



# Estimation of H II Bubble Size Distribution from 21 cm Power Spectrum with Artificial Neural Networks

Hayato Shimabukuro<sup>1,2</sup> , Yi Mao<sup>2</sup> , and Jianrong Tan<sup>2,3</sup> 

<sup>1</sup> Yunnan University, SWIFAR, No. 2 North Green Lake Road, Kunming 650500, China; [shimabukuro@ynu.edu.cn](mailto:shimabukuro@ynu.edu.cn) (HS)

<sup>2</sup> Department of Astronomy, Tsinghua University, Beijing 100084, China; [ymao@tsinghua.edu.cn](mailto:ymao@tsinghua.edu.cn) (YM)

<sup>3</sup> Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, United States of America

Received 2021 December 14; revised 2022 January 12; accepted 2022 January 17; published 2022 February 28

## Abstract

The bubble size distribution of ionized hydrogen regions probes information about the morphology of H II bubbles during reionization. Conventionally, the H II bubble size distribution can be derived from the tomographic imaging data of the redshifted 21 cm signal from the epoch of reionization, which, however, is observationally challenging even for upcoming large radio interferometer arrays. Given that these interferometers promise to measure the 21 cm power spectrum accurately, we propose a new method, which is based on artificial neural networks, to reconstruct the H II bubble size distribution from the 21 cm power spectrum. We demonstrate that reconstruction from the 21 cm power spectrum can be almost as accurate as being directly measured from the imaging data with fractional error  $\lesssim 10\%$ , even with thermal noise at the sensitivity level of the Square Kilometre Array. Nevertheless, the reconstruction implicitly exploits the modeling in reionization simulations, and hence the recovered H II bubble size distribution is not an independent summary statistic from the power spectrum, and should be used only as an indicator for understanding H II bubble morphology and its evolution.

*Key words:* methods: data analysis – methods: numerical – (cosmology:) dark ages – reionization – first stars – (cosmology:) diffuse radiation – cosmology: theory

## 1. Introduction

The epoch of reionization (EOR) is a unique period of time in cosmic evolution, during which ultraviolet (UV) and X-ray photons emitted from the first luminous objects (e.g., first stars and galaxies) ionize hydrogen atoms first in the surrounding intergalactic medium (IGM) and form bubbles of H II regions, and eventually these H II bubbles fill the whole Universe by  $z \simeq 6$  (e.g., Fan et al. 2006).

To unveil the nature of cosmic reionization, the cosmic 21 cm background has emerged as a promising probe of the EOR. The 21 cm line of atomic hydrogen results from the hyperfine transition due to spin coupling (Scott & Rees 1990; Madau et al. 1997). The tomographic images of 21 cm brightness temperature can directly tell the spatial distribution of H II bubbles and the complete history of cosmic reionization (e.g., Furlanetto et al. 2006; Pritchard & Loeb 2012). However, making three-dimensional (3D) 21 cm maps requires high sensitivity and spatial resolution, so it is technically extremely difficult. Instead, ongoing large radio interferometer array experiments, e.g., the Giant Metrewave Radio Telescope (GMRT)<sup>4</sup>, the LOw Frequency Array (LOFAR)<sup>5</sup>, the Murchison Widefield Array (MWA)<sup>6</sup> and

the Precision Array for Probing the Epoch of Reionization (PAPER)<sup>7</sup>, have first attempted to detect the 21 cm power spectrum from the EOR, a two-point statistic of 21 cm brightness temperature fluctuations (e.g., Furlanetto et al. 2006; Pritchard & Loeb 2012), and have put upper limits on the 21 cm power spectrum (Paciga et al. 2013; Yatawatta et al. 2013; Tingay et al. 2013; Patil et al. 2014, 2017; Parsons et al. 2014; Jelić et al. 2014; Ali et al. 2015; Dillon et al. 2015; Jacobs et al. 2015; Pober et al. 2015; Mertens et al. 2020). Furthermore, upcoming experiments such as the Square Kilometre Array (SKA)<sup>8</sup> (Mellema et al. 2013; Koopmans et al. 2015) and the Hydrogen Epoch of Reionization Array (HERA)<sup>9</sup> (DeBoer et al. 2017) promise to measure the 21 cm power spectrum from the EOR for the first time and with high sensitivity (The HERA Collaboration et al. 2021).

Understanding the morphology and topology of ionized bubbles (Mellema et al. 2015; Kulkarni et al. 2016, 2017; Hassan et al. 2018) is a key question related to the EOR. While the topological features of the 21 cm maps can be described by Minkowski functionals (Gleser et al. 2006; Lee et al. 2008; Friedrich et al. 2011; Hong et al. 2014; Yoshiura et al. 2017; Chen et al. 2019), the morphology of ionized regions can be quantified by measuring the size distribution of H II bubbles (Zahn et al. 2007;

<sup>4</sup> <http://gmrt.ncra.tifr.res.in>

<sup>5</sup> <http://www.lofar.org>

<sup>6</sup> <http://www.mwatelescope.org>

<sup>7</sup> <http://eor.berkeley.edu>

<sup>8</sup> <http://www.skatelescope.org>

<sup>9</sup> <http://reionization.org>

Mesinger & Furlanetto 2007; Zahn et al. 2011; Pan & Barkana 2012; Majumdar et al. 2014; Lin et al. 2016), which can be measured from the 21 cm maps (Kakiichi et al. 2017; Giri et al. 2018a, 2018b). For example, Kakiichi et al. (2017) suggested a novel technique called “granulometry” for such purpose, based on the idea that granulometry counts the number of ionized bubbles when their sizes are smaller than some threshold.

However, conventional methods for measuring the H II bubble size distribution require high signal-to-noise ratio imaging data of the redshifted 21 cm signal obtained with upcoming radio interferometers such as the SKA. While it is indeed one of the major science goals for the SKA (Koopmans et al. 2015), the 21 cm imaging is observationally more challenging than the 21 cm power spectrum measurements. This is because it will take significantly more integration time to reduce the thermal noise at individual pixels in the 21 cm images, in order to compensate the information loss in the process of Fourier transform of visibility data to obtain the imaging maps. But before the 21 cm imaging data become available, can we learn more information about reionization from the 21 cm power spectrum? Specifically, can we reconstruct the H II bubble size distribution from the 21 cm power spectrum measurements?

The 21 cm power spectrum and H II bubble size distribution are distinct statistical quantities, so from an informational point of view, one quantity cannot be used to infer the other directly, if no additional information is utilized. However, if we employ reionization simulations based on underlying reionization modeling, which can predict both observables from the same set of model parameters (“reionization parameters”), then this underlying reionization modeling essentially provides additional information on the connection between the 21 cm power spectrum and H II bubble size distribution. In principle, we can first obtain the best fit values of reionization parameters constrained by the 21 cm power spectrum (Greig & Mesinger 2015, 2017a, 2018; Shimabukuro & Semelin 2017), and then the H II bubble size distribution can be inferred by running the reionization simulation with the best fit reionization parameters. The disadvantage of this *indirect* approach is that the degeneracies in reionization parameters may bias the best fitting parameter inference—because such estimations are explicitly model-dependent—and thus result in errors in the estimations of H II bubble size distribution.

Recently, machine learning techniques have been widely applied to 21 cm cosmology in three regimes—parameter estimation (Shimabukuro & Semelin 2017; Gillet et al. 2019; Hassan et al. 2020; Zhao et al. 2022), emulation (Kern et al. 2017; Schmit & Pritchard 2018; Jennings et al. 2019) and classification (Hassan et al. 2019). These examples of applications demonstrate that the Artificial Neural Network (ANN) technique can easily establish the connection between two multi-dimensional variables, or “vectors”, if

they are correlated. In this paper, we propose a new method wherein the H II bubble size distribution is reconstructed *directly* from the 21 cm power spectrum using the ANN. Basically, the networks that connect the input (the 21 cm power spectrum) and the output (the H II bubble size distribution) are trained to match the predicted output to their true values, relying on a large number of simulation samples. Since the intermediate step of reionization parameter inference is bypassed, in principle, the reconstruction of H II bubble size distribution with this direct, data-driven, method can be more accurate than the aforementioned indirect approach—we shall test this point herein.

Note that this ANN-based method is *implicitly* model-dependent—the training data sets are based on reionization simulations and their modeling. When this method is applied to future 21 cm power spectrum observational data, caution should be taken about the consequence of the model-dependence—the reconstructed H II bubble size distribution is *not* an independent summary statistic from the power spectrum, and therefore should not be used for reionization parameter inference. Instead, the reconstructed H II bubble size distribution should be regarded only as an indicator for understanding the H II bubble morphology and its evolution.

