

From ANN to BNN: Inferring reionization parameters using uncertainty-aware emulators of 21-cm summaries

Yashrajsinh Mahida,^a Sanjay Kumar Yadav,^a Suman Majumdar,^a Leon Noble,^a Chandra Shekhar Murmu,^{a,b,c} Saswata Dasgupta,^{d,e} Sohini Dutta,^f Himanshu Tiwari,^{g,h} Abinash Kumar Shawⁱ

^aDepartment of Astronomy, Astrophysics & Space Engineering,
Indian Institute of Technology Indore,
Indore 453552, India

^bAstrophysics Research Centre of the Open University (ARCO),
The Open University of Israel,
1 University Road, POBox808, Ra'anana 4353701, Israel

^cDepartment of Natural Sciences,
The Open University of Israel,
1 University Road, POBox808, Ra'anana 4353701, Israel

^dInstitute of Astronomy,
University of Cambridge, Cambridge, UK

^eKavli Institute for Cosmology,
University of Cambridge, Cambridge, UK

^fJodrell Bank Centre for Astrophysics, University of Manchester, Manchester M13 9PL, UK

^gInternational Centre for Radio Astronomy Research (ICRAR),
Curtin University,
Bentley, Perth, WA, Australia

^hCommonwealth Scientific and Industrial Research Organisation (CSIRO),
Space & Astronomy,
Bentley, WA, Australia

ⁱDepartment of Computer Science, University of Nevada Las Vegas, 4505 S. Maryland Pkwy.,
Las Vegas, NV 89154, USA

E-mail: yhmahida@gmail.com

Abstract. Inferring astrophysical parameters from radio interferometric observations of the redshifted 21-cm signal from the Epoch of Reionization (EoR) is a challenging, yet crucial. A Bayesian inference pipeline for reionization parameter estimation, forward models the signal statistic, usually the power spectrum, and compares it to the observed statistic. However, the 21-cm signal coming from EoR is expected to be highly non-Gaussian in nature; therefore, we need to use higher-order statistics, e.g., bispectrum. Moreover, the forward modeling of the signal and its statistics for a varying set of model parameters requires rerunning the numerical simulations a large number of times, which is computationally very expensive and time-consuming. To overcome this challenge, many artificial neural network (ANN) based emulators have been introduced, which produce the 21-cm signal summary statistics in a fraction of the time, given input astrophysical parameters. However, ANN emulators have a fundamental drawback: they can only produce point-value predictions; thus, they fail to capture the uncertainty associated with their own predictions. Therefore, when such emulators are used in the Bayesian inference pipeline, they cannot naturally propagate their prediction uncertainties to the estimated model parameters. To address this problem, we have developed Bayesian neural network (BNN) emulators for the 21-cm signal statistics, which provide the posterior distribution of the predicted signal statistics, including their own prediction uncertainty. We use these BNN emulators in our Bayesian inference pipeline to infer the EoR parameters through 21-cm summaries of the mock observation of 21-cm signal with telescopic noise for 1000 hr of SKA-LOW observation. We show that BNN emulators are able to capture the prediction uncertainty for the 21-cm power spectrum and bispectrum, and using these emulators in the inference pipeline provides better and tighter constraints on them. To check the robustness of the emulators, we systematically reduced the training dataset and showed that, for smaller training datasets, BNN outperforms the ANN emulators. We also show that using the bispectrum as a summary statistic gives better constraints on EoR parameters than the power spectrum.

Keywords: Machine learning, reionization, non-gaussianity, Bayesian reasoning, cosmological simulations, intergalactic media

Contents

1	Introduction	1
2	Reionization 21-cm Signal Simulation	3
3	Estimating Signal Statistics	4
3.1	Power Spectrum	4
3.2	Bispectrum	5
4	Emulators: ANN vs BNN	6
4.1	Artificial Neural Network (ANN)	6
4.2	Bayesian Neural Network (BNN)	7
5	Emulation of Signal Statistics	8
5.1	Power spectrum emulation	9
5.2	Bispectrum emulation	11
6	Parameter Estimation	16
6.1	Estimation of Error Covariance	16
6.2	Results	17
7	Summary and Discussion	19

1 Introduction

The goal of modern cosmology is to understand the evolutionary history of the Universe as it has gone through different epochs of evolution. The phase of the Universe, when the first luminous sources started to form, is known as Cosmic Dawn (CD). The subsequent phase, in which the radiation from these luminous sources begins to ionize the neutral hydrogen (HI) of the intergalactic medium (IGM), is known as the Epoch of Reionization (EoR). These epochs are still not very well understood. Our current understanding of CD-EoR prominently comes from different indirect observations such as the Cosmic Microwave Background (CMB) [1–3], the absorption spectra of high redshift quasars [4–6], and the luminosity function and clustering properties of Lyman- α emitters [7–9], which cannot provide solid answers to essential questions about CD-EoR such as exact duration of these epochs, formation and properties of the ionizing sources, the morphology of ionized regions and its evolution. We require direct observations of CD-EoR to understand these epochs better.

Multiple efforts are going on for the direct observation of the EoR using the 21-cm emission line as a signal. As most of the IGM is made up of neutral hydrogen, the redshifted 21-cm signal from the hyperfine spin-flip transition of neutral hydrogen works as a tracer of the ionized region during EoR [10, 11]. There are many ongoing radio interferometers, such as uGMRT [12], HERA [13], LOFAR [14], MWA [15], and the upcoming SKAO [16], that are capable of detecting the 21-cm signal from EoR. Even though we are waiting for the direct detection of this signal, we already have upper limits on the spherically averaged 21-cm power spectrum from these interferometers [12–15, 17–21]. Recently, [22] has provided the upper limits on the 21-cm bispectrum from MWA observations.

The redshifted 21-cm signal coming from the EoR is known to be highly non-Gaussian [23–25]. This non-Gaussianity arises due to the underlying non-linear matter density distribution which is traced by the ionizing sources, and the evolving ionized regions. The power spectrum, which gives the variance of the signal at different k -modes, is an optimal statistic to describe any Gaussian field completely. Since the redshifted 21-cm signal is highly non-Gaussian, the power spectrum alone cannot capture all the information available in the signal. To extract the non-Gaussian information present in this signal, we need higher-order statistics such as the bispectrum [26–39]. Bispectrum, the Fourier conjugate of the three-point correlation function, has more information about the field than the power spectrum since it can capture the correlations present between the different scales. Therefore, it can provide more information about the astrophysical processes that drive the evolution of this non-Gaussianity. This extra information can help to put tighter constraints on the astrophysical parameters of the EoR models. In [39], it is demonstrated that considering the bispectrum for all unique triangle configurations as a summary statistic can provide tighter constraints on the EoR model parameters compared to the power spectrum.

One of the primary goals of the EoR observations through different radio interferometric arrays is to estimate the reionization model parameters from the observed redshifted 21-cm signal statistics through a Bayesian inference framework [19, 40–44]. Such a parameter estimation pipeline functions by having many random walkers scout the unknown multi-dimensional parameter space by estimating the target observable statistics at every step, which can typically be within the range of $10^4 - 10^6$ steps. Generally, one uses radiative transfer or seminumerical simulations to model these observable statistics; however, these simulations can be computationally very expensive and time-consuming to use in such a Bayesian inference pipeline. One way to overcome this bottleneck is to use neural network-based emulators for the target statistics. Once these emulators are trained on a sufficiently large and optimally sampled simulated signal database to predict the target signal statistic, they can predict the same for an unknown set of parameters in a few seconds. Thus, such emulators can be an ideal candidate to use in the inference pipelines where we need fast prediction of the observables iteratively. Recently, there have been several works on developing neural networks to emulate different EoR 21-cm signal statistics [39, 45–48]. Additionally, there have been efforts towards developing neural networks to directly predict reionization model parameters from signal statistics or signal maps [49–55]. Most of the neural networks used in these works have a fundamental drawback, i.e., they can only produce point values of the target statistics. Further, Neural networks are prone to overfitting, and thus it is very difficult to tune hyperparameters optimally for a given training dataset. Therefore, if a neural network is not trained robustly enough, it can make erroneous predictions without any warning flags. Since it can only predict point values, there is no direct way to quantify the uncertainty associated with its prediction. When such neural network architectures are used to make emulators used in the Bayesian inference pipeline, these prediction uncertainties are not naturally propagated to the estimation of the model parameters. One way to propagate these errors is using GPR as described in [56]. However, emulator errors propagated in this fashion is constant throughout the parameter space, which is not always true. Thus, it is very important to quantify these prediction uncertainties for every step in the parameter space.

To address these shortcomings of deterministic emulators, we have developed a Bayesian neural network (BNN) based emulator for EoR 21-cm observables. Bayesian neural networks use Bayesian inference to train the network for a given training dataset. The key difference between BNN and ANN is that a BNN learns the distribution of the weights and biases

(i.e., network parameters) instead of point values, as is the case for an ANN. Thus, when we draw a prediction from a BNN, it gives us the distribution of the predicted quantity corresponding to the distribution of its tuned network parameters. Therefore, we can directly quantify the uncertainties associated with the network predictions. Further, we can propagate these prediction uncertainties through our Bayesian inference pipeline to get a more robust estimation of the reionization parameters. There have been a couple of works related to the use of BNN for the estimation of the EoR parameters [57, 58]. Both of these works mainly focus on using the BNN for simulation-based inference (SBI) to directly estimate the EoR parameters. In this article, we train a BNN (via a database of simulated EoR 21-cm signal) to emulate the signal statistics and then use these emulators in an MCMC Bayesian inference pipeline to estimate the EoR parameters. [58] has compared an ANN emulator-based MCMC pipeline with SBI to estimate the parameters. They found that the former systematically performs better than the latter (including BNN). We have developed a BNN-emulator-based parameter estimation pipeline and shown that the BNN-based emulator can provide better and more robust constraints compared to an ANN-based emulator in an MCMC inference pipeline because of the quantification of the prediction uncertainties. For this work, we have developed a BNN-based emulator for both the power spectrum and the bispectrum of the 21-cm signal. We compare the performance of the BNN emulator with the corresponding ANN emulator in constraining the EoR parameters. We also systematically reduce the size of the training dataset and compare the performance of both types of emulators for the reduced training datasets to understand the minimum amount of training data required to get robust estimations. As mentioned earlier, the redshifted 21-cm signal from the EoR is highly non-Gaussian; thus, the use of a higher-order statistic, e.g., bispectrum, can provide better constraints on the reionization parameters compared to the power spectrum ([39] has demonstrated it with an ANN emulator). Therefore, we compare the performance of the EoR 21-cm power spectrum and bispectrum in constraining the EoR model parameters when both of them are emulated by two robustly trained BNN emulators.

This article is organized as follows: In section 2, we describe the details of the simulation used to generate the 21-cm signal maps used as the training dataset. Section 3 describes how we estimate the signal statistics from the reionization maps. In section 4, the drawbacks of the ANN-based emulators and how BNN-based emulators help to overcome those drawbacks are discussed. In section 5, we have described ANN and BNN emulator architectures and their performance for both power spectrum and bispectrum. In section 6, we compare the results of parameter estimations provided by ANN and BNN emulators. Finally, in section 7, we summarize and discuss our findings.

2 Reionization 21-cm Signal Simulation

We use a semi-numerical simulation based on excursion set formalism [59] to generate the redshifted 21-cm brightness temperature field of the EoR. This process involves three main steps. First, we generate the dark matter (DM) density field using a dark matter-only particle-mesh (PM) N -body code [60, 61] at redshift $z = 7$. We simulated our DM density field on 3072^3 grids with the spatial grid resolution of 0.07 cMpc (comoving Mpc), which makes our simulation volume $[215.04 \text{ cMpc}]^3$ and DM particle mass resolution of $1.09 \times 10^8 M_\odot$. We use the Friends-of-Friends (FoF) [61, 62] algorithm to identify the collapsed halos within the distribution of DM particles using a fixed linking length of 0.2 times the mean interparticle distance. The criterion to identify the collapsed halo is that it should

have at least 10 DM particles, which results in a minimum halo mass of $1.09 \times 10^9 M_\odot$. For the final step, we apply an excursion set-based algorithm[59] to the DM field and the halo catalog to generate the ionization field using the semi-numerical code **ReionYuga**¹ [63–65]. This code assumes that the hydrogen density perfectly follows the underlying DM density and that all the DM halos host the sources that produce the ionizing photons. There are three free parameters in this simulation framework: 1) $M_{\text{h,min}}$: This parameter sets the lower cutoff for the mass of the halo that participates in the reionization process. Only halos with a mass higher than $M_{\text{(h,min)}}$ have the sources that produce ionizing photons, 2) N_{ion} : We consider that the number of ionizing photons produced by the sources are proportional to the mass of the host DM halo and this dimensionless proportionality constant is N_{ion} . The N_{ion} is a parameter that essentially quantifies the efficiency of the ionizing sources. 3) R_{mfp} : This parameter represents the mean free path of ionizing photons in the IGM. For a detailed discussion of the reionization parameters, interested readers may refer to [63–65]. All the length scales mentioned in this article are in comoving units, so hereafter we drop the cMpc notation in favor of Mpc.

