Peering towards Cosmic Dawn
Proceedings IAU Symposium No. 333, 2018
V. Jelić & T. van der Hulst, eds.

doi:10.1017/S174392131701081X

Analysing 21cm signal with artificial neural network

Hayato Shimabukuro and Benoit Semelin

Sorbonne Universités, UPMC, LERMA, Observatoire de Paris, PSL research university, CNRS, F-75014, Paris, France
email: hayato.shimabukuro@obspm.fr

Abstract. The 21cm signal at epoch of reionization (EoR) should be observed within next decade. We expect that cosmic 21cm signal at the EoR provides us both cosmological and astrophysical information. In order to extract fruitful information from observation data, we need to develop inversion method. For such a method, we introduce artificial neural network (ANN) which is one of the machine learning techniques. We apply the ANN to inversion problem to constrain astrophysical parameters from 21cm power spectrum. We train the architecture of the neural network with 70 training datasets and apply it to 54 test datasets with different value of parameters. We find that the quality of the parameter reconstruction depends on the sensitivity of the power spectrum to the different parameter sets at a given redshift and also find that the accuracy of reconstruction is improved by increasing the number of given redshifts. We conclude that the ANN is viable inversion method whose main strength is that they require a sparse extrapolation of the parameter space and thus should be usable with full simulation.

Keywords. cosmology: theory, cosmology: diffuse radiation.

1. Introduction

Epoch of Reionization (EoR) is one of the milestones in cosmic history. However, the detailed physics of the EoR is currently poor understood because no observation has yet probed entire of the EoR yet. Recently, some first-generation radio interferometers have been attempting to detect statistically the 21cm signal from the EoR, such as the Murchison Wide field Array (MWA), the LOw Frequency ARray (LOFAR) and the Precision Array for Probing the Epoch of Reionization (PAPER). Furthermore, future instruments such as the Square Kilometre Array (SKA) and Hydrogen Epoch of Reionization Array (HERA) are designed to detect the 21cm signal power spectrum with higher signal to noise ratio and at higher redshift, during the Cosmic Dawn. To maximize the scientific return of the upcoming observations it is important to establish systematic procedures to derive constraints on the EoR modeling parameters from the observed 21-cm data. Recently, several works have studied how such constraints can be obtained by exploring the EoR parameter space with semi-numerical simulations. For example, Fisher analysis (e.g. Kubota *et al.* 2016; Shimabukuro *et al.* 2017) and Bayesian parameter inference such as the Markov Chain Monte Carlo (MCMC) approach (e.g. Greig & Mesinger 2015) have been applied.

In this work, we suggest a new approach for parameter reconstruction based on a machine learning method. The main purpose of machine learning is to find approximate functions that, given the input produce the desired outputs. This is achieved by “learning” from training datasets with known inputs and outputs (e.g. Ball & Brunner 2010; Bloom & Richards 2012). In our work, using a simple astrophysical parameterization of the EoR, we apply artificial neural network (ANN), which is one of the machine learning methods, to reconstruct the parameter values from the 21cm power spectrum data.

2. Cosmic 21cm signal

2.1. *Introduction of 21cm power spectrum*

The brightness temperature for the 21 cm signal is given by

$$\delta T_b(\nu) = \frac{T_{\rm S} - T_\gamma}{1+z}(1-e^{-\tau_{\nu_0}}) \sim 27 x_{\rm H}(1+\delta_m) \times \left(\frac{H}{dv_r/dr + H}\right)\left(1-\frac{T_\gamma}{T_{\rm S}}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_m h^2}\right)^{1/2}\left(\frac{\Omega_b h^2}{0.023}\right)\left(\frac{\Omega_b h}{0.031}\right)[\mathrm{mK}].$$

Here, $T_{\rm S}$ and T_γ represent the local spin temperature of the IGM and the CMB temperature, respectively. τ_{ν_0} is the local optical depth in the 21cm rest frame frequency $\nu_0 = 1420.4$ MHz, $x_{\rm H}$ is the local neutral fraction of the hydrogen gas, $\delta_m(\mathbf{x}, z) \equiv \rho/\bar{\rho} - 1$ is the evolved matter overdensity, dv_r/dr is the local gradient of the gas velocity along the line of sight and $H(z)$ is the Hubble parameter. All quantities are evaluated at redshift $z = \nu_0/\nu - 1$.