The rest of this paper is organized as follows. In Section 2, we describe the modeling of cosmic reionization, the 21 cm signal, and the bubble size distribution. In Section 3, we outline the ANN technique. We show our results in Section 4, and give concluding remarks in Section 5.

## 2. Simulation Data Preparation

### 2.1. Reionization Simulations

We perform semi-numerical simulations of reionization with the publicly available code 21cmFAST (Mesinger et al. 2011). This code is based on the semi-numerical treatment of cosmic reionization and thermal history of the IGM. It quickly generates the fields of density, velocity, ionization field, spin temperature and 21 cm brightness temperature on a grid. This code utilizes the excursion-set approach (Furlanetto et al. 2004) to identify ionized regions. Specifically, cells inside a spherical region are identified as ionized, if the number of ionizing photons in that region is larger than that of neutral hydrogen atoms or  $f_{\text{coll}}(\mathbf{x}, R, z) \geq \zeta^{-1}$ . Here,  $\zeta$  is the ionizing efficiency,  $f_{\text{coll}}(\mathbf{x}, R, z)$  is the collapsed fraction smoothed over a sphere with radius  $R$  and center at  $\mathbf{x}$  and redshift  $z$ . The smoothing scale  $R$  proceeds from large to small radius until the above condition is satisfied. If this does not happen with  $R$  down to the cell size, then the cell at  $\mathbf{x}$  is considered as partially ionized with the ionized fraction of  $\zeta f_{\text{coll}}(\mathbf{x}, R_{\text{cell}}, z)$ . While this formalism is based on several simplified assumptions, the ionized field obtained by this formalism is in reasonably good agreement

with that generated with full radiative transfer simulations (Zahn et al. 2011).

Our simulations were performed on a cubic box of 200 comoving Mpc on each side, with  $256^3$  grid cells. We apply the Latin Hypercube Sampling method (McKay et al. 1979) to scan the EOR parameter space, with one realization for each set of parameter values. This method is designed to be more efficient than the naive exhaustive grid-based search. To sample  $N$  points in an  $n$ -dimensional parameter space, we first divide the parameter space into  $N^n$  equal interval grids, and then choose a set of parameters from each row and column exclusively at the Latin Hypercube of the parameter space, so there are totally  $N$  points chosen. While there are several designs that satisfy that condition, we use the maximum Latin Hypercube algorithm that maximizes the minimum distance between the pairs (Morris & Mitchell 1995), which prevents highly clustered regions and ensures homogeneous sampling.

Our EOR model is parametrized with three parameters as follows. (1)  $\zeta$ , the *ionizing efficiency*.  $\zeta = f_{\text{esc}} f_* N_\gamma / (1 + \bar{n}_{\text{rec}})$  (Furlanetto et al. 2004, 2006), which is a combination of several parameters related to ionizing photons. Here,  $f_{\text{esc}}$  is the fraction of ionizing photons escaping from galaxies into the IGM,  $f_*$  is the fraction of baryons locked in stars,  $N_\gamma$  is the number of ionizing photons produced per baryon in stars and  $\bar{n}_{\text{rec}}$  is the mean recombination rate per baryon. The values of these parameters are very uncertain at high redshift (Gnedin et al. 2008; Wise & Cen 2009). In our data set, we explore the range of  $5 \leq \zeta \leq 100$ .

(2)  $T_{\text{vir}}$ , the *minimum virial temperature of haloes that host ionizing sources*. Typically,  $T_{\text{vir}}$  is about  $10^4$  K, corresponding to the temperature above which atomic cooling becomes effective. In our data set, we explore the range of  $10^4 \leq T_{\text{vir}} \leq 10^5$  K.

(3)  $R_{\text{mfp}}$ , the *mean free path of ionizing photons*. The propagation of ionizing photons through the ionized IGM strongly depends on the presence of absorption systems, and the sizes of ionized regions are determined by the balance between the sinks and sources of ionizing photons (see, e.g., McQuinn et al. 2011). This process is modeled by the mean free path of ionizing photons,  $R_{\text{mfp}}$  (Sobacchi & Mesinger 2014), i.e., the typical distance traveled by photons inside ionized regions before they are absorbed.  $R_{\text{mfp}}$  is determined by the number density of Lyman-limit systems and the optical depth of ionizing photons to them. In our data set, we explore the range of  $2 \leq R_{\text{mfp}} \leq 20$  Mpc in comoving scales.

In this paper, we adopt the standard  $\Lambda$ CDM cosmology with fixed values of cosmological parameters based on the Planck 2016 results (Planck Collaboration et al. 2016),  $(h, \Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s) = (0.678, 0.308, 0.0484, 0.692, 0.815, 0.968)$ .

## 2.2. 21 cm Power Spectrum

The 21 cm brightness temperature is given by (e.g., Mellema et al. 2013),

$$\begin{aligned} \delta T_b(\mathbf{x}, z) = & 27 x_{\text{HI}}(\mathbf{x}, z) [1 + \delta_m(\mathbf{x}, z)] \left( 1 - \frac{T_\gamma(z)}{T_S(\mathbf{x}, z)} \right) \\ & \times \left( 1 + \frac{1}{H(z)} \frac{dv_{\parallel}}{dr_{\parallel}} \right)^{-1} \left( \frac{1+z}{10} \frac{0.15}{\Omega_m h^2} \right)^{1/2} \left( \frac{\Omega_b h^2}{0.023} \right) [\text{mK}]. \end{aligned} \quad (1)$$

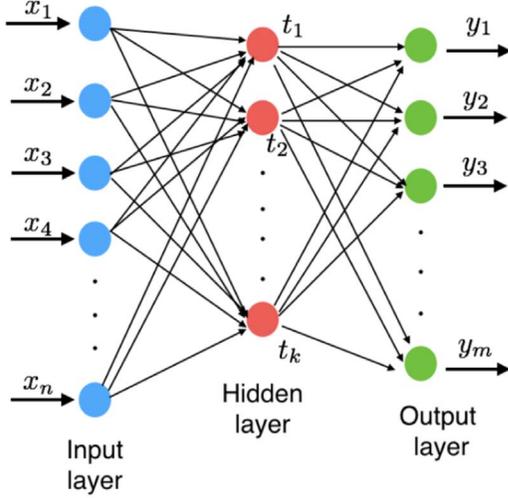
Here,  $T_S$  is the spin temperature of the IGM,  $T_\gamma$  is the cosmic microwave background (CMB) temperature and  $dv_{\parallel}/dr_{\parallel}$  is a peculiar velocity along the line of sight.  $x_{\text{HI}}$  is neutral fraction of the hydrogen atom gas, and  $\delta_m(\mathbf{x}, z) \equiv \rho/\bar{\rho} - 1$  is matter density fluctuations. All variables are evaluated at the redshift  $z = \nu_0/\nu - 1$ . We focus on the regime in which the gas has been significantly heated, so that  $T_S \gg T_\gamma$ . For simplicity, we compute the 21 cm signal without accounting for the redshift space distortion.

The simplest observable that radio interferometer arrays can measure is the 21 cm power spectrum which characterizes the fluctuations in the 21 cm brightness temperature. The 21 cm power spectrum is defined by (e.g., Furlanetto et al. 2006)  $\langle \delta T_b(\mathbf{k}) \delta T_b(\mathbf{k}') \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_{21}(k)$ . We also use the dimensionless 21 cm power spectrum,  $k^3 P_{21}(k)/2\pi^2$ .

## 2.3. H II Bubble Size Distribution

In this subsection, we briefly describe how we measure the H II bubble size distribution from the 3D ionization field map directly. While several different methods have been suggested to measure the bubble size and its distribution (e.g., Mesinger & Furlanetto 2007; Zahn et al. 2007, 2011; Majumdar et al. 2014; Lin et al. 2016), there is no consensus on method, due to the fact that the connectivity in 3D ionized regions is highly irregular and complex. In this paper, after the 3D ionized fraction field is obtained from the reionization simulation using 21cmFAST, the bubble size distribution can be measured from the map of ionization field with the method employed by 21cmFAST (for details, please refer to Furlanetto et al. (2004), Zahn et al. (2007), Mesinger & Furlanetto (2007), Zahn et al. (2011), Mesinger et al. (2011)). Specifically, after a pixel of an ionized region is randomly chosen, the distance from this pixel to the nearest pixel of a neutral region along a random direction is measured. This Monte-Carlo procedure is repeated  $10^7$  times, after which the bubble size distribution can be obtained by taking the volume-weighted average (Zahn et al. 2007; Mesinger & Furlanetto 2007). The probability distribution function (PDF) is

$$\text{PDF}(R) = R \frac{dn}{dR} = \frac{dn}{d \log R}, \quad (2)$$



**Figure 1.** Typical architecture of an ANN. The architecture of the ANN consists of an input layer, a hidden layer and an output layer of neurons. Each neuron in a layer is connected to the neurons in the next layer.

where  $n$  is the number of bubbles with the bubble size in the range from  $R$  to  $R+dR$ . Note that the PDF is normalized to unity.

