PAPER

Probing ultra-light axion dark matter from 21 cm tomography using Convolutional Neural Networks

To cite this article: Cristiano G. Sabiu *et al* JCAP01(2022)020

View the article online for updates and enhancements.

Probing ultra-light axion dark matter from 21 cm tomography using Convolutional Neural Networks

Cristiano G. Sabiu,[a] **Kenji Kadota,**[b,c,d,e] **Jacobo Asorey**[f] **and Inkyu Park**[a,g]

[a]Natural Science Research Institute, University of Seoul,
 163 Seoulsiripdaero, Dongdaemun-gu, Seoul 02504, Republic of Korea

[b]School of Fundamental Physics and Mathematical Sciences,
 Hangzhou Institute for Advanced Study, UCAS, Hangzhou 310024, China

[c]International Centre for Theoretical Physics Asia-Pacific (ICTP-AP),
 Beijing/Hangzhou, China

[d]Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS),
 Daejeon 34051, Republic of Korea

[e]Kobayashi-Maskawa Institute for the Origin of Particles and the Universe,
 Nagoya University, Nagoya 464-8602, Japan

[f]Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas (CIEMAT),
 Av. Complutense, 40, Madrid 28040, Spain

[g]Department of Physics, University of Seoul,
 163 Seoulsiripdaero, Dongdaemun-gu, Seoul 02504, Republic of Korea

 E-mail: csabiu@uos.ac.kr, kadota@ucas.ac.cn, jacobo.asorey@ciemat.es,
 icpark@uos.ac.kr

Abstract. We present forecasts on the detectability of Ultra-light axion-like particles (ULAP) from future 21 cm radio observations around the epoch of reionization (EoR). We show that the axion as the dominant dark matter component has a significant impact on the reionization history due to the suppression of small scale density perturbations in the early universe. This behavior depends strongly on the mass of the axion particle.

Using numerical simulations of the brightness temperature field of neutral hydrogen over a large redshift range, we construct a suite of training data. This data is used to train a convolutional neural network that can build a connection between the spatial structures of the brightness temperature field and the input axion mass directly. We construct mock observations of the future Square Kilometer Array survey, SKA1-Low, and find that even in the presence of realistic noise and resolution constraints, the network is still able to predict the input axion mass. We find that the axion mass can be recovered over a wide mass range with a precision of approximately 20%, and as the whole DM contribution, the axion can be detected using SKA1-Low at 68% if the axion mass is $M_X < 1.86 \times 10^{-20}$ eV although this can decrease to $M_X < 5.25 \times 10^{-21}$ eV if we relax our assumptions on the astrophysical modeling by treating those astrophysical parameters as nuisance parameters.

Contents

1 Introduction

Despite the remarkable success of the cold dark matter model in ΛCDM cosmology, the nature of dark matter, such as its mass and interaction rate, remains unknown. Among the popular candidates is the weakly interacting massive particle (WIMP), however its lack of detection from a variety of experimental searches motivates us to explore the wider range of dark matter properties beyond, for instance, those typical in the weak scale supersymmetry (the conventional values for the thermal WIMPs are the mass of $\sim \mathcal{O}(10^{2-3})\,\text{GeV}$ and the thermally averaged annihilation cross section $\langle \sigma v \rangle \sim \mathcal{O}(10^{-26})\,\text{cm}^3/\text{s}$ [1]. As an alternative to the conventional ΛCDM, the ultra-light particle (whose mass can be as light as $\sim 10^{-21}\,\text{eV}$ and sometimes dubbed 'fuzzy dark matter') has gained the revived interest whose wave-like nature is characterized by its de Broglie wavelength below which the fluctuation growth is suppressed due to 'quantum pressure' [2–5].

Such a matter fluctuation suppression at small scales is motivated from the discrepancies between the ΛCDM simulations and the observations, such as the simulation's stellar velocity dispersion larger than observed ones, even though more detailed analysis of baryonic processes would be required to study those small scale structures. Another motivation for such a light particle comes from the symmetries which could exist in the early Universe. An ultra-light particle can well be a remnant of the symmetry which was restored at a high energy scale, such as a pseudo-Nambu-Goldstone boson (pNGB) arsing from the spontaneous symmetry breaking of a continuous global symmetry. Axion is a well known example which arises due to the spontaneous breaking of Peccei-Quinn (PQ) symmetry introduced to solve the QCD strong-CP problem [6–12]. Axion can also serve as a promising dark matter candidate [13–27], and, more generally, the axion-like particles where the ultra-light mass and coupling are treated independently have gained growing interests as an alternative to the conventional ΛCDM [such scenarios with a wide range of axion mass e.g. from $10^{-33}\,\text{eV}$ to $10^{-10}\,\text{eV}$ are sometimes referred to as string axiverse; 2, 3, 18, 28–31].

We call such ultra-light scalar particles which can constitute the whole or non-negligible fraction of the dark matter the axion in this paper. The observational signals of small scale suppression such as Ly-alpha and 21 cm signals have been explored which can potentially lead to the stringent bounds on the axion mass [$> 20 \times 10^{-22}$ eV, 32]. Such suppression can also lead to the delay of epoch of reionization (EoR), which is the focus of our paper. EoR whose details still remain to be explored is among the main targets for the forthcoming 21 cm experiments such as the Square Kilometer Array (SKA) [33].

The goal of this paper is to demonstrate the feasibility of machine learning to extract the axion particle mass from the tomographic maps of the differential brightness temperature of neutral hydrogen. Convolutional Neural Networks (CNN) in particular are well suited to image analysis and have been shown to perform well on simulated 21 cm data [34], thus we utilize them in this work to study the spatial distribution of the 21 cm signal over a wide range of redshift. We use the `21cmFAST` code to produce the differential brightness temperature maps which constitute the training and testing data sets for our studies [35].

In section 2 we explain how our 21 cm maps are constructed and how the axion DM model is incorporated into this framework. We also take into account observational effects and noise associated with the future SKA design. In section 3, we introduce our machine learning methodology, define the specific architecture for our convolutional neural network and design our training and test data sets. Section 4 shows the ability of our machine learning algorithm to constrain the axion mass parameter and to differentiate our axion model (where the axion constitutes the entirety of dark matter) and the conventional ΛCDM model without axion.

Throughout this paper and unless otherwise stated the fiducial cosmological model is chosen to be flat ΛCDM, with $\Omega_M = 0.315$, $\Omega_b = 0.049$, $h = 0.67$, and $\sigma_8 = 0.829$ corresponding to the best fit Planck 2015 measurements [36].

2 Modeling the 21 cm signal

Neutral hydrogen in its ground state has a hyperfine splitting which corresponds a frequency of 1420 MHz. The relative occupations of these two levels is determined by the spin temperature T_s through,

$$\frac{n_1}{n_0} = 3e^{-\frac{T_\star}{T_s}},\qquad(2.1)$$

where $T_\star = 0.068$ K is the temperature corresponding to the energy difference between these two levels and n_1 and n_0 are the number densities of neutral hydrogen in the excited and ground states respectively.

The frequency of the redshifted 21 cm signal lies in the radio range and within the Rayleigh Jean's limit the intensity of this spectral line can be quantified by the brightness temperature T_b. This is observed as an offset from the CMB temperature and is called the differential brightness temperature δT_b. For an observed frequency ν, corresponding to a redshift z at a given point in space x [37],

$$\delta T_b(\nu, x) \approx 27 x_H(x)(1 + \delta(x)) \left(1 - \frac{T_\gamma(z)}{T_s(x)}\right) \frac{\Omega_b h^2}{0.023} \left(\frac{1+z}{10} \frac{0.15}{\Omega_M h^2}\right)^{1/2} [\text{mK}] \qquad (2.2)$$

where x_H is the fraction of neutral hydrogen, $\delta = \bar{\rho}/\rho - 1$ is the over-density relative to the mean density $\bar{\rho}$, and T_γ is the CMB temperature.

2.1 21 cm signal

We simulate the 21 cm maps using the semi-numerical simulation code `21cmFAST` [35, 38]. This code generates a Gaussian initial density field from an input power spectrum P_k and evolves it according to first order perturbation theory. It generates the density δ, spin temperature T_s, the fraction of neutral hydrogen x_H and the differential brightness temperature δT_b at every grid point within the simulation box at a given redshift. It can also output the data in a lightcone format where one axis acts as a redshift dimension.