We usually use the 21cm power spectrum in order to evaluate statistical property of 21cm fluctuations. We define the 21cm power spectrum as

$$\langle \delta T_b(\mathbf{k}) \delta T_b(\mathbf{k}^{'}) \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}^{'}) P_{21}(\mathbf{k}). \tag{2.1}$$

In our context, we use the *dimensional* 21cm power spectrum, $k^3 P(k)/2\pi^2$. We use the **21cmFAST** to generate the 21cm PS for a given set of astrophysical parameters (Mesinger *et al.* (2011)). We performed simulations in a 200 Mpc^3 comoving box with 300^3 grid cells for a wide range of EoR parameters described in the next section. In our calculation, we use the 21cm power spectrum in the range 0.06 Mpc$^{-1} \leqslant k \leqslant 1.4$ Mpc^{-1} divided into 14 bins.

2.2. *EoR model parameters*

It is common to characterize EoR models with parameters and then examine the effect of changing the parameters on the 21cm signal. We employ three key parameters (ζ : the ionizing efficiency, $T_{\rm vir}$: the minimum virial temperature of haloes producing ionizing photons, $R_{\rm mfp}$: the mean free path of ionizing photons) which are often used in the semi-numerical approach (please read Mesinger *et al.* (2011) in detail).

3. Artificial neural networks (ANNs)

ANNs are one of the machine learning methods and are a mathematical model inspired by the natural neuron network in our brain. The main purpose of ANNs is to construct approximate functions which associate input data with output data. In order to construct such a function, the ANN has to learn from " *training data*". The architecture of a simple class of ANN consists of three layers: the input layer, the hidden layer and the output layer. Each of them has a number of neurons as shown in Fig.1. In a more general case, we could choose the number of hidden layers and the number of neurons at each layer arbitrarily. In our study, we use 1 hidden layer and the number of neuron at input layers have 42 neurons corresponding to the bins of 21cm PS in the case of multiple redshift ($14 \times 3 = 42$). We prepare 70 training datasets and 54 test datasets which consist of the 21cm PS as input and EoR parameters as output. We use *back propagation* algorithm

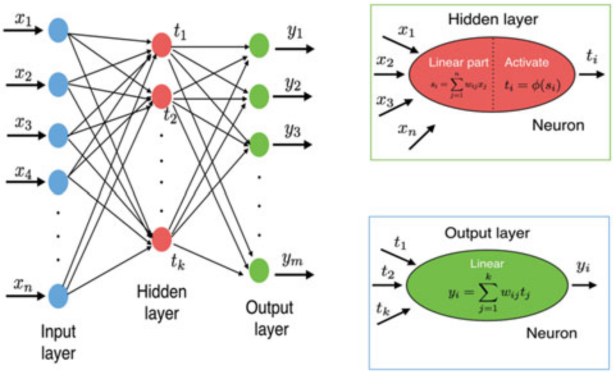

Figure 1. Typical architecture of an artificial neural network. The architecture of the ANN consists of an input layer, a hidden and an output layer of neurons. Each neuron connects the neurons in the next layer.

to train architecture of neural networks (please read Shimabukuro & Semelin (2017) if readers are interested in our design for ANNs).

4. Result

We show our main result. In Fig.2, we show the EoR parameters obtained by the ANN as function of true values used in the simulations, using the 21cm PS at $z = 9$, 10, 11, including both thermal noise and sample variance. As you can see, the recovered parameters are good agreement with expected values. In order to evaluate this agreement qualitatively, we use the normalised root mean square error (RMSE) defined by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_{\mathrm{train}}} \sum_{i=1}^{N_{\mathrm{train}}} X^2} \tag{4.1}$$

Here, $X = (\theta_{\mathrm{ANN}} - \theta_{\mathrm{data}})/\theta_{\mathrm{data}}$ (with θ one of $R_{\mathrm{mfp}}, T_{\mathrm{vir}}$ or, ζ). θ_{ANN} and θ_{data} are the EoR parameters evaluated by the ANN and from the training data, respectively. Smaller values of RMSE means that the difference between recovered and expected values is small. In table.1, we show RMSE for obtained parameters computed by the 21cm PS at single redshift ($z = 9$) without any noise (1st column) or with both thermal noise and cosmic variance (2nd column), multiple redshift case ($z = 9$,10,11) with both noises (3rd column). We also show RMSE in the case that the number of training datasets are reduced from $N_{\mathrm{train}} = 70$ to $N_{\mathrm{train}} = 20$ (4th column). As the number of redshifts of the 21cm PS increase, the RMSE becomes smaller for each parameter (compare 2nd column and 3rd column). We also see that the RMSE is smaller if we reduce the number of training datasets. From these facts, we find that the parameter reconstruction by the ANN is improved if we use the 21cm PS at different multiple redshifts and increase the number of training datasets.

