

Constraining the intergalactic medium at $z \approx 9.1$ using LOFAR Epoch of Reionization observations

R. Ghara^{1,2,3*}, S. K. Giri^{1,4*}, G. Mellema¹, B. Ciardi⁵, S. Zaroubi^{2,3,6}, I. T. Iliev⁷, L. V. E. Koopmans⁶, E. Chapman⁸, S. Gazagnes⁶, B. K. Gehlot^{6,9}, A. Ghosh^{10,11,12}, V. Jelić¹³, F. G. Mertens^{6,14}, R. Mondal⁷, J. Schaye¹⁵, M. B. Silva¹⁶, K. M. B. Asad¹⁷, R. Kooistra^{6,18}, M. Mevius¹⁹, A. R. Offringa^{6,19}, V. N. Pandey^{6,19} and S. Yatawatta¹⁹

Affiliations are listed at the end of the paper

Accepted 2020 February 13. Received 2020 February 10; in original form 2019 December 20

ABSTRACT

We derive constraints on the thermal and ionization states of the intergalactic medium (IGM) at redshift ≈ 9.1 using new upper limits on the 21-cm power spectrum measured by the LOFAR radio telescope and a prior on the ionized fraction at that redshift estimated from recent cosmic microwave background (CMB) observations. We have used results from the reionization simulation code GRIZZLY and a Bayesian inference framework to constrain the parameters which describe the physical state of the IGM. We find that, if the gas heating remains negligible, an IGM with ionized fraction $\gtrsim 0.13$ and a distribution of the ionized regions with a characteristic size $\gtrsim 8 h^{-1}$ comoving megaparsec (Mpc) and a full width at half-maximum (FWHM) $\gtrsim 16 h^{-1}$ Mpc is ruled out. For an IGM with a uniform spin temperature $T_s \gtrsim 3$ K, no constraints on the ionized component can be computed. If the large-scale fluctuations of the signal are driven by spin temperature fluctuations, an IGM with a volume fraction $\lesssim 0.34$ of heated regions with a temperature larger than CMB, average gas temperature 7–160 K, and a distribution of the heated regions with characteristic size 3.5–70 h^{-1} Mpc and FWHM of $\lesssim 110 h^{-1}$ Mpc is ruled out. These constraints are within the 95 per cent credible intervals. With more stringent future upper limits from LOFAR at multiple redshifts, the constraints will become tighter and will exclude an increasingly large region of the parameter space.

Key words: radiative transfer – galaxies: formation – intergalactic medium – cosmology: theory – dark ages, reionization, first stars – X-rays: galaxies.

1 INTRODUCTION

The Epoch of Reionization (EoR) is one of the least understood chapters in the history of our Universe. The formation of the first luminous sources initiated the transition of the cold and neutral intergalactic medium (IGM) into a hot and ionized state. This transition had a significant impact on the later stages of structure formation through various feedback mechanisms (see e.g. Ciardi & Ferrara 2005 for a review). Although we know that reionization took place, very few facts about it are known with certainty (see e.g. Morales & Wyithe 2010; Pritchard & Loeb 2012; Zaroubi 2013; Barkana 2016, for reviews).

Theoretical models suggest that ionizing ultraviolet (UV) photons from the first sources created localized ionized regions, which over time grew in size, started to overlap and, as an increasing number of sources formed, led to a complete reionization of the IGM. Observations of high-redshift ($z \gtrsim 6$) quasar absorption spectra suggest that complete reionization was reached around redshift ≈ 6 (e.g. Fan et al. 2006; Mortlock et al. 2011; Venemans et al. 2015;

Bañados et al. 2018). On the other hand, the measurement of the Thomson optical depth from the observation of cosmic microwave background (CMB) (Planck Collaboration VI 2018) suggests that the probable period of this event lies at redshift $\lesssim 10$ (Choudhury & Ferrara 2006; Mitra, Choudhury & Ferrara 2011, 2012). However, the details of the reionization process such as the exact timing of the EoR, the morphology of the H I distribution in the IGM, and the properties of early sources are still poorly known.

The redshifted 21-cm signal from neutral hydrogen in the IGM is the most promising probe of the EoR, as it has the ability to reveal many of the unknown facts about this epoch. Inspired by its potential, several radio telescopes such as the Low Frequency Array (LOFAR)¹ (van Haarlem et al. 2013; Patil et al. 2017), the Precision Array for Probing the Epoch of Reionization (PAPER)² (Parsons et al. 2014; Kolopanis et al. 2019), the Murchison Widefield Array (MWA)³ (Bowman et al. 2013; Barry et al. 2019), and the

¹<http://www.lofar.org/>

²<http://eor.berkeley.edu/>

³<http://www.mwatelescope.org/>

* E-mail: ghara.raghunath@gmail.com (RG); sambit.giri@gmail.com (SKG)

Hydrogen Epoch of Reionization Array (HERA)⁴ (DeBoer et al. 2017) have invested considerable amounts of observing time to detect this signal. Due to their limited sensitivity, these radio interferometers aim to measure the statistical fluctuations of the signal. The planned Square Kilometre Array (SKA)⁵ will in addition be able to produce actual tomographic images of the distribution of the signal on the sky (Mellema et al. 2015; Ghara et al. 2017). Beside these large radio interferometers, single antenna experiments such as EDGES (Bowman & Rogers 2010), EDGES2 (Monsalve et al. 2017; Bowman et al. 2018), SARAS (Patra et al. 2015), SARAS2 (Singh et al. 2017), BigHorns (Sokolowski et al. 2015), SciHi (Voytek et al. 2014), and LEDA (Price et al. 2018) are being used to attempt a detection of the sky-averaged 21-cm signal and its evolution with redshift.

In spite of all these efforts, so far no undisputed detection of the 21-cm signal from the EoR has been made. The main reason for this is that the signal is several orders of magnitude weaker than the Galactic and extragalactic foregrounds at these frequencies (see e.g. Shaver et al. 1999; Jelić et al. 2008). Moreover, the signal low amplitude implies long integration times are required to exceed the instrumental noise. Although there exist accurate methods to subtract (Harker et al. 2009; Chapman et al. 2013, 2016; Bonaldi & Brown 2015; Mertens, Ghosh & Koopmans 2018), suppress (Datta, Bharadwaj & Choudhury 2007; Majumdar, Bharadwaj & Choudhury 2012; Ghara, Choudhury & Datta 2016), or avoid (Datta, Bowman & Carilli 2010; Liu, Parsons & Trott 2014) the foregrounds, these only work if the sky signal has been measured with high fidelity over the time of observation. This then requires exquisite calibration of the system as any leftover artefacts from strong sources will make a measurement impossible (Barry et al. 2016; Patil et al. 2017). This implies calibrating the many hardware components of the telescope (see e.g. Kern et al. 2019) while a further complication is added by the presence of the temporally and spatially varying ionosphere (see e.g. Mevius et al. 2016).

Recently, Bowman et al. (2018) have claimed a detection of the sky-averaged 21-cm signal at $z \approx 17$ in observations with the EDGES2 low-band antenna. These results are debated (e.g. Draine & Miralda-Escudé 2018; Hills et al. 2018; Bradley et al. 2019; Singh & Subrahmanyan 2019), but if true would challenge our theoretical understanding of the early universe as explanations for its strength require either a previously unknown cooling mechanism (see e.g. Tashiro, Kadota & Silk 2014; Barkana 2018; Fialkov, Barkana & Cohen 2018; Muñoz & Loeb 2018) or a radio background other than the CMB (Feng & Holder 2018; Fialkov & Barkana 2019).

Other attempts have to date only provided upper limits on the expected signal. While global signal experiments probe the average brightness temperature, experiments with radio interferometers constrain the power spectrum of the expected 21-cm signal. Observations with the GMRT (Paciga et al. 2013) provided the very first upper limit, which was a 2σ value of $(248)^2$ mK² for $k = 0.50$ h Mpc⁻¹ at $z = 8.6$. Later PAPER and MWA produced additional upper limits (Parsons et al. 2014; Ali et al. 2015; Barry et al. 2019). Note that the PAPER collaboration initially reported a strong upper bound (Beardsley et al. 2016) that was later revised to a weaker upper bound (Cheng et al. 2018; Kolopanis et al. 2019). The first LOFAR upper limit on the power spectrum of the signal obtained from one night observation was $(79.6)^2$ mK² at $k = 0.053$ h Mpc⁻¹ and a redshift between 9.6 and 10.6 (Patil et al. 2017). Recently, upper limits were provided for even higher redshifts. Gehlot et al. (2019)

placed upper limits on the power spectrum in the redshift range $z = 19.8$ – 25.2 using observations with the LOFAR-Low Band Antenna array and Eastwood et al. (2019) placed upper limits at $z \approx 18.4$ using observations with the Owens Valley Radio Observatory Long Wavelength Array (OVRO-LWA).⁶

Mertens et al. (2020) have provided the second LOFAR upper limit on the 21-cm power spectrum at redshift ≈ 9.1 using 10 nights of observations. At $k = 0.1$ h Mpc⁻¹, the 2σ upper limit is $(106.65)^2$ mK², a factor of ≈ 8 improvement at the same k -scale over the value obtained from one night of observations (Patil et al. 2017) and the best upper limit so far on the large-scale power spectrum at redshift ≈ 9 . The results give upper limits for a range of k values but only at one redshift, $z \approx 9.1$. In this paper, we explore scenarios for the EoR that can be ruled out by these new upper limits. As they are about an order of magnitude higher than the most popular theoretical predictions, we can expect that only extreme models will be ruled out. Similar analyses were performed by Pober et al. (2015) and Greig, Mesinger & Pober (2016) for the earlier upper limits from PAPER, which were reported in Ali et al. (2015).

Extracting astrophysical and cosmological information from 21-cm observations is not straightforward as, in addition to the cosmological dependence, the characteristics of the expected signal depend crucially on specific properties of the early sources and their redshift evolution. While UV photons from such sources are mostly absorbed during ionization of H I in surrounding regions, X-ray photons, due to their longer mean free path (MFP), penetrate further into the neutral gas and increase its temperature (see e.g. Madau, Meiksin & Rees 1997; Zaroubi & Silk 2005; Zaroubi et al. 2007). At the same time, Ly α photons from the same sources determine the coupling strength of the H I spin temperature with the gas temperature. In view of this complexity, an exploration of many theoretical models of the expected 21-cm signal is necessary to interpret the results from radio observations. Such signal prediction algorithms are often combined with a Bayesian inference framework, such as the Markov chain Monte Carlo (MCMC), to explore and constrain the reionization parameters (e.g. Greig & Mesinger 2015; Greig & Mesinger 2017; Cohen et al. 2019; Park et al. 2019). This is the approach we will use here, relying on the GRIZZLY code (Ghara, Choudhury & Datta 2015a; Ghara et al. 2018) to generate the reionization scenarios and models for the 21-cm signal.

Since the codes that are used to generate the 21-cm signal use source parameters as input, the results from such Bayesian inference frameworks typically constrain these source parameters. However, it should be realized that the 21-cm observations themselves characterize the state of the IGM and do not contain any direct information about the source properties. It is perfectly possible that many different source models lead to the same 21-cm signal, especially if one only has information from a single redshift, as is the case for the latest LOFAR upper limits. As explained in more detail below, we will therefore have a strong focus on IGM parameters such as the average ionized fraction, average spin temperature, volume fraction of ‘heated region’ (i.e. partially ionized regions with gas temperature larger than the CMB temperature) and size distributions of ionized and heated regions. We give much less weight to the source parameters, which however are still needed by the models to generate the 21-cm signals.

Our paper is structured as follows. In Section 2, we describe the basic methodology to prepare the Bayesian framework used to interpret the observations. We present our results in Section 3 and

⁴<https://reionization.org/>

⁵<http://www.skatelescope.org/>

⁶<http://www.tauceti.caltech.edu/LWA/>

discuss them from the point of view of other observations in Section 4, before concluding in Section 5. The cosmological parameters as used in this study are $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_B = 0.044$, $h = 0.7$, consistent with the Wilkinson Microwave Anisotropy Probe (WMAP) results (Hinshaw et al. 2013). These are the same as used in the N -body simulation used in this paper. Within the error bars, these are consistent with Planck 2015 results (Planck Collaboration XIII 2016) that are used in Mertens et al. (2020). Note that all distances and scales used in this study are in comoving coordinates.

2 METHODOLOGY

Here, we introduce the approach employed to generate the 21-cm signal and compare it with the observational upper limit.

2.1 Cosmological simulations

We use the GRIZZLY code (Ghara et al. 2015a, 2018) to generate brightness temperature maps at redshift ≈ 9.1 . This algorithm requires cosmological density fields and halo catalogues as input. These are retrieved from results of the PRACE⁷ project PRACE4LOFAR, which was run specifically for the purpose of providing several cosmological simulations for the interpretation of LOFAR data. Here, we use the largest volume, of length $500 h^{-1}$ comoving megaparsec (Mpc) (see e.g. Giri et al. 2019a, b). This corresponds to a field of view of $4.27^\circ \times 4.27^\circ$ at redshift ≈ 9.1 which is comparable to the LOFAR primary beam of $\approx 4^\circ$. The cosmological N -body simulation was run using CUBEP³M (Harnois-Déraps et al. 2013) with 6912^3 particles and a mass resolution of $4.05 \times 10^7 M_\odot$. Haloes were identified on the fly with a spherical overdensity halo finder (Watson et al. 2013), and only those with masses $10^9 M_\odot$ and higher, i.e. resolved with at least ≈ 25 particles, were used. More details on the PRACE4LOFAR simulations can be found in Dixon et al. (2016).

The GRIZZLY simulations are run on gridded versions of the density fields from which the haloes have been removed as they are not part of the IGM and their effect is captured by the assumptions of the source model through the photon escape fraction. We use 300^3 grids for the results in this paper. The smallest k -scale that can be probed with this resolution is $\approx 1.9 h \text{ Mpc}^{-1}$ (corresponding to scale $\approx 3.3 h^{-1} \text{ Mpc}$). The smallest scale probed in Mertens et al. (2020), which is $\approx 0.4 h \text{ Mpc}^{-1}$ (corresponding to scale $\approx 15 h^{-1} \text{ Mpc}$), remains within the Nyquist limit of our simulation and free from the aliasing effect (Mao et al. 2012).

2.2 Modelling the 21-cm signal using GRIZZLY

The differential brightness temperature, δT_b , of the 21-cm signal can be expressed as (see e.g. Madau et al. 1997; Furlanetto, Oh & Briggs 2006b)

$$\delta T_b(\mathbf{x}, z) = 27 x_{\text{HI}}(\mathbf{x}, z) [1 + \delta_B(\mathbf{x}, z)] \left(\frac{\Omega_B h^2}{0.023} \right) \times \left(\frac{0.15}{\Omega_m h^2} \frac{1+z}{10} \right)^{1/2} \left(1 - \frac{T_\gamma}{T_S(\mathbf{x}, z)} \right) \text{ mK}, \quad (1)$$

where the quantities x_{HI} , δ_B , and $T_\gamma(z) = 2.73 (1+z) \text{ K}$ denote the neutral hydrogen fraction, baryonic density contrast, and the CMB temperature, respectively, each at position \mathbf{x} and redshift z . T_S represents the spin temperature of hydrogen in the IGM. In this paper, we will consider the dimensionless power

Table 1. The Δ_{21}^2 upper limits at 1σ level at redshift ≈ 9.1 from LOFAR observations for different k -bins (Mertens et al. 2020).

k ($h \text{ Mpc}^{-1}$)	$\Delta_{21}^2(k)$ (mK^2)	$\Delta_{21,\text{err}}^2$ (mK^2)
0.075	(58.97) ²	(30.26) ²
0.100	(95.21) ²	(33.98) ²
0.133	(142.17) ²	(39.98) ²
0.179	(235.80) ²	(51.81) ²
0.238	(358.95) ²	(64.00) ²
0.319	(505.26) ²	(87.90) ²
0.432	(664.23) ²	(113.04) ²

spectrum of the brightness temperature, i.e. $\Delta^2(k) = k^3 P(k)/2\pi^2$. The spherically averaged power spectrum $P(k)$ can be expressed as $\langle \delta \hat{T}_b(\mathbf{k}) \delta \hat{T}_b^*(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}') P(k)$, where $\delta \hat{T}_b(\mathbf{k})$ denotes the Fourier component of $\delta T_b(\mathbf{x})$ at wavenumber k .

