **451**, 266 (2015).
- [92] H. Shimabukuro, S. Yoshiura, K. Takahashi, S. Yokoyama, and K. Ichiki, 21 cm line bispectrum as a method to probe cosmic dawn and epoch of reionization, *Mon. Not. R. Astron. Soc.* **458**, 3003 (2016).
- [93] D. Sarkar, S. Majumdar, and S. Bharadwaj, Modelling the post-reionization neutral hydrogen (HI) 21-cm bispectrum, *Mon. Not. R. Astron. Soc.* **490**, 2880 (2019).
- [94] R. Durrer, M. Jalilvand, R. Kothari, R. Maartens, and F. Montanari, Full-sky bispectrum in redshift space for 21 cm intensity maps, *J. Cosmol. Astropart. Phys.* **12** (2020) 003.
- [95] S. Cunnington, C. Watkinson, and A. Poursidou, The H I intensity mapping bispectrum including observational effects, *Mon. Not. R. Astron. Soc.* **507**, 1623 (2021).
- [96] M. Kamran, S. Majumdar, R. Ghara, G. Mellema, S. Bharadwaj, J. R. Pritchard, R. Mondal, and I. T. Iliev, Probing IGM physics during cosmic dawn using the redshifted 21-cm bispectrum, *arXiv:2108.08201*.
- [97] D. Karagiannis, R. Maartens, and L. F. Randrianjanahary, Cosmological constraints from the power spectrum and

- bispectrum of 21 cm intensity maps, *J. Cosmol. Astropart. Phys.* **11** (2022) 003.
- [98] M. Tegmark, How to measure CMB power spectra without losing information, *Phys. Rev. D* **55**, 5895 (1997).
- [99] A. Liu and M. Tegmark, A method for 21 cm power spectrum estimation in the presence of foregrounds, *Phys. Rev. D* **83**, 103006 (2011).
- [100] J. S. Dillon, A. Liu, C. L. Williams, J. N. Hewitt, M. Tegmark, E. H. Morgan, A. M. Levine, M. F. Morales, S. J. Tingay, G. Bernardi *et al.*, Overcoming real-world obstacles in 21 cm power spectrum estimation: A method demonstration and results from early Murchison Widefield Array data, *Phys. Rev. D* **89**, 023002 (2014).
- [101] A. Liu and J. R. Shaw, Data analysis for precision 21 cm cosmology, *Publ. Astron. Soc. Pac.* **132**, 062001 (2020).
- [102] S. Cunnington, L. Wolz, A. Pourtsidou, and D. Bacon, Impact of foregrounds on HI intensity mapping cross-correlations with optical surveys, *Mon. Not. R. Astron. Soc.* **488**, 5452 (2019).
- [103] M. Kolopanis, D. C. Jacobs, C. Cheng, A. R. Parsons, S. A. Kohn, J. C. Pober, J. E. Aguirre, Z. S. Ali, G. Bernardi, R. F. Bradley *et al.*, A simplified, lossless re-analysis of PAPER-64, *Astrophys. J.* **883**, 133 (2019).
- [104] Y. Qin, V. Poulin, A. Mesinger, B. Greig, S. Murray, and J. Park, Reionization inference from the CMB optical depth and E-mode polarization power spectra, *Mon. Not. R. Astron. Soc.* **499**, 550 (2020).
- [105] J. Silk, A. Di Cintio, and I. Dvorkin, Galaxy formation, *Proceedings of the International School of Physics "Enrico Fermi"* **186**, 137 (2014).
- [106] M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel, and L. Hernquist, A model for cosmological simulations of galaxy formation physics, *Mon. Not. R. Astron. Soc.* **436**, 3031 (2013).
- [107] R. S. Somerville and R. Davé, Physical models of galaxy formation in a cosmological framework, *Annu. Rev. Astron. Astrophys.* **53**, 51 (2015).
- [108] A. Pillepich, V. Springel, D. Nelson, S. Genel, J. Naiman, R. Pakmor, L. Hernquist, P. Torrey, M. Vogelsberger, R. Weinberger *et al.*, Simulating galaxy formation with the IllustrisTNG model, *Mon. Not. R. Astron. Soc.* **473**, 4077 (2018).
- [109] M. Vogelsberger, F. Marinacci, P. Torrey, and E. Puchwein, Cosmological simulations of galaxy formation, *Nat. Rev. Phys.* **2**, 42 (2020).
- [110] A. Wetzel, C. C. Hayward, R. E. Sanderson, X. Ma, D. Angles-Alcazar, R. Feldmann, T. K. Chan, K. El-Badry, C. Wheeler, S. Garrison-Kimmel *et al.*, Public data release of the FIRE-2 cosmological zoom-in simulations of galaxy formation, *Astrophys. J. Suppl. Ser.* **265**, 44 (2023).