Within `21cmFAST` there are several astrophysical parameters that affect δT_b across the reionisation epoch, defined when the first energetic sources begin to ionize neutral hydrogen until we are left with a fully ionized plasma. The rate of production of UV photons is determined by the ionization efficiency of star-forming galaxies, which depends on the fraction of UV photons that escape the galaxy, f_{esc}, the fraction of gas in stars f_{star}, the number of ionizing photons per stellar baryon N_γ and the number of times a hydrogen atom recombines on average n_{rec}. However, within 21cmFAST these 4 variables are encapsulated within a single parameter ζ,

$$\zeta = 30 \left(\frac{f_{esc}}{0.15} \right) \left(\frac{f_\star}{0.1} \right) \left(\frac{N_\gamma}{4000} \right) \left(\frac{2}{1 + n_{rec}} \right). \tag{2.3}$$

Within the code ionized regions are chosen when $\zeta f_{coll}(x, z) \geq 1$ [39], where f_{coll} is the collapse fraction and can be computed using the extended Press-Schechter model [40, 41]. In this work we will set $\zeta = 30$ as a fiducial value however in section 4.2 we will relax this assumption and show how uncertainties on this quantity will impact our results.

Another important parameter is the minimum virial temperature of halos producing ionizing photons during the EoR. Typically, T_{vir} is chosen to be 10^4 K, corresponding to the temperature above which atomic cooling becomes effective to trigger gravo-thermal instability and thus star formation. Throughout this work we adopt $T_{vir} = 10^4$ K as a fiducial value, although in section 4.2 we will investigate a range of values to quantify its effect on our estimation of axion mass. The mean free path of ionizing photons, R_{mfp}, is another astrophysical parameter of `21cmFAST`. In this work we kept it fixed at its default value of 15 Mpc.

2.2 Axion dark matter

We generate linear matter power spectra for axion dark matter models using the `AxionCAMB`[1] code a modified version of the cosmological Boltzmann code, `CAMB`. Assuming our fiducial cosmological model parameters we can see the effect of varying the axion particle mass on the power spectrum in figure 1. The suppression of power at large k will have the effect of smoothing small scale structures. For lighter masses this effect is more prominent but as the mass increases the power spectrum becomes almost indistinguishable from the standard ΛCDM model over the scales shown in this figure. The suppression is expected to show up for the scales smaller than the de Broglie wavelength, which is inversely proportional to the momentum, $\lambda = 2\pi/p \sim 0.1\,[\text{kpc}](10^{-21}\,[\text{eV}]/M_X)(10^2\,[\text{km/s}]/v)$.

We now proceed to run `21cmFAST` with axion DM power spectra. In figure 2 we see the effect that the axion mass has on the formation of ionized regions. Lighter axion masses suppress the growth of ionized 'bubbles', delaying reionization, compared to higher axion masses and to standard ΛCDM. This is due to the power spectrum suppression in the presence of axions as can be seen in figure 1. The suppression of density fluctuations has a significant impact on the collapse fraction of gas [42–45] and pushes the reionization to lower redshift.
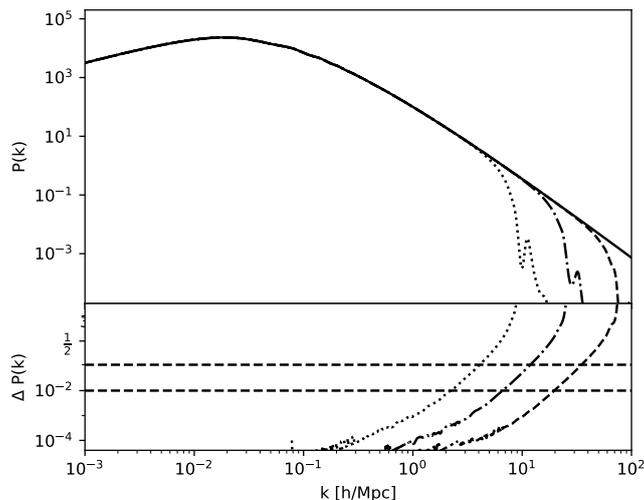
---

[1]https://github.com/dgrin1/axionCAMB.

Figure 1. The linear matter power spectrum at $z = 0$ for ΛCDM without axion (black solid line) and for axion scenarios with various axion masses: $M_X = 10^{-20}$ eV (dashed line), $M_X = 10^{-21}$ eV (dot-dashed) and $M_X = 10^{-22}$ eV (dotted). In the lower panel we display the fractional difference in the power spectra for each axion model compared to ΛCDM with a 1% and 10% reduction shown as the horizontal black dashed lines.

If we measure the reionization redshift, z_{reion} as the redshift at which half of the hydrogen becomes ionized we can plot what this looks like as a function solely of the axion mass. In figure 3 we show z_{reion} as a function of the axion mass and see that at larger masses ($\approx 10^{-19.5}$ eV) the reionization redshift is very close to that of ΛCDM (dashed line).

For lighter masses this ionization redshift is pushed to much lower redshift and at $M_X = 10^{-23}$ eV $z_{\text{reion}} \approx 4$. We consider the axion mass range of 10^{-23}–10^{-19} eV in the following numerical analysis for the illustration purpose, even though those small axion masses $M_X \lesssim 10^{-22.5}$ eV are excluded because the reionization should be complete by $z \sim 6$ [46].

2.3 SKA1-low mock observations

The full width at half-maximum (FWHM) of the point spread function (PSF) of an interferometric radio system can be written as,

$$\Delta\theta = \frac{\lambda}{B}, \tag{2.4}$$

where B is the maximum baseline and λ is the wavelength. For our interest in observing the redshifted 21 cm signal the above equation acquires a redshift dependence with $\lambda = \lambda_{21}(1+z)$. In this work we adopt the SKA1-Low design, which will have a central core of antennae with radius ~ 500 m.

Thus $\Delta\theta$ varies from $\sim 7.2'$, corresponding to a comoving angular diameter distance of 3.1 Mpc/h, at $z = 4.0$, to $21.8'$, with a scale of 4.4 Mpc/h, at $z = 14$. We use this varying resolution to smooth our simulated maps in angular space at each redshift/frequency slice.

In order to simulate realistic observations, we also consider the instrumental (thermal) noise and its impact on the observed 21 cm signal. Following Alonso et al. [47] we add onto
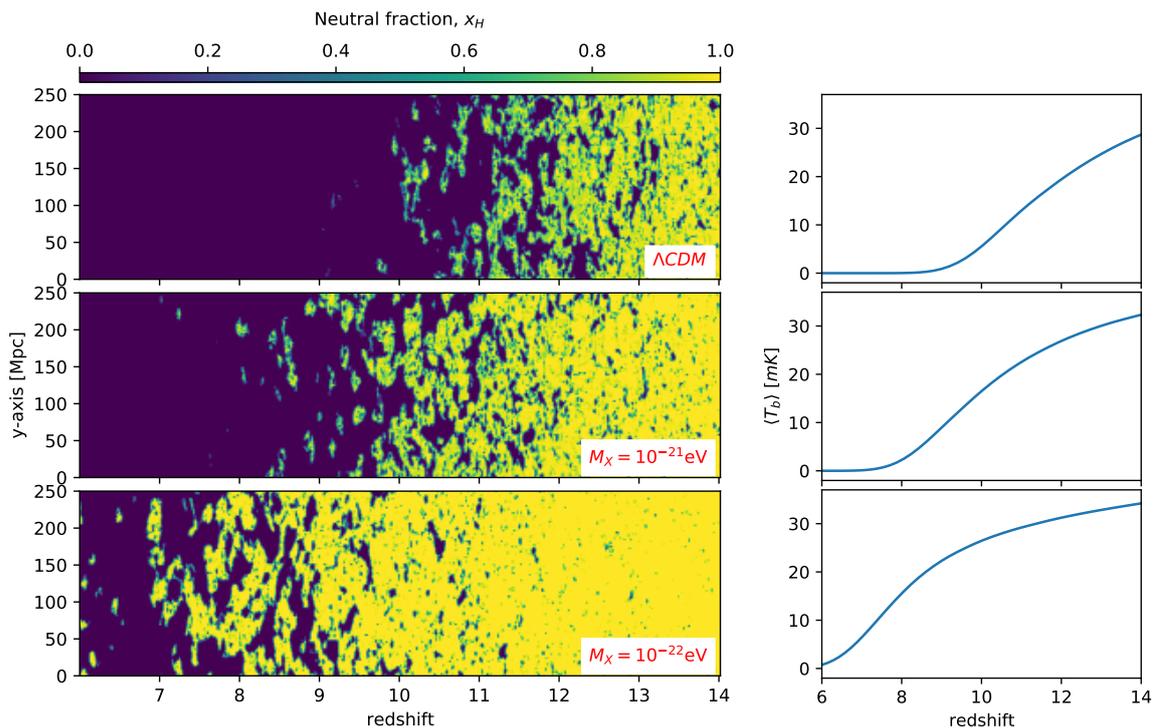
Figure 2. *Left*: simulated spatial distribution of the fraction of neutral hydrogen over the redshift range $6 < z < 14$ for three models. *Right*: the global brightness temperature, $\langle T_{\rm b} \rangle$, as a function of redshift for three models. In the top panels we show the ΛCDM case while the middle and bottom panels show axion models with particle masses of 10^{-21} and 10^{-22} eV respectively. The fiducial model is flat ΛCDM, with $\Omega_M = 0.315$, $\Omega_{\rm b} = 0.049$, $h = 0.67$, $\sigma_8 = 0.829$, $T_{\rm vir} = 10^4$ K and $\zeta = 30$.
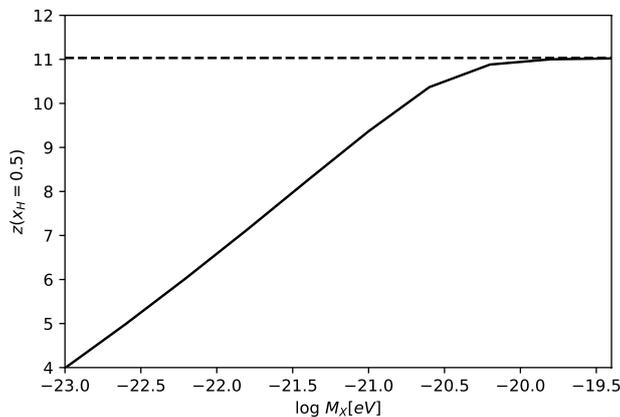