The GRIZZLY algorithm is based on a one-dimensional radiative transfer scheme and is an independent implementation of the BEARS algorithm described by Thomas & Zaroubi (2008, 2011), Thomas et al. (2009), and Krause et al. (2018). It approximates the transfer of photons by assuming that the effect from individual sources is isotropic and can therefore be pre-calculated as radial profiles around each source. The algorithm corrects for overlap by ensuring that the total ionized volume of the region created by multiple sources is the correct one. This approach makes the code very fast, a requirement necessary for parameter studies such as the one we perform here.

Ghara et al. (2018) presented a detailed comparison between the performance of this code and the full three-dimensional radiative transfer code C²RAY (Mellema et al. 2006b). We found that although GRIZZLY employs a range of approximations, its results agree with those of the full radiative transfer quite well, while being at least 10^5 times faster. In Appendix A, we give a brief outline of this code, while we refer the reader to the original papers (Ghara et al. 2015a, 2018) for a more detailed and complete description of the algorithm. Note that we have not included redshift space distortions while evaluating the power spectrum for different model parameters, as their impact remains rather small during the EoR, when ionization fluctuations dominate the power spectrum of δT_b (Jensen et al. 2013; Ghara et al. 2015a; Majumdar et al. 2016).

The upper limits from Mertens et al. (2020) at scales $k = 0.075$ and $0.1 h \text{ Mpc}^{-1}$ are $\Delta^2 = (58.97)^2$ and $(95.21)^2 \text{ mK}^2$, respectively (see also Table 1). Before proceeding, it should be realized that these values are rather high compared to the power spectrum at redshift ≈ 9.1 predicted by various standard reionization scenarios, such as in Mellema et al. (2006a), Iliiev et al. (2007), Greig & Mesinger (2015), Ghara, Datta & Choudhury (2015b), Hassan et al. (2016), Bolgar et al. (2018), and Ross et al. (2019). For example, the predicted power spectra at $z \approx 9$ at $k = 0.1 h \text{ Mpc}^{-1}$ are found to be $\lesssim 10^3 \text{ mK}^2$. Models that can be excluded by these upper limits therefore have to be quite extreme.

As the lowest upper limit is for the largest scales, our aim is to identify scenarios that produce large amplitudes for the large-scale fluctuations. Spatial fluctuations in the 21-cm signal can only be caused by spatial fluctuations in x_{HI} , δ_B , or T_S (see equation 1). Previous studies have shown that the fluctuations in δ_B are small on the scales measured by LOFAR (e.g. Peebles 1993). We therefore consider two different scenarios to identify models with either large x_{HI} and/or T_S fluctuations. In the first scenario, we assume a uniform T_S , so that the large-scale fluctuations of the signal are mostly driven by fluctuations in x_{HI} . In the second scenario, we relax the uniform

⁷Partnership for Advanced Computing in Europe: <http://www.prace-ri.eu/>.

Table 2. Overview of the source parameters used in GRIZZLY, their explored ranges as well as for which models these are used as input parameters.

Source parameters	Description	Explored range	Corresponding models
ζ	Ionization efficiency	$[10^{-2}, 10^{2.5}]$	Varied in both the uniform and non-uniform T_S models
M_{\min}	Minimum mass of the UV emitting haloes	$[10^9 M_\odot, 10^{12} M_\odot]$	Varied in the uniform T_S model Fixed to $10^9 M_\odot$ in the non-uniform T_S model
$M_{\min, X}$	Minimum mass of the X-ray emitting haloes	$[10^9 M_\odot, 10^{12} M_\odot]$	Used and varied only for the non-uniform T_S model
f_X	X-ray heating efficiency	$[0.1, 10]$	Used and varied only for the non-uniform T_S model
α	Spectral index of the X-ray spectrum	Fixed to 1.2 (fiducial) or 0.3	Used only for the non-uniform T_S model

T_S assumption and allow sources of heating to create local regions of high T_S . In this case, the large-scale fluctuations are predominantly sourced by fluctuations in T_S . These two scenarios will be discussed in detail later in Sections 3.1 and 3.2.

To calculate the evolution of the IGM for these scenarios, GRIZZLY needs to characterize the source properties with a range of parameters. The following are used in our study (also listed in Table 2).

(i) Ionization efficiency (ζ): The rate of ionizing photons per unit stellar mass escaping from a halo is given by $\dot{N}_i = \zeta \times 2.85 \times 10^{45} \text{ s}^{-1} M_\odot^{-1}$. This value corresponding to $\zeta = 1$ derives from the model galaxy spectrum employed when calculating x_{HII} and T_K , which has been produced with the publicly available code PEGASE2⁸ (Fioc & Rocca-Volmerange 1997). Note that the emission rate of the ionizing photons is assumed to be proportional to the halo mass. We refer the reader to Ghara et al. (2015a) for more details. We calculate the stellar mass of a halo using $M_* = f_* \times \frac{\Omega_B}{\Omega_m} \times M_{\text{halo}}$, where f_* is the star formation efficiency, fixed at 0.02 (Behroozi & Silk 2015; Sun & Furlanetto 2016). The parameter ζ combines all the degeneracies from various quantities related to the star formation rate and the emission rate of ionizing photons from the sources, as well as their escape fraction into the IGM. The case $\zeta = 1$ corresponds to a star formation efficiency of 2 per cent and an escape efficiency of 100 per cent, but also to a star formation efficiency of 20 per cent and an escape efficiency of 10 per cent. We vary ζ in both scenarios considered in this paper.

(ii) Minimum mass of the UV emitting haloes (M_{\min}): In the above parametrization of the ionizing efficiency, the number of ionizing photons escaping from a halo depends linearly on its mass. However, below a certain minimum mass radiative and mechanical feedback can severely reduce the star formation efficiency (see e.g. Hasegawa & Semelin 2013; Dawoodbhoy et al. 2018). We model this by introducing M_{\min} as the minimum mass of haloes from which ionizing photons escape into the IGM. As with ζ , this parameter represents different physical processes, not only feedback but for example also very low escape fractions from lower mass haloes (see e.g. Gnedin, Kravtsov & Chen 2008; Sharma et al. 2016). Due to the mass resolution of our N -body simulation (see Section 2.1), the lowest value for M_{\min} is $10^9 M_\odot$. Although haloes of lower masses could contribute, as we will see below, the LOFAR results are not able to constrain such low values. In general, one expects star formation in haloes with mass $\lesssim 10^9 M_\odot$ to be suppressed due to radiative feedback (see e.g. Wise et al. 2014; Dixon et al. 2016). Note that we do not employ radiative feedback in this study as M_{\min} remains $\gtrsim 10^9 M_\odot$ for the scenarios considered here. We vary M_{\min} in the uniform T_S scenario while fix it to $10^9 M_\odot$ in the non-uniform T_S model.

(iii) Minimum mass of X-ray emitting halo ($M_{\min, X}$): In addition to the stellar contributions, GRIZZLY can also include heating and ionization from X-ray sources such as quasars, high-mass X-ray binaries, etc. As not all star hosting haloes are necessarily substantial X-ray sources, we use the minimum mass of dark matter haloes that contain X-ray sources as a separate parameter. This allows us to include scenarios in which the X-ray source population deviates from the population of galaxies. We consider and vary this parameter only in the non-uniform T_S model.

(iv) X-ray heating efficiency (f_X) and spectral index (α): The spectrum of an X-ray source at energy E is modelled as a power law, i.e. $I_q(E) = A_q E^{-\alpha}$, where α is the spectral index. The normalization constant A_q is determined such that the X-ray luminosity per stellar mass is $3.4 \times 10^{34} f_X \text{ erg s}^{-1} M_\odot^{-1}$, where f_X is a parameter. This implies a rate of X-ray photons per unit stellar mass emitted from a halo $\dot{N}_x = f_X \times 8.47 \times 10^{43} \text{ s}^{-1} M_\odot^{-1}$. The value of \dot{N}_x for $f_X = 1$ is \sim two orders of magnitude larger than the measurements of high-mass X-ray binaries in local star-forming galaxies in 0.5–8 keV band (Mineo, Gilfanov & Sunyaev 2012). We assume that the UV band spans the range 13.6–100 eV, while the X-ray band goes from 100 eV to 10 keV. We vary f_X while we keep α fixed at 1.2 (fiducial) or 0.3 in the non-uniform T_S model.

2.3 Derived IGM parameters

As mentioned above, although GRIZZLY uses astrophysical source parameters to generate brightness temperature maps, the main goal of this work is to infer the IGM properties at $z \approx 9.1$ from the new LOFAR upper limit. At this epoch, the IGM is expected to consist of H II regions embedded in a (partially) neutral medium. The signal from such gas is in emission ($\delta T_b > 0$), in absorption ($\delta T_b < 0$), or zero depending on its spin temperature T_S . In addition to δ_B , two major sources of the spatial fluctuations of the signal are fluctuations in ionized fraction x_{HII} and spin temperature T_S .

If the signal is dominated by x_{HII} fluctuations, the maximum power spectrum obtained from a model depends not only on the volume-averaged ionized fraction (\bar{x}_{HII}) and spin temperature, but also on the size distribution of the H II bubbles (e.g. Furlanetto, McQuinn & Hernquist 2006a). We will therefore study the latter by characterizing the probability distribution function (PDF) of the sizes of H II regions with $R_{\text{peak}}^{\text{HII}}$ and $\Delta R_{\text{FWHM}}^{\text{HII}}$, which represent the size at which the PDF has a maximum and the full width at half-maximum (FWHM), respectively. Fig. 1 shows an example of such a distribution.

Similarly, in the presence of spin temperature fluctuations, the power spectrum of the 21-cm signal also depends on the size distribution of the heated regions (i.e. regions with $T_K > T_\gamma$) besides the average gas temperature \bar{T}_K , fraction of volume occupied by the heated regions f_{heat} , and mass-averaged brightness temperature (δT_b). Similarly to the PDF of H II regions, we will characterize the size

⁸<http://www2.iap.fr/pegase/>

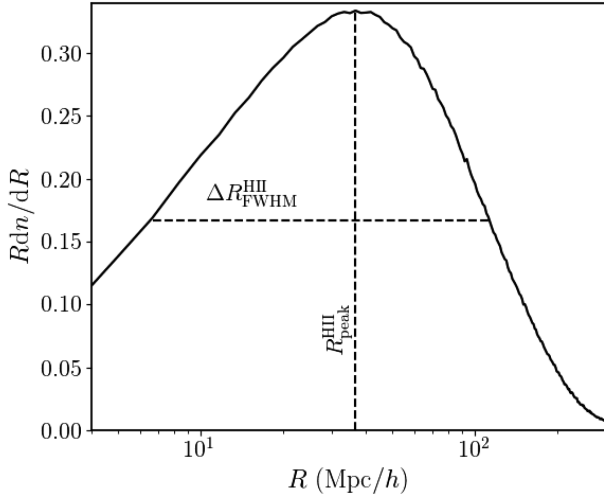


Figure 1. The PDF of the ionized regions of size R estimated using the MFP method. This ionization state is the same as shown in the left-hand panel of Fig. 2 which corresponds to the parameter choice $\zeta = 50$, $M_{\min} = 3 \times 10^{10} M_{\odot}$. $R_{\text{peak}}^{\text{HII}}$ and $\Delta R_{\text{FWHM}}^{\text{HII}}$ represent the size of the H II bubbles at which the PDF becomes maximum and the FWHM of the distribution, respectively.

distribution of the heated regions adopting the parameters $R_{\text{peak}}^{\text{heat}}$ and $\Delta R_{\text{FWHM}}^{\text{heat}}$.

There exists no unique way to characterize the size distribution of a complex three-dimensional structure such as the distribution of H II/heated regions in the IGM. We refer the reader to Friedrich et al. (2011), Lin et al. (2016), and Giri et al. (2018a) for an overview of the various methods that can be used. In this work, we will use a Monte Carlo based approach, namely the MFP method, first proposed by Mesinger & Furlanetto (2007). In the MFP algorithm, we randomly select a point inside the region of interest (e.g. H II regions) and shoot a ray in a random direction until it reaches the boundary of the region. The length of the ray is recorded. When this process is repeated numerous times, the PDF of the recorded lengths provides the PDF of the regions of interest. Here, we use 10^7 rays shot on the fly during the GRIZZLY simulation.

A list of parameters used to characterize the IGM is given below (also in Table 3):

- (i) Volume-averaged ionized fraction ($\overline{x_{\text{HII}}}$).
- (ii) Volume-averaged gas temperature in the partially ionized IGM with $x_{\text{HII}} < 0.5$ ($\overline{T_{\text{K}}}$).
- (iii) Uniform spin temperature of the IGM (T_{S}). Note that $1-(T_{\gamma}/T_{\text{S}})$ form will be used rather than T_{S} .
- (iv) Mass-averaged differential brightness temperature ($\overline{\delta T_{\text{b}}}$).
- (v) Volume fraction of heated regions with $T_{\text{K}} > T_{\gamma}$ (f_{heat}).
- (vi) $R_{\text{peak}}^{\text{HII}}$ ($R_{\text{peak}}^{\text{heat}}$): Size of the H II (heated) regions at which the PDF of the sizes peaks.
- (vii) $\Delta R_{\text{FWHM}}^{\text{HII}}$ ($\Delta R_{\text{FWHM}}^{\text{heat}}$): FWHM of the PDF of the sizes of the H II (heated) regions.

Note that we do not model the signal directly in terms of these IGM parameters, these are rather *derived* quantities from the simulations.

2.4 GRIZZLY emulator

Although GRIZZLY is fast and efficient, for parameter estimation with a Bayesian inference framework where hundreds of thousands of models may be needed, the use of GRIZZLY can become

computationally too expensive. We therefore adopt an alternative approach. First, we emulate the power spectra derived from GRIZZLY simulations using the machine learning algorithm known as Gaussian Process Regression (GPR; Rasmussen & Williams 2006). The power spectrum emulator is used to interpolate within the parameter space and evaluate the power spectrum for parameter values which have not been simulated. For a description on how to emulate EoR simulations with GPR, we refer the reader to Kern et al. (2017) and Jennings et al. (2019). We use the GPR module provided in the PYTHON package SCIKIT-LEARN (Pedregosa et al. 2011). We determine the values for the hyperparameters for GPR using cross-validation, a process which prevents overfitting of the model (e.g. Franklin 2005; Cawley & Talbot 2010). We have used 10-fold cross-validation (Kohavi 1995) to construct the emulators.

Given a set of parameters as described in the previous section, we have configured GRIZZLY to generate the spherically averaged power spectrum for the k -bins of the LOFAR data. However, as we will see later, not all the data points from the upper limit of the power spectrum are useful for this analysis. We therefore only use power spectrum amplitudes at scales $k \lesssim 0.15 h \text{ Mpc}^{-1}$ to build up our emulator, more specifically at $k = 0.075, 0.1, \text{ and } 0.13 h \text{ Mpc}^{-1}$. We quantify the accuracy of the emulators with their mean squared error (MSE).⁹ In order to test the accuracy, we calculate the MSE for the testing set. The testing set is independent of the data set used for emulation. The MSE of the emulators for predicting the 21-cm power spectrum is found to be less than 10 per cent.

We combine this emulator with the MCMC module available in the EMCEE PYTHON package (Foreman-Mackey et al. 2013) to explore the parameter space of different scenarios. As we are interested in the IGM parameters, we also construct emulators for mapping the source parameters to the IGM parameters. The MSE of these emulators is less than 5 per cent.

2.5 Bayesian inference framework

As described in the previous section, we combine the GRIZZLY emulator with an MCMC algorithm to explore the parameter space for different scenarios and to constrain them using the observed upper limits. The probability of any parameter value θ , i.e. the posterior $p(\theta|\mathbf{x})$, given some observation \mathbf{x} , is defined by Bayes' theorem as

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta) p(\theta), \quad (2)$$

where $p(\theta)$ represents the prior on the parameter values. The quantity $p(\mathbf{x}|\theta)$, also known as the likelihood \mathcal{L} , gives the probability of any observation given certain parameters. It should be kept in mind that the likelihood cannot be defined by the formal χ^2 method as the observed power spectrum is an upper limit only. Therefore, here we define the likelihood as follows.

Let us denote the observed power spectrum $\Delta_{\text{o}}^2(k_i)$ by $\Delta_{21}^2(k_i) \pm \Delta_{21,\text{err}}^2(k_i)$, while the model power spectrum estimated using the emulator for a set of parameters θ is denoted by $\Delta_{\text{m}}^2(k_i, \theta)$. Two major sources of uncertainty on the modelled large-scale power spectrum are: (i) error from the emulators themselves, (ii) sample variance that increases at larger scales. The combined error remains

⁹The MSE of the emulator is defined as (e.g. Jennings et al. 2019)

$$\text{MSE} = \left\langle \left(\frac{Q_{\text{true}} - Q_{\text{emulated}}}{Q_{\text{true}}} \right)^2 \right\rangle,$$

where $\langle \rangle$ represents the mean estimate. The quantities Q_{true} and Q_{emulated} are true and emulated values, respectively.