5. Summary

In this work, we applied artificial neural network (ANN) to estimate EoR parameters from the 21cm power spectrum (PS). We used 21cmFAST to produce 21cm PS for different values of the following parameters: R_{mfp}, ζ and T_{vir}. We ran 70 simulations, that provided us with 70 training datasets made of the 21cm PS and the corresponding EoR parameter values. With these datasets, we trained the ANN, which consists of an input layer, one hidden layer and an output layer. We found that the accuracy of parameters

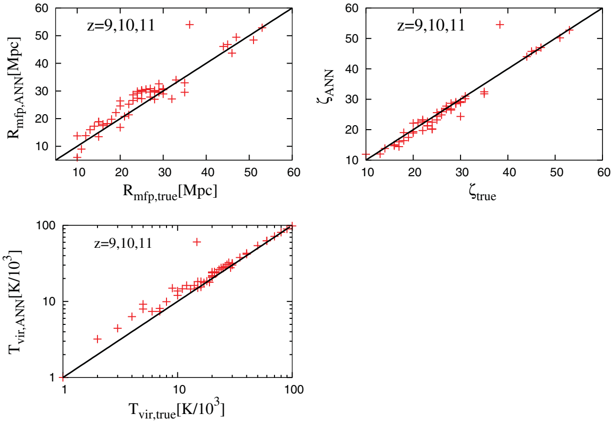

Figure 2. Comparing the EoR model parameter values computed by the ANN against the values used in the simulation using the PS at $z = 9, 10$, and 11 as input data and including both thermal noise and sample variance.

	$\chi^2_{\mathrm{wo/noise}}$	$\chi^2_{\mathrm{w/noise}}$	$\chi^2_{\mathrm{w/noise,zevolution}}$	$\chi^2_{\mathrm{w/noise,reduced}}$
R_{mfp}	0.228	0.258	0.172	0.262
ζ	0.271	0.288	0.168	0.290
$\log(T_{\mathrm{vir}})$	0.027	0.038	0.019	0.029

Table 1. The mean chi-square values for the EoR model parameter values computed by the ANN from the PS as functions of the values used in the simulations. We show mean chi-square values for four types of data. The first column of this table is the chi-square values when we do not include any noise, in second column we take thermal noise and sample variance into account. We use the 21cm PS at $z = 9$ for both cases. The third column is when we take redshift evolution($z = 9, 10, 11$) into account and the fourth column is when we take both thermal noise and sample variance into account and we reduce the number of training data samples.

computed by the ANN was improved compared with those obtained based on the PS at a single redshift if we used multiple redshift. This is simply that the neural network takes advantage in its learning process of the fact that the input hold more information and that the 21cm PS is sensitive to different parameters at different redshifts. The accuracy of parameters was also improved if we increased the number of training datasets.

One of the remarkable points of our study is that we can recover the EoR parameters from the power spectrum using the ANN with good accuracy although we only use 70 training datasets. This number is quite small compared with that required for MCMC (e.g. 10^5 realizations of the PS are used in 21CMMC Greig & Mesinger (2015)).

References

Ball, N. M. & Brunner, R. J. 2010, *International Journal of Modern Physics D*, 19, 1049

Bloom, J. S. & Richards, J. W. 2012, *Advances in Machine Learning and Data Mining for Astronomy*, 89

Greig, B. & Mesinger, A. 2015, *MNRAS*, 449, 4246

Kubota, K., Yoshiura, S., Shimabukuro, H., & Takahashi, K. 2016, *PASJ*, 68, 61

Shimabukuro, H., Yoshiura, S., Takahashi, K., Yokoyama, S., & Ichiki, K. 2017, *MNRAS*, 468, 1542

Shimabukuro, H. & Semelin, B. 2017, *MNRAS*, 468, 3869

Mesinger. A, Furlanetto. S, & Cen. R, 2011, *MNRAS*, 411, 955

https://doi.org/10.1017/S174392131701081X Published online by Cambridge University Press