### 3. Artificial Neural Networks

In this section, we briefly describe the architecture of the ANN. The ANN is a machine learning technique inspired by the natural neuron networks in a human brain. It can be regarded as approximate functions that associate the input data with the output data. By repeated “training” with a set of simulation data (a.k.a. “training data”), the ANN can optimize itself in terms of its capability of predicting the output for a new set of data (a.k.a. “test data”). A typical ANN has a simple architecture that consists of three layers, as illustrated in Figure 1—an input layer, a hidden layer and an output layer, with each layer having a number of neurons. More generally, the number of hidden layers and that of neurons at each layer can be chosen arbitrarily.

In our paper, for example, we set 14 neurons at the input layer, corresponding to the number of  $k$ -bins in the 21 cm power spectrum at each redshift. In the output layer, we set 212 neurons, which is the number of radius bins in the H II size PDF. We set up five hidden layers, each of which contains 212 neurons.

The ANN works as follows. The input data  $\{x_j\}$  are fed to the neurons in the input layer. The  $i^{\text{th}}$  neuron  $s_i$  in the first hidden layer is connected to the  $j^{\text{th}}$  neuron in the input layer linearly with an associated weight  $w_{ij}^{(1)}$ , i.e.,

$$s_i = \sum_{j=1}^n w_{ij}^{(1)} x_j, \quad (3)$$

where  $n$  is the dimension of the input data. In the hidden layer, the  $i^{\text{th}}$  neuron is then activated by an activation function  $\phi(s)$ , i.e., the output of this neuron  $t_i = \phi(s_i)$ . Generally, the activation function is a nonlinear function. We employ the sigmoid function  $\phi(s) = 1/(1 + e^{-s})$ , because it has nice properties that it saturates to constant values when  $|s|$  is large, and that it is a smooth and differentiable function. Neurons in the next hidden layer are linearly connected to the activated neurons in the previous hidden layer, and then activated by  $\phi(s)$  in a similar manner. Thanks to the nonlinear activation function, a trained ANN can approximate any function, in principle. The output data in the output layer is a linear combination of the activated neurons in the last hidden layer,

$$y_i = \sum_{j=1}^k w_{ij}^{(L)} t_j, \quad (4)$$

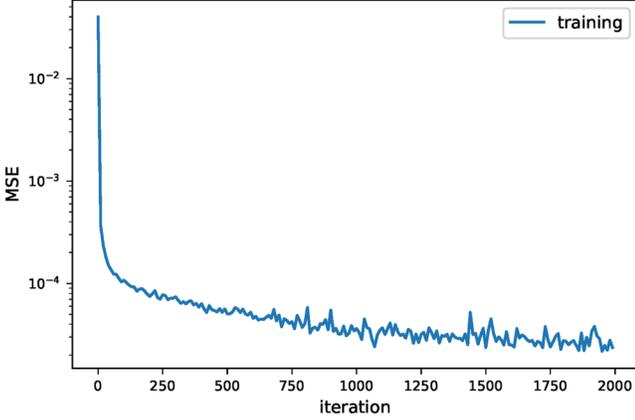
where  $k$  is the number of neurons in the last hidden layer, and  $L - 1$  is the number of all hidden layers. Note that the output data in the output layer are not activated.

The ANN trains its weights in such a manner that, for a set of training data with known values of input and output vectors, the output data generated by the networks are sufficiently close to the true values. Quantitatively, the weights are adjusted to minimize the cost function which is defined as

$$E = \sum_{\alpha=1}^{N_{\text{train}}} E_{\alpha} = \sum_{\alpha=1}^{N_{\text{train}}} \left[ \frac{1}{2} \sum_{i=1}^m (y_{i,\alpha} - d_{i,\alpha})^2 \right], \quad (5)$$

where  $N_{\text{train}}$  is the number of training data sets,  $m$  is the number of neurons at the output layer, and  $y$  and  $d$  are the network-generated and the true values of output data, respectively. We need to compute the partial derivative of  $E$  with respect to the individual weights  $w_{ij}^{(l)}$  and find the local minimum of  $E$  using gradient descent. For this purpose, we employ the *back propagation algorithm* to compute the trained weights (Rumelhart et al. 1986). The number of iterations for this algorithm should be large enough to ensure the convergence of results. Once we have trained the network weights using the training samples, we can make predictions for the output data for test samples, or apply the network to observational data.

In our paper, the input data are the 21 cm power spectrum  $P_{21}(k, z)$  at some redshift  $z$ , with the wavenumber ranging from  $k = 0.12$  to  $1.1 \text{ Mpc}^{-1}$  in 14 logarithmic  $k$ -bins (unless noted otherwise). We choose to avoid the larger-scale modes ( $k < 0.1 \text{ Mpc}^{-1}$ ) because of foreground contamination (e.g., Liu et al. 2014). The output data are the H II bubble size distribution PDF( $R$ ) at the corresponding redshift  $z$ , with the bubble size radius distributed in the range of  $0.78 \leq R \leq 1000 \text{ Mpc}$  in  $N_{\text{radius}} = 212$  logarithmic  $R$ -bins. Our data sets consist of  $N = 1000$  realizations of simulations, with one realization for each set of values in the EOR parameter space  $\{\zeta, T_{\text{vir}}, R_{\text{mfp}}\}$ . For each realization, we sample the data in 50 different redshifts in the range of  $z = 7-12$  (i.e.,  $\Delta z = 0.1$ ).



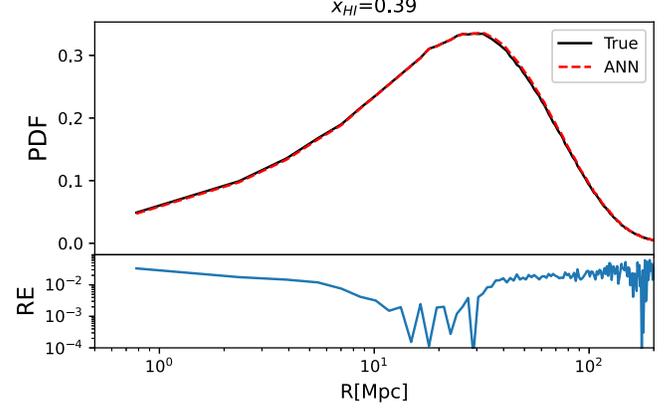
**Figure 2.** MSE evaluated for training samples as a function of the iteration number.

Thus our total data sets contain 50,000 samples of input and output data. We first use  $N_{\text{train}} = 48,000$  random samples as the training data (with 9600 samples as the validation data sets) to train our neural network. After that, we apply the trained network to 2000 test samples of 21 cm power spectra, and generate 2000 H II bubble size distributions. These network-generated PDFs can be compared with the actual PDFs that are computed from the ionized maps directly (dubbed with “ANN” and “true” in figures or subscripts of quantities throughout this paper, respectively). For the purpose of illustration, unless noted otherwise, we choose a reference case with parameter values  $\zeta = 52.0$ ,  $T_{\text{vir}} = 4.5 \times 10^4$  K and  $R_{\text{mfp}} = 18.3$  Mpc, consistent with current observational constraints on the reionization history (e.g., Greig & Mesinger 2017b). Note that the allowed regime of mean neutral fraction is  $0.01 \leq \bar{x}_{\text{HI}} \leq 0.99$  in our data sets and the data with  $\bar{x}_{\text{HI}} = 0$  and  $\bar{x}_{\text{HI}} = 1$  are excluded from the samples.

For training the networks, we test the convergence of the back propagation algorithm and plot the mean squared error (MSE)

$$\text{MSE} = \frac{1}{N_{\text{train}} N_{\text{radius}}} \sum_{\alpha=1}^{N_{\text{train}}} \sum_{i=1}^{N_{\text{radius}}} [\text{PDF}_{\text{ANN}}(R_{i,\alpha}) - \text{PDF}_{\text{true}}(R_{i,\alpha})]^2 \quad (6)$$

as a function of the iteration number for this algorithm in Figure 2. Here  $\text{PDF}(R_{i,\alpha})$  represents the number of bubbles with size  $R_{i,\alpha}$  in the  $i^{\text{th}}$   $R$ -bin for the  $\alpha^{\text{th}}$  training sample. We find that the MSE converges to much below  $10^{-4}$  after 2000 iterations, corresponding to a numerical absolute error of 0.01 in the value of PDF, so we set 2000 back propagation iterations for all computations throughout this paper. This means that the PDF generated by our ANN has a numerical limit of 0.01, below which the PDF can be dominated by numerical errors.