We generate our training data set by varying these parameters in the following range - $M_{\text{h,min}}(\times 10^8 M_\odot) \in [10, 800]$, $N_{\text{ion}} \in [10, 200]$, and R_{mfp} (Mpc) $\in [1.12, 40.32]$. We use a uniform grid sampling (please refer to [39] to know more about the effect of sampling) in this 3D parameter space corresponding to the given parameter ranges, which allows us to sample a maximum of 7200 parameter sets, for which we simulate the EoR 21-cm brightness temperature maps at redshift $z = 7$. Note that when comparing with a realistic observation via a radio interferometric array, one should ideally simulate the signal for the whole range of target frequency channels (or redshifts) to include the lightcone effect [31, 66, 67] and for a volume representing the observing volume constituted by the field-of-view of the array multiplied with the comoving depth determined by the frequency or redshift range. This will require one to simulate the signal for a large number of coeval cubes at different redshifts within the observing redshift range and then constitute the signal lightcone by stitching the slices from coeval cubes. Performing this task for 7200 parameter sets will require a significant amount of computing time and storage space for those coeval and lightcone cubes. Thus, for the purpose of the development of these emulators, we focus on a specific redshift and simulate one coeval cube centred at that redshift while varying the EoR parameters for building the signal maps for the training and validation dataset. This ensures that we capture the variation of the astrophysical parameters within our training set while not going overboard with the computing overhead. In the future, we wish to upgrade our emulators by training them on signal lightcones with adequate computing resources.

3 Estimating Signal Statistics

3.1 Power Spectrum

The power spectrum is one of the most popular Fourier statistics to describe any field. For the 21-cm signal, the power spectrum can be expressed as:

$$\langle \Delta_b(\mathbf{k}) \Delta_b(\mathbf{k})^* \rangle = V P(\mathbf{k}), \quad (3.1)$$

where $\Delta_b(\mathbf{k})$ is the Fourier transform of the brightness temperature field, and V is the volume of the simulation box. This power spectrum gives the power of the field for a particular

¹<https://github.com/rajeshmondal18/ReionYuga>

k -mode. We estimate the bin-averaged power spectrum from our 3D simulation box by considering a spherical bin of radii between k and $k + dk$. Thus, power spectrum of the i^{th} bin is expressed as:

$$\bar{P}_i(k) = \frac{1}{V N_k} \sum_{\mathbf{k} \in i} \langle \Delta_b(k) \Delta_b(-k) \rangle, \quad (3.2)$$

where N_k is the total number of k modes inside the i^{th} bin. For this work, we have used the normalized dimensionless power spectrum expressed as,

$$\Delta^2(k) = \left(\frac{k^3}{2\pi^2} \right) \bar{P}(k) \quad (3.3)$$

3.2 Bispectrum

The bispectrum is the Fourier conjugate of the three-point correlation function. For the 21-cm brightness temperature field, the bispectrum $B(k_1, k_2, k_3)$ is defined as:

$$\langle \Delta_b(k_1) \Delta_b(k_2) \Delta_b(k_3) \rangle = V \delta_{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, 0} B(k_1, k_2, k_3), \quad (3.4)$$

where $\Delta_b(\mathbf{k})$ is the Fourier transform of the brightness temperature field, and V is the volume of the simulation box. The Kronecker delta $\delta_{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, 0}$ equals to one when the condition $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ is met, which ensures that only closed k -triangles contribute to the bispectrum; otherwise, the Kronecker delta is zero.

We estimate the binned bispectrum,

$$\bar{B}_i(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{V N_{\text{tri}}} \sum_{[\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0] \in i} \langle \Delta_b(\mathbf{k}_1) \Delta_b(\mathbf{k}_2) \Delta_b(\mathbf{k}_3) \rangle, \quad (3.5)$$

from the 3D simulation box following the algorithm presented in [26, 27]. Here, N_{tri} represents the total number of statistically independent closed k -triangles in the i^{th} bin. The bispectrum depends on the shape and size of the triangles in Fourier space. We parametrize our bispectrum via two additional parameters that help to characterize the shape of the triangle in the Fourier space:

$$n = \frac{k_2}{k_1}, \quad \text{and} \quad \cos \theta = -\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}, \quad (3.6)$$

where k_1 and k_2 are the magnitudes of the vectors \mathbf{k}_1 and \mathbf{k}_2 , respectively, and θ is the angle between them. To estimate the bispectrum of the field, we are identifying the triangles in the Fourier space, with side lengths being the magnitude of the three wave vectors. There can be multiple triangles with the same shapes, depending on the values of these vectors. To identify all the unique triangle configurations, we apply the following uniqueness conditions presented in [68]

$$k_1 \geq k_2 \geq k_3, \quad (3.7)$$

$$0.5 \leq n \leq 1.0, \quad (3.8)$$

$$0.5 \leq \cos \theta \leq 1.0. \quad (3.9)$$

Similar to the power spectrum, we use a normalized dimensionless bispectrum following [27],

$$\Delta^3(k_1, n, \cos \theta) = \left(\frac{k_1^6 n^3}{(2\pi^2)^2} \right) \bar{B}(k_1, n, \cos \theta). \quad (3.10)$$

4 Emulators: ANN vs BNN

4.1 Artificial Neural Network (ANN)

An artificial neural network (ANN) is the fundamental neural network architecture of Deep Learning. The fundamental building block of ANN architecture is called a “Neuron” or a “Node”. Multiple neurons make a layer, and these fully interconnected layers make a neural network. ANN consists of three layers: Input, Hidden, and Output. Information is processed as it flows through layers of the network, with each neuron in each layer making decisions based on its input and contributing to the final output. The learning or training of NN consists of two processes, namely forward propagation and backward propagation.

In forward propagation, the training data is given to the network via the input layer and passes through the hidden layers to the output layer, generating an output. Every neuron in the network has a point value of weight and bias associated with it. Each neuron takes the input from all the neurons of the previous layer; then it takes the linear combination of all the values with the weight and bias of the corresponding neuron. The activation function takes this linear combination as input, and the output of the activation function goes to the next layer. The activation function introduces nonlinearity to the network. When the output layer receives the results from the preceding hidden layer, it predicts the output. This predicted output is compared to the true output using a *loss function*. Our goal is to minimize the value of the loss function, i.e., the weight and bias of all neurons should be tuned in such a manner that the network is able to predict the output values with minimal loss of some desired quantity. To achieve this, based on the value of the loss function, one has to update the values of every weight and bias. This task is taken care of by the backward propagation.

In backward propagation, the loss (or difference) between predicted and actual output values of training data is calculated using some predefined loss function. The weight and bias of each neuron are updated during each iteration of the training, such that the loss tends to zero. The task of determining the updated values for each neuron is taken care of by the *optimization algorithms*. One of the most common optimization algorithms is the gradient descent algorithm [69]. The basic idea behind gradient descent is to update the parameters in the direction opposite to the gradient of the loss function with respect to the parameters. This is done because the gradient points toward the steepest increase of the function, so moving in the opposite direction should lead to a decrease in the function value. The gradients of the loss with respect to each weight and bias are computed using the chain rule of calculus, allowing the network to understand how each parameter contributes to the error. The algorithm then adjusts the weights and biases to reach the minimum of the loss function. After weights and biases have been changed, again, forward propagation happens, and the loss is calculated. According to this, the values of weight and biases are updated. This iterative process is called training the network. It continues until we get the values of weights and biases such that we achieve the minimum desired loss. For a detailed discussion of ANN, interested readers may refer to [70].

Drawbacks of ANN

ANN architecture and training process have many hyperparameters, and it is a challenging task to tune those hyperparameters to get the combination for which we can achieve the minimum loss. So, for ANNs, we find the set of hyperparameters for which we get minimum loss, and corresponding to that, we have minimum loss point estimate values for ANN model parameters (i.e., weights and biases of each neuron), which means we have a single point

values of weight and bias for each neuron, tuned using the backpropagation algorithm. So, when we get a prediction from the ANN, it is also a single-value prediction. The problem with single-value prediction is that the predicted value depends on the quality of the data in your training dataset, and we do not know the uncertainty related to this prediction. The quantification of uncertainties related to deep learning models is very crucial [71]. There are fundamental uncertainties related to deep learning models: Aleatoric and Epistemic uncertainty [72–74]. Aleatoric uncertainty is inherent to the data itself. It comes from the underlying randomness or just the lack of information in the data. This type of uncertainty can not be reduced just by providing more training data to the network. On the other hand, Epistemic uncertainty is the uncertainty related to the model, which tries to explain or, in our case, tries to emulate the data itself. This comes from uncertainties in the neural network model architecture (mainly from choices of hyperparameters).

One aims to make an emulator for a given summary statistic and use it to do the Bayesian inference to constrain the reionization model parameters. Since ANNs give point value predictions, they do not quantify all the uncertainties associated with them, which we mentioned above. There is no way to know how good (or bad) the prediction is in unknown regions of parameter space, as the ANN does not provide the measure of uncertainty in its predictions, and it can confidently predict highly inaccurate values. Use of such an emulator for Bayesian inference may suffer from a lack of trust in the inferred parameters, as there are unknown uncertainties from emulators propagating in the inference pipeline over which we do not have any control. Therefore, we need models that have some measure of the uncertainty associated with their predictions.

4.2 Bayesian Neural Network (BNN)

To overcome the drawbacks of ANNs, we need to consider models capable of quantifying uncertainty related to their predictions. One way to do this is to incorporate Bayesian statistics into the deep learning model. The fundamental idea behind Bayesian statistics is to use a probabilistic approach to quantify the uncertainties in the inferred parameters. Therefore, we want to train neural networks through the Bayesian inference approach. In this approach, we consider network parameters, e.g., weights and biases, as random variables and try to infer them given some prior dataset [75]. These types of neural networks are known as Bayesian neural networks (BNN) [76–79]. BNNs have some apparent advantages over ANNs. They are trained through the Bayesian approach, so they naturally incorporate the probabilistic way to quantify uncertainties. They are less prone to overfitting and underfitting the data. They can also improve by distinguishing between epistemic and aleatoric sources of uncertainty [74]. Furthermore, due to their probabilistic nature of training, they can, in principle, be trained even for a very small training dataset without overfitting. For predictions, they provide a large epistemic uncertainty for the out-of-distribution (OOD) points instead of confidently giving the wrong outputs like ANNs. All of these make BNN trustworthy and reliable for making a robust emulator to perform parameter estimation.

In BNNs, the network weights and biases are considered random variables. Given some training dataset $\mathcal{D} : \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where (x, y) is a pair of input and output. Consider w be the set of parameter values that belong to \mathbb{R}^N space, N = number of parameters. We set some prior $p(w)$ for the distribution of the parameters w . The posterior distribution of parameters following Bayes’ rule will be:

$$p(w|\mathcal{D}) = \frac{p(\mathcal{D}|w) p(w)}{p(\mathcal{D})}, \quad (4.1)$$

where $p(\mathcal{D}|w)$ is the likelihood and $p(\mathcal{D})$ is the model evidence. For simplicity, we will assume a Gaussian prior for the parameters. So, after marginalizing the posterior $p(w|\mathcal{D})$, we will get the Gaussian distribution of each parameter in terms of their mean and variance.

The training problem of BNN is now reduced to a parameter inference problem. There are two main computational approaches to infer the posterior [76, 78]: 1) Sampling techniques like Markov Chain Monte Carlo (MCMC) and 2) Approximation methods like Variational Inference. Even though variational inference techniques are less expensive to perform than traditional MCMC methods, in this paper, we have used the MCMC method to train our BNNs simply because they are quite accurate compared to approximation methods. Our BNNs do not have very complex architectures, which makes them ideal candidates for training using MCMC without too much computational overhead. The main advantage of sampling methods is that when we increase the number of iterations, we asymptotically should converge to the true posterior. To train our BNN models, we use the No-U-Turn sampling (NUTS) technique [80]. The NUTS algorithm is an advanced version of the Hamiltonian Monte Carlo (HMC) sampling technique. The HMC sampling uses the Hamiltonian equations to explore the parameter space using a random variable and its auxiliary momentum vector [81]. It requires a time step and a total number of steps to perform the numerical integration. The NUTS algorithm learns this step size during the burn-in and automates the whole calculation without deciding the step size at each iteration.

Once the training is done for a given dataset and the posterior distribution over the model weights, i.e., $p(w|\mathcal{D})$ is obtained, it can then generate the predictions for some new unknown input data point x^* by marginalizing the likelihood with the posterior distribution.

$$p(y^*|x^*, \mathcal{D}) = \int p(y^*|x^*, w) p(w|\mathcal{D}) dw. \quad (4.2)$$

As mentioned above, in this specific work, all parameters are in terms of the Gaussian mean and variance. Hence, we get a BNN prediction in terms of the mean and variance of the predicted quantity.

5 Emulation of Signal Statistics

To build the training and test sets for our emulators, we first estimate the target signal statistics, i.e., power spectrum and bispectrum, from the simulated 21-cm maps for 7200 parameter sets as described in Sec 2. Out of these 7200 signal estimates, we set aside randomly selected 50 samples as the test set to check the performance of our emulators, and out of the remaining data, 80% is used for training, and 20% is used to validate the corresponding network’s performance. We have used the KERAS² [82] library based on the TensorFlow³ [83] package to develop the ANN-based emulators and PYRO⁴ [84] probabilistic programming library based on the PyTorch⁵ [85] package to develop BNN-based emulators. Initially, we train our emulators for the entire dataset and tune the network to enhance their

²<https://keras.io/>

³<https://www.tensorflow.org/>

⁴<https://pyro.ai/>

⁵<https://pytorch.org/>

prediction accuracy and relative errors. However, to test the robustness of these emulators, we gradually reduce the training data set volume and try to identify the minimum number of points in the parameter space that are required for getting an optimally trained emulator, both for the power spectrum and bispectrum.