- [111] S. G. Murray, B. Greig, A. Mesinger, J. B. Muñoz, Y. Qin, J. Park, and C. A. Watkinson, 21 cmFAST v3: A Python-integrated C code for generating 3D realizations of the cosmic 21 cm signal, *J. Open Source Software* **5**, 2582 (2020).
- [112] D. Sarkar and E. D. Kovetz, Measuring the cosmic expansion rate using 21-cm velocity acoustic oscillations, *Phys. Rev. D* **107**, 023524 (2023).
- [113] G. Shmueli, D. Sarkar, and E. D. Kovetz, Mitigating the optical depth degeneracy in the cosmological measurement of neutrino masses using 21-cm observations, *Phys. Rev. D* **108**, 083531 (2023).
- [114] C. Mason, M. Trenti, and T. Treu, The galaxy UV luminosity function before the epoch of reionization, *Astrophys. J.* **813**, 21 (2015).
- [115] J. Park, A. Mesinger, B. Greig, and N. Gillet, Inferring the astrophysics of reionization and cosmic dawn from galaxy luminosity functions and the 21-cm signal, *Mon. Not. R. Astron. Soc.* **484**, 933 (2019).
- [116] A. Rudakovskiy, A. Mesinger, D. Savchenko, and N. Gillet, Constraints on warm dark matter from UV luminosity functions of high- $z$  galaxies with Bayesian model comparison, *Mon. Not. R. Astron. Soc.* **507**, 3046 (2021).
- [117] M. D. McKay, R. J. Beckman, and W. J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* **21**, 239 (1979).
- [118] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow *et al.*, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015), Software available from: <https://www.tensorflow.org/>.
- [119] F. Chollet *et al.*, Keras (2015), Software available from: <https://keras.io>.
- [120] K. Fukushima, Visual feature extraction by a multilayered network of analog threshold elements, *IEEE Trans. Syst. Sci. Cybern.* **5**, 322 (1969).
- [121] A. Spurio Mancini, D. Piras, J. Alsing, B. Joachimi, and M. P. Hobson, CosmoPower: Emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys, *Mon. Not. R. Astron. Soc.* **511**, 1771 (2022).
- [122] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, [arXiv:1502.03167](https://arxiv.org/abs/1502.03167).
- [123] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [124] B. W. O'Shea and M. L. Norman, Population III star formation in a Lambda CDM universe. 2. Effects of a photo-dissociating background, *Astrophys. J.* **673**, 14 (2008).
- [125] C. A. Mason, A. Fontana, T. Treu, K. B. Schmidt, A. Hoag, L. Abramson, R. Amorin, M. Bradac, L. Guaita, T. Jones *et al.*, Inferences on the timeline of reionization at  $z \sim 8$  from the KMOS lens-amplified spectroscopic survey, *Mon. Not. R. Astron. Soc.* **485**, 3947 (2019).
- [126] T. R. Choudhury, A. Paranjape, and S. E. I. Bosman, Studying the Lyman- $\alpha$  optical depth fluctuations at  $z \sim 5.5$  using fast semi-numerical methods, *Mon. Not. R. Astron. Soc.* **501**, 5782 (2021).
- [127] E. Visbal, Z. Haiman, and G. L. Bryan, Limits on Population III star formation in minihaloes implied by Planck, *Mon. Not. R. Astron. Soc.* **453**, 4456 (2015).
- [128] A. Fialkov, R. Barkana, E. Visbal, D. Tseliakhovich, and C. M. Hirata, The 21-cm signature of the first stars during the Lyman-Werner feedback era, *Mon. Not. R. Astron. Soc.* **432**, 2909 (2013).
- [129] E. Visbal, Z. Haiman, B. Terrazas, G. L. Bryan, and R. Barkana, High-redshift star formation in a time-dependent Lyman-Werner background, *Mon. Not. R. Astron. Soc.* **445**, 107 (2014).

- [130] C. Safranek-Shrader, M. Agarwal, C. Federrath, A. Dubey, M. Milosavljevic, and V. Bromm, Star formation in the first Galaxies I: Collapse delayed by Lyman-Werner radiation, *Mon. Not. R. Astron. Soc.* **426**, 1159 (2012).
- [131] M. Ricotti, N. Y. Gnedin, and J. M. Shull, Feedback from galaxy formation: Production and photodissociation of primordial molecular hydrogen, *Astrophys. J.* **560**, 580 (2001).
- [132] K. Ahn, P. R. Shapiro, I. T. Iliev, G. Mellema, and U. L. Pen, The inhomogeneous background of hydrogen-molecule dissociating radiation during cosmic reionization, *Astrophys. J.* **695**, 1430 (2009).