**Figure 3.** (Top) H II bubble size distribution measured from the ionization field (black solid line) and that reconstructed from the 21 cm power spectrum by the ANN (red dashed line) at  $\bar{x}_{\text{HI}} = 0.39$  for our fiducial test EOR model. The KL divergence in this case is  $D_{\text{KL}} = 9.00 \times 10^{-5}$ . (Bottom) relative error (“RE”) of the ANN-reconstructed PDF with respect to the “true” bubble size distribution. We cut it off at the radius wherein the PDF is smaller than 0.01 (the numerical limit set by our ANN).

The networks are tested by evaluating the accuracy of the recovered H II bubble size distribution in terms of the Kullback-Leibler (KL) divergence (Kullback & Leibler 1951). The KL divergence is useful in quantifying the similarity between two PDFs  $P_i$  and  $Q_i$  (here  $i$  is the index for the data points). It is defined as

$$D_{\text{KL}}(P||Q) = \sum_i P_i \log \left( \frac{P_i}{Q_i} \right). \quad (7)$$

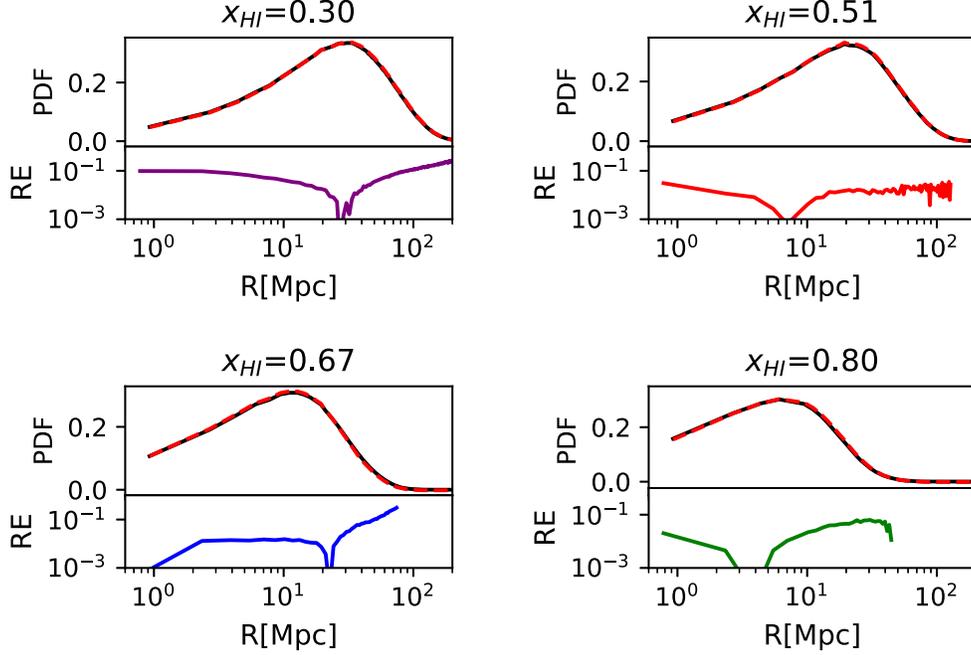
$D_{\text{KL}}$  is close to zero if two PDFs are similar. In our case,  $P_i$  and  $Q_i$  represent  $\text{PDF}_{\text{true}}(R_i, \alpha)$  and  $\text{PDF}_{\text{ANN}}(R_i, \alpha)$  for a given data sample  $\alpha$ , respectively.

## 4. Results

We apply our trained ANN to the test samples, and, in this section, show the results of reconstructing the H II bubble size distribution PDF from the 21 cm power spectrum, as compared with the PDF measured from the ionized fraction field. We first assume that the input 21 cm power spectrum is the pure signal from the simulation, and will later consider the scenario wherein the input power spectrum contains the thermal noise from radio interferometers.

### 4.1. ANN Recovery with Pure Signal

In the absence of thermal noise, the ANN-reconstructed H II bubble PDF is compared with the PDF from the ionized fraction field in Figure 3, for our fiducial test model at the stage when the mean neutral fraction  $\bar{x}_{\text{HI}} = 0.39$ . The KL divergence in this case is  $9.00 \times 10^{-5}$ , and the relative error of the reconstruction, i.e., systematic error using the ANN, is

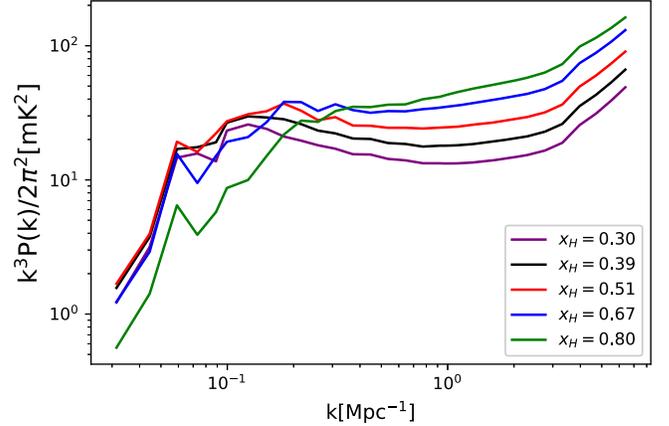


**Figure 4.** Same as Figure 3 but for different stages of reionization at  $\bar{x}_{\text{HI}} = 0.30, 0.51, 0.67, 0.80$ . Larger  $\bar{x}_{\text{HI}}$  corresponds to the earlier stage of reionization. The KL divergence is  $D_{\text{KL}} = 2.85 \times 10^{-5}, 3.54 \times 10^{-5}, 1.44 \times 10^{-3}$  and  $2.32 \times 10^{-4}$ , respectively.

$10^{-3} - 10^{-1}$ . This demonstrates that the reconstruction method works well. Note that the PDF generated by our ANN has a numerical limit of 0.01, below which the PDF is comparable to the numerical error set by the number of iterations in the back propagation algorithm, so we only calculate the relative error of the reconstructed PDF at the radii wherein the PDF  $\geq 0.01$ .

We further test the accuracy of reconstruction at different stages of the reionization,  $\bar{x}_{\text{HI}} = 0.30, 0.51, 0.67$  and  $0.80$ , in Figure 4. As reionization proceeds from high to low  $\bar{x}_{\text{HI}}$ , the peak of the PDF shifts from small to large radii, due to the growth of ionized bubbles. This is consistent with the evolution of the peak of the 21 cm power spectrum which shifts from large to small  $k$ , as depicted in Figure 5. The KL divergence for the reconstruction of PDF is very close to zero ( $D_{\text{KL}} \sim 10^{-4}$ ), and the relative error is  $\lesssim 10\%$  for  $R \lesssim 100$  Mpc at all time.

Furthermore, we test the reconstruction for different test models of reionization (see Table 1). For the same mean  $x_{\text{HI}}$ , the PDFs and 21 cm power spectra can vary for different models of reionization. Figure 6 affirms that at the same mean  $\bar{x}_{\text{HI}} = 0.39$ , our fiducial model (Model 1) has the largest radius of the peak PDF, while the peak radius for Model 2 is the smallest. This is consistent with the fact that the peak of the 21 cm power spectrum appears in the smallest  $k$  for Model 1 and the largest  $k$  for Model 2, as displayed in the right panel of Figure 6. We compare the reconstruction among three different

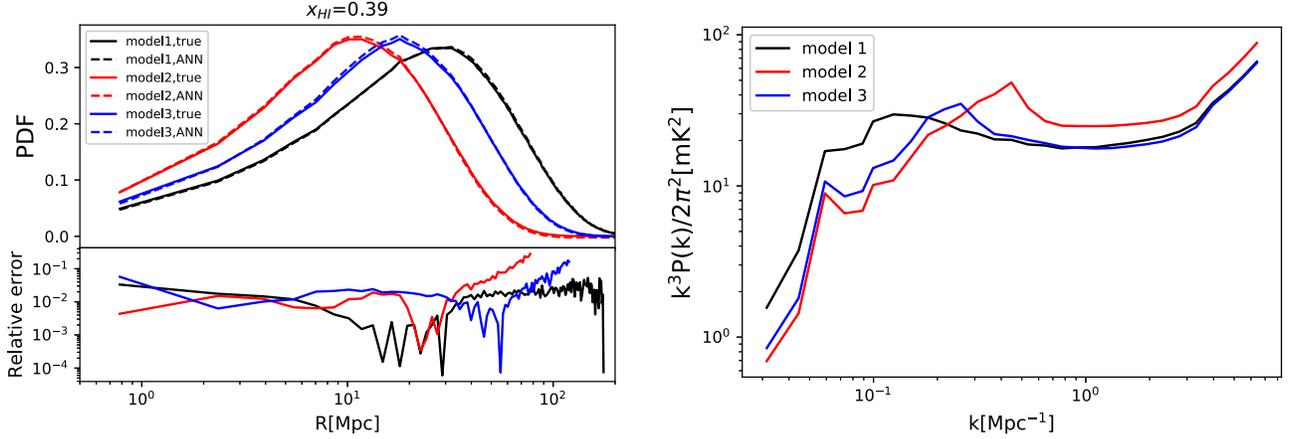


**Figure 5.** The 21 cm power spectrum at different stages of reionization,  $\bar{x}_{\text{HI}} = 0.30/0.39/0.51/0.67/0.80$  (purple/black/red/blue/green), respectively, for our fiducial test EOR model.