Table 3. An overview of the IGM parameters considered in this paper. Except for T_S in the case of the uniform T_S model, all of these are *derived* from the simulation results. We explore a range $[-12:1]$ for $1 - T_\gamma/T_S$. The last column refers to the models in which such a quantity is considered.

IGM parameters	Description	Corresponding models
$\overline{x_{\text{H II}}}$	Volume-averaged ionized fraction	Uniform and non-uniform T_S models
$\overline{T_K}$	Volume-averaged gas temperature in the partially ionized IGM with $x_{\text{H II}} < 0.5$	Non-uniform T_S model
$1 - T_\gamma/T_S$	T_γ and T_S are the CMB and spin temperature	Uniform T_S model
$\overline{\delta T_b}$	Mass-averaged differential brightness temperature	Uniform and non-uniform T_S models
f_{heat}	Volume fraction of regions with temperature larger than T_γ	Non-uniform T_S model
$R_{\text{peak}}^{\text{H II}}$	Size at which the PDF of the size distribution of the H II regions peaks	Uniform T_S model
$R_{\text{peak}}^{\text{heat}}$	Size at which the PDF of the size distribution of the heated regions peaks	Non-uniform T_S model
$R_{\text{FWHM}}^{\text{H II}}$	FWHM of the PDF of the size distribution of the H II regions	Uniform T_S model
$R_{\text{FWHM}}^{\text{heat}}$	FWHM of the PDF of the size distribution of the heated regions	Non-uniform T_S model

$\lesssim 10$ per cent for the scales considered in this study. Thus, we assume a 10 per cent modelling error, $\sigma_m(k_i) = 0.1 \times \Delta_m^2(k_i, \theta)$. This error is always larger than the sampling error from the simulation. The total variance $\sigma^2 = \Delta_{21, \text{err}}^4 + \Delta_{m, \text{err}}^4$ includes the errors from the observation and simulations. For an upper limit, we define the likelihood $\mathcal{L}(\theta)$ for a model with parameters θ as (see Appendix B for the derivation)

$$\mathcal{L}(\theta) = \prod_i \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\Delta_{21}^2(k_i) - \Delta_m^2(k_i, \theta)}{\sqrt{2}\sigma(k_i)} \right) \right]. \quad (3)$$

This expression results in the following key behaviour:

- (i) If the model power spectrum, $\Delta_m^2(k_i, \theta)$, as estimated using the emulator for a set of parameters θ is larger than the observed power spectrum, $\Delta_{21}^2(k_i) + \sigma(k_i)$, within at least one k -bin k_i , $\mathcal{L}(\theta)$ approaches 0 and that model is ruled out by the upper limit.
- (ii) If $\Delta_m^2(k_i, \theta)$ is smaller than $\Delta_{21}^2(k_i) - \sigma(k_i)$ for *all* k -bins, $\mathcal{L}(\theta)$ approaches 1, and that model is consistent with the upper limit.
- (iii) In case the above two conditions do not hold, the likelihood estimated from equation (3) remains between 0 and 1.

In this work, we aim to find models that are excluded by the measured upper limit. Thus, we use $\mathcal{L}_{\text{ex}}(\theta) = 1 - \mathcal{L}(\theta)$ in the MCMC analysis as the likelihood of a set of parameters θ to be excluded by the upper limit.¹⁰

In addition to this, we use a prior on the ionized fraction estimated from the measured Thomson scattering optical depth in Planck Collaboration VI (2018). As we do not have any prior information about the redshift evolution of the average ionized fraction, $\overline{x_{\text{H II}}}$, we estimate the maximum value which is possible at redshift ≈ 9.1 as follows. If we assume that $\overline{x_{\text{H II}}}$ increases or remains constant with time, a Thomson scattering optical depth $\tau = 0.054 \pm 0.007$ translates into a maximum ionized fraction 0.57 ± 0.24 at redshift ≈ 9.1 . Here, we thus use $\overline{x_{\text{H II}, \text{max}}}(z = 9.1) = 0.81$ as the maximum possible value for $\overline{x_{\text{H II}}}$ at redshift ≈ 9.1 . This corresponds to a scenario in which the universe is neutral at $z > 9.1$, has a constant ionized fraction in the range $9.1 > z > 6$ and reionization ends suddenly at redshift ≈ 6 . Note that this is unlikely to be a realistic scenario, as $\overline{x_{\text{H II}}}$

is expected to gradually increase to 1 with time. While studies such as Pober et al. (2015), Greig et al. (2016), and Monsalve et al. (2019) consider model-dependent reionization histories and estimate $x_{\text{H II}}$ by comparing the estimated τ for the models with the measured τ from the CMB observations, here we use $\overline{x_{\text{H II}, \text{max}}}(z = 9.1) = 0.81$ as a model-independent conservative upper limit of the ionized fraction at $z \approx 9.1$.

3 RESULTS

Now we apply the parameter estimation framework described in the previous section to the upper limits from Mertens et al. (2020) (also given in Table 1). As described before, we will discuss two scenarios. While the first one considers ionized patches in a uniform T_S IGM, the second one also includes T_S fluctuations. We present our results in the following sections.

3.1 Ionized patches and a uniform T_S

In this section, we focus on the scenario in which the large-scale modes are caused by the presence of ionized regions, and assume a uniform spin temperature with a value T_S (see e.g. Ali et al. 2015 and Pober et al. 2015 for previous papers adopting a uniform T_S model). These ionized regions are expected to be photoheated to a temperature $T_K \approx 10^4$ K and emit no signal as $x_{\text{H I}} \approx 0$. Here, T_S represents the spin temperature of the neutral part of hydrogen in the IGM. The sizes and spatial distribution of the ionized regions are determined by the astrophysical parameters ζ and M_{min} . Therefore, this model has three parameters ζ , M_{min} , and $1 - T_\gamma/T_S$ which we will explore.

We further assume the existence of a uniform Ly α background which fully couples T_S to the kinetic temperature T_K , and thus a uniform T_S implies a uniform T_K for the neutral IGM. The lowest value of T_K is obtained in the complete absence of heating processes, when adiabatic cooling due to cosmological expansion gives $T_K = 2.1$ K at $z \approx 9.1$ for our choice of cosmological parameters (calculated using CMBFAST; Zaldarriaga & Seljak 2000).¹¹ Higher values for T_K can be caused by heating through X-rays. To obtain a uniform

¹⁰Note that, following the same calculation shown in Appendix B, one can also directly estimate the likelihood of set of parameters θ to be excluded as $\mathcal{L}_{\text{ex}}(\theta) = \prod_i \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\Delta_{21}^2(k_i) - \Delta_m^2(k_i, \theta)}{\sqrt{2}\sigma(k_i)} \right) \right]$.

¹¹Here we do not consider any additional cooling mechanisms, such as the interaction between baryons and cold dark matter particles, which have been proposed to explain the recent EDGES low-band observations of the global signal at $z \approx 17$ (Barkana 2018; Bowman et al. 2018) nor additional sources

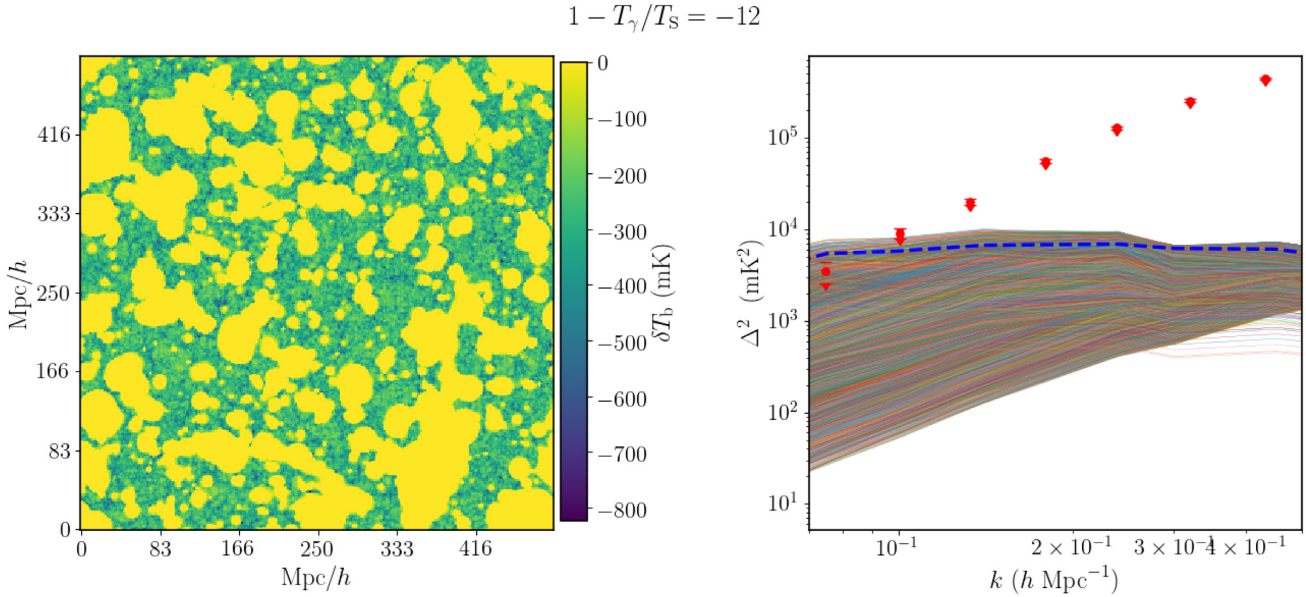


Figure 2. *Left-hand panel:* a slice through the brightness temperature cube at $z \approx 9.1$ for the parameter choice $\zeta = 50$, $M_{\min} = 3 \times 10^{10} M_{\odot}$, and $1 - T_{\gamma}/T_S = -12$. The averaged ionized fraction of this map is 0.55. *Right-hand panel:* the curves show the power spectra of the 21-cm brightness temperature as a function of scale for 8556 different combinations of ζ and M_{\min} . We assume $1 - T_{\gamma}/T_S = -12$, which corresponds to a uniform $T_S = 2.1$ K. The red points with error bars (2σ) show the upper limits from the 10-night observations with LOFAR (Mertens et al. 2020). The dashed blue curve refers to the spherically averaged power spectrum of the brightness temperature cube from which the slice in the left-hand panel has been extracted. The model power spectra shown in the right-hand panel are also used to build an emulator of the power spectrum using the GPR.

distribution, this heating will have to be driven by very hard rather than soft X-ray photons (see e.g. Fialkov, Barkana & Visbal 2014; Pacucci et al. 2014). Since the spin temperature appears in the differential brightness temperature expression (equation 1) as $1 - T_{\gamma}/T_S$, we use this, rather than T_S , as a parameter in our study. For $T_S \gg T_{\gamma}$, $1 - T_{\gamma}/T_S$ approaches 1, while for the lowest value of $T_S = 2.1$ K, $1 - T_{\gamma}/T_S \approx -12$. We therefore explore the range $[-12, 1]$.

Since we assume that $1 - T_{\gamma}/T_S$ is constant, the power spectrum scales by $(1 - T_{\gamma}/T_S)^2$ at all wavenumbers. Therefore, we train our GPR emulator only to generate power spectra for different combinations of ζ and M_{\min} while keeping $1 - T_{\gamma}/T_S = 1$. For ζ and M_{\min} , we select the ranges $[10^{-2} - 10^{2.5}]$ and $[10^9 - 10^{12} M_{\odot}]$, respectively. The total number of GRIZZLY models used for training the emulator is 8556.

We first illustrate the outcome of this set of GRIZZLY models in Fig. 2. The left-hand panel shows a 2D slice from the brightness temperature cube for the case $\zeta = 50$, $M_{\min} = 3 \times 10^{10} M_{\odot}$, and $1 - T_{\gamma}/T_S = -12$. This combination of parameters produces an ionization map with large H II bubbles with characteristic size larger than several tens of Mpc and a volume-averaged ionized fraction of 0.55. The corresponding power spectrum is plotted as a thick dashed curve in the right-hand panel of Fig. 2, together with the other 8555 models from the training set. All these curves assume the minimal value of $1 - T_{\gamma}/T_S = -12$.

The red points in the right-hand panel of Fig. 2 denote the current LOFAR upper limits on Δ^2 with 1σ error bars. Clearly, some of the models have a power spectrum amplitude larger than the upper limits at the larger scales. These results also show that scales with $k \gtrsim 0.15 h \text{ Mpc}^{-1}$ do not significantly constrain the models, which is why, as mentioned in Section 2.4, we only use the lowest three

k values to build the emulator and to calculate the likelihood in the MCMC framework.

3.1.1 GRIZZLY and IGM parameters

Fig. 3 shows the dependence of the power spectra at scale $k = 0.075 h \text{ Mpc}^{-1}$ on the parameters ζ and M_{\min} obtained from the training set. The left-hand and right-hand panels of the figure correspond to $1 - T_{\gamma}/T_S = -12$ and -9 , respectively. The solid curves in both panels represent the contours corresponding to the upper limit at this scale, i.e. $\Delta^2 = (58.97)^2 \text{ mK}^2$. One can easily see that a significant part of the parameter space can be ruled out by this upper limit alone for $1 - T_{\gamma}/T_S = -12$. However, the volume of parameter space which can be excluded rapidly shrinks for higher values of $1 - T_{\gamma}/T_S$, and almost no constraints can be set for $1 - T_{\gamma}/T_S \gtrsim -8$.

Fig. 3 also shows that the section of parameter space covering $\zeta \gtrsim 10$ and $10^{9.8} M_{\odot} \lesssim M_{\min} \lesssim 10^{11} M_{\odot}$, which produces a highly ionized IGM is disfavoured. In fact, the excluded parameter space remains close to the parameter space which completes reionization by redshift ≈ 9.1 , which corresponds to the region in white at the bottom right corner of both panels.

In the left-hand panel of Fig. 4, we plot the average ionized fraction $\overline{x_{\text{H II}}}$ as a function of the ζ and M_{\min} values we have explored. One can easily see that the models with the largest amplitude of the large-scale power spectrum correspond to an ionized fraction ≈ 0.5 . This is expected as at this stage of the reionization process, the typical dimension of the bubbles becomes comparable to the size of the scale of interest (see e.g. Ghara et al. 2015a). As the ionized fraction approaches 1, the power spectrum decreases and becomes negligible at the end of the reionization process due to the paucity of neutral hydrogen. One can see that for $1 - T_{\gamma}/T_S = -12$, the excluded parameter space corresponds to average ionized fractions

of excess radio background as considered in studies such as Feng & Holder (2018) and Fialkov & Barkana (2019).

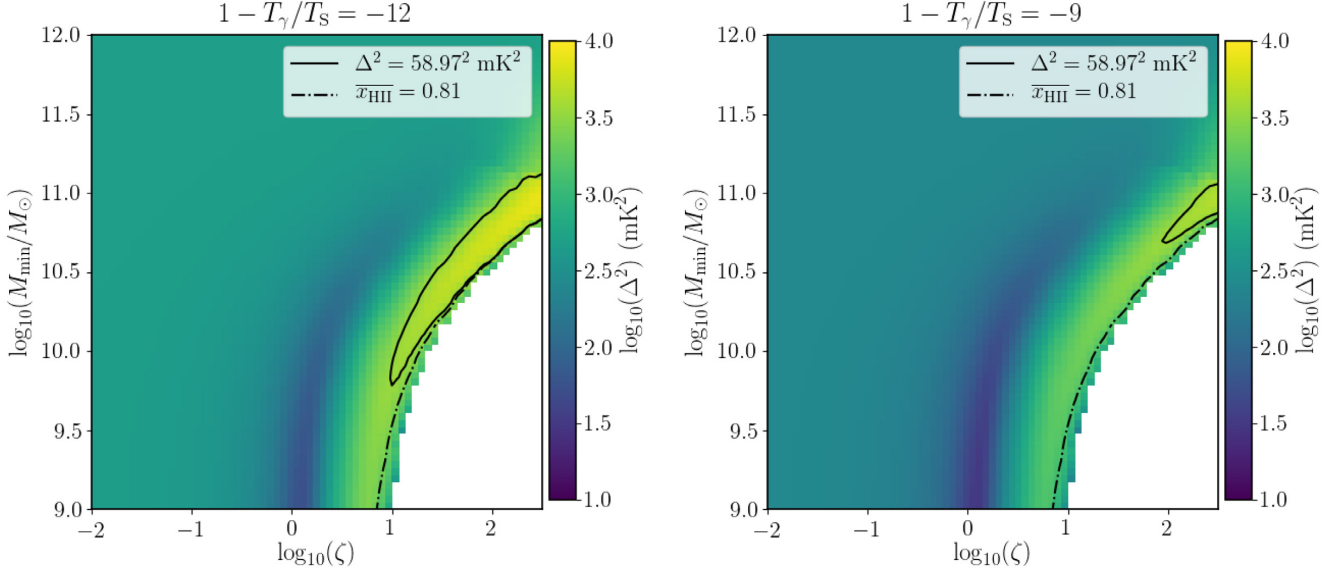


Figure 3. Power spectra at scale $k = 0.075 h \text{ Mpc}^{-1}$ from the training set as a function of ζ and M_{min} . We assume $1 - T_\gamma/T_s = -12$ and -9 for the left-hand and right-hand panels, respectively, i.e. $T_s = 2.1$ and 2.73 K at $z \approx 9.1$. The white region at the right bottom corner of the panels corresponds to a fully ionized IGM. The solid contours in both panels represent the upper limit constraint from LOFAR at scale $k = 0.075 h \text{ Mpc}^{-1}$, i.e. $\Delta^2 = (58.97)^2 \text{ mK}^2$. For a deterministic observation, the region enclosed by the solid contour will be excluded. The dash-dotted lines denote the contour for $\bar{x}_{\text{HII}} = 0.81$.