5.1 Power spectrum emulation

We have developed emulators to predict the 21-cm signal power spectrum for three input reionization parameters. The networks are trained to predict the power spectrum for 7 k -bins in the range $[0.11-2.75] \text{ Mpc}^{-1}$. This range is chosen in accordance with the range of k -modes that upcoming telescopes will be able to probe. Both ANN and BNN emulator architectures have 3 nodes for three input parameters as the input layer and 7 nodes corresponding to the power spectrum values for each k -bin in the output layer. The detailed architecture of the ANN-based emulator is provided in table 1. As one can see, the architecture of the ANN-based emulators has more layers and nodes compared to the BNN-based emulator, which means BNN-based emulators are able to capture more information out of data than the ANN-based emulators with a smaller architecture. We have used ELU activation for both ANN and BNN emulators. The ANN emulator is trained for 300 epochs with a learning rate of $\sim 10^{-3}$. The BNN-based emulator uses Bayesian inference to train the emulator given a training dataset. So they have different sets of hyperparameters than the ANN emulators. For training the BNN, one needs to specify the total number of steps for the MCMC chain to run and the likelihood function. We use the multivariate Gaussian log-likelihood as the likelihood function and train our model for 1000 MCMC steps, with the initial 500 steps serving as warm-up steps. The number of steps is selected based on two main criteria: convergence and training time. More steps take a longer time to train the model. So, we found that this number of steps was enough to train the model properly. If we consider more steps, the quality of the model does not improve considerably compared to the increased training time.

The performance of both of these emulators on an unseen test data for parameters $M_{(\text{h,min})} = 467.37 \times 10^8 M_{\odot}$, $N_{\text{ion}} = 181.05$, $R_{\text{mfp}} = 12.65 \text{ Mpc}$ and mass average neutral fraction $\bar{x}_{\text{HI}} = 0.377$ is shown in the figure 1. This particular choice of the parameter is to demonstrate the performance of our emulators at a low neutral fraction ($\bar{x}_{\text{HI}} = 0.377$), which was absent in [39] due to a lack of training dataset. The true simulated values of the power spectrum, shown in black, are compared with the power spectrum predicted by the ANN-based emulator (in red) along with the mean prediction of the power spectrum by the BNN-based emulator (in blue). We also show the prediction uncertainty associated with the BNN-based emulator with 3σ confidence in the blue shaded region. We can see that the mean values of the power spectrum predicted by the BNN-based emulator are similar to or better than the point estimation for most of the k -bins provided by the ANN-based emulator. For all k -bins, the mean values predicted by the BNN-based emulator are within 3σ of the true power spectrum values, whereas there is no way to quantify the same for the power spectrum predicted by the ANN-based emulator.

Furthermore, as mentioned earlier, we perform the robustness study by systematically decreasing the total training dataset. The initial training was done using 7150 training dataset. We randomly sample from this complete dataset to generate the smaller training dataset. Here, we show the results for the 1000 and 500 datasets, which we found as the lowest number of training datasets for which both ANN-based and BNN-based emulators were performing considerably well. Performance of emulators trained on these smaller datasets for

$[M_{h, \min}(10^8 M_{\odot}) = 467.32, N_{\text{ion}} = 181.05, R_{\text{mfp}}(\text{Mpc}) = 12.65, x_{\text{HI}} = 0.377]$

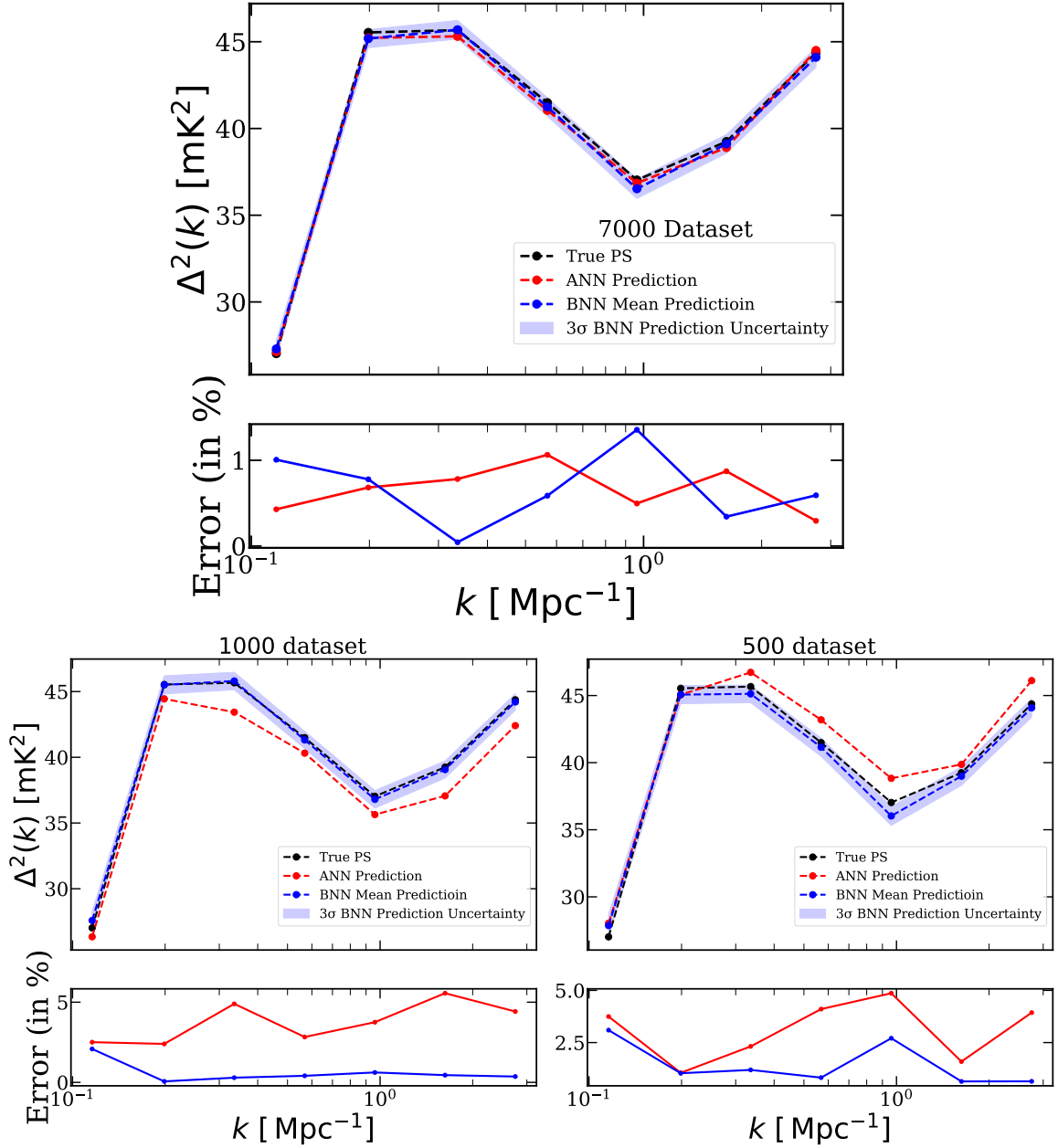


Figure 1. Plot shows the performance of ANN and BNN power spectrum emulators trained on different dataset sizes. The top plot shows the results for the complete dataset, and the bottom plots show the results for 1000 (left) and 500 (right) datasets. Here, red and blue lines correspond to the power spectrum prediction of the ANN and BNN emulators, respectively. The blue shaded region is the 3 σ uncertainty associated with the mean prediction by BNN. The true power spectrum is shown with the black line. The bottom panel of each plot shows the percentage errors of the ANN prediction in red and the BNN mean prediction in blue.

the same unseen reionization parameters is shown in figure 1. We can see that as we decrease the training dataset, the point prediction of the ANN-based emulator deviates considerably from the true values of the power spectrum. This happens because for smaller training datasets, ANN-based emulators tend to overfit the training data and thus perform poorly on the unseen test data. Meanwhile, the BNN-based emulators perform better, as we can see that the uncertainties associated with the BNN predictions have increased; however, the true values are still within the 3σ confidence of the mean prediction.

Layer	Nodes	Activation	Layer	Nodes	Activation
Input Layer	3	-	Input Layer	3	-
Hidden Layer 1	1024	ELU	Hidden Layer 1	50	ELU
Hidden Layer 2	512	ELU	Hidden Layer 2	100	ELU
Hidden Layer 3	256	ELU	Hidden Layer 3	50	ELU
Hidden Layer 4	128	ELU	Output Layer	7	-
Hidden Layer 5	64	ELU			
Hidden Layer 6	32	ELU			
Output Layer	7	-			
ANN-based emulator			BNN-based emulator		

Table 1. Model architecture of power spectrum emulators: ANN-based emulator (Left) and BNN-based emulator (Right)

5.2 Bispectrum emulation

We consider the bispectrum corresponding to a total of 9 k_1 -bins in the range $[0.14 - 2.37] \text{ Mpc}^{-1}$. Similarly to the power spectrum, this range of k_1 -bins is also chosen following the range of k -bins that upcoming telescopes will be able to probe. For each k_1 -bin, we have 11 bins for n and 50 for the $\cos\theta$ values. That gives us bispectrum values for a total of 550 bins of triangle configurations. However, as mentioned earlier, if we consider only the unique triangle shapes ($n \cos\theta \geq 1$), the number of triangle configurations reduces to 328 for each k_1 -bin. Furthermore, unlike the power spectrum, which is always positive, the bispectrum can have both positive and negative values. Since we have a total of 9 k_1 -bins, we want to develop the emulator that predicts the sign and magnitude of a total of $(9 \times 328 =)$ 2952 triangle configurations of the bispectrum. The values of the bispectrum have a high dynamic range depending on the k_1 -bins, and a single ANN architecture cannot capture the high dynamic range and sign variability of all k_1 -bins for all parameters of the training dataset. So, to overcome this problem, we made nine different ANN emulators corresponding to 9 k_1 -bins, and each ANN emulator is trained to predict the magnitude and sign of the bispectrum via regression and classification, respectively. There are two main architectures for our ANN emulator. Both architectures have an input layer with three neurons and two output layers having 328 neurons, one for sign and one for magnitude. We have used the ReLU activation function and mean squared error for the loss function during the training of all emulators. The learning rate is $\sim 10^{-3}$ for all emulators, but exact values differ for each emulator. The detailed architectures of these ANN emulators are provided in table 2. The prediction of ANN trained on a full dataset compared to the simulated bispectrum is shown in figure 2.

We train the BNN-based bispectrum emulator on the same datasets. Similar to the power spectrum, we require less depth in our BNN networks compared to ANN-based net-

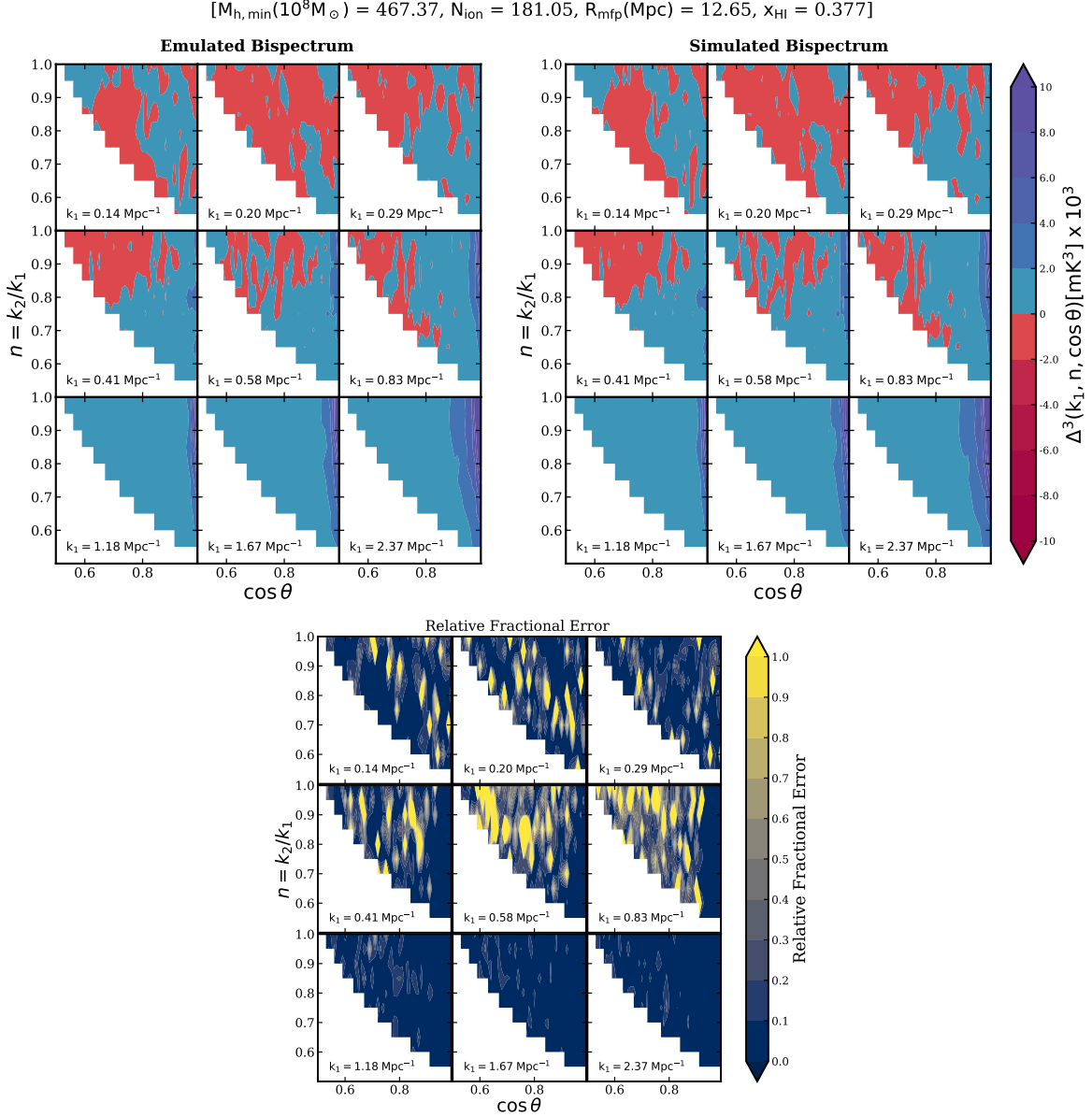


Figure 2. Bispectrum prediction by ANN emulators. The top left panel shows the emulated bispectrum, and the right panel shows the simulated bispectrum for the same set of EoR parameters. The relative fractional error between the simulated and emulated bispectrum is shown in the lower panel. The islands of the high relative fractional errors mostly correspond to the errors in the sign prediction.