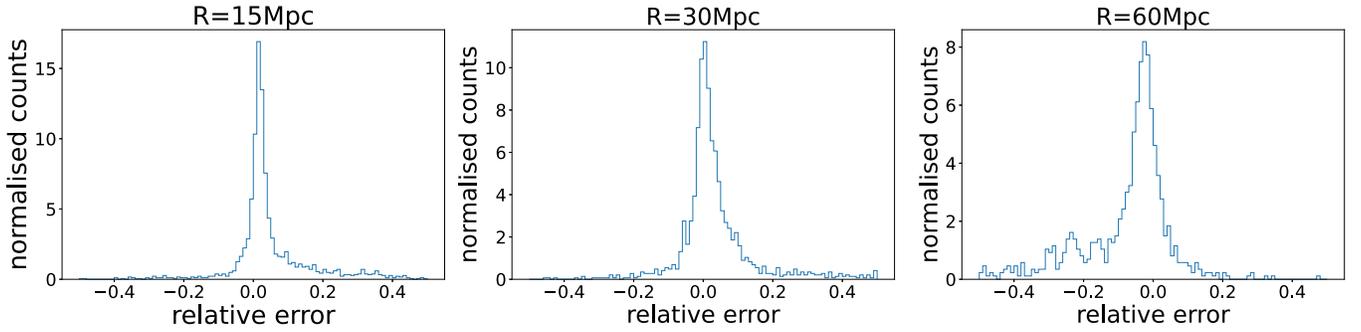
**Table 1**

Parameter Values of Reionization Models used for Test Samples in Figure 6

	$\zeta$	$T_{\text{vir}} (\times 10^4 \text{ K})$	$R_{\text{mfp}} (\text{Mpc})$
Model 1 (fiducial)	52.0	4.5	18.3
Model 2	93.2	3.8	5.1
Model 3	49.0	3.1	9.1



**Figure 6.** (Left) Same as Figure 3 but for three different test models given in Table 1 (black/red/blue curves for Model 1/2/3, respectively). The KL divergence for Models 1, 2 and 3 is  $D_{\text{KL}} = 6.78 \times 10^{-5}$ ,  $4.77 \times 10^{-3}$  and  $2.35 \times 10^{-3}$ , respectively. (Right) The 21 cm power spectrum at the same fixed  $\bar{x}_{\text{HI}} = 0.39$  for these models.



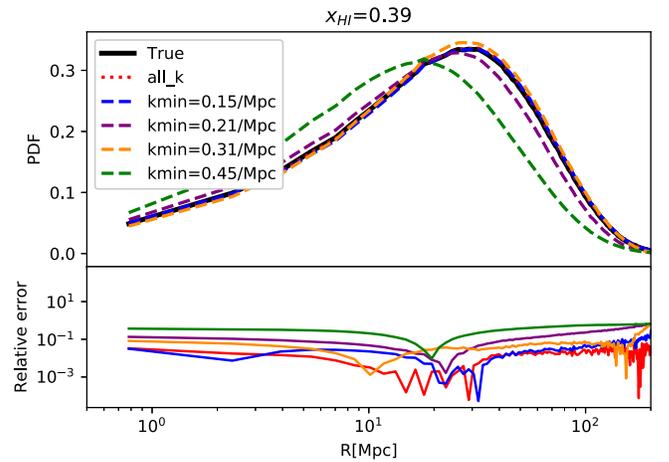
**Figure 7.** Distribution of the relative errors of bubble size distribution from all test models at some fixed bubble radii  $R = 15$  (left), 30 (middle) and 60 Mpc (right).

reionization models at the same fixed  $\bar{x}_{\text{HI}} = 0.39$  in the left panel of Figure 6, and find good accuracy for all cases. This indicates that the ANN can distinguish different H II bubble size distributions even if the ionization field is at the same global mean stage.

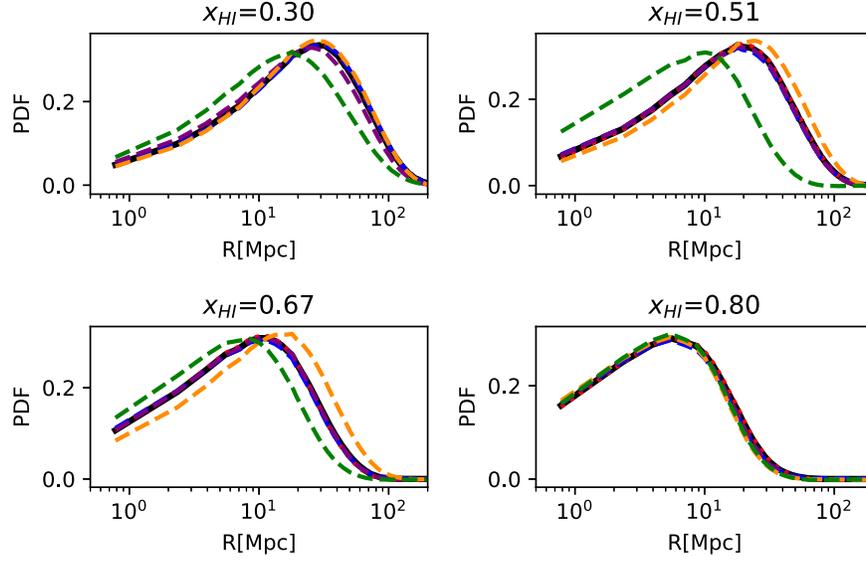
To evaluate the accuracy for all test data (with different EOR models and at different redshifts), we plot the distribution histogram of relative error of the reconstructed PDF with respect to the “true” PDF, for some fixed bubble sizes, in Figure 7. In most cases, the relative error is  $< 10\%$ . This demonstrates that the H II bubble size distribution can be recovered successfully with good accuracy from the 21 cm power spectrum using the ANN technique, regardless of the stages of reionization and reionization models.

#### 4.2. Scale Dependence of ANN Recovery

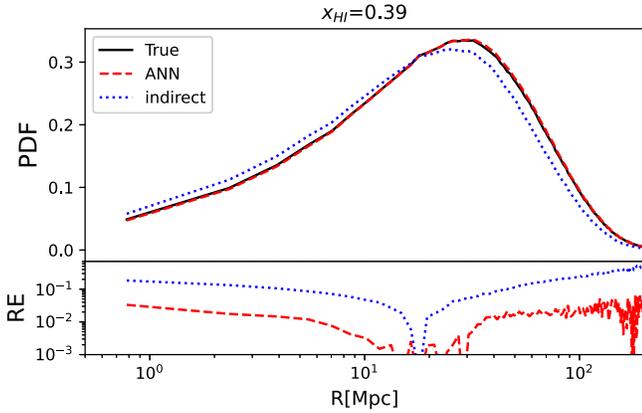
In practical observations, the large-scale power may be lost due to the removal of foreground contamination. Therefore, it is important to understand how the minimum wavenumber  $k_{\text{min}}$  of the 21 cm power spectrum affects the reconstruction of the bubble size distribution, and test this convergence using simulation data. In Figure 8, we compare the H II bubble size



**Figure 8.** Same as Figure 3 but using only part of the information  $k_{\text{min}} \leq k \leq 1.1 \text{ Mpc}^{-1}$  in the power spectrum, with  $k_{\text{min}} = 0.15, 0.21, 0.31$  and  $0.45 \text{ Mpc}^{-1}$  (blue/purple/orange/green dashed, respectively), in comparison with the fiducial case of using all modes  $k_{\text{min}} = 0.12 \text{ Mpc}^{-1}$  (red dotted line). The KL divergence is  $D_{\text{KL}} = 9.00 \times 10^{-5}$ ,  $4.34 \times 10^{-4}$ ,  $9.77 \times 10^{-4}$ ,  $1.31 \times 10^{-3}$  and  $5.77 \times 10^{-2}$  for  $k_{\text{min}} = 0.12, 0.15, 0.21, 0.31$  and  $0.45 \text{ Mpc}^{-1}$ , respectively.