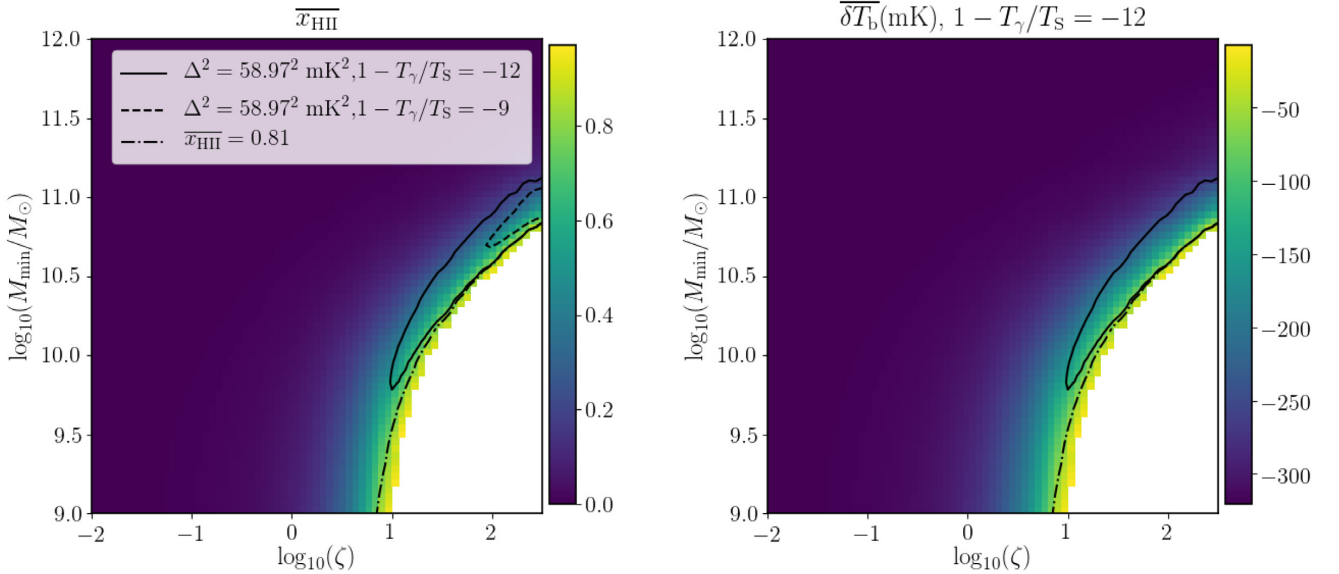


Figure 4. Averaged ionized fraction (left-hand panel) and brightness temperature (right-hand panel) at $z \approx 9.1$ from the training set as a function of ζ and M_{min} . We assume $1 - T_\gamma/T_s = -12$ in the right-hand panel. The white regions at the right bottom corners of the panels correspond to a fully ionized IGM. The contours in both panels represent the upper limit from LOFAR at scale $k = 0.075 h \text{ Mpc}^{-1}$, i.e. $\Delta^2 = (58.97)^2 \text{ mK}^2$, for $1 - T_\gamma/T_s = -12$ (solid) and -9 (dashed). For a deterministic observation, the region enclosed by the contours will be excluded. The dash-dotted lines denote the contour $\bar{x}_{\text{HII}} = 0.81$.

$\gtrsim 0.2$. It is interesting to note that, coincidentally, in this scenario the parameter space excluded by the LOFAR upper limit shares the same boundary at $\bar{x}_{\text{HII}} \approx 0.81$ with the parameter space excluded by the CMB Thomson scattering optical depth constraint on the maximum possible value of ionized fraction at redshift ≈ 9.1 (dash-dotted line, see Section 2.5).

The right-hand panel of Fig. 4 shows the value of the average brightness temperature $\overline{\delta T_b}$ as a function of ζ and M_{min} for $1 - T_\gamma/T_s = -12$. $\overline{\delta T_b}$ falls between $\approx -300 \text{ mK}$ (fully neutral) and zero (fully ionized). We find that the excluded parameter space is

concentrated around $\overline{\delta T_b} \gtrsim -250 \text{ mK}$. This is due to the fact that the average ionized fraction remains $\gtrsim 0.2$ for the excluded parameter space for the case of the lowest spin temperature.

Fig. 5 shows the dependences of $R_{\text{peak}}^{\text{HII}}$ and $\Delta R_{\text{FWHM}}^{\text{HII}}$ on ζ and M_{min} . Clearly, the most probable size of the bubbles and the FWHM increase with increasing ζ and decreasing M_{min} , as the average ionized fraction increases (also see Giri et al. 2018a; Giri, Mellema & Ghara 2018b). As expected, the parameter space that is excluded has preferentially a large characteristic size of the ionized regions. More specifically, the part of the parameter space that can be

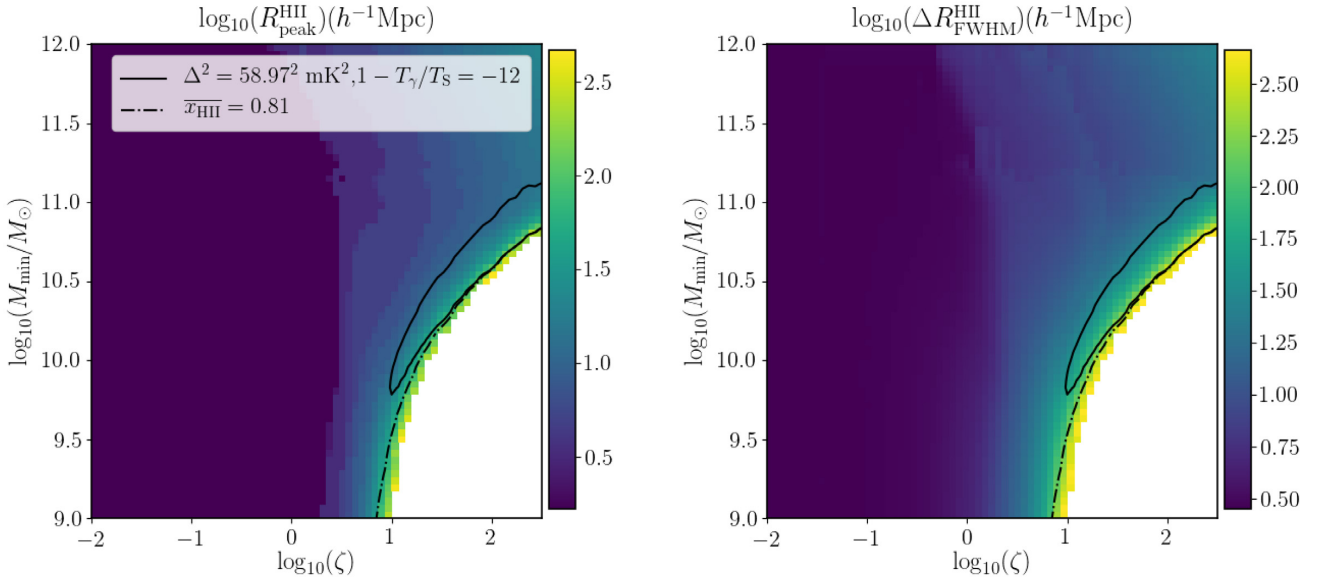


Figure 5. The distribution of $R_{\text{peak}}^{\text{HII}}$ (left-hand panel) and $\Delta R_{\text{FWHM}}^{\text{HII}}$ (right-hand panel) over the parameter space of ζ and M_{min} for $z \approx 9.1$. $R_{\text{peak}}^{\text{HII}}$ and $\Delta R_{\text{FWHM}}^{\text{HII}}$ represent the radius at which the probability distribution of the sizes of the ionized regions has maximum amplitude and the FWHM of that distribution, respectively. We use the MFP method to estimate the bubble size distribution. The region in white in the bottom right of the panels corresponds to fully ionized IGM at redshift ≈ 9.1 . The solid curves show the contours of $\Delta^2 = (58.97)^2 \text{ mK}^2$ for $1 - T_\gamma/T_S = -12$. For a deterministic observation, the region enclosed by the solid contour will be excluded. The dash-dotted line shows the contour for $\overline{x_{\text{HII}}} = 0.81$.

ruled out for $1 - T_\gamma/T_S = -12$ shows $R_{\text{peak}}^{\text{HII}} \gtrsim 10 h^{-1} \text{ Mpc}$ and $\Delta R_{\text{FWHM}}^{\text{HII}} \gtrsim 30 h^{-1} \text{ Mpc}$.

To test whether the results for the derived IGM quantities could be sensitive to the choices for the source model, we also explored a source model in which the ionizing emissivity depends non-linearly on the halo mass. As can be seen in Appendix C, changing the source model affects the constraints on the source parameters but reproduces the same constraints on the derived IGM parameters, illustrating that these constitute the more robust results of our study.

3.1.2 MCMC results

Up to this point, we have only explored the implications of the LOFAR upper limits using the results from GRIZZLY for slices through selected parameter spaces. In this section, we employ our parameter estimation framework that includes the emulator results and an MCMC algorithm. The aim is to explore the full parameter space and find the probability that the models are ruled out by the current upper limit from LOFAR. We use 20 walkers and 10^6 steps for this MCMC analysis. We checked the convergences of the MCMC chains using the integrated autocorrelation time as suggested in Goodman & Weare (2010) and find that the chains are well converged for this number of steps.

The likelihood used for this MCMC analysis is given by equation (3). In addition, we use a flat prior on $\overline{x_{\text{HII}}}(z = 9.1) \leq 0.81$. Fig. 6 shows the posterior distribution of the parameters ζ , M_{min} , and $1 - T_\gamma/T_S$, with the solid and dashed curves indicating the 68 per cent and 95 per cent credible intervals¹² of the excluded models within the range of parameters considered here. As expected, and as already suggested by Fig. 3, higher values of ζ ($\gtrsim 10$) and lower values of M_{min} (in particular $10^{9.8} M_\odot \lesssim M_{\text{min}} \lesssim 10^{11} M_\odot$) are more likely to

¹²We estimate the credible intervals of our posterior distributions by the approach based on computing the highest density interval (see e.g. Hyndman 1996).

be excluded as they result in higher ionization and thus a large-scale power spectrum more likely to exceed the observed one. Similarly, a colder IGM is more likely to be ruled out than a hotter IGM, as the former increases the signal strength.

We use a separate emulator to estimate the IGM parameters $\overline{x_{\text{HII}}}$, $R_{\text{peak}}^{\text{HII}}$, and $\Delta R_{\text{FWHM}}^{\text{HII}}$ for this scenario from the same set of GRIZZLY source parameters as used in our MCMC framework. This emulator is constructed using the same method as described in Section 2.4. Fig. 7 shows the posterior distribution of the IGM parameters. The constraints on the excluded IGM parameters are also listed in Table 4. Clearly, an IGM with $\overline{x_{\text{HII}}} \approx 0.13 - 0.74$, $(1 - T_\gamma/T_S) \lesssim -8.5$, and H II bubble distribution characterized by $R_{\text{peak}}^{\text{HII}} \approx 8 - 58 h^{-1} \text{ Mpc}$ and $\Delta R_{\text{FWHM}}^{\text{HII}} \approx 16 - 185 h^{-1} \text{ Mpc}$ is ruled out within 95 per cent credible intervals. This part of the parameter space corresponds to $-250 \text{ mK} \lesssim \overline{\delta T_b} \lesssim -55 \text{ mK}$. Note that the excluded parameter space requires satisfying all of the above-quoted conditions. These results are in agreement with our findings in Section 3.1.1. However, it is also clear from Figs 6 and 7 that tighter constraints on the power spectrum are required to put any bounds on source and IGM parameters with this analysis if the IGM is not very cold.

3.2 Spin temperature fluctuations

In this section, we relax the uniform T_S assumption and consider the scenario in which X-ray sources cause partial ionization and heating of the IGM. However, we will not vary all five GRIZZLY parameters ζ , M_{min} , $M_{\text{min}, X}$, f_X , and α (see Section 2.2). Instead, we fix the values of M_{min} and α and only retain the remaining three parameters. This choice is motivated by a preliminary study suggesting that the LOFAR upper limits provide very weak constraints on M_{min} and α .

We set $M_{\text{min}} = 10^9 M_\odot$, i.e. the lowest dark matter halo mass provided by our N -body results (Section 2.1). This means that, unlike the previous scenario, all haloes contribute to the ionization of the neutral hydrogen in the IGM. All the haloes also emit Ly α photons,

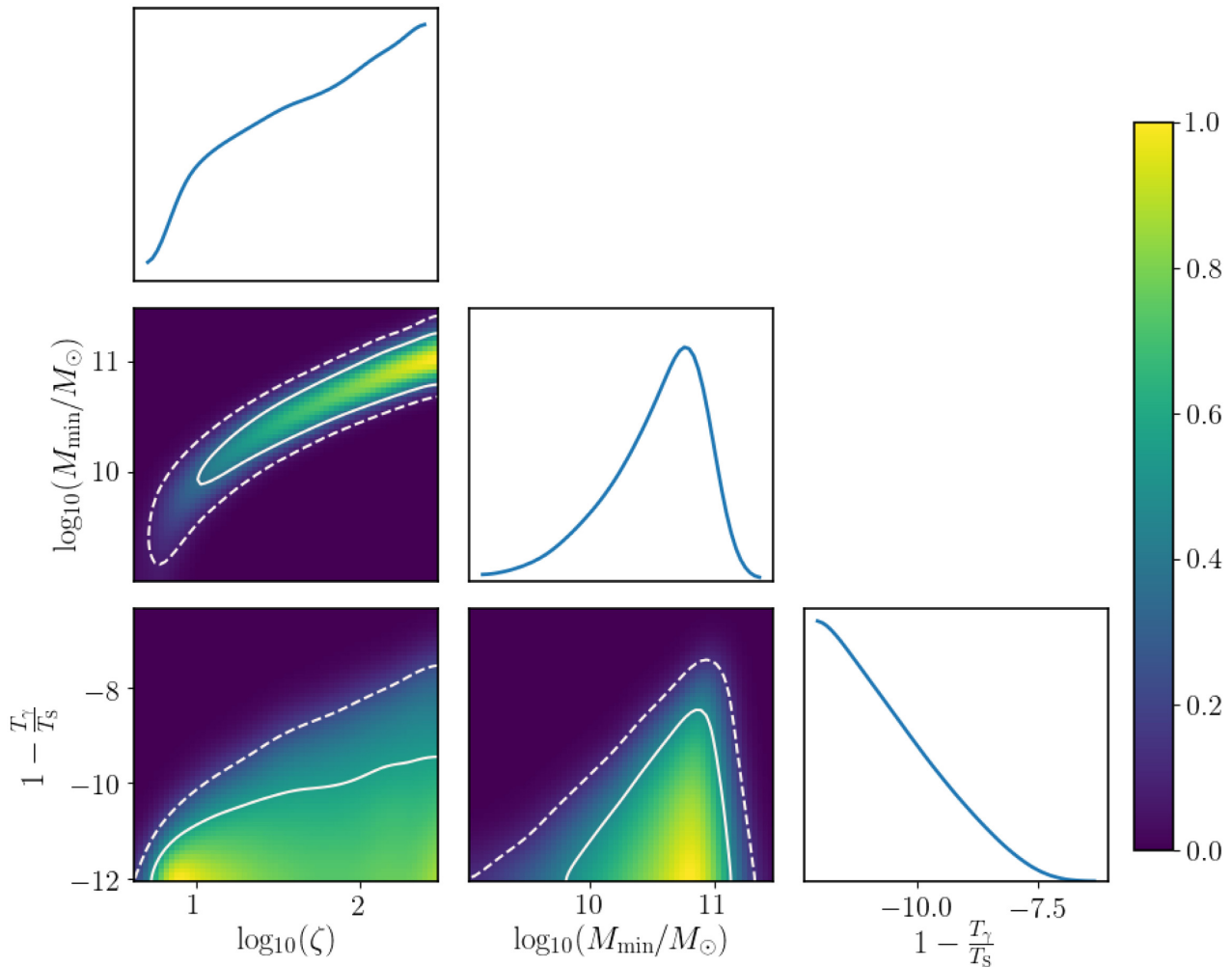


Figure 6. Constraints on the three GRIZZLY parameters of the uniform T_S scenario (see Section 3.1) from the MCMC analysis using the LOFAR upper limit for $z \approx 9.1$. The colour bar shows the probability that models are ruled out. The solid and dashed curves show the 68 and 95 per cent credible intervals of the ruled out models. The diagonal panels show the marginalized probability distribution by which each parameter value as used in the MCMC analysis is ruled out.

building a strong Ly α background. We thus assume that the Ly α coupling is saturated (in other word, $T_S = T_K$) in this scenario. The value of the X-ray spectral index α is uncertain and dependent on the properties of the X-ray sources. For X-ray sources such as quasars and mini-quasars, the spectrum can be very steep, with $\alpha \gtrsim 1$ (Vignali, Brandt & Schneider 2003; Gallerani et al. 2017; Martocchia et al. 2017), while for high-mass X-ray binaries, the observed spectral index can be as small as $\alpha \approx 0.2$ (Mineo et al. 2012; Islam et al. 2019). In this study, we assume $\alpha = 1.2$. Below we discuss the effect of different α on our results.