Figure 3. The redshift of reionization (i.e. when the volume averaged $\langle x_{\rm H} \rangle = 0.5$) as a function of axion mass (black line). The redshift of reionization for the ΛCDM without axion is shown as dashed line.

our simulated maps a Gaussian random field with a r.m.s. variation of,

$$\sigma_{\rm noise} = T_{\rm sys} \sqrt{\frac{4\pi f_{\rm sky}}{\Omega_{\rm beam} N_{\rm dish} t_{\rm int} \Delta\nu}}. \tag{2.5}$$
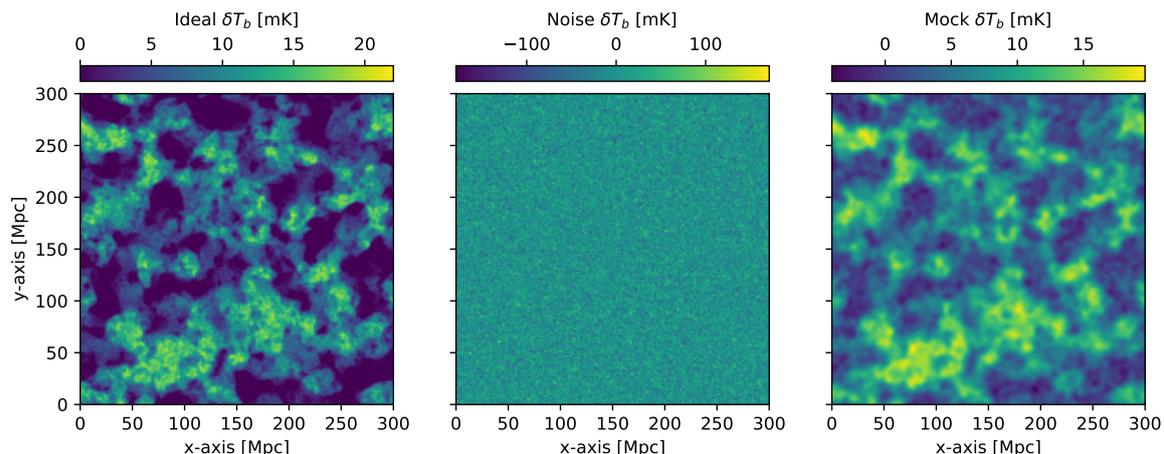
Figure 4. *Left*: the simulated 21 cm differential brightness temperature at z=10 in a comoving slice of 30 Mpc. *Middle*: a realisation of a noise map. *Right*: same as left plot with noise and smoothing applied.

where N_{dish} is the number of antenna, $\Omega_{\mathrm{beam}} = 1.133\theta_{\mathrm{beam}}^2$ is the beam size, t_{int} is the integration time taken to be 1,000 hours, $\Delta\nu$ is the frequency bandwidth, which is $\sim 1\,\mathrm{MHz}$, $f_{\mathrm{sky}} = 0.02$ is the fraction of sky observed and the system temperature is $T_{\mathrm{sys}} = T_{rx} + T_{\mathrm{gal}}$ is the combination of the receiver temperature $T_{rx} = 0.1T_{\mathrm{gal}} + 40\,\mathrm{K}$ and the contribution from our own galaxy at frequency ν, which is modelled as, $T_{\mathrm{gal}} = 25\left(\frac{\nu}{408\,\mathrm{MHz}}\right)^{-2.75}\mathrm{K}$ [33, i.e. the SKA Red book].

The observed 21 cm signal includes a type of contamination known as the foreground emission, mostly dominated by synchrotron emission from our own galaxy. This type of contamination is usually removed following different techniques [48]. As the foreground removal usually affects more the larger scale modes of the temperature maps [49, 50], we leave this for a future study and assume for this work a perfect foreground removal.

In figure 4 we show a noiseless 'ideal' map output at $z = 10$ in a 30 cMpc slice (left panel). We then create a Gaussian noise realization corresponding to this design (middle panel), which does not display any spatial structure. In the right hand panel of figure 4 we see the effect of adding the noise to the ideal image and then smoothing the map with a Gaussian filter with a size corresponding to eq. (2.4). We will use this procedure to construct realistic observations in the next section.

3 Methodology

The methodology that we adopt is a forward modeling approach using simulations and machine learning to regress to the parameters of interest.

3.1 Simulation set

We create 400 simulations with axion masses (sampled uniformly in log-space) in the range 10^{-23}–$10^{-19}\,\mathrm{eV}$ while keeping the cosmological parameters fixed at the Planck15 fiducial values. Each simulation is generated from a unique Gaussian random realisation of the input power spectra, and thus our training and test data do not share similar spatial structures. This is crucial in general regression problems using neural networks as we do not want the

network to learn specific localised features of the fields. We chose the upper cutoff in the mass range considering that the difference in the power spectrum for $M_X \sim 10^{-19}$ eV and ΛCDM is almost indistinguishable even at scales of $k \approx 100h/$Mpc. The lower mass was chosen more arbitrarily, although we note from figure 3 that the redshift of reionization for $M_X \sim 10^{-23}$ eV is so low that this model would have already been ruled out by observations of the Gunn-Peterson effect in high redshift Quasars [46]. We mention also that axion dark matter with masses in the range 10^{-33} eV $\leq M_X \leq 10^{-24}$ eV are excluded by the Planck CMB data [51, 52], while Lyman-alpha forest data excludes $M_X < 2 \times 10^{-21}$ eV [32].

The simulations are run with a box size of 300 Mpc and a grid size of 256. We output the simulation in a lightcone which has 1,500 tomographic bins equally spaced in comoving distance between z=4 and 14. We reinterpret the redshift dimension as frequency and downgrade the resolution to 128 bins of width equivalent to $\Delta\nu = 1.4$ MHz by taking the mean of the pixel values. We then downgrade the angular dimensions to 64×64 by averaging the pixels.

We also create a set of realistic noisy simulations by applying the noise and Gaussian smoothing from eqs. (2.5) & (2.4), respectively, before we downgrade the dimensions of the data cube.

3.2 Machine learning approach

We implement a machine learning approach known as convolution neural networks (CNN) to build a connection between the spatial distribution of the simulated 21 cm signal and the input parameters of the model.