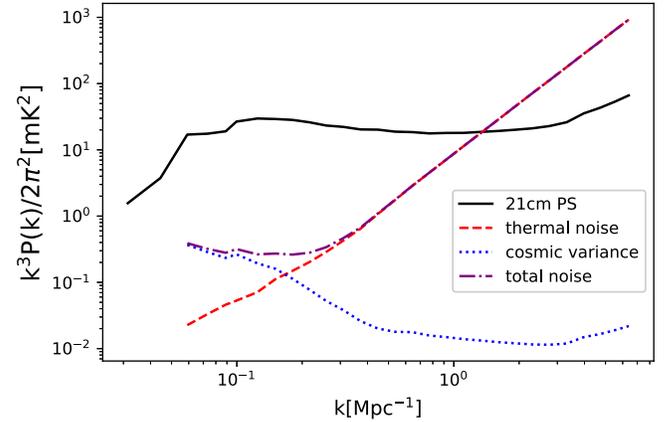


**Figure 9.** Same as the top panel of Figure 8 but for different stages of reionization at  $\bar{x}_{\text{HI}} = 0.30, 0.51, 0.67, 0.80$ .



**Figure 10.** Same as Figure 3 but the reconstruction is made by the *indirect* approach (blue dotted) as well as recovered directly by the 21 cm power spectrum with the ANN (red dashed).

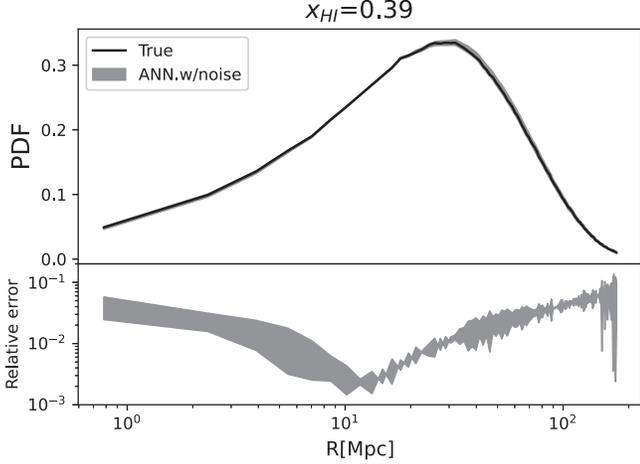
distributions recovered by the ANN using the 21 cm power spectrum with varying  $k_{\min} = 0.12$  (all bins), 0.15, 0.21, 0.31 and  $0.45 \text{ Mpc}^{-1}$ , at  $\bar{x}_{\text{HI}} = 0.39$  for our fiducial test EOR model, while the maximum wavenumber  $k_{\max} = 1.1 \text{ Mpc}^{-1}$  is fixed. We find that the reconstructed PDF from  $k_{\min} = 0.15 \text{ Mpc}^{-1}$  is almost indistinguishable from that using all modes, and the KL divergence is just as good. This value of  $k_{\min} = 0.15 \text{ Mpc}^{-1}$  is consistent with the peak in the 21 cm power spectrum that contains information on the characteristic bubble size. Also, losing more large-scale information (i.e., enlarging  $k_{\min}$ ) can hurt the reconstruction and result in a relative error larger than 10%. This implies that the large-scale information in the 21 cm power spectrum, particularly at the peak of power, is indeed



**Figure 11.** The 21 cm power spectrum signal (black solid), the total noise power spectrum for the configuration of SKA1 (purple dot-dashed) including the contributions from thermal noise (red dashed) and cosmic variance (blue dotted), for the fiducial test model at  $\bar{x}_{\text{HI}} = 0.39$ .

essential for reconstructing the PDF. We further test the scale dependence at different stages of reionization in Figure 9. We find that the largest possible  $k_{\min}$ , which compromises to give a good reconstruction, can depend on the stage of reionization, because the scale of power spectrum should be large enough, compared to the typical bubble size at that moment. For example, powers with  $k_{\min} = 0.21 \text{ Mpc}^{-1}$  can also give as good reconstruction of PDF as powers of all modes, before reionization proceeds halfway ( $\bar{x}_{\text{HI}} \geq 0.5$ ).

For the complete information, we also vary  $k_{\max}$  with fixed  $k_{\min} = 0.12 \text{ Mpc}^{-1}$ , and find that the recovered PDF and KL divergence only weakly depend on  $k_{\max}$ .



**Figure 12.** Same as Figure 3 but the reconstruction with the ANN is from the 21 cm power spectrum with total noise including cosmic variance and thermal noise (with the shaded region representing the  $1\sigma$  confidence region). The mean and variance of PDF are computed from 10 realizations of random noise.

#### 4.3. Comparison with Indirect Reconstruction Approach

In Section 1, we commented that there could be an alternative, *indirect*, reconstruction approach, which is to infer the H II bubble size distribution from the prediction of the underlying reionization simulation with the reionization model parameters that are best fit by the 21 cm power spectrum. This indirect approach is more interpretable and intuitive than the direct ANN-based reconstruction. In this subsection, we compare the accuracies of both approaches in Figure 10. We find that while the indirect approach can outline the approximate shape of the PDF, it can only capture the characteristics in a biased manner, in terms of both location and height of the PDF peak. Quantitatively, the indirect approach results in an error of tens of per cent. This should be due to the degeneracies in reionization parameters that can result in large errors in best fitting the parameter values, which, in turn, leads to errors in the H II bubble size distribution in the indirect approach. In comparison, the direct, ANN-based approach can have an error within 10%, and is therefore more accurate than the indirect approach.

#### 4.4. ANN Recovery with Thermal Noise

In previous subsections, we assume that the input 21 cm power spectrum is the pure signal from simulations. In practical observations, however, the measurements of 21 cm power spectrum contain random noise. For large radio interferometer arrays like the SKA, the noise is dominated by thermal noise at small scales, but cosmic variance becomes important at large scales. In this subsection, we will take into account both thermal noise and cosmic variance, and investigate the effect of

the noise power spectrum on the reconstruction of H II bubble size distribution.

The thermal noise power spectrum for a mode  $k$  is given by McQuinn et al. (2006); Mao et al. (2008, 2013)

$$P_{\text{th, 1 mode}}(k, \mu) = d_A^2 y \frac{\Omega}{t} \frac{T_{\text{sys}}^2}{\bar{n}(L_{k_\perp}) A_e}. \quad (8)$$

Here  $d_A(z)$  is the comoving angular diameter distance at  $z$ ,  $y(z) \equiv \lambda_{21}(1+z)^2/H(z)$ , where  $\lambda_{21} = \lambda(z)/(1+z) = 0.21$  m and  $H(z)$  is the Hubble parameter at  $z$ .  $\Omega = \lambda^2/A_e$  is solid angle spanning the field of view and  $t$  is the total integration time.  $T_{\text{sys}}$  is the system temperature of antenna, which is the sum of the receiver temperature of  $\sim 100$  K and the sky temperature  $T_{\text{sky}} = 60(\nu/300 \text{ MHz})^{-2.55}$  K. Compact layout of the radio interferometer array can repeatedly measure one visibility mode, thereby reducing the thermal noise.  $\bar{n}(L_{k_\perp}) A_e$  denotes the number of redundant baselines  $L_{k_\perp}$  corresponding to  $k_\perp$  within the baseline area equal to the effective area per station  $A_e$ . The thermal noise for the mode  $k$  depends on the projection of  $k$  on the sky plane  $k_\perp = k\sqrt{1-\mu^2}$ , where  $\mu = \cos\theta$  and  $\theta$  is the angle between the mode  $k$  and the line of sight.