The remaining three parameters constitute our parameter space. For ζ we keep the same range used in the previous scenario, while we vary $M_{\min, X}$ between 10^9 and $10^{12} M_\odot$, and f_X between 0.1 and 10. As we will see below, this choice covers the regime that is interesting from the point of view of the current LOFAR upper limits. As the run time of the simulations with spin temperature fluctuations is much longer than in the previous scenario, we initially cover the parameter space with a coarser grid. We then visually identify the part of the parameter space that provides a large amplitude of the large-scale power spectrum and fine sample only that region to increase the accuracy of the emulator. We thus end up using only 1495 power

spectra generated using GRIZZLY to train our GPR emulator for this scenario.

In Fig. 8, we show a slice from the brightness temperature map corresponding to the scenario with $\zeta = 0.1$, $M_{\min, X} = 3 \times 10^{11} M_\odot$, and $f_X = 2$. The average ionized fraction remains ≈ 0.01 due to the small value of ζ . The average volume fraction of heated regions of this map is also small (≈ 0.1) as $M_{\min, X}$ is large, and thus only a few of the massive haloes contribute to the heating. While in the previous scenario the patchiness of the signal was due to the ionized regions only, now it is also sourced by the heated regions around the sources.

The thick dashed curve in the right-hand panel of Fig. 8 refers to the power spectrum of the δT_b map shown in the left-hand panel, together with the 1494 other power spectra used to build the three-parameter emulator of Δ^2 for this scenario. Similar to the previous case, we find that the large-scale power spectra of some of the extreme models are larger than the LOFAR upper limits, which are shown by the red data points and their limits in the right-hand panel of the figure.

3.2.1 GRIZZLY and IGM parameters

Fig. 9 shows the power spectrum at scale $k = 0.075 h \text{ Mpc}^{-1}$ in the 2D parameter space of f_X and $M_{\min, X}$. Note that unlike Fig. 3,

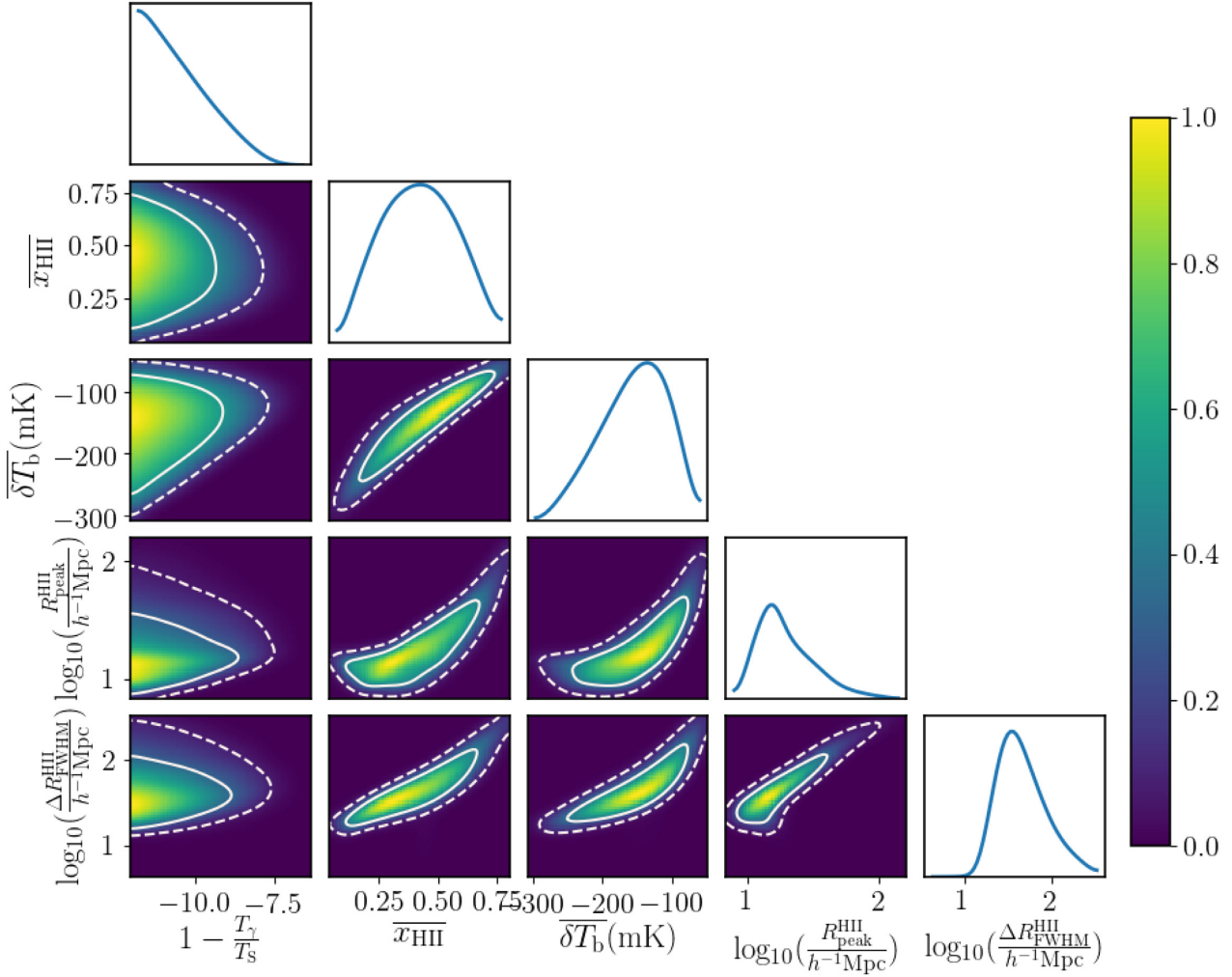


Figure 7. Similar to Fig. 6, but this shows constraints on the IGM parameters at $z \approx 9.1$ in the uniform T_S scenario. The colour bar shows the probability that models are ruled out. The solid and dashed curves correspond to the 68 and 95 per cent credible intervals of the ruled out models. The marginalized probability distributions of the IGM parameters are shown in the diagonal panels.

Table 4. Constraints from the MCMC analysis on the IGM parameters of the uniform T_S scenario at $z \approx 9.1$. Note that our analysis excludes the parameter space that satisfies all the conditions given in this table.

IGM parameters of uniform T_S scenario	Prior	68 per cent credible interval of the excluded models	95 per cent credible interval of the excluded models
\bar{x}_{HII}	Flat in [0, 0.81]	[0.24, 0.60]	[0.13, 0.74]
$1 - T_\gamma/T_S$	Flat in [-12, 1]	$[-\infty, -10.21]$	$[-\infty, -8.50]$
T_S (K)	Flat in [2.1, ∞]	[0, 2.435]	[0, 2.874]
δT_b (mK)	-	[-189.31, -87.65]	[-251.23, -56, 75]
$R_{\text{peak}}^{\text{HII}}$ ($h^{-1}\text{Mpc}$)	-	[9.89, 24.55]	[7.55, 58.07]
$\Delta R_{\text{FWHM}}^{\text{HII}}$ ($h^{-1}\text{Mpc}$)	-	[21.88, 70.79]	[16.37, 184.93]

where the power spectra were derived from the GRIZZLY simulations, here they are evaluated directly with the emulator. In this plot we fix $\zeta = 0.1$, which ensures a small average ionized fraction at $z = 9.1$ ($\bar{x}_{\text{HII}} = 0.01$). Clearly, the power spectrum remains the lowest and constant for a combination of a high value of f_X and a low value of $M_{\text{min}, X}$. In this case, heating of the partially ionized gas in the IGM due to X-rays becomes very efficient, raising the gas temperature above the CMB ($T_K \gg T_\gamma$) and rendering δT_b independent of the spin temperature. On the other hand, the heating of the gas in the IGM

remains inefficient for a combination of small f_X and high $M_{\text{min}, X}$. As expected, the power spectrum for such models (top left corner of the Figure) is larger than the power spectrum for the heated IGM (bottom right corner of the Figure).

One can see that the spin temperature fluctuations are efficient around the diagonal of the parameter space, starting from small values of f_X and $M_{\text{min}, X}$. Specifically, the combination of high $M_{\text{min}, X}$ and f_X enhances the large-scale power spectrum. In this combination, the heated/emission regions around the rare X-ray emitting sources

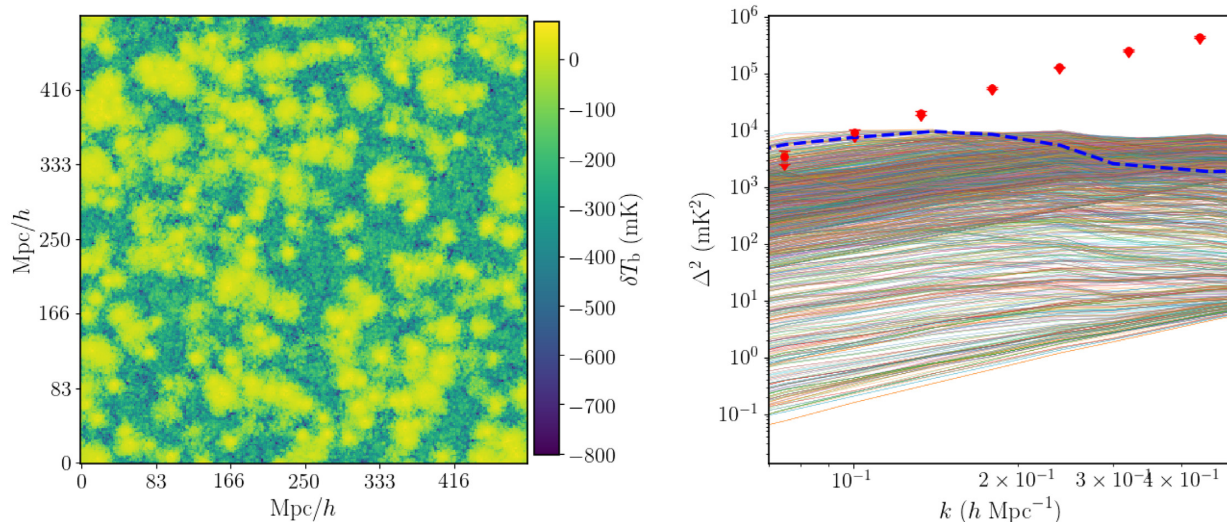


Figure 8. *Left-hand panel:* a slice through the brightness temperature cube for the parameter choice $\zeta = 0.1$, $M_{\min} = 10^9 M_{\odot}$, $M_{\min, X} = 3 \times 10^{11} M_{\odot}$, $f_X = 2$, and $\alpha = 1.2$. The average ionized fraction of this map is 0.01, while the average volume fraction of the heated regions is 0.1. *Right-hand panel:* the curves show the power spectra of the 21-cm brightness temperature as a function of scale for 1495 different combinations of parameters. The red points with error bars show the upper limits from the 10-night observations with LOFAR (Mertens et al. 2020). The blue dashed curve represents the power spectrum of the brightness temperature cube from which the slice in the left-hand panel has been extracted.

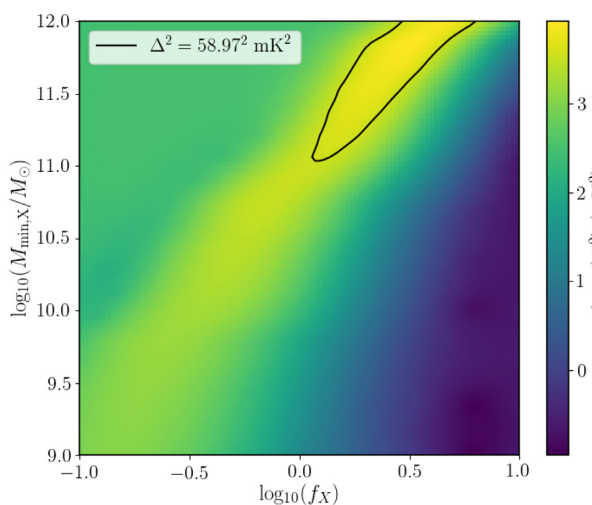


Figure 9. Spherically averaged power spectra of the 21-cm signal from $z \approx 9.1$ at scale $0.075 h \text{ Mpc}^{-1}$ as a function of $M_{\min, X}$ and f_X . These power spectra are generated using an emulator which is trained with 1495 models from GRIZZLY. This plot corresponds to $\zeta = 0.1$, $M_{\min} = 10^9 M_{\odot}$, and $\alpha = 1.2$. The solid contour represents the upper limit constraint from LOFAR at scale $0.075 h \text{ Mpc}^{-1}$, i.e. $\Delta^2 = (58.97)^2 \text{ mK}^2$. For a deterministic observation, the region enclosed by the solid contour will be excluded.

remain isolated in the background absorption signal (see e.g. the left-hand panel of Fig. 8). Also, the partial ionization and heating of the IGM far away from the X-ray emitting sources remain small for a high value of $M_{\min, X}$. We also plot the contour corresponding to the LOFAR current upper limit at scale $k = 0.075 h \text{ Mpc}^{-1}$, i.e. $(58.97)^2 \text{ mK}^2$. Clearly, some parts of the parameter space with the combination of high $M_{\min, X}$ ($\gtrsim 10^{11} M_{\odot}$) and f_X ($\gtrsim 1$) are ruled out with high confidence.

Next, we will consider the global parameters of this scenario. Note that to estimate the IGM parameters, we use an emulator different from the one used for the source parameters. In Fig. 10, we show

the average temperature ($\overline{T_K}$) of regions with ionized fraction smaller than 0.5, the volume fraction of heated regions f_{heat} , and the average brightness temperature $\overline{\delta T_b}$. As expected, $\overline{T_K}$ remains small for a combination of high $M_{\min, X}$ and low f_X , which also keeps f_{heat} low. In this case the average signal remains in absorption, as shown in the right-hand panel of the figure. On the other hand, $\overline{T_K}$ is high for the opposite case of a low $M_{\min, X}$ and a high f_X , for which f_{heat} approaches 1 and $\overline{\delta T_b}$ becomes positive. The parameter space excluded by the LOFAR upper limit at scale $k = 0.075 h \text{ Mpc}^{-1}$ is shown by the solid curves in all panels. It corresponds to $10 \text{ K} \lesssim \overline{T_K} \lesssim 100 \text{ K}$, $f_{\text{heat}} \lesssim 0.3$, and $-200 \text{ mK} \lesssim \overline{\delta T_b} \lesssim -100 \text{ mK}$.