In figure 5 we show a schematic of our CNN model architecture. It takes as input a $64 \times 64 \times 128$ lightcone data cube where each voxel has dimension 4.68 cMpc $\times 4.68$ cMpc $\times 1.48$ Mhz (*cMpc* means co-moving Mpc). The input is 3-dimensioanlly convolved with 16 ($2 \times 2 \times 2$ sized) filter then pooled down to half the spatial dimension using the maximum value in ($2 \times 2 \times 2$) voxels. This process is repeated three more times with 32, 64 and 128 filters. Before each convolution the feature maps are batch normalised to stabilise the distribution. After the convolution layers the data structure has shape $4 \times 4 \times 8 \times 128$ voxels, these are then flattened into a 1D vector and passed to a conventional neural network that has 3 fully connected layers of 512, 128 and 64 neurons respectively. The final layer has a single output representing our model parameter, M_X. Between each fully connected layer a dropout procedure is added which randomly zeros 15% of the weights during training, preventing over-fitting and is an efficient way of performing model averaging within neural networks. Our CNN model is built using `Tensorflow/KERAS`.[2] [53]

We split our simulations into 3 samples: 80% for training the network, 10% for validating the performance of the network (but which are not used directly to tune the network weights), and 10% for testing the fully trained network.

We perform data augmentation and expand the training set by a factor of two by introducing random flips and rotations of ±90 and 180 degrees in the 2 angular dimensions, while keeping the frequency/redshift axis unchanged. This allows the network to learn a more generalised spatial representation of the data.

In figure 6 we see the training performance of the network in the left panel. The training and validation training loss track each other fairly well after 10 epochs. This reassures us that the network is not over-fitting and not learning only the specific training data.
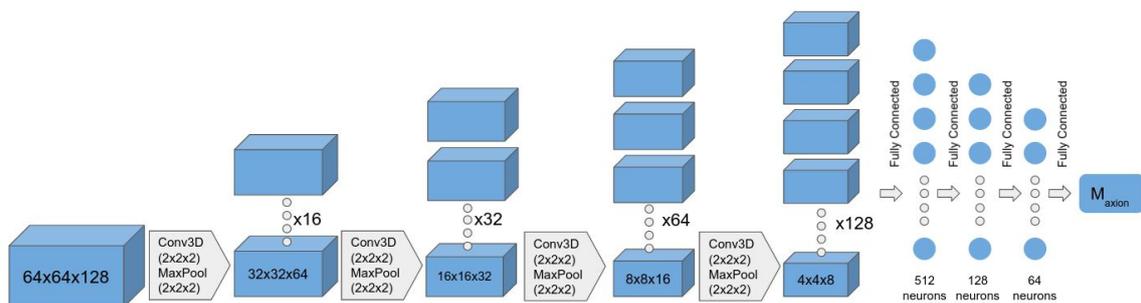
---

[2]https://www.tensorflow.org/.

Figure 5. The architecture of our convolutional neural network. The input data cube has $64 \times 64 \times 128$ voxels. The four 3-dimensional convolution layers have 16, 32, 64 and 128 filters, respectively, each using a $2 \times 2 \times 2$ filter size and a stride of $1 \times 1 \times 1$. Beside each convolution layer, a batch normalization layer is added to normalize the distribution (enhancing the stability) we then add a max pooling layer to decrease the size of the output. After the four convolution layers the elements are 'flattened' to a 1-d vector, and passed to three dense fully connected neural network layers with 512, 128, 64 neurons, and finally a single output prediction for the axion mass. Between each fully connected layer a dropout layer is added to randomly zero 15% of the weights during training.
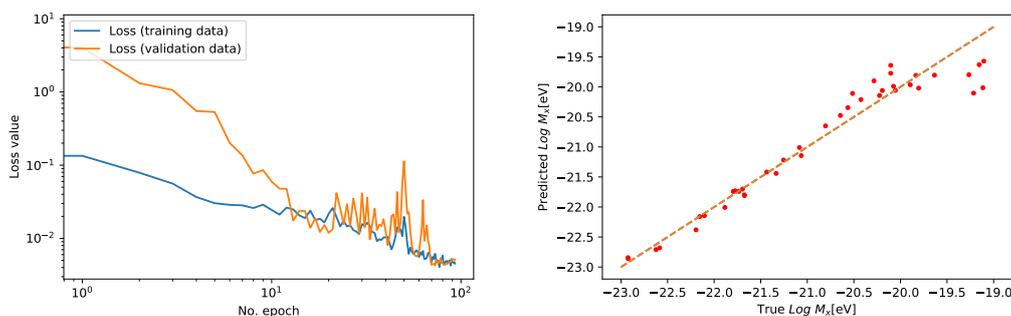

Figure 6. *Left*: the training MSE loss metric of our network over 100 epochs. The blue line denotes the loss of the training data while the orange line shows the loss computed on the validation data. *Right*: the network prediction on the testing data vs. the input truth are shown as red points and the dashed line shows the one-to-one, target prediction.

In the right hand panel of figure 6 we see the predictions of the trained network performance on the test data which has not been previously seen by the network. The predictions for M_X follow closely to the one-to-one line which give us confidence that the network has learned a generalised mapping between the 3-dimensional brightness temperature field and the axion mass, under our model assumptions. It should be noted that the predicted M_X matches the input truth over a wide range in axion mass, but begins to deviate when the true axion mass $M_X \gtrsim 10^{-20}$ eV. The location of this plateau may be interpreted as the limit of the CNN to differentiate between the mock observations with axion masses above this mass scale. This is expected because the such a large axion mass leads to the suppression at a scale smaller than the experimentally observable scale. The limitation of CNN is also illustrated in the following section.
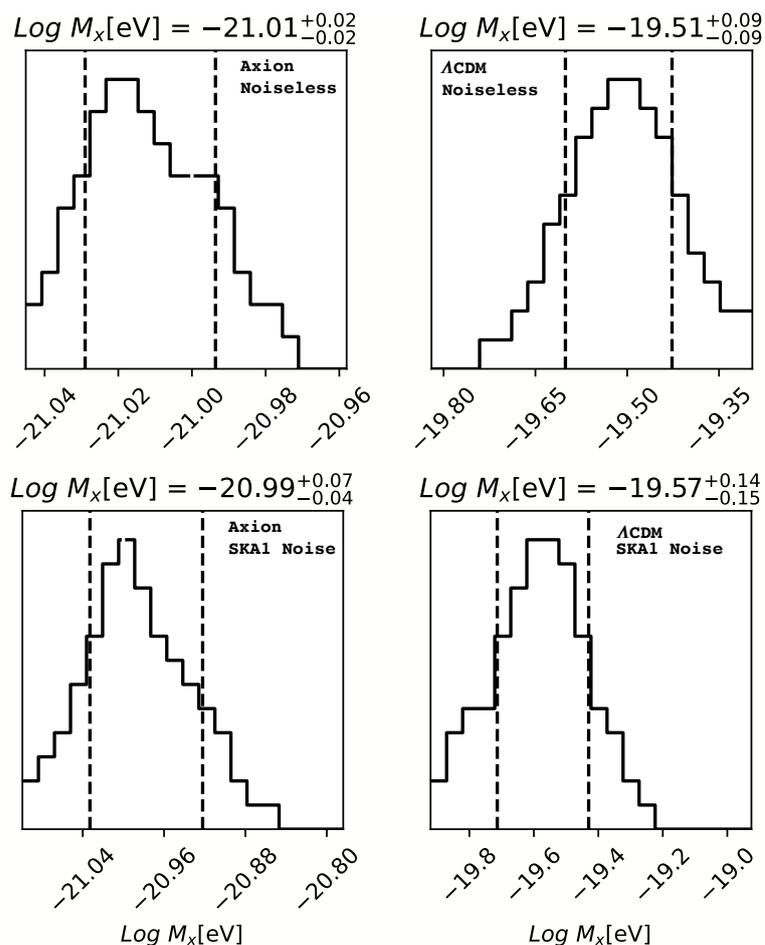
Figure 7. *Top left*: the distribution of our CNN's predictions on 100 noiseless simulations fixed axion mass of 10^{-21} eV (denoted by the vertical blue line). The vertical dashed lines show the 1-σ (68%) credible interval. *Top right*: same as the left plot but the simulations are created using a conventional ΛCDM scenario without axions, for which the CNN could predict the axion mass because the CNN is trained only with the data including axions. *Bottom*: as above but now each simulation has a random Gaussian noise realisation added to its signal and it is convolved with a smoothing length applicable to SKA1-Low core design computed through eq. (2.4).

4 Results

4.1 Axion mass prediction

In the previous section we trained a CNN to predict the axion mass from a realistic mock observations with the SKA1-Low 21 cm survey (with and without noise and smoothing).

We now proceed to determine the accuracy of these predictions and test the ability of the network to differentiate the conventional ΛCDM from axion dark matter models. We create 100 simulation each for 4 scenarios: a conventional ΛCDM scenario and an axion dark matter scenario with $M_X = 10^{-21}$ eV, both with and without noise + smoothing.