- [133] A. Fialkov, Supersonic relative velocity between dark matter and baryons: A review, *Int. J. Mod. Phys. D* **23**, 1430017 (2014).
- [134] R. Barkana, The rise of the first stars: Supersonic streaming, radiative feedback, and 21-cm cosmology, *Phys. Rep.* **645**, 1 (2016).
- [135] D. Tseliakhovich and C. Hirata, Relative velocity of dark matter and baryonic fluids and the formation of the first structures, *Phys. Rev. D* **82**, 083520 (2010).
- [136] J. Bovy and C. Dvorkin, Low-mass suppression of the satellite luminosity function due to the supersonic baryon-cold-dark-matter relative velocity, *Astrophys. J.* **768**, 70 (2013).
- [137] A. Stacy, V. Bromm, and A. Loeb, Effect of streaming motion of baryons relative to dark matter on the formation of the first stars, *Astrophys. J. Lett.* **730**, L1 (2011).
- [138] A. Fialkov, R. Barkana, D. Tseliakhovich, and C. M. Hirata, Impact of the relative motion between the dark matter and baryons on the first stars, *Mon. Not. R. Astron. Soc.* **424**, 1335 (2012).
- [139] F. Schmidt, Effect of relative velocity and density perturbations between baryons and dark matter on the clustering of galaxies, *Phys. Rev. D* **94**, 063508 (2016).
- [140] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC Hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [141] F. Feroz, M. P. Hobson, and M. Bridges, MultiNest: An efficient and robust Bayesian inference tool for cosmology and particle physics, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2009).
- [142] Y. Qin (private communication).
- [143] S. Singh, J. Nambissan T., R. Subrahmanyam, N. Udaya Shankar, B. S. Girish, A. Raghunathan, R. Somashekar, K. S. Srivani, and M. Sathyanarayana Rao, On the detection of a cosmic dawn signal in the radio background, *Nat. Astron.* **6**, 607 (2022).
- [144] J. D. Bowman and A. E. E. Rogers, A lower limit of  $dz > 0.06$  for the duration of the reionization epoch, *Nature (London)* **468**, 796 (2010).
- [145] R. A. Monsalve, B. Greig, J. D. Bowman, A. Mesinger, A. E. E. Rogers, T. J. Mozdzen, N. S. Kern, and N. Mahesh, Results from EDGES High-Band: II. Constraints on parameters of early galaxies, *Astrophys. J.* **863**, 11 (2018).
- [146] R. A. Monsalve, A. Fialkov, J. D. Bowman, A. E. E. Rogers, T. J. Mozdzen, A. Cohen, R. Barkana, and N. Mahesh, Results from EDGES high-band: III. New constraints on parameters of the early universe, *Astrophys. J.* **875**, 67 (2019).
- [147] T. Fragos, B. Lehmer, M. Tremmel, P. Tzanavaris, A. Basu-Zych, K. Belczynski, A. Hornschemeier, L. Jenkins, V. Kalogera, A. Ptak *et al.*, X-ray binary evolution across cosmic time, *Astrophys. J.* **764**, 41 (2013).
- [148] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial networks, [arXiv:1406.2661](https://arxiv.org/abs/1406.2661).
- [149] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (2011), Vol. 22, pp. 1237–1242, [10.5591/978-1-57735-516-8/IJCAI11-210](https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-210).
- [150] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, Going deeper with convolutions, [arXiv:1409.4842](https://arxiv.org/abs/1409.4842).
- [151] D. Sarkar, J. Flitter, and E. D. Kovetz, Exploring delaying and heating effects on the 21-cm signature of fuzzy dark matter, *Phys. Rev. D* **105**, 103529 (2022).
- [152] E. D. Kovetz, V. Poulin, V. Gluscevic, K. K. Boddy, R. Barkana, and M. Kamionkowski, Tighter limits on dark matter explanations of the anomalous EDGES 21 cm signal, *Phys. Rev. D* **98**, 103529 (2018).
- [153] E. D. Kovetz, I. Cholis, and D. E. Kaplan, Bounds on ultralight hidden-photon dark matter from observation of the 21 cm signal at cosmic dawn, *Phys. Rev. D* **99**, 123511 (2019).
- [154] J. Flitter and E. D. Kovetz, Closing the window on fuzzy dark matter with the 21-cm signal, *Phys. Rev. D* **106**, 063504 (2022).
- [155] T. Adi, S. Libanore, H. A. G. Cruz, and E. D. Kovetz, Constraining primordial magnetic fields with line-intensity mapping, *J. Cosmol. Astropart. Phys.* **09** (2023) 035.
- [156] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, [arXiv:1207.0580](https://arxiv.org/abs/1207.0580).