In this scenario, the size distribution of the heated regions is more relevant than the size distribution of the ionized regions. Similarly to the size distribution of the ionized regions considered in the previous scenario, here we analyse the size distribution of the heated regions, characterizing the PDF with two parameters, namely $R_{\text{peak}}^{\text{heat}}$ and $\Delta R_{\text{FWHM}}^{\text{heat}}$, which represent the size of the heated regions at which the PDF becomes maximum and FWHM of the PDF, respectively. Fig. 11 shows the distribution of $R_{\text{peak}}^{\text{heat}}$ and $\Delta R_{\text{FWHM}}^{\text{heat}}$, suggesting that the characteristic size of the heated regions incSingh & Subrahmanyan 2reases with increasing f_X . The white regions represent an IGM fully heated above the CMB temperature. The parameter space in the range $R_{\text{peak}}^{\text{heat}} \approx 5 - 20 h^{-1} \text{ Mpc}$ and $\Delta R_{\text{FWHM}}^{\text{heat}} \approx 10 - 30 h^{-1} \text{ Mpc}$ is the one excluded by the LOFAR upper limit at $z = 9.1$. Note that the excluded parameter space requires satisfying all of the above-mentioned conditions.

3.2.2 MCMC results

Next, we explore the three-dimensional parameter space, i.e. ζ , $M_{\min, X}$ and f_X , using MCMC to find models that are ruled out by the current LOFAR upper limit. Similar to our previous scenario, we have used 20 walkers and 10^6 steps for the MCMC analysis and checked the convergence of the chains. Note that we use a flat prior on $\overline{x_{\text{HI}}}(z = 9.1) \leq 0.81$. The outcome of the analysis is summarized in

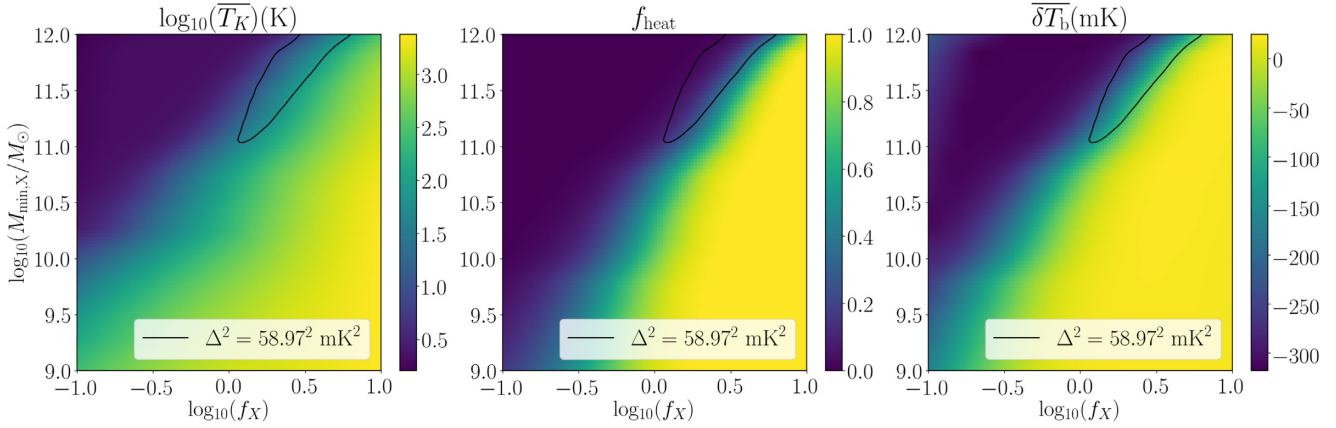


Figure 10. Average gas temperature for regions with ionized fraction less than 0.5 (left-hand panel), volume fraction of the heated regions (middle panel), and average brightness temperature at redshift ≈ 9 as a function of $M_{\text{min}, X}$ and f_X . Here $\zeta = 0.1$. The solid curve represents the contour corresponding to $\Delta^2 = (58.97)^2 \text{ mK}^2$ at scale $0.075 h \text{ Mpc}^{-1}$ which is the LOFAR upper limit on the spherically averaged power spectrum. For a deterministic observation, the region enclosed by the solid contours will be excluded.

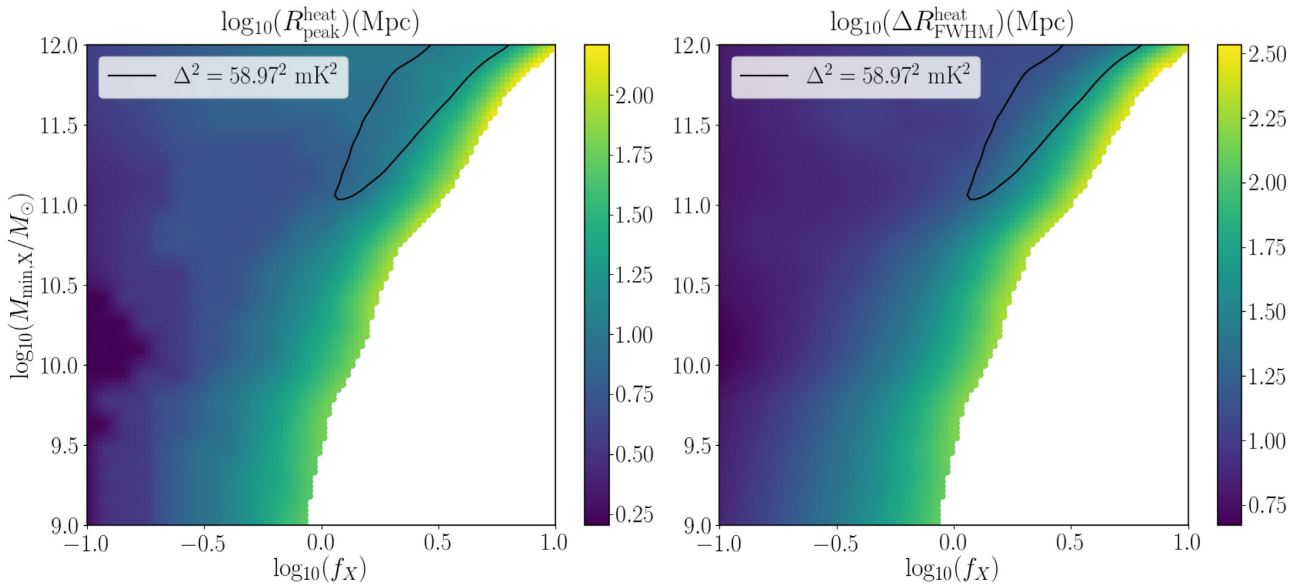


Figure 11. Distribution of $R_{\text{peak}}^{\text{heat}}$ (left-hand panel) and $\Delta R_{\text{FWHM}}^{\text{heat}}$ (right-hand panel) at $z \approx 9.1$ as a function of the parameters f_X and $M_{\text{min}, X}$ for $\zeta = 0.1$. $R_{\text{peak}}^{\text{heat}}$ and $\Delta R_{\text{FWHM}}^{\text{heat}}$ represent the radius at which the probability distribution of the sizes of the heated regions (i.e. regions with $T_K > T_\gamma$) has maximum amplitude and the FWHM of that distribution, respectively. We use the MFP method to estimate the size distribution of the heated regions. The region in white in the bottom right of the panels corresponds to IGM fully heated above the CMB temperature. The solid curve represents the contour corresponding to $\Delta^2 = (58.97)^2 \text{ mK}^2$ which is the LOFAR upper limit on the power spectrum at scale $0.075 h \text{ Mpc}^{-1}$. For a deterministic observation, the region enclosed by the solid contours will be excluded.

Fig. 12. Clearly, a high emissivity of X-ray photons ($f_X \gtrsim 0.3$) with a large $M_{\text{min}, X}$ ($\gtrsim 10^{10} M_\odot$) is the most likely to be excluded within the 68 per cent credible intervals by LOFAR alone. This combination of parameter values results in large heated regions around rare massive haloes embedded in a cold IGM. On the other hand, the combination of large f_X and a small $M_{\text{min}, X}$ causes more uniform heating and thus it reaches more easily the $T_S \gg T_\gamma$ condition where the power spectrum remains lower than the measured one. Similarly, a very small value of f_X yields almost no heating and coincides with the scenario discussed in the previous section. In such models, a larger value of ζ is more likely to be ruled out as we have also seen in the previous scenario. Therefore, we see a second ruled out region in the parameter space shown in Fig. 12.

Next, we will constrain the IGM parameters of this non-uniform T_S scenario, and show the posterior distribution of the IGM parameters in Fig. 13. These results are also listed in Table 5. Clearly, two regimes of the parameter space are likely to be excluded. The first one has large H II regions in a poorly heated IGM, which is the configuration already discussed in the previous section. In this case, least likely values of the IGM parameters are: $0.5 \lesssim \bar{x}_{\text{H II}} \lesssim 0.6$, $\overline{T_K} \lesssim 3.55 \text{ K}$ with $f_{\text{heat}} \approx 0$. The second part of the parameter space which is likely to be excluded corresponds to large heated regions with: $\bar{x}_{\text{H II}} \lesssim 0.08$, $7 \text{ K} \lesssim \overline{T_K} \lesssim 160 \text{ K}$, $-234 \lesssim \overline{\delta T_b} \lesssim -65 \text{ mK}$, $f_{\text{heat}} \lesssim 0.34$, $3.5 h^{-1} \text{ Mpc} \lesssim R_{\text{peak}}^{\text{heat}} \lesssim 70 h^{-1} \text{ Mpc}$, and $\Delta R_{\text{FWHM}}^{\text{heat}} \lesssim 110 h^{-1} \text{ Mpc}$. These limits correspond to 95 per cent credible intervals as shown in Fig. 13.

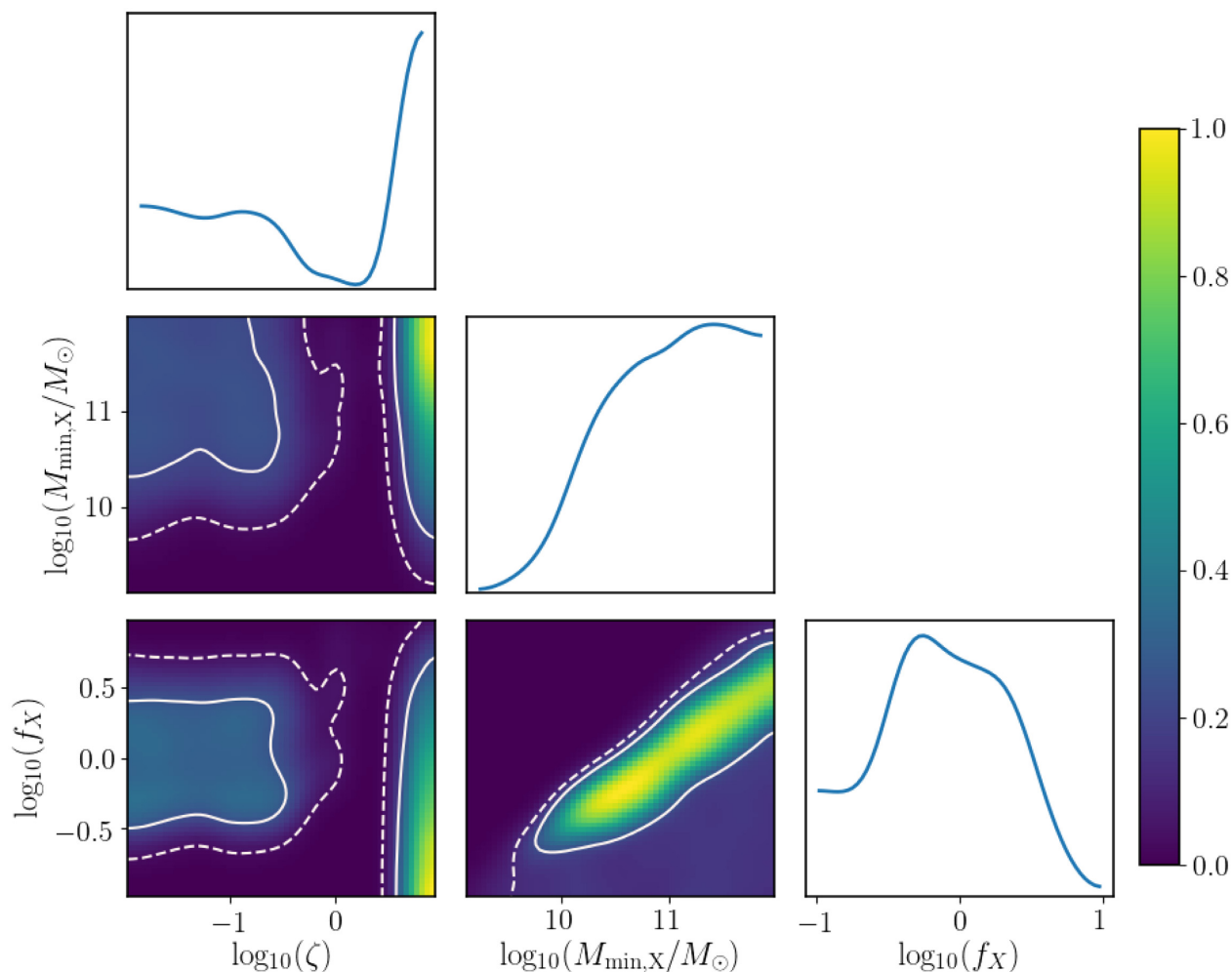


Figure 12. Constraints on the three parameters of our second scenario with non-uniform T_S fluctuation models presented in Section 3.2 from the MCMC analysis using the LOFAR upper limit at $z \approx 9.1$. The colour bar shows the probability that models are ruled out. The solid and dashed curves show the 68 and 95 per cent credible intervals of the ruled out models. The diagonal panels show the marginalized probability distribution for the parameters used in the MCMC analysis in this scenario.

Up to this point we have only considered $\alpha = 1.2$. A less steep SED with a smaller value of α contains a smaller number of soft X-ray photons and a larger number of hard X-ray photons. Thus, the heating due to an X-ray spectrum with smaller α is less patchy than that from a steeper spectrum (see e.g. Pacucci et al. 2014; Das et al. 2017; Islam et al. 2019), resulting in a smaller amplitude of the large-scale power spectrum of the signal. We have verified that for $\alpha = 0.3$ the results are similar to those obtained with $\alpha = 1.2$, except that the contour of the excluded region (see Fig. 9) shrinks towards higher $M_{\min,X}$ values and it shifts slightly towards higher values of f_X .

4 DISCUSSION

We have considered two extreme scenarios, one in which fluctuations at large scales are driven by large ionized regions in a uniform spin temperature IGM, and the other in which they are driven by large heated regions in a non-uniform spin temperature IGM. One question that naturally arises is whether there exist other models capable of exceeding the LOFAR upper limits which are not covered by the two scenarios we have explored. As fluctuations in the 21-cm signal are induced by ionization and/or spin temperature fluctuations, it seems

hard to come up with alternative scenarios which can be excluded without invoking non-standard physics.

A second question is whether the extreme cases considered are in any way realistic or whether they are already excluded by other observations. We have limited ourselves to deriving constraints from the LOFAR upper limits at $z = 9.1$ and have not added information from other redshifts, apart from a very conservative upper limit on the ionized fraction based on the Thomson scattering optical depth derived from the Planck results. This has been a conscious choice as using data from multiple redshifts requires assumptions about the evolution of the source properties which, given the small constraining power of the LOFAR upper limits, does not seem justified. However, it is still possible to apply a minimal check on the models that we find to be excluded by the LOFAR upper limits.

We first consider the scenario in which the excluded models require a fairly large value for \bar{x}_{HII} . The results from Mitra, Choudhury & Ferrara (2015) show that the combined constraints from Planck and $z > 6$ quasar spectra imply that $\bar{x}_{\text{HII}} \lesssim 0.4$ at $z = 9.1$. Although this limits the constraining power of the LOFAR upper limits, the latter is still unique in excluding some models, as we have found cases with $\bar{x}_{\text{HII}} \approx 0.3$ and $1 - T_\gamma/T_S = -12$ which violate them (see Fig. 4).