In the top panels of figure 7 we show the distribution of M_X predictions for the $M_X = 10^{-21}$ eV simulations on the left and for ΛCDM without axions on the right, un-

der noiseless/ideal conditions. We find that the recovered axion mass for the $M_X = 10^{-21}$ eV simulations is $M_X = 0.98^{+0.04}_{-0.05} \times 10^{-21}$ eV, while the ΛCDM simulations have an inferred axion mass of $M_X = 3.09^{+0.71}_{-0.58} \times 10^{-20}$ eV. The network was trained exclusively with simulations that include axions, therefore the inferred axion mass applying the trained CNN to the ΛCDM data without axions corresponds to the upper mass limit of the CNN's ability to detect axion dark matter.

In the bottom panels of figure 7 we show similar results but now under realistic noise conditions. As we might expect the distribution is broadened and the inferred uncertainty is increased. In the bottom left we see that the mass of the axion is estimated to be $M_X = 1.02^{+0.18}_{-0.09} \times 10^{-21}$ eV, which again recovers the true input mass. In the bottom right we see that the ΛCDM simulations are estimated to have an axion mass of $M_X = 2.69^{+1.02}_{-0.78} \times 10^{-20}$ eV (again this is the mass range for which the CNN cannot differentiate between axion models and the conventional ΛCDM).

We can infer from this that the axion as the whole dark matter contribution can be detected using SKA1-Low at 68% if the axion mass is $M_X < 1.86 \times 10^{-20}$ eV. Here we assume that the distribution of predicted axion masses of the standard ΛCDM simulations represents the null hypothesis and a one-tailed (left) test encapsulates the significance of the axion detection.

We can also compare the optical depth from these models with the value derived from Planck CMB data. The Planck 2018 measurement of the electron scattering optical depth is $\tau_e = 0.0522 \pm 0.0080$ [specifically TT+LowE, 54] which is consistent with our simulated models that give $\tau_e = 0.051$ for standard ΛCDM and $\tau_e = 0.041$ for an axion model with mass $M_X = 10^{-21}$ eV, corresponding to deviations of 0.15σ and 1.4σ respectively.

However it is important to note that our results so far have assumed a fiducial cosmological model and specific choices of astrophysical parameters $T_{\rm vir}$ and ζ. While Planck and other data sets provide a tight prior which supports our assumed cosmological model, we may worry that our assumed values of $T_{\rm vir}$ and ζ, which are not well understood, may bias our recovered axion masses as discussed in the following section.

4.2 Parameter degeneracies

We now relax our assumptions of the astrophysical parameters and check for possible degeneracies between the predicted axion mass, $T_{\rm vir}$ and ζ.

In the left panels of figure 8 we see the effect of varying $T_{\rm vir}$ on the spatial distribution of the neutral fraction of hydrogen over a wide redshift range of $6 < z < 14$. Higher $T_{\rm vir}$ causes the reionization to occur at lower redshift because the halo masses corresponding to that temperature require additional time to grow.

In the right panels of figure 8 we investigate the sensitivity of ζ on the neutral fraction. We see that for larger ζ the reionization starts earlier which is not too surprising given the nature of eq. (2.3) since most of its terms contribute to the heating of the intergalactic medium. Conversely decreasing the ionizing efficiency will prolong and delay the reionization. It is important to note from figure 8 and figure 2 that the effect of varying the axion mass has a visually similar effect to varying either of the 2 astrophysical parameters, $T_{\rm vir}$ and ζ. This could pose a problem for constraining the axion mass if this degeneracy is significant.

The suppression of density fluctuations caused by axions pushes the reionization to lower redshift. We can see this clearly in figure 9, where we show the redshift evolution of the neutral fraction in two axion models and in the standard ΛCDM model. If we set $T_{\rm vir} = 4 \times 10^4$ K in the ΛCDM model so that the reionization redshift matches that of the axion model with
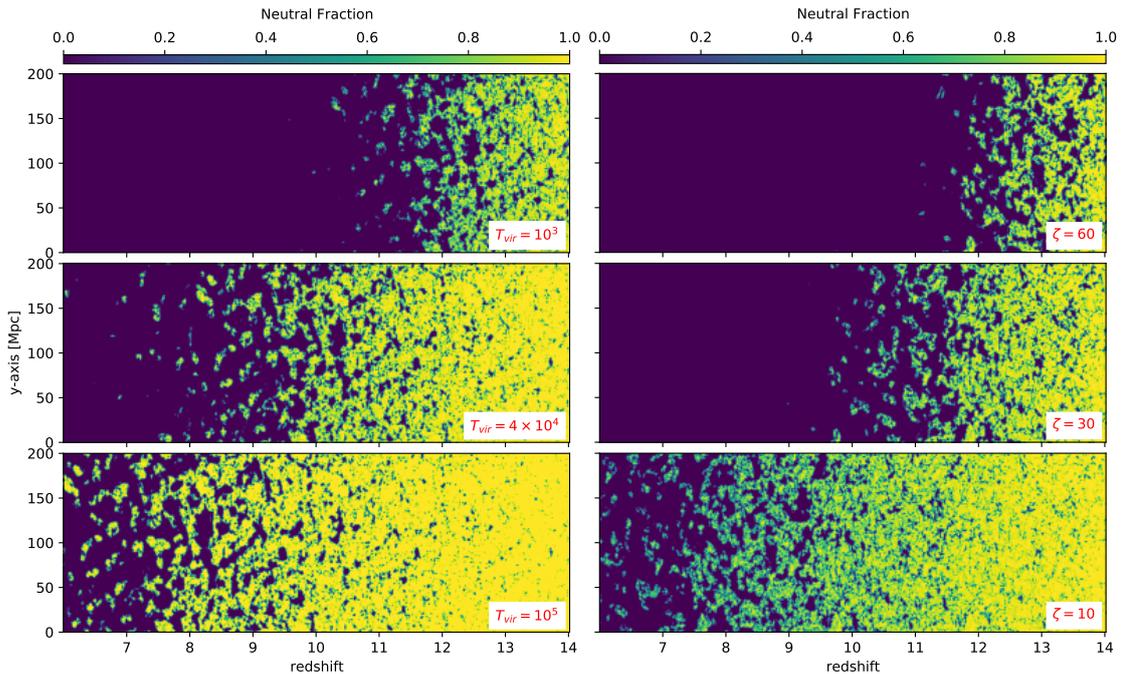
Figure 8. Simulated spatial distribution of the fraction of neutral hydrogen over the redshift range $6 < z < 14$ for various models. Unless otherwise stated the fiducial model is flat ΛCDM, with $\Omega_M = 0.315$, $\Omega_{\rm b} = 0.049$, $h = 0.67$, $\sigma_8 = 0.829$, $T_{\rm vir} = 10^4$ K and $\zeta = 30$. *Left column*: the parameter $T_{\rm vir}$ is varied from the fiducial model with values of $10^3, 4 \times 10^4, 10^5$ K from top to bottom respectively. *Right column*: the parameter ζ is varied from the fiducial model with values of 60, 30, and 10 from top to bottom respectively.

$M_X = 10^{-21}$ eV and $T_{\rm vir} = 10^4$ K, we see that their reionization histories track each other very closely, but not exactly. The higher $T_{\rm vir}$ ΛCDM model has lower $x_{\rm H}$ at high redshift and higher $x_{\rm H}$ at lower redshift compared to the axion model. Likewise we can find the appropriate value of $\zeta (= 15)$ in the fiducial ΛCDM model where its reionization redshift matches that of the axion $M_X = 10^{-21}$ eV model. Once again its neutral fraction follows very closely to each other but there is a clear difference which evolves with redshift indicating that tomographic observations may be able to break the degeneracy that exists between these parameters. It is worth noting from figure 9 that the rate of change of reionization is different depending on the value of ζ, with smaller ζ prolonging the reionization epoch. On the other hand $T_{\rm vir}$ has negligible effect on the width of the transition but it shifts the redshift. We can also see that the deviation between axion and standard ΛCDM is larger at higher redshift while they abruptly converge at lower redshift as the universe completes its reionization. This is because axion DM will suppress small scale structures which are formed at higher redshift, whereas at lower redshift there are more larger structures forming so axion suppression is less important.