works to capture more information about the bispectrum. Similar to ANN bispectrum emulators, we train 9 BNN models for each k_1 -bin. Each BNN has an input layer with three neurons and only one output layer with 328 neurons; since we use Bayesian inference, we do not separate signs and magnitudes for BNNs. All BNNs have three hidden layers. We have used the ELU activation function for all the BNN architectures. We perform a total of 2000 MCMC steps, with the initial 500 steps being the warm-up steps to train the emulators. The performance of the BNN-based emulator is shown in figure 3. We compare the prediction of

Layer	Nodes	Activation	Layer	Nodes	Activation
Input Layer	3	-	Input Layer	3	-
Hidden Layer 1	2255	ReLU	Hidden Layer 1	2255	ReLU
Hidden Layer 2	1524	ReLU	Hidden Layer 2	1524	ReLU
Hidden Layer 3	1008	ReLU	Hidden Layer 3	1008	ReLU
Output Layer 1 (Magnitude)	328	-	Hidden Layer 4	956	ReLU
Output Layer 2 (Sign)	328	-	Output Layer 1 (Magnitude)	328	-
			Output Layer 2 (Sign)	328	-

Table 2. ANN model architecture of the two bispectrum emulators: left for the first four k_1 -bins and right for the remaining five k_1 -bins.

Layer	Nodes	Activation
Input Layer	3	-
Hidden Layer 1	128	ELU
Hidden Layer 2	256	ELU
Hidden Layer 3	512	ELU
Output Layer	328	-

Table 3. BNN model architecture of bispectrum emulator for all k_1 -bins.

both the emulator using the relative fractional error. For both cases, there are islands of very high (some even more than 100%) as evident from the figures 2 and 3. These high relative fractional errors are mostly due to the wrong sign prediction by the emulators, as they occur for the triangles where there are clear sign variations in the emulated bispectrum. Also, the k_1 -bins where there is no sign change, i.e., $k_1 = (1.18, 1.67, 2.37)$ Mpc^{-1} , do not have these islands of high relative fractional error as bispectrum values for all the triangles are positive. BNN predicted bispectrum in general has fewer of these islands than the ANN prediction, which shows that training the neural network through Bayesian inference improves the performance of the network.

Similar to the power spectrum, we checked the robustness of the BNN-based bispectrum emulators. For this, we trained our neural networks for a smaller dataset of 500 and 1000 randomly sampled parameter sets from the full dataset. The numbers of hidden layers are [128, 256] and [52, 128] for the 1000 and 500 datasets, respectively. The predictions by these networks are compared to the networks trained on the complete data, as shown in figure 4. Here we have shown the predictions for two triangle configurations: equilateral and squeezed-limit triangles. The top row shows the prediction of the equilateral triangles (left) and squeezed-limit triangles (right) given by both ANN-based and BNN-based emulators. The shaded regions show the 3σ errors associated with the prediction of the BNN-based emulators. The same is shown for emulators trained with 1000 and 500 datasets in the second and third rows, respectively. We can see that the 3σ errors associated with the BNN emulator prediction of the equilateral triangles are greater than the squeezed-limit triangle because the equilateral triangles have more sign flips for the bispectrum values, whereas the

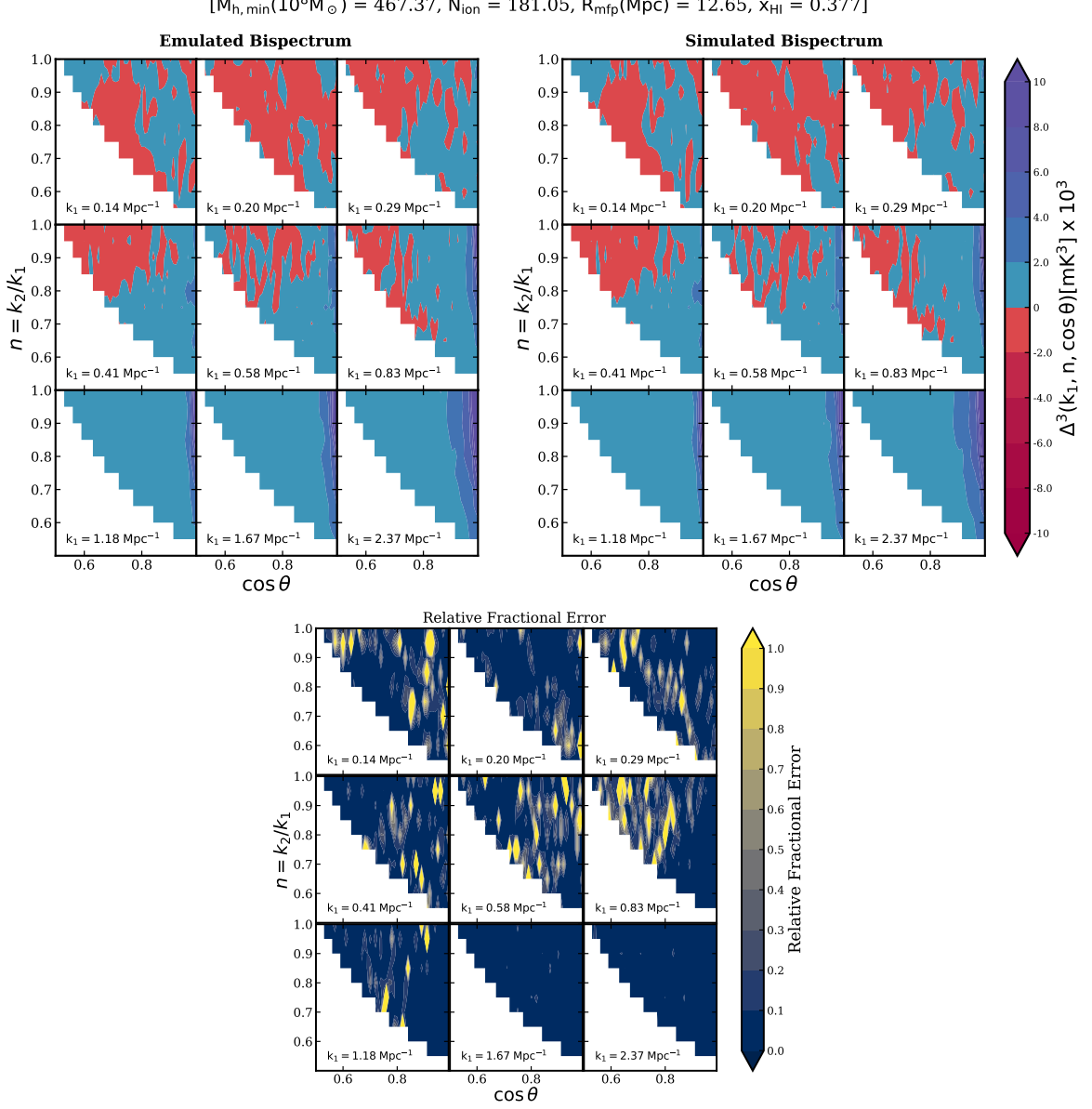


Figure 3. Bispectrum prediction by BNN emulators. The top left panel shows the mean values of the emulated bispectrum, and the right panel shows the simulated bispectrum for the same set of EoR parameters. The relative fractional error between the simulated and emulated bispectrum is shown in the lower panel. We can see that the island of high relative errors is overall decreased from the ANN prediction.

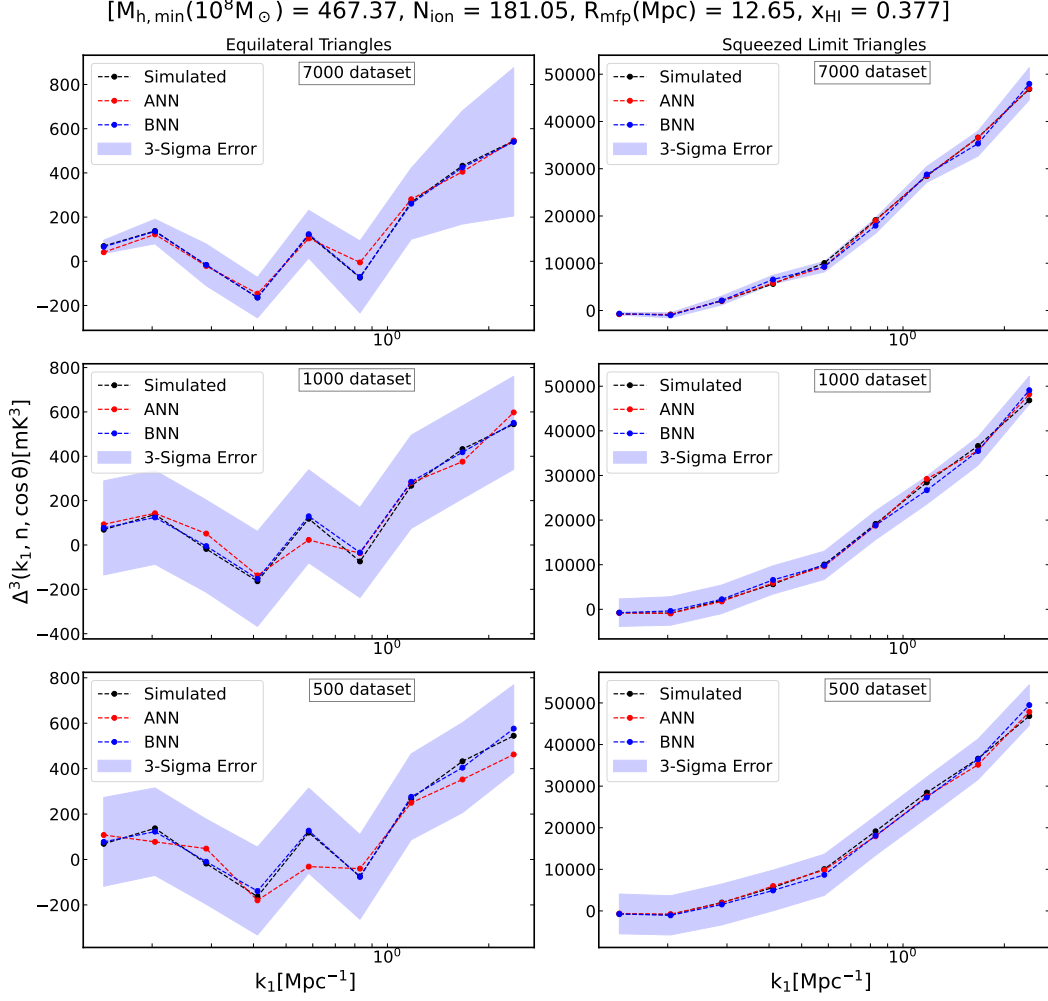


Figure 4. Line plots for the two triangle configurations of the bispectrum. The left column shows emulator predictions for equilateral triangle configurations, and the right column shows predictions for squeezed-limit triangle configurations. The red line shows the bispectrum prediction by the ANN, and the blue line shows the prediction by the BNN, with the blue shaded region representing the 3σ error on the BNN prediction. Both predictions are compared with the simulated bispectrum presented with the black line. The rows represent the total number of training datasets used to train the corresponding emulators.

squeezed-limit bispectra remain positive always. Therefore, it is easier for emulators to train and predict the squeezed-limit bispectrum, and consequently, the errors associated with its prediction are less compared to the equilateral bispectrum. The BNN prediction errors also increase with a decrease in the training dataset for both triangle configurations. We can also see from the figure 4 that the mean predictions of the BNN emulators (in blue) are better than the point predictions of the ANN emulators (in red), even though both are trained on the same dataset. This fact is also evident in the relative fractional error plot shown in figure 2 and figure 3.