The thermal noise for the spherically averaged power spectrum over a  $k$ -shell is given by Lidz et al. (2011)

$$P_{\text{thermal}}(k) = \left[ \sum_{\mu} \frac{N_c(k, \mu)}{P_{\text{th, 1 mode}}^2(k, \mu)} \right]^{-1/2}. \quad (9)$$

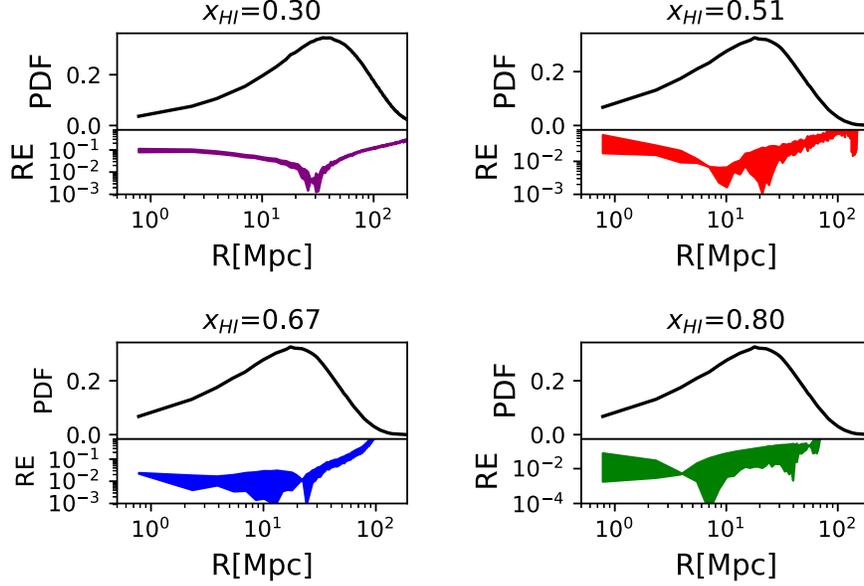
$N_c(k, \mu)$  is the number of modes in the ring with  $\mu$  on the spherical  $k$ -shell with the logarithmic step size  $\delta k/k = \epsilon$ ,  $N_c = \epsilon k^3 \Delta\mu \times \text{vol}/4\pi^2$ , and  $\text{vol}$  is the survey volume of the sky. The sum here accounts for the noise reduction by combining independent modes. Thus it runs over the upper half shell with positive  $\mu$  since the brightness temperature is a real-valued field, and only half of the Fourier modes are independent.

The cosmic variance for the 21 cm power spectrum is estimated by

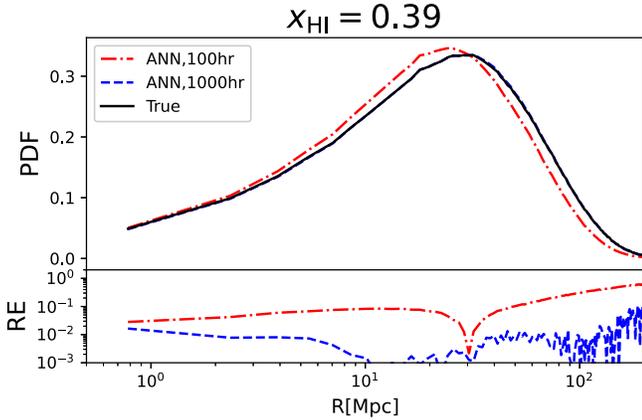
$$P_{\text{cv}}(k) = \frac{1}{\sqrt{N_{\text{modes}}}} P_{21}(k), \quad (10)$$

where  $N_{\text{modes}} = \epsilon k^3 \times \text{vol}/4\pi^2$  is the number of modes in the upper half  $k$ -shell.

In this paper, we consider an experiment similar to the low-frequency array of SKA Phase 1 (SKA1). Specifically, we assume a configuration such that 224 stations are compactly laid out in the core with 1000 m in diameter, and the minimum baseline between stations is 60 m. We assume that the field of view of a single primary beam is full width at half maximum (FWHM)  $\sim 3.5^\circ$  at  $z \sim 8$ , the effective area per station  $A_e \approx 421 \text{ m}^2$  at  $z \sim 8$ , the total integration time is 1000 hr, the bandwidth of a redshift-bin is 10 MHz and the step size of a  $k$ -bin is  $\epsilon = \delta k/k = 0.1$ . Figure 11 plots the total noise power spectrum  $P_N(k) = P_{\text{thermal}}(k) + P_{\text{cv}}(k)$  as well as the respective



**Figure 13.** Same as Figure 12 but for different stages of reionization at  $\bar{x}_{\text{HI}} = 0.30, 0.51, 0.67, 0.80$ .



**Figure 14.** Same as Figure 12 but for different integration time of 100 hr (red dot-dashed) and 1000 hr (blue dashed). We only show one realization here, for illustrative purpose.

contributions from cosmic variance and thermal noise. Our result is consistent with previous studies, e.g., Koopmans et al. (2015). For SKA1, the cosmic variance is always negligible compared to the signal, and the thermal noise is smaller than the signal for  $k \leq 1 \text{ Mpc}^{-1}$ . In other words, the 21 cm signal dominates over the noise except at small scales. Since the reconstruction is not sensitive to  $k_{\text{max}}$ , as shown in Section 4.2, we expect that the reconstructed PDF should not be significantly affected by the noise.

We model the measured 21 cm power spectrum as  $P(k) = P_{21}(k) + N(k)$ , where  $P_{21}(k)$  is the 21 cm power spectrum signal, and  $N(k)$  is a random draw from a Gaussian

probability distribution with zero mean and variance equal to the square of total noise power spectrum  $P_N^2(k)$ . For each test EOR model, we generate 10 independent realizations from the total noise power spectrum as the input data for the ANN, and from the 10 different outputs of reconstructed PDFs, compute the mean and variance. Figure 12 displays the  $1\sigma$  confidence level region of the 10 different outputs of reconstructed PDFs for the fiducial test EOR model at  $\bar{x}_{\text{HI}} = 0.39$ . We also show the evolution of the reconstruction in Figure 13. We find that even if the total noise is accounted for at the sensitivity level of SKA1, the reconstruction of H II bubble size distribution with the ANN still works well at the relative error level of 10% (except for large radii  $\gtrsim 100 \text{ Mpc}$ ). This finding is confirmed for other test EOR models.

Nevertheless, if we reduce the total integration time from 1000 hr to 100 hr, then the thermal noise will be much larger. As a result, as depicted in Figure 14, the reconstruction of H II bubble size distribution will cause significant systematic error, which is  $\gtrsim 10\%$  at scales smaller than the peak of PDF, but can be tens of per cent at larger scales.

## 5. Summary and Conclusions

In this paper, we propose a new, ANN-based method to reconstruct the H II bubble size distribution directly from the 21 cm power spectrum. Our method will allow tracing the evolution of H II bubble size distribution during cosmic reionization when only 21 cm power spectrum measurements will be available, but direct measurements of bubble size distribution cannot be done without 3D 21 cm imaging data.

Nevertheless, our reconstruction method implicitly exploits the modeling in reionization simulations, and hence the recovered H II bubble size distribution is not an independent summary statistic from the power spectrum, and should be used only as an indicator for understanding H II bubble morphology and its evolution.

We train our neural networks with 48,000 training data sets and tested the networks with 2000 test data sets. These data sets are generated by varying EOR parameters for 1000 realizations with the semi-numerical code `21cmFAST`. We use the 21 cm power spectrum for  $k=0.12\text{--}1.1\text{ Mpc}^{-1}$  in 14  $k$ -bins as the input of the networks, and generate the H II bubble size distribution PDF( $R$ ) for  $R=0.78\text{--}1000\text{ Mpc}$  in 212  $R$ -bins as the output, at  $z=7\text{--}12$ . We train the weights of ANN using the back propagation algorithm.

We apply the trained networks to the test data sets to test the accuracy of recovery. We demonstrate that the recovered H II bubble size distribution can be almost as accurate as that directly measured from the ionization map with the fractional error  $<10\%$  for  $R \lesssim 100\text{ Mpc}$  at all stages of reionization, with the KL divergence  $D_{\text{KL}} \ll 1$  at all time. This result is generic for a number of EOR models.

We further investigate the main contributions to the reconstruction, and find that the large-scale modes are particularly important. The reconstruction results are sensitive to the minimum wavenumber cutoff  $k_{\text{min}}$ , while weakly depending on the maximum wavenumber cutoff  $k_{\text{max}}$ . The  $k_{\text{min}}$  should correspond to the scale that is much larger than the typical bubble size, so it depends on the stage of reionization. For the early and middle stages ( $\bar{x}_{\text{HI}} \geq 0.5$ ),  $k_{\text{min}}$  must be smaller than  $0.21\text{ Mpc}^{-1}$ , in order for the reconstruction results to converge. For the later stage, e.g., at  $\bar{x}_{\text{HI}} \geq 0.39$ ,  $k_{\text{min}}$  should be smaller than  $0.15\text{ Mpc}^{-1}$ .

In principle, the reconstruction of PDF can be achieved alternatively with an indirect approach, which is to first constrain the best fit values of reionization parameters from the 21 cm power spectrum measurements and then obtain the H II bubble size distribution from the prediction of reionization simulations using the best fit reionization parameters. However, this indirect approach can result in an error of tens of per cent in the reconstructed PDF, in comparison with the  $<10\%$  error in our direct, ANN-based method.