In figure 10 we now take a closer look at the seemingly degenerate models depicted in figure 9, i.e i) an axion model with $M_X = 10^{-21}$ eV and fiducial astrophysical parameters ii) a fiducial ΛCDM model with a high value of $T_{\rm vir} = 4 \times 10^4$ K and iii) a fiducial ΛCDM model with a low value of $\zeta = 15$. In the top panels of figure 10 we show the 2-dimensional spatial distribution of the differential brightness temperature field within a thin slice of ~ 2 cMpc at
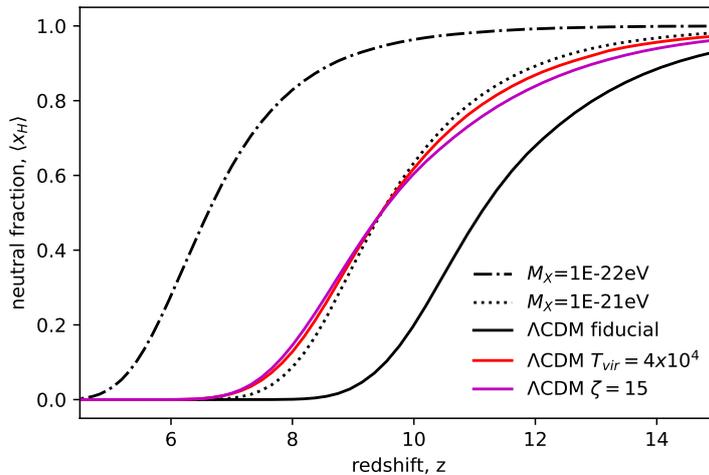
Figure 9. The Neutral fraction as a function of redshift for a standard ΛCDM model with fiducial values of $T_{\rm vir} = 10^4 K$ and $\zeta = 30$ (solid black) and two ultra light axion DM models with $M_X = 10^{-21}$ eV (dotted) and $M_X = 10^{-22}$ eV (dot-dashed). We also show two alternative ΛCDM models without axions i) with a high value of $T_{\rm vir} = 4 \times 10^4 K$ (solid red) and ii) with a low value of $\zeta = 15$ (solid magenta), these model values were chosen so as to match the reionization redshift of the $M_X = 10^{-21}$ eV axion model.

$z = 9.5$ when $\langle x_{\rm H} \rangle = 0.5$. The 3 maps are visually similar and difficult to differentiate. If, instead, we focus on the 1D distribution of pixel values in the right hand panel, the high $T_{\rm vir}$ ΛCDM and axion models are almost indistinguishable which may be a cause for concern in our goal to constrain the axion mass. The low ζ ΛCDM model shows a deviation from the axion model using this 1$^{\rm st}$ order statistical description of the field.

Remembering from figure 9 that the difference in the mean neutral fraction, $\langle x_{\rm H} \rangle$, between these models has a redshift evolution, we now look at the field at an evolved redshift of $z = 8$ in the bottom panels of figure 10. As we might have expected the different redshift evolutions of these models result in visually different fields. These can be more quantitatively understood from the 1D distributions on the right. In both the upper and lower right hand panels of figure 9, the high $T_{\rm vir}$ ΛCDM model peaks at a higher brightness temperature compared to the low ζ model. This is because high $T_{\rm vir}$ preferentially selects heavier halos for the ionization sources, while a low value of ζ applies a low ionizing efficiency to all halo masses. Thus the axion model is closer to high $T_{\rm vir}$ ΛCDM because the power suppression wipes out smaller structures that don't reach their $T_{\rm vir}$ threshold.

We proceed, once again, to constrain the axion mass, although we now relax our assumptions regarding the astrophysical modeling and no longer fix the values of ζ and $T_{\rm vir}$. We thus aim to essentially marginalise over these nuisance parameters rather than constrain them.[3]

We repeat the analysis of the previous section by creating a further 400 simulations that span not only axion mass but also the astrophysical parameters in the range $\zeta = 10$–60 and $T_{\rm vir} = 10^3$–10^5 K. The three model parameters $\{\log M_X, \log T_{\rm vir}, \zeta\}$ were sampled linearly

---

[3]See however, for instance, Kapahtia et al. [55] for the attempts to constrain these astrophysical parameters, given similar observations to those we consider here.
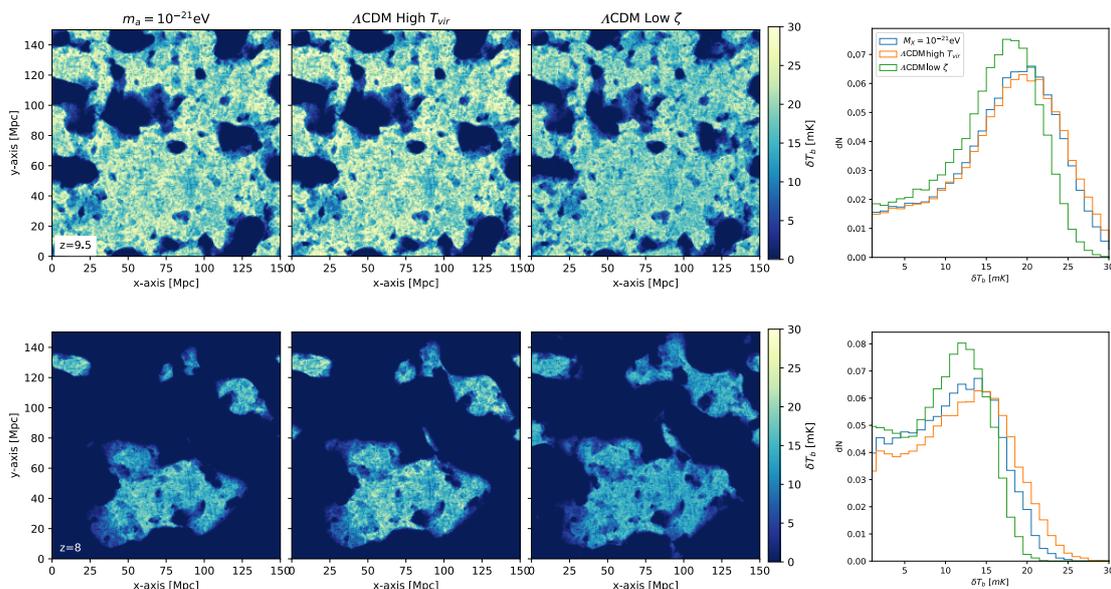
Figure 10. The left three panels we show 2-dimensional slices through the simulation cube for various quantities and redshifts. On the right we show the 1-D histogram of cell values associated with those images. *Top*: the differential brightness temperature of the 21 cm signal at $z = 9.5$ for i) an axion DM model with fiducial $T_{\rm vir}$ and ζ ii) a fiducial ΛCDM model with a high $T_{\rm vir} = 4 \times 10^4 K$ and iii) a fiducial ΛCDM model with a low value of $\zeta = 15$. These 3 'degenerate' models are the same as those shown in figure 9. *Bottom*: the same as above but now at $z = 8.0$.

from a latin hypercube design, which has been shown to be an optimal space filling strategy that avoids repeating similar parameter values [56, 57].[4]

We again train two networks, one with the ideal simulations and one with the SKA1-Low noise simulations. As before we create 100 simulations each for 4 scenarios: a ΛCDM model and axion DM with $M_X = 10^{-21}$ eV, both with and without noise+smoothing. These sets of simulations will test the uncertainty of the axion mass prediction within the trained networks.

In the top panels of figure 11 we show the distribution of M_X predictions for the $M_X = 10^{-21}$ eV simulations on the left and for ΛCDM on the right, under noiseless/ideal conditions. We find that the recovered axion mass for the $M_X = 10^{-21}$ eV simulations is $M_X = 1.48^{+0.86}_{-0.43} \times 10^{-21}$ eV, while the ΛCDM simulations have an inferred axion mass of $M_X = 1.70^{+0.87}_{-0.70} \times 10^{-21}$ eV.

In the bottom panels of figure 11 we show similar results but now under realistic noise conditions parameterized by the Gaussian noise given in eq. (2.5) and the resolution of eq. (2.4). As we might expect the distribution of predicted axion masses is broadened and the inferred uncertainty is increased. In the bottom left we see that the mass of the axion is estimated to be $M_X = 1.35^{+0.99}_{-0.46} \times 10^{-21}$ eV, which again recovers the true input mass. In the bottom right we see that the ΛCDM simulations are estimated to have an axion mass of $M_X = 1.32^{+1.07}_{-0.49} \times 10^{-20}$ eV. We can infer from this that the axion as the whole DM contribution can be detected using SKA1-Low at 68% if the axion mass is $M_X < 5.25 \times 10^{-21}$ eV.