6 Parameter Estimation

Here, we constrain the reionization parameters using Bayesian inference and use ANN and BNN-based emulators for the power spectrum and bispectrum as forward models in the Bayesian inference pipeline to get the posterior of the parameters. Since we are interested in parameter estimation for a given reionization model, we can safely ignore model evidence as a normalizing constant. Then, posterior becomes proportional to the product of the likelihood $p(\mathcal{D}|\theta, \mathcal{M})$ and the prior $p(\theta|\mathcal{M})$, where θ is a parameter vector, \mathcal{D} is observed data and \mathcal{M} is model. We have assumed the uniform prior over all three parameters. A multivariate Gaussian log-likelihood is defined as:

$$\ln \mathcal{L} = \ln p(\mathcal{D}|\theta, \mathcal{M}) = -\frac{1}{2}[(\mathbf{P}_{\text{ref}} - \mu)^T (\Sigma)^{-1} (\mathbf{P}_{\text{ref}} - \mu)] - \frac{1}{2} \ln(2\pi |\Sigma|). \quad (6.1)$$

Here \mathbf{P}_{ref} denotes the observed data (for our case, it can be mock power spectrum or bispectrum data), μ denotes our model prediction (prediction of ANN or BNN emulators for respective statistic), and Σ is the error covariance.

6.1 Estimation of Error Covariance

Prior estimation of errors in the inference pipeline is necessary for robust likelihood estimation. The error covariance matrix has information about errors associated with the observation and, in the case of BNN-based emulators, errors associated with the model prediction. The observed 21-cm signal from EoR has uncertainties due to many observational effects, such as foreground contamination, sample variance, calibration errors, and instrumental noise. We consider only sample variance and instrumental noise for our analysis, assuming other contaminations are well-modeled and removed from the observed signal. Considering the 21-cm signal and system noise will not be correlated with each other, we can write the total covariance as:

$$\Sigma = \sigma_{\text{T}}^2 = \sigma_{\text{SV}}^2 + \sigma_{\text{N}}^2, \quad (6.2)$$

σ_{SV}^2 and σ_{N}^2 are covariance contributions from sample variance and system noise, respectively. These covariance calculations are different for the power spectrum and bispectrum. For the power spectrum, we calculate the sample variance from the binned power spectrum following:

$$\sigma_{\text{SV}}^2(P_i) = \frac{\left(\Delta_b^2(k_i)\right)^2}{N_{k_i}}, \quad (6.3)$$

where $\Delta_b^2(k_i)$ and N_{k_i} are the power and total sample of k -bins in the i^{th} bin, respectively. Similar to the power spectrum, the sample variance of the bispectrum is obtained from the bin-averaged bispectrum $\bar{B}_i(k_1, n, \cos \theta)$ from,

$$\sigma_{\text{SV}}^2(B) = \frac{\left(\Delta_b^3(k_1, n, \cos \theta)\right)^2}{N_{\text{tri}}}, \quad (6.4)$$

here $\Delta_b^3(k_1, n, \cos \theta)$ is normalized binned bispectrum, and N_{tri} is the total number of triangles in i^{th} triangle bin. The instrumental noise contribution for the power spectrum can be calculated as follows:

$$\sigma_{\text{N}}^2(P_i) = 2 \frac{V_f}{V_P} \left(\Delta_{\text{N}}^2(k_i)\right)^2. \quad (6.5)$$

Here $V_f = 2\pi^3/V_s$ is the volume of the fundamental cell in the Fourier domain, where V_s is the total survey volume, $V_P = 4\pi k_i \Delta k_i$ is the shell volume of the i^{th} bin with Δk_i bin width and $\Delta_N^2(k_i) = k_i^3 P_N(k_i)/(2\pi^2)$ is normalized noise power in the i^{th} bin. The noise power spectrum $P_N(k_i)$ for a radio interferometer is given as [86, 87]:

$$P_N(k, z) = \frac{T_{\text{sys}}^2(z) \chi^2(z) r_\nu(z) \lambda^4(z)}{A_{\text{eff}}^2 t_{\text{obs}} N_{\text{pol}} \Sigma(\mathbf{U}, z) \nu_{21\text{cm}}}. \quad (6.6)$$

Here, $\lambda(z) = 21 \times (1 + z)$ cm and $\nu_{21\text{cm}} = 1420$ MHz are the wavelengths at redshift z and restframe wavelength of the 21-cm signal. The $\chi(z)$ is comoving distance to redshift z , and $r_\nu = (c/H(z))(1 + z)^2$. The system temperature T_{sys} can be calculated using $T_{\text{sys}}(\nu) = 100 + 300(150 \text{ MHz}/\nu)^{2.55}$ K [88]. We consider the number of polarizations (N_{pol}) to be 2. The effective collecting area of each antenna is given as $A_{\text{eff}} = A_{\text{eff}}(\nu_{\text{crit}}) \times \epsilon(\nu)$, where $\epsilon(\nu)$ is defined as [86]

$$\epsilon(\nu) = \begin{cases} (\nu_{\text{crit}}/\nu)^2, & \text{if } \nu > \nu_{\text{crit}} \\ 1, & \text{otherwise} \end{cases} \quad (6.7)$$

The baseline density of system $\Sigma(\mathbf{U}, z) = N_a^2/(2\pi U_{\text{max}}^2)$ is considered to be constant within the core radius. N_a is the total number of such antenna and u_{max} is the maximum baseline in the units of $\lambda(z)$. We follow [89] for the variance calculation of the bispectrum. The variance in the bispectrum due to system noise is computed using [90, 91],

$$\sigma_N^2(B) \approx s_B \frac{V_f}{V_B} P_N(k_1, z) P_N(k_2, z) P_N(k_3, z). \quad (6.8)$$

Here, V_f and $P_N(k, z)$ are the same volumes of the fundamental cell and noise power spectrum as mentioned above for the power spectrum. The values of s_B are 1, 2 and 6 for general, isosceles and equilateral k -bins triangles, respectively, and $V_B \approx 8\pi^2 k_1 k_2 k_3 \Delta k_1 \Delta k_2 \Delta k_3$, where k_1, k_2 and k_3 are the length of the sides of the bispectrum triangle in the k -space bin, and the bin width of those k values are $\Delta k_1, \Delta k_2, \Delta k_3$, respectively. For this work, we have considered 1000 hr of SKA-LOW observation and all the parameters values related to instrumental noise covariance for both power spectrum and bispectrum are taken as follows: $A_{\text{eff}}(\nu_{\text{crit}})$ is taken to be 962 m² with ν_{crit} being 110 MHz [92]. We consider core radius of $R_{\text{max}} = 2$ Km with total antenna $N_a = 296$ [93]. For the calculation of V_B for the bispectrum, the relative k -bin size is taken as $\Delta k/k \sim 1$.

6.2 Results

Here, we present and compare the estimation of the reionization parameters provided by both types of emulators. We have shown the corner plots that compare the posterior of parameters estimated via ANN and BNN emulators. This comparison is done using emulators trained on three different sizes of datasets (as discussed in Sec 5) for both the power spectrum and the bispectrum. Furthermore, we also compare the estimations obtained by the power spectrum and the bispectrum using the BNN-based emulator.

The reionization parameter estimation using power spectrum as a summary statistic is shown in figure 5. The top panel shows the corner plot of the parameter estimation obtained via emulators trained on the complete datasets, whereas the two plots in the bottom panel show the estimation obtained using emulators trained on 1000 datasets (left) and 500 datasets (right). The diagonal panels in each of the plots show the marginalized distribution of the parameters given using the ANN-based emulator (in orange) and the BNN-based emulator (in

$$[M_{\text{h,min}}(10^8 M_{\odot}) = 467.37, N_{\text{ion}} = 181.05, R_{\text{mfp}}(\text{Mpc}) = 12.65, \bar{x}_{\text{HI}} = 0.377]$$

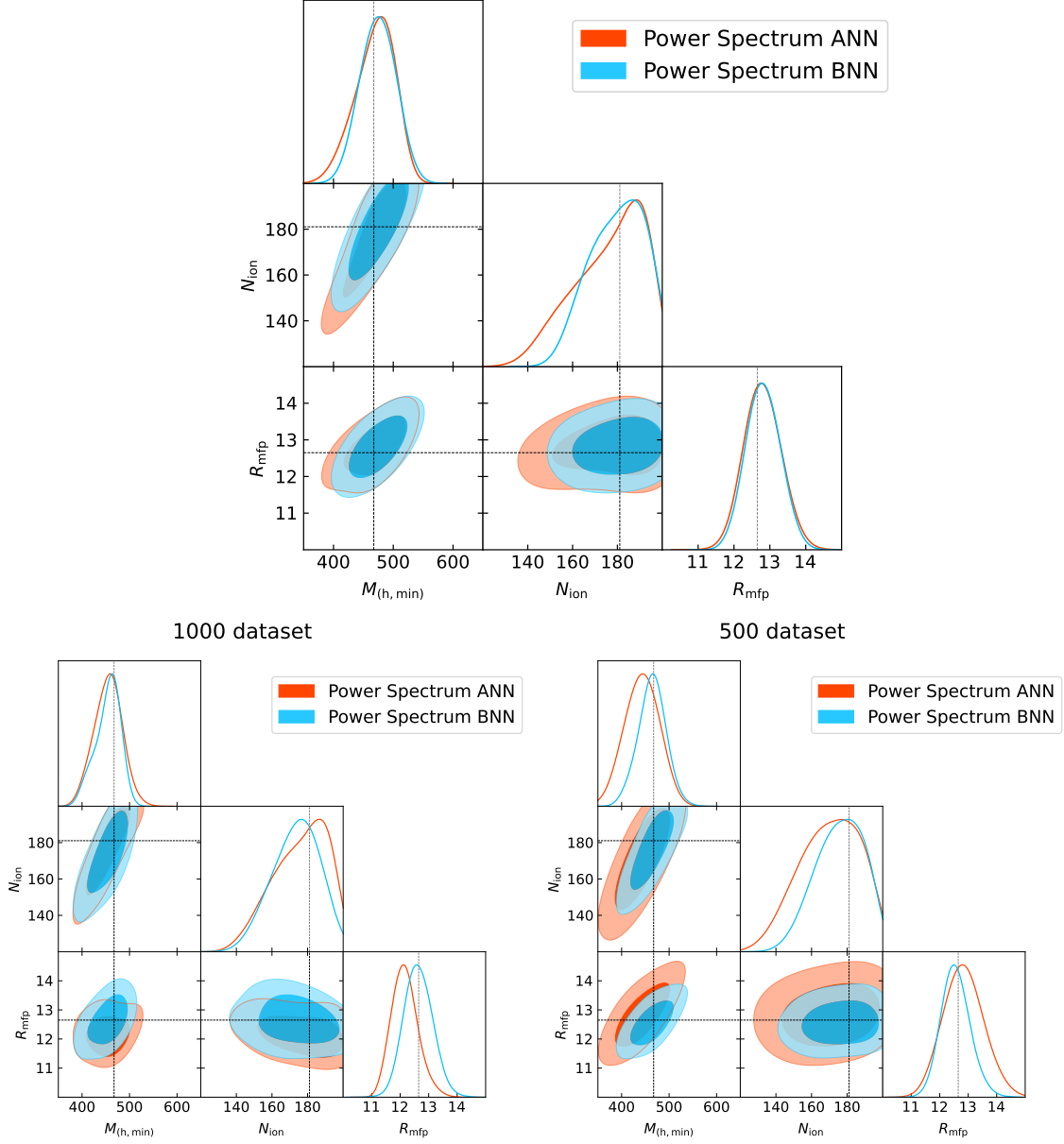


Figure 5. Posteriors of reionization parameters obtained using power spectrum as a summary statistic, emulated using ANN-based emulator (orange) and BNN-based emulator (sky blue). The top panel shows the constraints from emulators trained on the complete (7000) dataset. In the bottom panel, the left plot shows constraints from emulators trained on 1000 data points, and the right one from 500 data points. The true reionization parameters $M_{\text{h,min}} = 467.37 \times 10^8 M_{\odot}$, $N_{\text{ion}} = 181.05$, $R_{\text{mfp}} = 12.65$ Mpc and $\bar{x}_{\text{HI}} = 0.377$ are shown using black dotted line.

skyblue), with true values represented using the dashed lines. The non-diagonal panels show the 2D joint distribution of the two parameters, where 1σ and 2σ contours are represented by dark and light-shaded regions, respectively, and true values are shown using vertical and horizontal lines. We can see that for all three cases, the BNN-based emulator provides

better constraints than the ANN. For smaller datasets, ANN-based emulators provide biased posteriors than BNN. This happens because ANNs are more prone to overfitting and provide biased predictions for smaller training datasets. On the other hand, BNNs do not give biased predictions but predict higher uncertainties when trained on a smaller training dataset, as shown in figure 4.