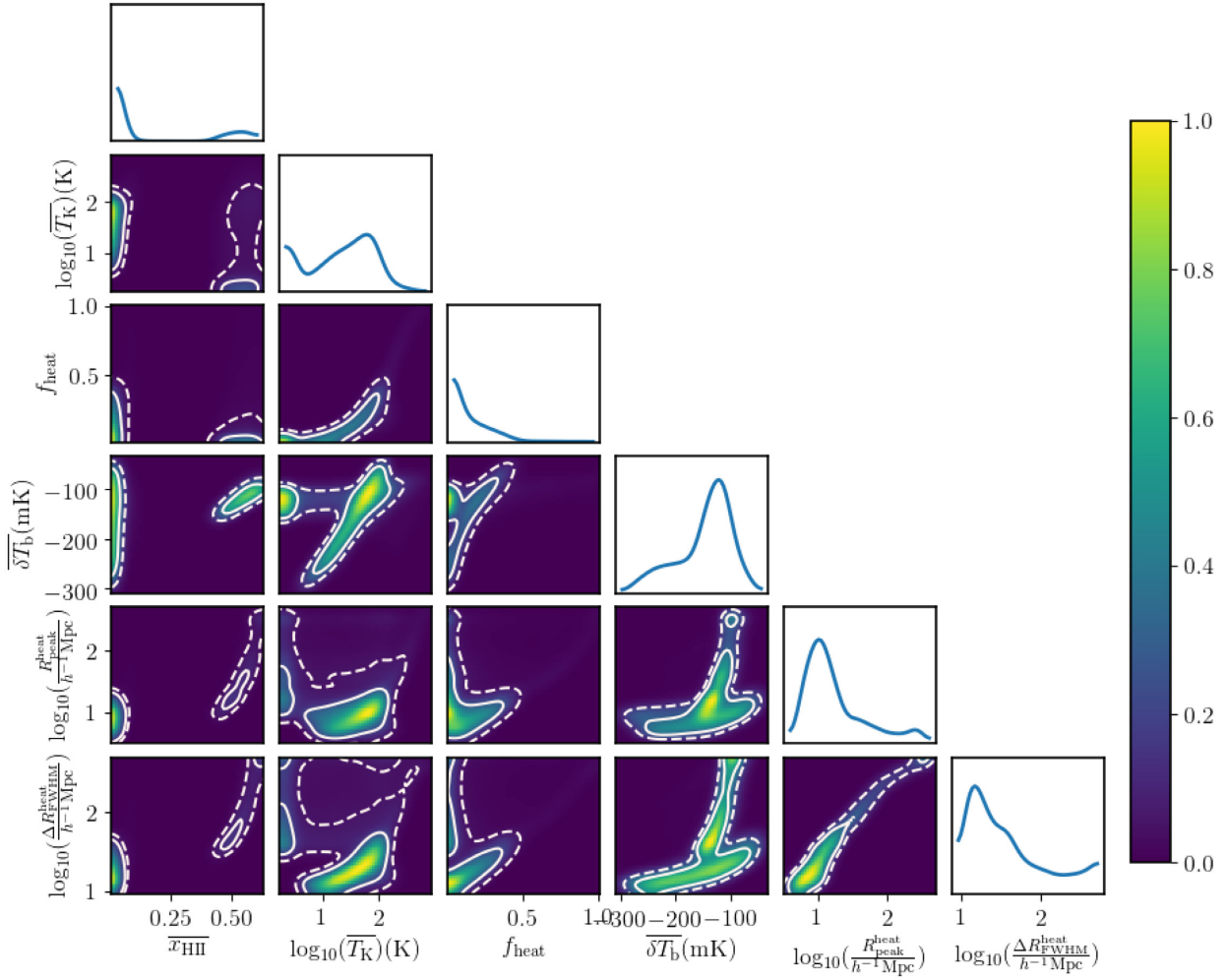


Figure 13. Constraints on the IGM parameters of the non-uniform T_S model presented in Section 3.2 from the MCMC analysis using the LOFAR upper limit at $z \approx 9.1$. The colour bar shows the probability that models are ruled out. The solid and dashed curves show the 68 and 95 per cent credible intervals of the ruled out models. The diagonal panels show the marginalized probability distribution for each of the IGM parameters considered in this scenario.

Table 5. Constraints from the MCMC analysis on the IGM parameters of the non-uniform T_S scenario at $z \approx 9.1$. Note that our analysis excludes the parameter space that satisfies all the conditions given in this table.

IGM parameters of non-uniform T_S scenario	Prior	68 per cent credible interval of the excluded models	95 per cent credible interval of the excluded models
$\overline{x_{\text{HI}}}$	Flat in [0, 0.81]	[0, 0.06], [0.50, 0.58]	[0, 0.08], [0.45, 0.62]
$\overline{T_K}$ (K)	Flat in [2.10, ∞)	[19.23, 115.61], [2.10, 2.32]	[7.41, 158.48], [2.10, 3.55]
f_{heat}	–	[0, 0.14]	[0, 0.34]
$\overline{\delta T_b}$ (mK)	–	[−154.50, −84.26]	[−234.15, −65.53]
$R_{\text{peak}}^{\text{heat}}$ (h^{-1} Mpc)	–	[5.32, 17.78]	[3.50, 69.82]
$\Delta R_{\text{FWHM}}^{\text{heat}}$ (h^{-1} Mpc)	–	[10.47, 38.01]	[0, 113.76]

Monsalve et al. (2017) presented phenomenological constraints on the evolution of the global 21-cm signal derived from EDGES high-band observations. These constraints are mostly about changes in the signal and are therefore very different from the single $z = 9.1$ upper limits used for our results. However, our approach does produce values for the global signal (see the right-hand panels in Figs 4 and 10), with excluded models lying in the range $-250 \text{ K} \lesssim \overline{\delta T_b} \lesssim -55 \text{ mK}$. These can be compared to the values in fig. 9 of

Monsalve et al. (2017), where the authors show that for a minimum value of -200 mK , the Δz for the FWHM of the entire absorption feature has to be above ≈ 5 . They also show that this lower limit is inconsistent with an end of reionization at $z \approx 6$. At face value this implies that the models excluded by the new LOFAR upper limits on the 21-cm power spectrum are also excluded by the EDGES high-band constraints on the evolution of the global signal. However, it should be kept in mind that the EDGES constraints are based

on an assumed Gaussian profile for the evolution of the global signal. Furthermore, the systematics for the EDGES results are not fully known (e.g. Hills et al. 2018; Singh & Subrahmanyan 2019).

This comparison to previous results shows that the new LOFAR upper limits exclude rather extreme models which were already unlikely in view of other observational constraints. However, it is important to point out that the LOFAR observations are of a very different character and thus contribute a new and independently obtained piece of the reionization puzzle. As we obtain more stringent upper limits and additional redshift points, the constraints will improve and start to rule out increasingly large regions of the parameter space.

5 SUMMARY AND CONCLUSIONS

In this paper, we have used the new LOFAR upper limit on the dimensionless spherically averaged power spectrum of the 21-cm signal from redshift ≈ 9.1 (Mertens et al. 2020) and investigated which reionization scenarios can be ruled out by it. The upper limits as obtained from 10 nights of observations are $(58.97)^2$ and $(95.21)^2 \text{ mK}^2$ at scales $k = 0.075$ and 0.1 h Mpc^{-1} , respectively. As these numbers are much larger than the amplitude of the power spectrum expected for standard reionization histories, we mainly focused on the extreme models that produce such high values for the large-scale power spectrum. However, our study also covers the usual range of the parameter space.

With the code GRIZZLY, we generated power spectra for thousands of models for different combinations of parameters namely, ionization efficiency (ζ), minimum mass of the UV emitting haloes (M_{\min}), minimum mass of X-ray emitting haloes ($M_{\min, X}$), and X-ray heating efficiency (f_X). On the basis of these results, we build emulators for different scenarios based on GPR that map source parameters to power spectra. These emulators combined with an MCMC framework are then used to constrain the source parameters at $z \approx 9.1$ using the observed upper limits. We also build emulators that map source parameters to IGM parameters, which are used to put constraints on the IGM parameters. We considered two extreme scenarios in which large-scale fluctuations of the signal are driven by (i) ionized regions embedded in an IGM with a uniform spin temperature, and (ii) spin temperature fluctuations.

As the 21-cm observations themselves characterize the state of the IGM, a major focus of this study is to constrain the thermal and ionization state of the IGM at $z \approx 9.1$ using these upper limits. We study the state of the IGM in terms of parameters such as the average ionization fraction ($\overline{x_{\text{HII}}}$), average gas temperature of the partially ionized IGM ($\overline{T_{\text{K}}}$), $(1 - T_{\gamma}/T_{\text{S}})$, mass-averaged brightness temperature ($\overline{\delta T_{\text{b}}}$), volume fraction of the heated region (f_{heat}), size of the H II (heated) regions at which the PDF of the sizes peaks $R_{\text{peak}}^{\text{HII}} (R_{\text{peak}}^{\text{heat}})$, and the FWHM of the PDFs $\Delta R_{\text{FWHM}}^{\text{HII}} (\Delta R_{\text{FWHM}}^{\text{heat}})$. The results of our study can be summarized as follows.

(i) In the uniform T_{S} scenario, we found that the models which can be ruled out by the upper limit have a high-UV photon emission rate. More specifically, the model with the coldest possible IGM, i.e. $T_{\text{S}} \simeq 2.1 \text{ K}$, requires an emission rate $\gtrsim 2.85 \times 10^{46} \text{ s}^{-1} M_{\odot}^{-1}$, which is 10 times larger than that predicted by population synthesis codes. At the same time, those models require a suppression of ionizing photons from haloes with mass $\lesssim 10^{9.8} M_{\odot}$.

(ii) A high emissivity of the UV photons renders the gas in the IGM largely ionized at the target redshift, so that ionized fractions $\overline{x_{\text{HII}}} \gtrsim 0.13$ are excluded within a 95 percent credible interval. At the same time, the H II bubbles required have to be few in number and

large in size. The characteristic size of the H II bubbles needs to be, $R_{\text{peak}}^{\text{HII}} \gtrsim 8 \text{ h}^{-1} \text{ Mpc}$, with an FWHM of the probability distribution of the size distribution larger than $16 \text{ h}^{-1} \text{ Mpc}$. This keeps the average brightness temperature of the excluded models $\gtrsim -250 \text{ mK}$. The size of the parameter space which can be excluded depends crucially on the value of T_{S} , as it decreases with increasing T_{S} and no constraints can be set for $T_{\text{S}} \gtrsim 3 \text{ K}$.

(iii) For the scenario where the large-scale fluctuations of the signal are driven by spin temperature fluctuations, we found that the models ruled out are those in which regions with temperature larger than CMB cover a volume fraction $\lesssim 0.34$ and at the same time are large with a characteristic size in the range $3.5 - 70 \text{ h}^{-1} \text{ Mpc}$ and a size distribution with an FWHM of $\lesssim 110 \text{ h}^{-1} \text{ Mpc}$. The average gas temperature of the partially ionized regions for these excluded models is $7-160 \text{ K}$, while the average brightness temperature lies in between -234 and -65 mK . The heated regions required for these excluded models are large in size and few in number at the same time. This implies that scenarios in which the heating is driven by fewer X-ray emitting sources hosted by the rare massive haloes ($M_{\min, X} \gtrsim 10^{10} M_{\odot}$) with a high emissivity of X-ray photons (X-ray luminosity $\gtrsim 10^{34} \text{ erg s}^{-1} M_{\odot}^{-1}$) are more likely to be ruled out by the current upper limit.

As the current upper limits on the 21-cm power spectrum are rather large and restricted to one redshift, the constraints on the IGM and source parameters that can be obtained are not yet very strong. However, they do illustrate the potential of this type of observations to characterize the state of the IGM and from this the properties of early sources in a redshift range which has not been yet well explored. We expect LOFAR to produce more stringent upper limits on the power spectrum both through analysing more of the available data (also at other redshifts) and improving the methods to deal with systematic effects. Combining these with other observables, such as the global 21-cm signal and observations of high- z galaxies using the present and next generation of ground-based and space telescopes such as the *James Webb Space Telescope*, the European Extremely Large Telescope, and the Atacama Large Millimetre Array, will give us a much deeper understanding of this crucial period in the history of the Universe.

ACKNOWLEDGEMENTS

The authors would like to thank Jens Jasche and Daniel J. Mortlock for useful discussions, and an anonymous referee for insightful comments. We acknowledge that the results in this paper have been achieved using the PRACE Research Infrastructure resources Curie based at the Très Grand Centre de Calcul (TGCC) operated by CEA near Paris, France and Marenostrum based in the Barcelona Supercomputing Center, Spain. Time on these resources was awarded by PRACE under PRACE4LOFAR grants 2012061089 and 2014102339 as well as under the Multiscale Reionization grants 2014102281 and 2015122822. We have used resources provided by the Swedish National Infrastructure for Computing (SNIC) (proposal number SNIC 2018/3-40) at PDC, Royal Institute of Technology, Stockholm. GM, RG, and SKG are thankful for support by Swedish Research Council grant 2016-03581. RG and SZ furthermore acknowledge support by the Israel Science Foundation (grant no. 255/18). VJ acknowledges support by the Croatian Science Foundation for a project IP-2018-01-2889 (LowFreqCRO). FGM and LVEK acknowledge support from a SKA-NL roadmap grant from the Dutch Ministry of OCW. EC acknowledges support from the Royal Society via the Dorothy Hodgkin Fellowship.