We see that if the astrophysical parameters are treated as nuisance variables their uncertainty does not bias our estimate of the axion mass, however it does degrade the precision to

---

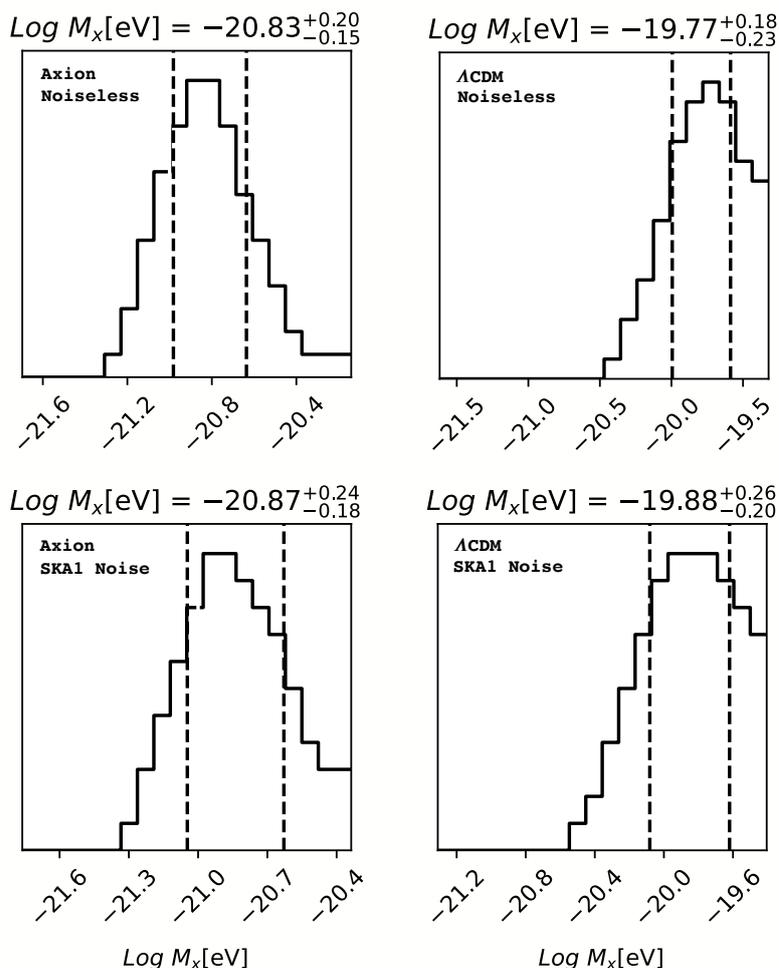[4]Implemented using the `PYTHON` package `pyDOE` https://pythonhosted.org/pyDOE/.

Figure 11. *Top left*: the distribution of our CNN's predictions on 100 noiseless simulations with varying $T_{\rm vir} = 10^3$–10^5 and $\zeta = 10$–60 and fixed axion mass of 10^{-21} eV (denoted by the vertical blue line). The vertical dashed lines show the 1-σ (68%) credible interval. *Top right*: same as the left plot but the simulations are created using ΛCDM model without axions. *Bottom*: as above but now each simulation has a random Gaussian noise realisation added to its signal and it is convolved with a smoothing length applicable to SKA1-Low core design computed through eq. (2.4).

which the axion mass is recovered. This translates to approximately a factor of 2(3) increase in the error of the axion mass for the noiseless (SKA1-Low noisy) simulated images. The tomographic information is of great help in disentangling axion mass estimates and the astrophysical parameters, and our results are encouraging for the future prospects on the axion search or more generally the search for a model beyond the conventional ΛCDM through the machine learning.

We also tested the sensitivity of the global 21 cm signal to the axion mass. Using the same simulations as test data we created a simpler 4 layer dense neural network with 1024 neurons and train it as before. We find that the network can indeed recover the axion mass, however the errors are approximately twice larger than those obtained via the full 3D 21 cm field.

5 Conclusions

In this work we created simulated maps of the future SKA1-Low 21 cm radio survey under an ultra light axion-like dark matter model. We found that the redshift at which the universe becomes half reionized depends strongly on the mass of the axion particle.

We created a set of simulations that sample a range of axion masses and astrophysical parameters. Using a machine learning approach, we trained a convolutional neural network to regress from the simulated 3-dimensional differential brightness temperature field directly to the axion mass, essentially marginalizing over the astrophysical parameters.

We found that the neural network could successfully recover the input truth values of the testing data over a wide range in axion masses with a precision of $\sim 20\%$. In order to assess the uncertainty on the recovered value we created a further 100 independent simulations each for an axion model and a standard CDM model and observed the distribution of their predictions from the pre-trained network. We found that for an axion mass of $M_X = 10^{-21}$ eV we could recover its mass with an uncertainty of $^{+5.94}_{-4.80} \times 10^{-21}$ eV.

We conclude that this approach could be used to detect axion dark matter using SKA1-Low at 68% if the axion mass is $M_X < 1.86 \times 10^{-20}$ eV although this can decrease to $M_X < 5.25 \times 10^{-21}$ eV if we relax our assumptions on the astrophysical modeling. These results however depend on the assumption of Planck 2015 cosmological parameters and the specific design parameters of the future SKA1-Low configurations.

While this work is encouraging in the search for axion dark matter, there are several important systematics that we did not consider here but leave for future work. These include the foreground contamination from radio emission in our own galaxy, redshift-space distortions and non-linear evolution of the density field.

References

[1] G. Jungman, M. Kamionkowski and K. Griest, *Supersymmetric dark matter*, *Phys. Rept.* **267** (1996) 195 [`hep-ph/9506380`] [ɪɴSPIRE].

[2] W. Hu, R. Barkana and A. Gruzinov, *Cold and fuzzy dark matter*, *Phys. Rev. Lett.* **85** (2000) 1158 [`astro-ph/0003365`] [ɪɴSPIRE].

[3] L. Hui, J.P. Ostriker, S. Tremaine and E. Witten, *Ultralight scalars as cosmological dark matter*, *Phys. Rev. D* **95** (2017) 043541 [arXiv:1610.08297] [InSPIRE].

[4] M. Safarzadeh and D.N. Spergel, *Ultra-light Dark Matter is Incompatible with the Milky Way's Dwarf Satellites*, arXiv:1906.11848 [InSPIRE].

[5] L. Hui, *Wave Dark Matter*, arXiv:2101.11735 [InSPIRE].

[6] R.D. Peccei and H.R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440 [InSPIRE].

[7] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223 [InSPIRE].

[8] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279 [InSPIRE].

[9] J. Preskill, M.B. Wise and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127 [InSPIRE].

[10] L.F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133 [InSPIRE].

[11] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137 [InSPIRE].

[12] P. Sikivie, *Experimental Tests of the Invisible Axion*, *Phys. Rev. Lett.* **51** (1983) 1415 [*Erratum ibid.* **52** (1984) 695] [InSPIRE].

[13] G. Raffelt and L. Stodolsky, *Mixing of the Photon with Low Mass Particles*, *Phys. Rev. D* **37** (1988) 1237 [InSPIRE].

[14] ADMX collaboration, *A SQUID-based microwave cavity search for dark-matter axions*, *Phys. Rev. Lett.* **104** (2010) 041301 [arXiv:0910.5914] [InSPIRE].

[15] K. Kadota, J.-O. Gong, K. Ichiki and T. Matsubara, *CMB probes on the correlated axion isocurvature perturbation*, *JCAP* **03** (2015) 026 [arXiv:1411.3974] [InSPIRE].

[16] J.-W. Lee, *Brief History of Ultra-light Scalar Dark Matter Models*, *EPJ Web Conf.* **168** (2018) 06005 [arXiv:1704.05057] [InSPIRE].

[17] CAST collaboration, *New CAST Limit on the Axion-Photon Interaction*, *Nature Phys.* **13** (2017) 584 [arXiv:1705.02290] [InSPIRE].

[18] K. Kadota, Y. Mao, K. Ichiki and J. Silk, *Cosmologically probing ultra-light particle dark matter using* 21 *cm signals*, *JCAP* **06** (2014) 011 [arXiv:1312.1898] [InSPIRE].

[19] K. Kelley and P.J. Quinn, *A Radio Astronomy Search for Cold Dark Matter Axions*, *Astrophys. J. Lett.* **845** (2017) L4 [arXiv:1708.01399] [InSPIRE].