Similar corner plots of the reionization parameter posteriors obtained using bispectrum as a summary statistic are shown in figure 6. We have used all unique triangle configurations to estimate the reionization parameters. As in the power spectrum, here also, there are three plots showing the estimation obtained via emulators trained on three training datasets. We see that BNN-based emulators provide a tighter estimation of parameters than ANN-based emulators. This is because the bispectrum is a complex statistic to emulate as compared to the power spectrum. Bispectrum has a high dynamic range and sign variation, which are difficult for ANN-based emulators to emulate, especially for smaller datasets. BNN-based emulators can predict these variations in sign and magnitude at the cost of assigning a large error in their predictions for a given triangle configuration, as shown in figure 4. Therefore, for large datasets, ANN emulators do a good job as their constraints are comparable to those provided by BNN emulators; however, the difference in their estimation is very prominent for smaller datasets. So, BNN-based emulators overall perform better in emulating the bispectrum than ANN-based emulators and thus, provide better constraints of reionization parameters. Furthermore, it is important to notice that BNN-based emulators provide the distribution of the summary statistics given the input parameters, whereas ANN-based emulators provide only a single value of statistics sampled from that distribution. So, the BNN-based emulator inherently carries more information about the emulated statistics than the ANN-based emulator. Moreover, as we see in the in Sec 2 figure 4, even the mean values predicted by the BNN-based emulators are more accurate than the point prediction of the ANN-based emulators, especially for the case of smaller training datasets and for the triangle configurations of bispectrum which varies in sign for different k_1 -bins. These result in a tighter constraint on reionization parameters using BNN-based emulators compared to ANN-based emulators for both power spectrum and bispectrum.

Additionally, we also compare the parameter constraining capabilities of two statistics, power spectrum and bispectrum, using the BNN-based emulator demonstrated in figure 7. It is evident from these plots that we get much tighter constraints on reionization parameters using the bispectrum as a summary statistic for all three cases. This is expected because, as discussed earlier, the bispectrum, being a higher-order (three-point) statistic, has more information content and consequently more constraining power than the power spectrum. Thus, considering the bispectrum for all unique triangle configurations ensures better constraints on the reionization parameters. These results are consistent with the findings of [39], who compared the constraining power of the power spectrum and bispectrum of the EoR 21-cm signal using ANN-based emulators alone.

7 Summary and Discussion

Many ongoing and upcoming radio interferometers, such as uGMRT, HERA, LOFAR, MWA, and SKAO, are trying to probe the EoR through the redshifted 21-cm signal with unprecedented accuracy and resolution. This 21-cm signal is a highly non-Gaussian signal; thus, probing this non-Gaussianity will provide a better understanding of the EoR physics. There-

$$[M_{h,\min}(10^8 M_\odot) = 467.37, N_{\text{ion}} = 181.05, R_{\text{mfp}}(\text{Mpc}) = 12.65, \bar{x}_{\text{HI}} = 0.377]$$

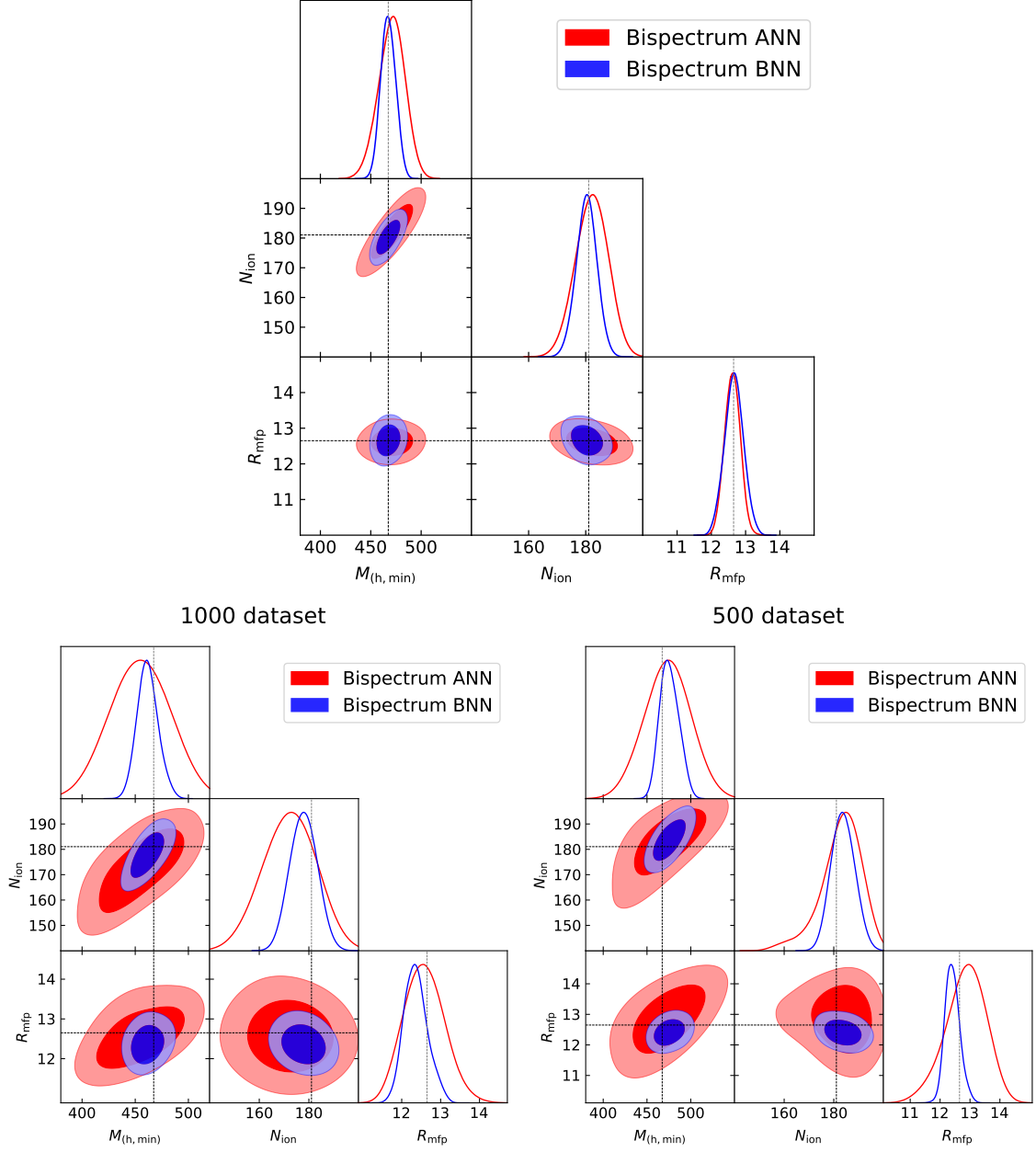


Figure 6. Posteriors of reionization parameters obtained using bispectrum as a summary statistic, emulated using ANN-based emulator (red) and BNN-based emulator (blue). The top panel shows the constraints from emulators trained on the complete dataset. In the bottom panel, the left plot shows constraints from emulators trained on 1000 data points, and the right one from 500 data points. The true reionization parameters $M_{h,\min} = 467.37 \times 10^8 M_\odot$, $N_{\text{ion}} = 181.05$, $R_{\text{mfp}} = 12.65$ Mpc and $\bar{x}_{\text{HI}} = 0.377$ are shown using black dotted line.

$$[M_{\text{h,min}}(10^8 M_{\odot}) = 467.37, N_{\text{ion}} = 181.05, R_{\text{mfp}}(\text{Mpc}) = 12.65, \bar{x}_{\text{HI}} = 0.377]$$

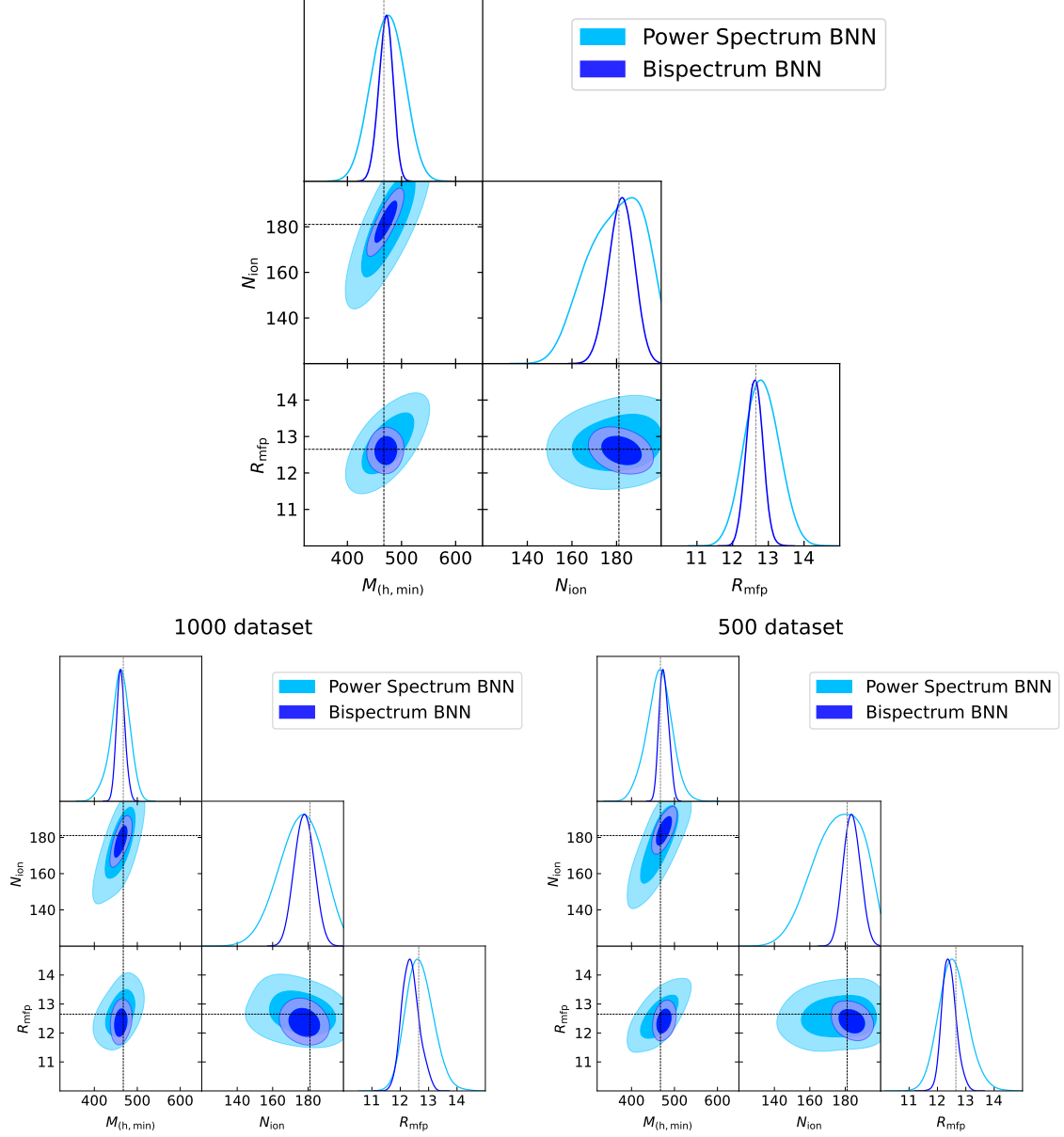


Figure 7. Comparison of constraining power of the power spectrum (sky blue) and bispectrum (blue) as a summary statistic emulated using BNN-based emulators for the true reionization parameters $M_{\text{h,min}} = 467.37 \times 10^8 M_{\odot}$, $N_{\text{ion}} = 181.05$, $R_{\text{mfp}} = 12.65$ Mpc and $\bar{x}_{\text{HI}} = 0.377$ shown using black dotted line.

fore, the bispectrum is an optimal statistic to study and constrain the EoR parameters than the power spectrum. We use a Bayesian inference pipeline to obtain the statistical inference of the reionization model parameters from these summary statistics. This inference technique requires modeling the 21-cm signal using numerical or seminumerical simulations, and rerunning these simulations a large number of times is computationally expensive. To overcome this problem, neural network-based emulators are used to emulate the 21-cm signal statistics given reionization parameters.

Previous works on reionization parameter estimation using Artificial neural network (ANN) emulators for 21-cm signal summary statistics has a fundamental drawback that these ANN-based emulators can produce only point value predictions of the target signal statistic; thus, they fail to capture the uncertainty related to their own predictions. Therefore, when such emulators are used in the Bayesian inference pipeline, they can not naturally propagate their prediction uncertainties to the estimated model parameters. To address this problem, we have developed Bayesian neural network (BNN) based emulators for the power spectrum and bispectrum of the redshifted 21-cm signal. The BNN-based emulators provide the mean prediction for the summary statistics as well as the uncertainties associated with their prediction, and we can propagate this uncertainty through our Bayesian inference pipeline. We have demonstrated that using BNN-based emulators, we get a more robust and trustworthy estimation of the reionization parameters compared to the ANN-based emulators.

In this work, we show that the BNN-based emulator performs better than the ANN-based emulator, especially for the bispectrum, as it is a more complex statistic to emulate. Moreover, we systematically reduce the size of the training dataset to check the robustness of the emulators on the smaller datasets. The performance of the emulators trained with 1000 and 500 datasets is shown in Sec 5. The prediction of the ANN-based emulators deviates considerably from the true values of the statistics, as it tends to overfit for the smaller dataset, whereas the BNN-based emulators provide better mean value predictions than the ANN. Also, the uncertainties associated with the BNN prediction increase with decreasing training dataset, and the true value of the statistics almost always lies within the 3σ of the mean prediction, whereas there is no quantification of the same for ANN emulators.