Our reconstruction is tested when the thermal noise and cosmic variance at the SKA1 noise level are applied to the 21 cm power spectrum. Since the total noise for SKA1 is subdominant for  $k \lesssim 1\text{ Mpc}^{-1}$ , assuming the integration time of 1000 hrs, our reconstruction results are not affected much by the noise, i.e., the recovered PDF agrees with that directly measured from the ionization map with the fractional error  $<10\%$  for the radii  $R \lesssim 100\text{ Mpc}$  at all stages of reionization.

Note that this fractional error of  $\lesssim 10\%$  refers to the difference in the reconstructed PDF with respect to the true value. These are the systematic errors using the ANN.

However, an estimation of statistical uncertainties is not performed in this paper. In principle, this can be done by neural network techniques such as the density-estimation likelihood-free inference (Alsing et al. 2018, 2019; Zhao et al. 2022). We defer the implementation of this technique to future work.

## Acknowledgments

This work is supported by the National SKA Program of China (Grant No. 2020SKA0110401), NSFC (Grant Nos. 12103044, 11821303 and 11850410429) and National Key R&D Program of China (Grant No. 2018YFA0404502). H.S. was also supported in part by grants from Yunnan University. We thank Rennan Barkana, Xuelei Chen, Anastasia Fialkov, Nicolas Gillet, Andrei Mesinger and Benjamin Wandelt for fruitful discussions and valuable feedbacks.

## ORCID iDs

Hayato Shimabukuro  <https://orcid.org/0000-0003-4850-0656>

Yi Mao  <https://orcid.org/0000-0002-1301-3893>

Jianrong Tan  <https://orcid.org/0000-0001-6161-7037>

## References

- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, *ApJ*, 809, 61
- Alsing, J., Charnock, T., Feeney, S., & Wandelt, B. 2019, *MNRAS*, 488, 4440
- Alsing, J., Wandelt, B., & Feeney, S. 2018, *MNRAS*, 477, 2874
- Chen, Z., Xu, Y., Wang, Y., & Chen, X. 2019, *ApJ*, 885, 23
- DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, *PASP*, 129, 045001
- Dillon, J. S., Tegmark, M., Liu, A., et al. 2015, *PhRvD*, 91, 023002
- Fan, X., Carilli, C. L., & Keating, B. 2006, *ARA&A*, 44, 415
- Friedrich, M. M., Mellema, G., Alvarez, M. A., Shapiro, P. R., & Iliev, I. T. 2011, *MNRAS*, 413, 1353
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, 433, 181
- Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, *ApJ*, 613, 1
- Gillet, N., Mesinger, A., Greig, B., Liu, A., & Ucci, G. 2019, *MNRAS*, 484, 282
- Giri, S. K., Mellema, G., Dixon, K. L., & Iliev, I. T. 2018a, *MNRAS*, 473, 2949
- Giri, S. K., Mellema, G., & Ghara, R. 2018b, *MNRAS*, 479, 5596
- Gleser, L., Nusser, A., Ciardi, B., & Desjacques, V. 2006, *MNRAS*, 370, 1329
- Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, *ApJ*, 672, 765
- Greig, B., & Mesinger, A. 2015, *MNRAS*, 449, 4246
- Greig, B., & Mesinger, A. 2017a, *MNRAS*, 472, 2651
- Greig, B., & Mesinger, A. 2017b, *MNRAS*, 465, 4838
- Greig, B., & Mesinger, A. 2018, *MNRAS*, 477, 3217
- Hassan, S., Andrianomena, S., & Doughty, C. 2020, *MNRAS*, 494, 5761
- Hassan, S., Davé, R., Mitra, S., et al. 2018, *MNRAS*, 473, 227
- Hassan, S., Liu, A., Kohn, S., & La Plante, P. 2019, *MNRAS*, 483, 2524
- Hong, S. E., Ahn, K., Park, C., et al. 2014, *JKAS*, 47, 49
- Jacobs, D. C., Pober, J. C., Parsons, A. R., et al. 2015, *ApJ*, 801, 51
- Jelić, V., de Bruyn, A. G., Mevius, M., et al. 2014, *A&A*, 568, A101
- Jennings, W. D., Watkinson, C. A., Abdalla, F. B., & McEwen, J. D. 2019, *MNRAS*, 483, 2907
- Kakiichi, K., Majumdar, S., Mellema, G., et al. 2017, *MNRAS*, 471, 1936
- Kern, N. S., Liu, A., Parsons, A. R., Mesinger, A., & Greig, B. 2017, *ApJ*, 848, 23
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, Vol. 1, Giardini Naxos, Italy, 9–13 June 2014

- Kulkarni, G., Choudhury, T. R., Puchwein, E., & Haehnelt, M. G. 2016, *MNRAS*, **463**, 2583
- Kulkarni, G., Choudhury, T. R., Puchwein, E., & Haehnelt, M. G. 2017, *MNRAS*, **469**, 4283
- Kullback, S., & Leibler, R. A. 1951, *Ann. Math. Statist.*, **22**, 79
- Lee, K.-G., Cen, R., Gott, J. R., III, & Trac, H. 2008, *ApJ*, **675**, 8
- Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, *ApJ*, **741**, 70
- Lin, Y., Oh, S. P., Furlanetto, S. R., & Sutter, P. M. 2016, *MNRAS*, **461**, 3361
- Liu, A., Parsons, A. R., & Trott, C. M. 2014, *PhRvD*, **90**, 023019
- Madau, P., Meiksin, A., & Rees, M. J. 1997, *ApJ*, **475**, 429
- Majumdar, S., Mellema, G., Datta, K. K., et al. 2014, *MNRAS*, **443**, 2843
- Mao, Y., D'Aloisio, A., Zhang, J., & Shapiro, P. R. 2013, *PhRvD*, **88**, 081303
- Mao, Y., Tegmark, M., McQuinn, M., Zaldarriaga, M., & Zahn, O. 2008, *PhRvD*, **78**, 023529
- McKay, M. D., Beckman, R. J., & Conover, W. J. 1979, *Technometrics*, **21**, 239
- McQuinn, M., Oh, S. P., & Faucher-Giguère, C.-A. 2011, *ApJ*, **743**, 82
- McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, *ApJ*, **653**, 815
- Mellema, G., Koopmans, L., Shukla, H., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, Vol. 1, Giardini Naxos, Italy, 9–13 June 2014
- Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2013, *ExA*, **36**, 235
- Mertens, F. G., Mevius, M., Koopmans, L. V. E., et al. 2020, *MNRAS*, **493**, 1662
- Mesinger, A., & Furlanetto, S. 2007, *ApJ*, **669**, 663
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, **411**, 955
- Morris, M. D., & Mitchell, T. J. 1995, *J. Stat. Plan. Inference*, **43**, 381
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, **433**, 639
- Pan, T., & Barkana, R. 2012, arXiv:1209.5751
- Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, *ApJ*, **788**, 106
- Patil, A. H., Yatawatta, S., Koopmans, L. V. E., et al. 2017, *ApJ*, **838**, 65
- Patil, A. H., Zaroubi, S., Chapman, E., et al. 2014, *MNRAS*, **443**, 1113
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13
- Pober, J. C., Ali, Z. S., Parsons, A. R., et al. 2015, *ApJ*, **809**, 62
- Pritchard, J. R., & Loeb, A. 2012, *RPPh*, **75**, 086901
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, *Natur*, **323**, 533
- Schmit, C. J., & Pritchard, J. R. 2018, *MNRAS*, **475**, 1213
- Scott, D., & Rees, M. J. 1990, *MNRAS*, **247**, 510
- Shimabukuro, H., & Semelin, B. 2017, *MNRAS*, **468**, 3869
- Sobacchi, E., & Mesinger, A. 2014, *MNRAS*, **440**, 1662
- The HERA Collaboration, Abdurashidova, Z., Aguirre, J. E., et al. 2021, arXiv:2108.02263
- Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, *PASA*, **30**, e007
- Wise, J. H., & Cen, R. 2009, *ApJ*, **693**, 984
- Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, *A&A*, **550**, A136
- Yoshiura, S., Shimabukuro, H., Takahashi, K., & Matsubara, T. 2017, *MNRAS*, **465**, 394
- Zahn, O., Lidz, A., McQuinn, M., et al. 2007, *ApJ*, **654**, 12
- Zahn, O., Mesinger, A., McQuinn, M., et al. 2011, *MNRAS*, **414**, 727
- Zhao, X., Mao, Y., Cheng, C., & Wandelt, B. D. 2022, *ApJ*, **926**, 151