[20] F.P. Huang, K. Kadota, T. Sekiguchi and H. Tashiro, *Radio telescope search for the resonant conversion of cold dark matter axions from the magnetized astrophysical sources*, *Phys. Rev. D* **97** (2018) 123001 [arXiv:1803.08230] [InSPIRE].

[21] A. Hook, Y. Kahn, B.R. Safdi and Z. Sun, *Radio Signals from Axion Dark Matter Conversion in Neutron Star Magnetospheres*, *Phys. Rev. Lett.* **121** (2018) 241102 [arXiv:1804.03145] [InSPIRE].

[22] K. Kadota, T. Kobayashi and H. Otsuka, *Axion inflation with cross-correlated axion isocurvature perturbations*, *JCAP* **01** (2016) 044 [arXiv:1509.04523] [InSPIRE].

[23] D. Harari and P. Sikivie, *Effects of a Nambu-Goldstone boson on the polarization of radio galaxies and the cosmic microwave background*, *Phys. Lett. B* **289** (1992) 67 [InSPIRE].

[24] M.A. Fedderke, P.W. Graham and S. Rajendran, *Axion Dark Matter Detection with CMB Polarization*, *Phys. Rev. D* **100** (2019) 015040 [arXiv:1903.02666] [InSPIRE].

[25] K. Kadota, J. Ooba, H. Tashiro, K. Ichiki and G.-C. Liu, *Cross-correlation between* 21 *cm radiation and CMB B modes from the cosmic birefringence in the presence of a light scalar field*, *Phys. Rev. D* **100** (2019) 063506 [arXiv:1906.00721] [INSPIRE].

[26] H. Shimabukuro, K. Ichiki and K. Kadota, *Constraining the nature of ultra light dark matter particles with the* 21 *cm forest*, *Phys. Rev. D* **101** (2020) 043516 [arXiv:1910.06011] [INSPIRE].

[27] D. Bak, J.-W. Lee and S. Park, *Effective cross section of fuzzy dark matter halos*, *J. Korean Phys. Soc.* **79** (2021) 582 [arXiv:2010.14738] [INSPIRE].

[28] P. Svrček and E. Witten, *Axions In String Theory*, *JHEP* **06** (2006) 051 [hep-th/0605206] [INSPIRE].

[29] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper and J. March-Russell, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530 [arXiv:0905.4720] [INSPIRE].

[30] L. Amendola and R. Barbieri, *Dark matter from an ultra-light pseudo-Goldsone-boson*, *Phys. Lett. B* **642** (2006) 192 [hep-ph/0509257] [INSPIRE].

[31] D.J.E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1 [arXiv:1510.07633] [INSPIRE].

[32] V. Iršič, M. Viel, M.G. Haehnelt, J.S. Bolton and G.D. Becker, *First constraints on fuzzy dark matter from Lyman-α forest data and hydrodynamical simulations*, *Phys. Rev. Lett.* **119** (2017) 031302 [arXiv:1703.04683] [INSPIRE].

[33] SKA collaboration, *Cosmology with Phase 1 of the Square Kilometre Array: Red Book 2018: Technical specifications and performance forecasts*, *Publ. Astron. Soc. Austral.* **37** (2020) e007 [arXiv:1811.02743] [INSPIRE].

[34] Y. Kwon, S.E. Hong and I. Park, *Deep-Learning Study of the* 21 *cm Differential Brightness Temperature During the Epoch of Reionization*, *J. Korean Phys. Soc.* **77** (2020) 49 [arXiv:2006.06236] [INSPIRE].

[35] A. Mesinger, S. Furlanetto and R. Cen, *21cmFAST: A Fast, Semi-Numerical Simulation of the High-Redshift* 21 *cm Signal*, *Mon. Not. Roy. Astron. Soc.* **411** (2011) 955 [arXiv:1003.3878] [INSPIRE].

[36] PLANCK collaboration, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13 [arXiv:1502.01589] [INSPIRE].

[37] S. Furlanetto, S.P. Oh and F. Briggs, *Cosmology at Low Frequencies: The* 21 *cm Transition and the High-Redshift Universe*, *Phys. Rept.* **433** (2006) 181 [astro-ph/0608032] [INSPIRE].

[38] A. Mesinger and S. Furlanetto, *Efficient Simulations of Early Structure Formation and Reionization*, *Astrophys. J.* **669** (2007) 663 [arXiv:0704.0946] [INSPIRE].

[39] S. Furlanetto, M. Zaldarriaga and L. Hernquist, *The Growth of HII regions during reionization*, *Astrophys. J.* **613** (2004) 1 [astro-ph/0403697] [INSPIRE].

[40] J.R. Bond, S. Cole, G. Efstathiou and N. Kaiser, *Excursion set mass functions for hierarchical Gaussian fluctuations*, *Astrophys. J.* **379** (1991) 440 [INSPIRE].

[41] C.G. Lacey and S. Cole, *Merger rates in hierarchical models of galaxy formation*, *Mon. Not. Roy. Astron. Soc.* **262** (1993) 627 [INSPIRE].

[42] M. Sitwell, A. Mesinger, Y.-Z. Ma and K. Sigurdson, *The Imprint of Warm Dark Matter on the Cosmological* 21 *cm Signal*, *Mon. Not. Roy. Astron. Soc.* **438** (2014) 2664 [arXiv:1310.0029] [INSPIRE].

[43] A. Sarkar, R. Mondal, S. Das, S.K. Sethi, S. Bharadwaj and D.J.E. Marsh, *The effects of the small-scale DM power on the cosmological neutral hydrogen (HI) distribution at high redshifts*, *JCAP* **04** (2016) 012 [arXiv:1512.03325] [INSPIRE].

[44] O. Nebrin, R. Ghara and G. Mellema, *Fuzzy Dark Matter at Cosmic Dawn: New* 21 *cm Constraints*, *JCAP* **04** (2019) 051 [arXiv:1812.09760] [INSPIRE].

[45] D. Jones, S. Palatnick, R. Chen, A. Beane and A. Lidz, *Fuzzy Dark Matter and the* 21 *cm Power Spectrum*, *Astrophys. J.* **913** (2021) 7 [arXiv:2101.07177] [inSPIRE].

[46] SDSS collaboration, *Evidence for Reionization at* $z \sim 6$*: Detection of a Gunn-Peterson trough in a* $z = 6.28$ *Quasar*, *Astron. J.* **122** (2001) 2850 [astro-ph/0108097] [inSPIRE].

[47] D. Alonso, P. Bull, P.G. Ferreira and M.G. Santos, *Blind foreground subtraction for intensity mapping experiments*, *Mon. Not. Roy. Astron. Soc.* **447** (2015) 400 [arXiv:1409.8667] [inSPIRE].

[48] M. Spinelli et al., *SKAO HI Intensity Mapping: Blind Foreground Subtraction Challenge*, arXiv:2107.10814 [inSPIRE].

[49] J. Asorey et al., *HIR4: cosmology from a simulated neutral hydrogen full sky using Horizon Run 4*, *Mon. Not. Roy. Astron. Soc.* **495** (2020) 1788 [arXiv:2001.00833] [inSPIRE].

[50] F. Shi et al., *HIR4: cosmological signatures imprinted on the cross-correlation between a* 21 *cm map and galaxy clustering*, *Mon. Not. Roy. Astron. Soc.* **499** (2020) 4613 [arXiv:2006.01407] [inSPIRE].

[51] R. Hlozek, D. Grin, D.J.E. Marsh and P.G. Ferreira, *A search for ultralight axions using precision cosmological data*, *Phys. Rev. D* **91** (2015) 103512 [arXiv:1410.2896] [inSPIRE].

[52] R. Hlozek, D.J.E. Marsh and D. Grin, *Using the Full Power of the Cosmic Microwave Background to Probe Axion Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 3063 [arXiv:1708.05681] [inSPIRE].

[53] M. Abadi et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, (2015) https://www.tensorflow.org/.

[54] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [arXiv:1807.06209] [inSPIRE].

[55] A. Kapahtia, P. Chingangbam, R. Ghara, S. Appleby and T.R. Choudhury, *Prospects of constraining reionization model parameters using Minkowski tensors and Betti numbers*, *JCAP* **05** (2021) 026 [arXiv:2101.03962] [inSPIRE].

[56] V. Eglajs and P. Audze, *New approach to the design of multifactor experiments*, *Probl. Dyn. Strengths* **35** (1977) 104.

[57] M. Mckay, R. Beckman and W. Conover, *A comparison of three methods for selecting vales of input variables in the analysis of output from a computer code*, *Technometrics* **21** (1979) 239.