We use these emulators in a Bayesian inference pipeline to obtain the constraints on the reionization parameters. We show that BNN-based emulators provide more robust constraints for both power spectrum (see figure 5) and bispectrum (see figure 6) for all three training datasets. BNN-based emulators provide the distribution of summary statistics based on input parameters, whereas ANN-based emulators provide simply a single value of statistics drawn from that distribution. Therefore, compared to ANN-based emulators, BNN-based emulators offer more information about the emulated statistics. Furthermore, as shown in the figures in Sec 2, even the mean values predicted by BNN-based emulators outperform the point prediction by ANN-based emulators, particularly when dealing with smaller training datasets and more sign flips for the bispectrum values. As a result, for both power spectrum and bispectrum, BNN-based emulators have better and more robust constraints on reionization parameters than ANN-based emulators. Moreover, we also show that we get better constraints on reionization parameters considering the bispectrum (all unique triangle configurations) as a summary statistic than the power spectrum for BNN-based emulators. This is in agreement with the results of [39], where the same results were obtained through ANN-based emulators.

For this work, we have used coeval cubes at redshift of $z = 7$ instead of lightcone cubes due to the limitation of computational resources to generate the training dataset for the

lightcones. Furthermore, we assume that the signal is obtained by removing the effects of foreground contamination, and it does not contain any residual foreground effects. However, we have included the sample variance and telescopic noise for the SKA-LOW-like telescope for 1000 hr of observation for both the power spectrum and the bispectrum in our study, and assume that other systematics effects are removed from the signal. This is a simplified assumption, so covariance would be higher than what we have assumed in this work. In future works, we plan to include the effects of lightcone, residual foreground, along with cosmic variance, and study their effects on the estimation of the reionization parameters using BNN-based bispectrum emulator.

Acknowledgments

YM acknowledges the financial support by the Department of Science and Technology, Government of India, through the INSPIRE Fellowship. SM would like to thank the Science and Engineering Research Board (SERB) and the Department of Science and Technology (DST), Government of India, for financial support through Core Research Grant No. CRG/2021/004025 titled “Observing the Cosmic Dawn in Multicolour using Next Generation Telescopes”. LN acknowledges the financial support by the Department of Science and Technology, Government of India, through the INSPIRE Fellowship. CSM would like to acknowledge financial support from the Council of Scientific and Industrial Research (CSIR) via a CSIR-SRF Fellowship (Grant No. 09/1022(0080)/2019-EMR-I) and from the ARCO Prize Fellowship.

References

- [1] E. Komatsu, K.M. Smith, J. Dunkley, C.L. Bennett, B. Gold, G. Hinshaw et al., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *The Astrophysical Journal Supplement* **192** (2011) 18 [[1001.4538](#)].
- [2] Planck Collaboration, R. Adam, N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck intermediate results. XLVII. Planck constraints on reionization history*, *Astronomy & Astrophysics* **596** (2016) A108 [[1605.03507](#)].
- [3] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *Astronomy & Astrophysics* **641** (2020) A6 [[1807.06209](#)].
- [4] R.H. Becker, X. Fan, R.L. White, M.A. Strauss, V.K. Narayanan, R.H. Lupton et al., *Evidence for Reionization at $z \sim 6$: Detection of a Gunn-Peterson Trough in a $z=6.28$ Quasar*, *The Astronomical Journal* **122** (2001) 2850 [[astro-ph/0108097](#)].
- [5] X. Fan, M.A. Strauss, D.P. Schneider, R.H. Becker, R.L. White, Z. Haiman et al., *A Survey of $z \sim 5.7$ Quasars in the Sloan Digital Sky Survey. II. Discovery of Three Additional Quasars at $z \sim 6$* , *The Astronomical Journal* **125** (2003) 1649 [[astro-ph/0301135](#)].
- [6] R. Barnett, S.J. Warren, G.D. Becker, D.J. Mortlock, P.C. Hewett, R.G. McMahon et al., *Observations of the Lyman series forest towards the redshift 7.1 quasar ULAS J1120+0641*, *Astronomy & Astrophysics* **601** (2017) A16 [[1702.03687](#)].
- [7] M. Ouchi, K. Shimasaku, H. Furusawa, T. Saito, M. Yoshida, M. Akiyama et al., *Statistics of 207 Ly α Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models*, *The Astrophysical Journal* **723** (2010) 869 [[1007.2961](#)].

- [8] Z.-Y. Zheng, J. Wang, J. Rhoads, L. Infante, S. Malhotra, W. Hu et al., *First Results from the Lyman Alpha Galaxies in the Epoch of Reionization (LAGER) Survey: Cosmological Reionization at $z \sim 7$* , *The Astrophysical Journal Letters* **842** (2017) L22 [[1703.02985](#)].
- [9] A.J. Taylor, L.L. Cowie, A.J. Barger, E.M. Hu and A. Songaila, *The Evolution of the Ultraluminous Ly α Luminosity Function over $z = 5.7$ - 6.6* , *The Astrophysical Journal* **914** (2021) 79 [[2104.13382](#)].
- [10] S.R. Furlanetto, S.P. Oh and F.H. Briggs, *Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe*, *Physics Reports* **433** (2006) 181 [[astro-ph/0608032](#)].
- [11] J.R. Pritchard and A. Loeb, *21 cm cosmology in the 21st century*, *Reports on Progress in Physics* **75** (2012) 086901 [[1109.6012](#)].
- [12] G. Paciga, J.G. Albert, K. Bandura, T.-C. Chang, Y. Gupta, C. Hirata et al., *A simulation-calibrated limit on the $H I$ power spectrum from the GMRT Epoch of Reionization experiment*, *Monthly Notices of the Royal Astronomical Society* **433** (2013) 639 [[1301.5906](#)].
- [13] HERA Collaboration, Z. Abdurashidova, J.E. Aguirre, P. Alexander, Z.S. Ali, Y. Balfour et al., *First Results from HERA Phase I: Upper Limits on the Epoch of Reionization 21 cm Power Spectrum*, *The Astrophysical Journal* **925** (2022) 221 [[2108.02263](#)].
- [14] F.G. Mertens, M. Mevius, L.V.E. Koopmans, A.R. Offringa, S. Zaroubi, A. Acharya et al., *Deeper multi-redshift upper limits on the epoch of reionisation 21 cm signal power spectrum from LOFAR between $z = 8.3$ and $z = 10.1$* , *Astronomy & Astrophysics* **698** (2025) A186 [[2503.05576](#)].
- [15] N. Barry, M. Wilensky, C.M. Trott, B. Pindor, A.P. Beardsley, B.J. Hazelton et al., *Improving the Epoch of Reionization Power Spectrum Results from Murchison Widefield Array Season 1 Observations*, *The Astrophysical Journal* **884** (2019) 1 [[1909.00561](#)].
- [16] L. Koopmans, J. Pritchard, G. Mellema, J. Aguirre, K. Ahn, R. Barkana et al., *The Cosmic Dawn and Epoch of Reionisation with SKA*, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 1, Apr., 2015, DOI [[1505.07568](#)].
- [17] B.K. Gehlot, F.G. Mertens, L.V.E. Koopmans, M.A. Brentjens, S. Zaroubi, B. Ciardi et al., *The first power spectrum limit on the 21-cm signal of neutral hydrogen during the Cosmic Dawn at $z = 20$ - 25 from LOFAR*, *Monthly Notices of the Royal Astronomical Society* **488** (2019) 4271 [[1809.06661](#)].
- [18] F.G. Mertens, M. Mevius, L.V.E. Koopmans, A.R. Offringa, G. Mellema, S. Zaroubi et al., *Improved upper limits on the 21 cm signal power spectrum of neutral hydrogen at $z \approx 9.1$ from LOFAR*, *Monthly Notices of the Royal Astronomical Society* **493** (2020) 1662 [[2002.07196](#)].
- [19] Z. Abdurashidova, J.E. Aguirre, P. Alexander, Z.S. Ali, Y. Balfour, R. Barkana et al., *HERA Phase I Limits on the Cosmic 21 cm Signal: Constraints on Astrophysics and Cosmology during the Epoch of Reionization*, *The Astrophysical Journal* **924** (2022) 51 [[2108.07282](#)].
- [20] A. Acharya, F. Mertens, B. Ciardi, R. Ghara, L.V.E. Koopmans and S. Zaroubi, *Revised LOFAR upper limits on the 21-cm signal power spectrum at $z \approx 9.1$ using machine learning and gaussian process regression*, *Monthly Notices of the Royal Astronomical Society* **534** (2024) L30 [[2408.10051](#)].
- [21] E. Ceccotti, A.R. Offringa, F.G. Mertens, L.V.E. Koopmans, S. Munshi, J.K. Chege et al., *First upper limits on the 21-cm signal power spectrum of neutral hydrogen at $z = 9.16$ from the LOFAR 3C196 field*, *arXiv e-prints* (2025) arXiv:2504.18534 [[2504.18534](#)].
- [22] S. Singh Gill, S. Bharadwaj, K.M.A. Elahi, S.K. Sethi and A.K. Patwa, *The EoR 21-cm Bispectrum at $z = 8.2$ from MWA data I: Foregrounds and preliminary upper limits*, *arXiv e-prints* (2025) arXiv:2507.04964 [[2507.04964](#)].

- [23] S. Bharadwaj and S.K. Pandey, *Probing non-Gaussian features in the HI distribution at the epoch of re-ionization*, *Monthly Notices of the Royal Astronomical Society* **358** (2005) 968 [[astro-ph/0410581](#)].
- [24] I.T. Iliev, G. Mellema, U.L. Pen, H. Merz, P.R. Shapiro and M.A. Alvarez, *Simulating cosmic reionization at large scales - I. The geometry of reionization*, *Monthly Notices of the Royal Astronomical Society* **369** (2006) 1625 [[astro-ph/0512187](#)].
- [25] G. Mellema, I.T. Iliev, U.-L. Pen and P.R. Shapiro, *Simulating cosmic reionization at large scales - II. The 21-cm emission features and statistical signals*, *Monthly Notices of the Royal Astronomical Society* **372** (2006) 679 [[astro-ph/0603518](#)].
- [26] S. Majumdar, J.R. Pritchard, R. Mondal, C.A. Watkinson, S. Bharadwaj and G. Mellema, *Quantifying the non-Gaussianity in the EoR 21-cm signal through bispectrum*, *Monthly Notices of the Royal Astronomical Society* **476** (2018) 4007 [[1708.08458](#)].
- [27] S. Majumdar, M. Kamran, J.R. Pritchard, R. Mondal, A. Mazumdar, S. Bharadwaj et al., *Redshifted 21-cm bispectrum - I. Impact of the redshift space distortions on the signal from the Epoch of Reionization*, *Monthly Notices of the Royal Astronomical Society* **499** (2020) 5090 [[2007.06584](#)].
- [28] M. Kamran, R. Ghara, S. Majumdar, R. Mondal, G. Mellema, S. Bharadwaj et al., *Redshifted 21-cm bispectrum - II. Impact of the spin temperature fluctuations and redshift space distortions on the signal from the Cosmic Dawn*, *Monthly Notices of the Royal Astronomical Society* **502** (2021) 3800 [[2012.11616](#)].
- [29] M. Kamran, S. Majumdar, R. Ghara, G. Mellema, S. Bharadwaj, J.R. Pritchard et al., *Probing IGM Physics during Cosmic Dawn using the Redshifted 21-cm Bispectrum*, *arXiv e-prints* (2021) [arXiv:2108.08201](#) [[2108.08201](#)].
- [30] M. Kamran, R. Ghara, S. Majumdar, G. Mellema, S. Bharadwaj, J.R. Pritchard et al., *Redshifted 21-cm bispectrum: impact of the source models on the signal and the IGM physics from the Cosmic Dawn*, *Journal of Cosmology and Astroparticle Physics* **2022** (2022) 001 [[2207.09128](#)].
- [31] R. Mondal, G. Mellema, A.K. Shaw, M. Kamran and S. Majumdar, *The Epoch of Reionization 21-cm bispectrum: the impact of light-cone effects and detectability*, *Monthly Notices of the Royal Astronomical Society* **508** (2021) 3848 [[2107.02668](#)].
- [32] 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*, *Monthly Notices of the Royal Astronomical Society* **458** (2016) 3003 [[1507.01335](#)].
- [33] A. Hutter, C.A. Watkinson, J. Seiler, P. Dayal, M. Sinha and D.J. Croton, *The 21 cm bispectrum during reionization: a tracer of the ionization topology*, *Monthly Notices of the Royal Astronomical Society* **492** (2020) 653 [[1907.04342](#)].
- [34] C.A. Watkinson, S.K. Giri, H.E. Ross, K.L. Dixon, I.T. Iliev, G. Mellema et al., *The 21-cm bispectrum as a probe of non-Gaussianities due to X-ray heating*, *Monthly Notices of the Royal Astronomical Society* **482** (2019) 2653 [[1808.02372](#)].
- [35] Q.-B. Ma, B. Ciardi, M.B. Eide, P. Busch, Y. Mao and Q.-J. Zhi, *Investigating X-Ray Sources during the Epoch of Reionization with the 21 cm Signal*, *The Astrophysical Journal* **912** (2021) 143 [[2103.09394](#)].
- [36] J. Raste, G. Kulkarni, C.A. Watkinson, L.C. Keating and M.G. Haehnelt, *The 21-cm bispectrum from neutral hydrogen islands at $z \gtrsim 6$* , *Monthly Notices of the Royal Astronomical Society* **529** (2024) 129 [[2308.09744](#)].
- [37] L. Noble, M. Kamran, S. Majumdar, C.S. Murmu, R. Ghara, G. Mellema et al., *Impact of the Epoch of Reionization sources on the 21-cm bispectrum*, *Journal of Cosmology and Astroparticle Physics* **2024** (2024) 003 [[2406.03118](#)].

- [38] C.A. Watkinson, B. Greig and A. Mesinger, *Epoch of reionization parameter estimation with the 21-cm bispectrum*, *Monthly Notices of the Royal Astronomical Society* **510** (2022) 3838 [2102.02310].
- [39] H. Tiwari, A.K. Shaw, S. Majumdar, M. Kamran and M. Choudhury, *Improving constraints on the reionization parameters using 21-cm bispectrum*, *Journal of Cosmology and Astropartical Physics* **2022** (2022) 045 [2108.07279].
- [40] R. Mondal, A. Fialkov, C. Fling, I.T. Iliev, R. Barkana, B. Ciardi et al., *Tight constraints on the excess radio background at $z = 9.1$ from LOFAR*, *Monthly Notices of the Royal Astronomical Society* **498** (2020) 4178 [2004.00678].
- [41] R. Ghara, S.K. Giri, G. Mellema, B. Ciardi, S. Zaroubi, I.T. Iliev et al., *Constraining the intergalactic medium at $z \approx 9.1$ using LOFAR Epoch of Reionization observations*, *Monthly Notices of the Royal Astronomical Society* **493** (2020) 4728 [2002.07195].
- [42] R. Ghara, S.K. Giri, B. Ciardi, G. Mellema and S. Zaroubi, *Constraining the state of the intergalactic medium during the Epoch of Reionization using MWA 21-cm signal observations*, *Monthly Notices of the Royal Astronomical Society* **503** (2021) 4551 [2103.07483].
- [43] B. Greig, A. Mesinger, L.V.E. Koopmans, B. Ciardi, G. Mellema, S. Zaroubi et al., *Interpreting LOFAR 21-cm signal upper limits at $z \approx 9.1$ in the context of high- z galaxy and reionization observations*, *Monthly Notices of the Royal Astronomical Society* **501** (2021) 1 [2006.03203].
- [44] R. Ghara, S. Zaroubi, B. Ciardi, G. Mellema, S.K. Giri, F.G. Mertens et al., *Constraints on the state of the intergalactic medium at $z \sim 8 - 10$ using redshifted 21 cm observations with LOFAR*, *Astronomy & Astrophysics* **699** (2025) A109 [2505.00373].
- [45] C.J. Schmit and J.R. Pritchard, *Emulation of reionization simulations for Bayesian inference of astrophysics parameters using neural networks*, *Monthly Notices of the Royal Astronomical Society* **475** (2018) 1213 [1708.00011].
- [46] W.D. Jennings, C.A. Watkinson, F.B. Abdalla and J.D. McEwen, *Evaluating machine learning techniques for predicting power spectra from reionization simulations*, *Monthly Notices of the Royal Astronomical Society* **483** (2019) 2907 [1811.09141].
- [47] 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*, *Monthly Notices of the Royal Astronomical Society* **527** (2024) 9977 [2201.08205].
- [48] D. Breitman, A. Mesinger, S.G. Murray, D. Prelogović, Y. Qin and R. Trotta, *21CMEMU: an emulator of 21CMFAST summary observables*, *Monthly Notices of the Royal Astronomical Society* **527** (2024) 9833 [2309.05697].
- [49] H. Shimabukuro and B. Semelin, *Analysing the 21 cm signal from the epoch of reionization with artificial neural networks*, *Monthly Notices of the Royal Astronomical Society* **468** (2017) 3869 [1701.07026].
- [50] N. Gillet, A. Mesinger, B. Greig, A. Liu and G. Ucci, *Deep learning from 21-cm tomography of the cosmic dawn and reionization*, *Monthly Notices of the Royal Astronomical Society* **484** (2019) 282 [1805.02699].
- [51] M. Choudhury, A. Datta and A. Chakraborty, *Extracting the 21 cm Global signal using artificial neural networks*, *Monthly Notices of the Royal Astronomical Society* **491** (2020) 4031 [1911.02580].
- [52] D. Prelogović, A. Mesinger, S. Murray, G. Fiameni and N. Gillet, *Machine learning astrophysics from 21 cm lightcones: impact of network architectures and signal contamination*, *Monthly Notices of the Royal Astronomical Society* **509** (2022) 3852 [2107.00018].
- [53] 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, *Monthly Notices of the Royal Astronomical Society* **512** (2022) 5010 [2112.13866].
- [54] A. Tripathi, A. Datta, M. Choudhury and S. Majumdar, *Extracting the Global 21-cm signal from Cosmic Dawn and Epoch of Reionization in the presence of Foreground and Ionosphere*, *Monthly Notices of the Royal Astronomical Society* **528** (2024) 1945 [2401.01935].
 - [55] M. Choudhury, R. Ghara, S. Zaroubi, B. Ciardi, L.V.E. Koopmans, G. Mellema et al., *Inferring IGM parameters from the redshifted 21-cm power spectrum using Artificial Neural Networks*, *Journal of Cosmology and Astroparticle Physics* **2025** (2025) 003 [2407.03523].
 - [56] 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*, *The Astrophysical Journal* **848** (2017) 23 [1705.04688].
 - [57] H.J. Hortúa, L. Malago and R. Volpi, *Constraining the Reionization History using Bayesian Normalizing Flows*, *arXiv e-prints* (2020) arXiv:2005.07694 [2005.07694].
 - [58] R. Meriot, B. Semelin and D. Cornu, *Comparison of Bayesian inference methods using the LORELI II database of hydro-radiative simulations of the 21-cm signal*, *Astronomy & Astrophysics* **698** (2025) A80 [2411.03093].
 - [59] S. Furlanetto, M. Zaldarriaga and L. Hernquist, *The Growth of HII regions during reionization*, *Astrophys. J.* **613** (2004) 1 [astro-ph/0403697].
 - [60] S. Bharadwaj and P.S. Srikant, *HI fluctuations at large redshifts: III — Simulating the signal expected at GMRT*, *Journal of Astrophysics and Astronomy* **25** (2004) 67 [astro-ph/0402262].
 - [61] R. Mondal, S. Bharadwaj, S. Majumdar, A. Bera and A. Acharyya, *The effect of non-Gaussianity on error predictions for the Epoch of Reionization (EoR) 21-cm power spectrum.*, *Monthly Notices of the Royal Astronomical Society* **449** (2015) L41 [1409.4420].
 - [62] M. Davis, G. Efstathiou, C.S. Frenk and S.D.M. White, *The evolution of large-scale structure in a universe dominated by cold dark matter*, *The Astrophysical Journal* **292** (1985) 371.
 - [63] T.R. Choudhury, M.G. Haehnelt and J. Regan, *Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Ly α forest data*, *Monthly Notices of the Royal Astronomical Society* **394** (2009) 960 [0806.1524].
 - [64] S. Majumdar, G. Mellema, K.K. Datta, H. Jensen, T.R. Choudhury, S. Bharadwaj et al., *On the use of seminumerical simulations in predicting the 21-cm signal from the epoch of reionization*, *Monthly Notices of the Royal Astronomical Society* **443** (2014) 2843 [1403.0941].
 - [65] R. Mondal, S. Bharadwaj and S. Majumdar, *Statistics of the epoch of reionization (EoR) 21-cm signal - II. The evolution of the power-spectrum error-covariance*, *Monthly Notices of the Royal Astronomical Society* **464** (2017) 2992 [1606.03874].
 - [66] K.K. Datta, G. Mellema, Y. Mao, I.T. Iliev, P.R. Shapiro and K. Ahn, *Light-cone effect on the reionization 21-cm power spectrum*, *Monthly Notices of the Royal Astronomical Society* **424** (2012) 1877 [1109.1284].
 - [67] K.K. Datta, H. Jensen, S. Majumdar, G. Mellema, I.T. Iliev, Y. Mao et al., *Light cone effect on the reionization 21-cm signal - II. Evolution, anisotropies and observational implications*, *Monthly Notices of the Royal Astronomical Society* **442** (2014) 1491 [1402.0508].
 - [68] S. Bharadwaj, A. Mazumdar and D. Sarkar, *Quantifying the Redshift Space Distortion of the Bispectrum I: Primordial Non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **493** (2020) 594 [2001.10243].
 - [69] S. Ruder, *An overview of gradient descent optimization algorithms*, *arXiv e-prints* (2016) arXiv:1609.04747 [1609.04747].
 - [70] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press (2016).

- [71] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh et al., *A review of uncertainty quantification in deep learning: Techniques, applications and challenges*, *Information Fusion* **76** (2021) 243.
- [72] A.D. Kiureghian and O. Ditlevsen, *Aleatory or epistemic? does it matter?*, *Structural Safety* **31** (2009) 105.
- [73] E. Hüllermeier and W. Waegeman, *Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods*, *Machine Learning* **110** (2021) 457.
- [74] J. Gawlikowski, C.R.N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng et al., *A survey of uncertainty in deep neural networks*, *Artificial Intelligence Review* **56** (2023) 1513.
- [75] C. Blundell, J. Cornebise, K. Kavukcuoglu and D. Wierstra, *Weight uncertainty in neural network*, in *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37 of *Proceedings of Machine Learning Research*, (Lille, France), pp. 1613–1622, PMLR, 07–09 Jul, 2015, <https://proceedings.mlr.press/v37/blundell15.html>.
- [76] J. Arbel, K. Pitas, M. Vladimirova and V. Fortuin, *A Primer on Bayesian Neural Networks: Review and Debates*, *arXiv e-prints* (2023) arXiv:2309.16314 [2309.16314].
- [77] T. Charnock, L. Perreault-Levasseur and F. Lanusse, *Bayesian neural networks*, Artificial Intelligence for High Energy Physics, pp. 663–713, WORLD SCIENTIFIC (2020), DOI.
- [78] L.V. Jospin, H. Laga, F. Boussaid, W. Buntine and M. Bennamoun, *Hands-on bayesian neural networks—a tutorial for deep learning users*, *IEEE Computational Intelligence Magazine* **17** (2022) 29.
- [79] A. Olivier, M.D. Shields and L. Graham-Brady, *Bayesian neural networks for uncertainty quantification in data-driven materials modeling*, *Computer Methods in Applied Mechanics and Engineering* **386** (2021) 114079.
- [80] M.D. Homan and A. Gelman, *The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo*, *Journal of Machine Learning Research* **15** (2014) 1593–1623.
- [81] R. Neal, *MCMC Using Hamiltonian Dynamics*, in *Handbook of Markov Chain Monte Carlo*, pp. 113–162, CRC Press (2011), DOI.
- [82] F. Chollet et al., “Keras.” <https://keras.io>, 2015.
- [83] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015.
- [84] E. Bingham, J.P. Chen, M. Jankowiak, F. Obermeyer, N. Pradhan, T. Karaletsos et al., *Pyro: Deep Universal Probabilistic Programming*, *Journal of Machine Learning Research* (2018) .
- [85] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, *arXiv e-prints* (2019) arXiv:1912.01703 [1912.01703].
- [86] P. Bull, P.G. Ferreira, P. Patel and M.G. Santos, *Late-time Cosmology with 21 cm Intensity Mapping Experiments*, *The Astrophysical Journal* **803** (2015) 21 [1405.1452].
- [87] A. Obuljen, E. Castorina, F. Villaescusa-Navarro and M. Viel, *High-redshift post-reionization cosmology with 21cm intensity mapping*, *Journal of Cosmology and Astropartical Physics* **2018** (2018) 004 [1709.07893].
- [88] G. Mellema, L.V.E. Koopmans, F.A. Abdalla, G. Bernardi, B. Ciardi, S. Daiboo et al., *Reionization and the Cosmic Dawn with the Square Kilometre Array*, *Experimental Astronomy* **36** (2013) 235 [1210.0197].
- [89] C.S. Murmu, K.K. Datta, S. Majumdar and T.R. Greve, *Impact of astrophysical scatter on the epoch of reionization $[H I]_{21}$ bispectrum*, *Journal of Cosmology and Astropartical Physics* **2024** (2024) 032 [2311.17062].

- [90] R. Scoccimarro, E. Sefusatti and M. Zaldarriaga, *Probing primordial non-Gaussianity with large-scale structure*, *Physical Review D* **69** (2004) 103513 [[astro-ph/0312286](#)].
- [91] M. Liguori, E. Sefusatti, J.R. Fergusson and E.P.S. Shellard, *Primordial Non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure*, *Advances in Astronomy* **2010** (2010) 980523 [[1001.4707](#)].
- [92] S.K. Giri, G. Mellema and R. Ghara, *Optimal identification of H II regions during reionization in 21-cm observations*, *Monthly Notices of the Royal Astronomical Society* **479** (2018) 5596 [[1801.06550](#)].
- [93] A. Mazumder, A. Datta, A. Chakraborty and S. Majumdar, *Observing the reionization: effect of calibration and position errors on realistic observation conditions*, *Monthly Notices of the Royal Astronomical Society* **515** (2022) 4020 [[2207.06169](#)].