REFERENCES

- Ali Z. S. et al., 2015, *ApJ*, 809, 61
- Bañados E. et al., 2018, *Nature*, 553, 473
- Barkana R., 2016, *Phys. Rep.*, 645, 1
- Barkana R., 2018, *Nature*, 555, 71
- Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, *MNRAS*, 461, 3135
- Barry N. et al., 2019, *ApJ*, 884, 1
- Beardsley A. P. et al., 2016, *ApJ*, 833, 102
- Behroozi P. S., Silk J., 2015, *ApJ*, 799, 32
- Bolgar F., Eames E., Hottier C., Semelin B., 2018, *MNRAS*, 478, 5564
- Bonaldi A., Brown M. L., 2015, *MNRAS*, 447, 1973
- Bowman J. D., Rogers A. E. E., 2010, *Nature*, 468, 796
- Bowman J. D. et al., 2013, *Publ. Astron. Soc. Aust.*, 30, e031
- Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, *Nature*, 555, 67
- Bradley R. F., Tauscher K., Rapetti D., Burns J. O., 2019, *ApJ*, 874, 153
- Cawley G. C., Talbot N. L., 2010, *J. Mach. Learn. Res.*, 11, 2079
- Chapman E. et al., 2013, *MNRAS*, 429, 165
- Chapman E., Zaroubi S., Abdalla F. B., Dulwich F., Jelić V., Mort B., 2016, *MNRAS*, 458, 2928
- Cheng C. et al., 2018, *ApJ*, 868, 26
- Choudhury T. R., Ferrara A., 2006, *MNRAS*, 371, L55
- Ciardi B., Ferrara A., 2005, *Space Sci. Rev.*, 116, 625
- Cohen A., Fialkov A., Barkana R., Monsalve R., 2019, preprint ([arXiv:1910.06274](https://arxiv.org/abs/1910.06274))
- Das A., Mesinger A., Pallottini A., Ferrara A., Wise J. H., 2017, *MNRAS*, 469, 1166
- Datta K. K., Bharadwaj S., Choudhury T. R., 2007, *MNRAS*, 382, 809
- Datta A., Bowman J. D., Carilli C. L., 2010, *ApJ*, 724, 526
- Dawoodbhoy T. et al., 2018, *MNRAS*, 480, 1740
- DeBoer D. R. et al., 2017, *PASP*, 129, 045001
- Dixon K. L., Iliev I. T., Mellema G., Ahn K., Shapiro P. R., 2016, *MNRAS*, 456, 3011
- Draine B. T., Miralda-Escudé J., 2018, *ApJ*, 858, L10
- Eastwood M. W. et al., 2019, *AJ*, 158, 84
- Fan X. et al., 2006, *AJ*, 132, 117
- Feng C., Holder G., 2018, *ApJ*, 858, L17
- Fialkov A., Barkana R., 2019, *MNRAS*, 486, 1763
- Fialkov A., Barkana R., Visbal E., 2014, *Nature*, 506, 197
- Fialkov A., Barkana R., Cohen A., 2018, *Phys. Rev. Lett.*, 121, 011101
- Fioc M., Rocca-Volmerange B., 1997, *A&A*, 326, 950
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Friedrich M. M., Mellema G., Alvarez M. A., Shapiro P. R., Iliev I. T., 2011, *MNRAS*, 413, 1353
- Franklin J., 2005, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Vol. 27, *The Mathematical Intelligencer*, p. 83
- Furlanetto S. R., McQuinn M., Hernquist L., 2006a, *MNRAS*, 365, 115
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006b, *Phys. Rep.*, 433, 181
- Gallerani S. et al., 2017, *MNRAS*, 467, 3590
- Gehlot B. K. et al., 2019, *MNRAS*, 488, 4271
- Ghara R., Choudhury T. R., Datta K. K., 2015a, *MNRAS*, 447, 1806
- Ghara R., Datta K. K., Choudhury T. R., 2015b, *MNRAS*, 453, 3143
- Ghara R., Choudhury T. R., Datta K. K., 2016, *MNRAS*, 460, 827
- Ghara R., Choudhury T. R., Datta K. K., Choudhuri S., 2017, *MNRAS*, 464, 2234
- Ghara R., Mellema G., Giri S. K., Choudhury T. R., Datta K. K., Majumdar S., 2018, *MNRAS*, 476, 1741
- Giri S. K., Mellema G., Dixon K. L., Iliev I. T., 2018a, *MNRAS*, 473, 2949
- Giri S. K., Mellema G., Ghara R., 2018b, *MNRAS*, 479, 5596
- Giri S. K., Mellema G., Aldheimer T., Dixon K. L., Iliev I. T., 2019a, *MNRAS*, 489, 1590
- Giri S. K., D'Aloisio A., Mellema G., Komatsu E., Ghara R., Majumdar S., 2019b, *J. Cosmol. Astropart. Phys.*, 2019, 058
- Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, *ApJ*, 672, 765
- Goodman J., Weare J., 2010, *Commun. Appl. Math. Comput. Sci.*, 5, 65
- Greig B., Mesinger A., 2015, *MNRAS*, 449, 4246
- Greig B., Mesinger A., 2017, *MNRAS*, 472, 2651
- Greig B., Mesinger A., Pober J. C., 2016, *MNRAS*, 455, 4295
- Harker G. et al., 2009, *MNRAS*, 397, 1138
- Harnois-Dérans J., Pen U.-L., Iliev I. T., Merz H., Emberson J. D., Desjacques V., 2013, *MNRAS*, 436, 540
- Hasegawa K., Semelin B., 2013, *MNRAS*, 428, 154
- Hassan S., Davé R., Finlator K., Santos M. G., 2016, *MNRAS*, 457, 1550
- Hills R., Kulkarni G., Meerburg P. D., Puchwein E., 2018, *Nature*, 564, E32
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Hyndman R. J., 1996, *Am. Stat.*, 50, 120
- Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007, *MNRAS*, 376, 534
- Islam N., Ghara R., Paul B., Choudhury T. R., Nath B. B., 2019, *MNRAS*, 487, 2785
- Jelić V. et al., 2008, *MNRAS*, 389, 1319
- Jennings W. D., Watkinson C. A., Abdalla F. B., McEwen J. D., 2019, *MNRAS*, 483, 2907
- Jensen H. et al., 2013, *MNRAS*, 435, 460
- Kern N. S., Liu A., Parsons A. R., Mesinger A., Greig B., 2017, *ApJ*, 848, 23
- Kern N. S., Parsons A. R., Dillon J. S., Lanman A. E., Fagnoni N., de Lera Acedo E., 2019, *ApJ*, 884, 105
- Kohavi R., 1995, in *Proc. 14th International Joint Conference on Artificial Intelligence - Volume 2. IJCAI'95*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, p. 1137
- Kolopanis M. et al., 2019, *ApJ*, 883, 133
- Krause F., Thomas R. M., Zaroubi S., Abdalla F. B., 2018, *New Astron.*, 64, 9
- Lin Y., Oh S. P., Furlanetto S. R., Sutter P. M., 2016, *MNRAS*, 461, 3361
- Liu A., Parsons A. R., Trott C. M., 2014, *Phys. Rev. D*, 90, 023019
- Madau P., Meiksin A., Rees M. J., 1997, *ApJ*, 475, 429
- Majumdar S., Bharadwaj S., Choudhury T. R., 2012, *MNRAS*, 426, 3178
- Majumdar S., Datta K. K., Ghara R., Mondal R., Choudhury T. R., Bharadwaj S., Ali S. S., Datta A., 2016, *JA&A*, 37, 32
- Mao Y., Shapiro P. R., Mellema G., Iliev I. T., Koda J., Ahn K., 2012, *MNRAS*, 422, 926
- Martocchia S. et al., 2017, *A&A*, 608, A51
- Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006a, *New Astron.*, 11, 374
- Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006b, *MNRAS*, 372, 679
- Mellema G., Koopmans L., Shukla H., Datta K. K., Mesinger A., Majumdar S., 2015, *Proc. Sci., HI Tomographic Imaging of the Cosmic Dawn and Epoch of Reionization with SKA, SISSA, Trieste, PoS#37*
- Mertens F. G., Ghosh A., Koopmans L. V. E., 2018, *MNRAS*, 478, 3640
- Mertens F. G. et al., 2020, *MNRAS*, 493, 1685
- Mesinger A., Furlanetto S., 2007, *ApJ*, 669, 663
- Mevius M. et al., 2016, *Radio Sci.*, 51, 927
- Mineo S., Gilfanov M., Sunyaev R., 2012, *MNRAS*, 419, 2095
- Mitra S., Choudhury T. R., Ferrara A., 2011, *MNRAS*, 413, 1569
- Mitra S., Choudhury T. R., Ferrara A., 2012, *MNRAS*, 419, 1480
- Mitra S., Choudhury T. R., Ferrara A., 2015, *MNRAS*, 454, L76
- Mondal R., Bharadwaj S., Majumdar S., Bera A., Acharyya A., 2015, *MNRAS*, 449, L41
- Monsalve R. A., Rogers A. E. E., Bowman J. D., Mozdzen T. J., 2017, *ApJ*, 835, 49
- Monsalve R. A., Fialkov A., Bowman J. D., Rogers A. E. E., Mozdzen T. J., Cohen A., Barkana R., Mahesh N., 2019, *ApJ*, 875, 67
- Morales M. F., Wyithe J. S. B., 2010, *ARA&A*, 48, 127
- Mortlock D. J. et al., 2011, *Nature*, 474, 616
- Muñoz J. B., Loeb A., 2018, preprint ([arXiv:1802.10094](https://arxiv.org/abs/1802.10094))
- Paciga G. et al., 2013, *MNRAS*, 433, 639
- Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, *MNRAS*, 443, 678
- Park J., Mesinger A., Greig B., Gillet N., 2019, *MNRAS*, 484, 933
- Parsons A. R. et al., 2014, *ApJ*, 788, 106
- Patil A. H. et al., 2017, *ApJ*, 838, 65

- Patra N., Subrahmanyan R., Sethi S., Udaya Shankar N., Raghunathan A., 2015, *ApJ*, 801, 138
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Peebles P. J. E., 1993, *Principles of Physical Cosmology*. Princeton Univ. Press, Princeton
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Planck Collaboration VI, 2018, preprint (arXiv:1807.06209)
- Pober J. C. et al., 2015, *ApJ*, 809, 62
- Price D. C. et al., 2018, *MNRAS*, 478, 4193
- Pritchard J. R., Loeb A., 2012, *Rep. Prog. Phys.*, 75, 086901
- Rasmussen C. E., Williams C. K. I., 2006, *Gaussian Processes for Machine Learning*. The MIT Press
- Ross H. E., Dixon K. L., Ghara R., Iliev I. T., Mellema G., 2019, *MNRAS*, 487, 1119
- Sharma M., Theuns T., Frenk C., Bower R., Crain R., Schaller M., Schaye J., 2016, *MNRAS*, 458, L94
- Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, *A&A*, 345, 380
- Singh S., Subrahmanyan R., 2019, *ApJ*, 880, 26
- Singh S. et al., 2017, *ApJ*, 845, L12
- Sokolowski M. et al., 2015, *Publ. Astron. Soc. Aust.*, 32, e004
- Sun G., Furlanetto S. R., 2016, *MNRAS*, 460, 417
- Tashiro H., Kadota K., Silk J., 2014, *Phys. Rev. D*, 90, 083522
- Thomas R. M., Zaroubi S., 2008, *MNRAS*, 384, 1080
- Thomas R. M., Zaroubi S., 2011, *MNRAS*, 410, 1377
- Thomas R. M. et al., 2009, *MNRAS*, 393, 32
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Venemans B. P. et al., 2015, *ApJ*, 801, L11
- Vignali C., Brandt W. N., Schneider D. P., 2003, *AJ*, 125, 433
- Voytek T. C., Natarajan A., Jáuregui García J. M., Peterson J. B., López-Cruz O., 2014, *ApJ*, 782, L9
- Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, *MNRAS*, 433, 1230
- Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, *MNRAS*, 442, 2560
- Zaldarriaga M., Seljak U., 2000, *ApJS*, 129, 431
- Zaroubi S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, *Astrophysics and Space Science Library*, Vol. 396, *The First Galaxies*. Springer-Verlag, Berlin, p. 45
- Zaroubi S., Silk J., 2005, *MNRAS*, 360, L64
- Zaroubi S., Thomas R. M., Sugiyama N., Silk J., 2007, *MNRAS*, 375, 1269

APPENDIX A: GRIZZLY

Here, we briefly describe the one-dimensional radiative transfer method used in the code GRIZZLY to simulate the redshifted 21-cm signal from the EoR. We refer the reader to Ghara et al. (2015a, 2018) for a more detailed description of the method. While the basic approach of GRIZZLY mainly follows the BEARS algorithm (Thomas & Zaroubi 2008; Thomas et al. 2009; Thomas & Zaroubi 2011), it differs slightly from the original method. Both codes avoid solving the one-dimensional radiative transfer equations on the fly and, instead, they use previously generated 1D ionization profiles to simulate an ionization field. Below, we briefly describe the steps used in GRIZZLY:

- (i) First, we generate a large number of 1D profiles of ionized fraction and kinetic temperature for different combinations of source parameter values. The parameters used for this are ionization efficiency, the ratio of X-ray and UV luminosities, X-ray spectral index, overdensity of the uniform background IGM and redshift. In this study, we assume that the age of the source is 10 Myr. For a given cosmology, these profiles need to be generated only once.
- (ii) Next, we determine the size of the H II regions in all the 1D profiles and create a list of their radii for different parameter values.

These are defined as the distance from the centre of the source at which the ionized fraction drops to 0.5.

- (iii) Given a halo with a certain mass and position, we first determine the corresponding UV luminosity. From this, we determine the size of the H II region around that halo using the density field and the list of radii as generated in the previous step. This is done iteratively as follows. We start with a small value of the radius, estimate the spherically averaged overdensity contained within it and look for the same combination of radius and overdensity in the pre-compiled list. If this is not found, we change the initial choice of the radius and continue the iteration until a match is obtained. This step is repeated for all haloes. The corresponding ionization profiles are used to generate the ionization field.

- (iv) When individual H II regions overlap, we estimate the number of photons in excess and distribute them around the surface of the overlapping regions so that all the photons are used for ionization.

- (v) We then generate the kinetic temperature field from the ionization field and a correlation of the ionized fraction and the gas temperature (for details, see Ghara et al. 2015a).

Given the value of the uniform T_S , the δT_b maps can be generated using the ionization maps and density field for our first scenario. For our second scenario, which assumes $T_S = T_K$, we use the ionization, density, and temperature maps to generate the δT_b maps following equation (1).

APPENDIX B: LIKELIHOOD FOR UPPER LIMIT OBSERVATIONS

Using Bayes theorem, we can write the posterior of our model parameters θ for simulating the model power spectrum $\Delta_m^2(\mathbf{k}, \theta)$ given the observed power spectrum $\Delta_o^2(\mathbf{k})$ as follows:

$$p(\theta|\Delta_o^2(\mathbf{k})) \propto p(\Delta_o^2(\mathbf{k})|\theta) p(\theta), \quad (\text{B1})$$

where the first and second term in the right hand side of the equation are the likelihood $\mathcal{L}(\theta|\Delta_o^2(\mathbf{k}))$ and prior, respectively.

If $\Delta_o^2(\mathbf{k})$ is a deterministic, a scenario is ruled out when the modelled power spectrum $\Delta_m^2(k_i)$ is above $\Delta_o^2(k_i)$ in any one wavenumber-bin k_i . We can write $\mathcal{L}(\theta|\Delta_o^2(k_i))$ as a Heaviside function $\mathcal{H}(\Delta_o^2(k_i) - \Delta_m^2(k_i))$. However, the $\Delta_o^2(k_i)$ is probabilistic with mean of $\Delta_{21}^2(k_i)$ and standard deviation of $\Delta_{21,\text{err}}^2(k_i)$. Therefore, we need to draw a $\Delta_a^2(k_i)$ from a normal distribution $\mathcal{N}(\Delta_{21}^2(k_i), \Delta_{21,\text{err}}^2(k_i))$ and calculate the probability of our model. Here, $\Delta_a^2(k_i)$ is a nuisance parameter over which we can marginalize to get the $\mathcal{L}(\theta|\Delta_o^2(k_i))$. Therefore, $\mathcal{L}(\theta|\Delta_o^2(k_i))$ can be written as follows:

$$\mathcal{L}(\theta|\Delta_o^2(k_i)) = \int_{-\infty}^{\infty} p(\Delta_o^2(k_i)|\Delta_a^2) p(\Delta_a^2(k_i)|\theta) d\Delta_a^2. \quad (\text{B2})$$

The value for $p(\Delta_a^2(k_i)|\theta)$ is a Heaviside function $\mathcal{H}(\Delta_a^2(k_i) - \Delta_m^2(k_i))$, while $p(\Delta_o^2(k_i)|\Delta_a^2(k_i))$ is defined by a $\mathcal{N}(\Delta_{21}^2(k_i), \Delta_{21,\text{err}}^2(k_i))$. Putting these functions into equation B2,

we get

$$\begin{aligned}
& \mathcal{L}(\theta | \Delta_0^2(k_i)) \\
&= \frac{1}{\sqrt{2\pi} \Delta_{21,\text{err}}^2} \int_{-\infty}^{\infty} \mathcal{H}(\Delta_a^2(k_i) - \Delta_m^2(k_i)) e^{-\frac{1}{2} \left(\frac{\Delta_a^2(k_i) - \Delta_m^2(k_i)}{\Delta_{21,\text{err}}^2} \right)^2} d\Delta_a^2 \\
&= \frac{1}{\sqrt{2\pi} \Delta_{21,\text{err}}^2} \int_{\Delta_m^2(k_i)}^{\infty} e^{-\frac{1}{2} \left(\frac{\Delta_a^2(k_i) - \Delta_m^2(k_i)}{\Delta_{21,\text{err}}^2} \right)^2} d\Delta_a^2 \\
&= \frac{1}{2} \left[1 + \text{erf} \left(\frac{\Delta_{21}^2(k_i) - \Delta_m^2(k_i)}{\sqrt{2} \Delta_{21,\text{err}}^2(k_i)} \right) \right], \tag{B3}
\end{aligned}$$

where $\text{erf}(x)$ is the error function. The power in various k bins can be correlated due to the non-Gaussian nature of the 21-cm signal (see e.g. Mondal et al. 2015). However, the current observation is noise dominated and it is not sensitive to the non-Gaussianity of the signal. The finite size of LOFAR stations will affect the uv tracks and therefore correlate the data in various k -bins. However, this effect is minor as the widths of the k -bins are large enough to minimize the correlation between the bins. Thus, the likelihood calculated above is mutually exclusive in each k bin. Therefore, the total likelihood is the product of the likelihoods at various k bins where we have observations. The likelihood of a parameter value θ is $(1 - \prod_i \mathcal{L}(\theta | \Delta_0^2(k_i)))$.

APPENDIX C: ROBUSTNESS OF THE CONSTRAINTS ON THE IGM PARAMETERS

We presented the constraints on the IGM parameters as our main results as the source parameters are model dependent. However, one may worry that the derived IGM parameters could somehow depend on the chosen source model. In this appendix, we consider a different source model to show the robustness of the constraints on the IGM parameters. We consider the results for the uniform T_S model as presented in Section 3.1.2. The original source model assumed a linear relation between stellar and halo mass: $M_\star \propto M_{\text{halo}}$, see Section 2.2. Here, we instead use $M_\star \propto M_{\text{halo}}^\beta$ where we choose $\beta = 1.2$, keeping the normalization constant the same as before. This source model implies that higher mass haloes contribute relatively more to ionization than in the original source model.

We follow the same method as described in Section 2.4 to develop separate emulators for this source model using 442 simulations and explore the same parameter space as in Section 3.1.2. We also use the same number of walkers and steps in the MCMC analysis. The constraints on the source parameters and the IGM parameters from the MCMC analysis are shown in Figs C1 and C2, respectively.

The constraints on the source parameters ζ and M_{min} obviously differ from the ones shown in Section 3.1.2. As the star formation rate in the modified source model is higher compared to the original one, the part of parameter space that is ruled out shifts towards lower ζ values. However, the constraints on the IGM parameters as shown in Fig. C2 remain similar to what we previously found (see Fig 7). This shows that the constraints on the IGM parameters as

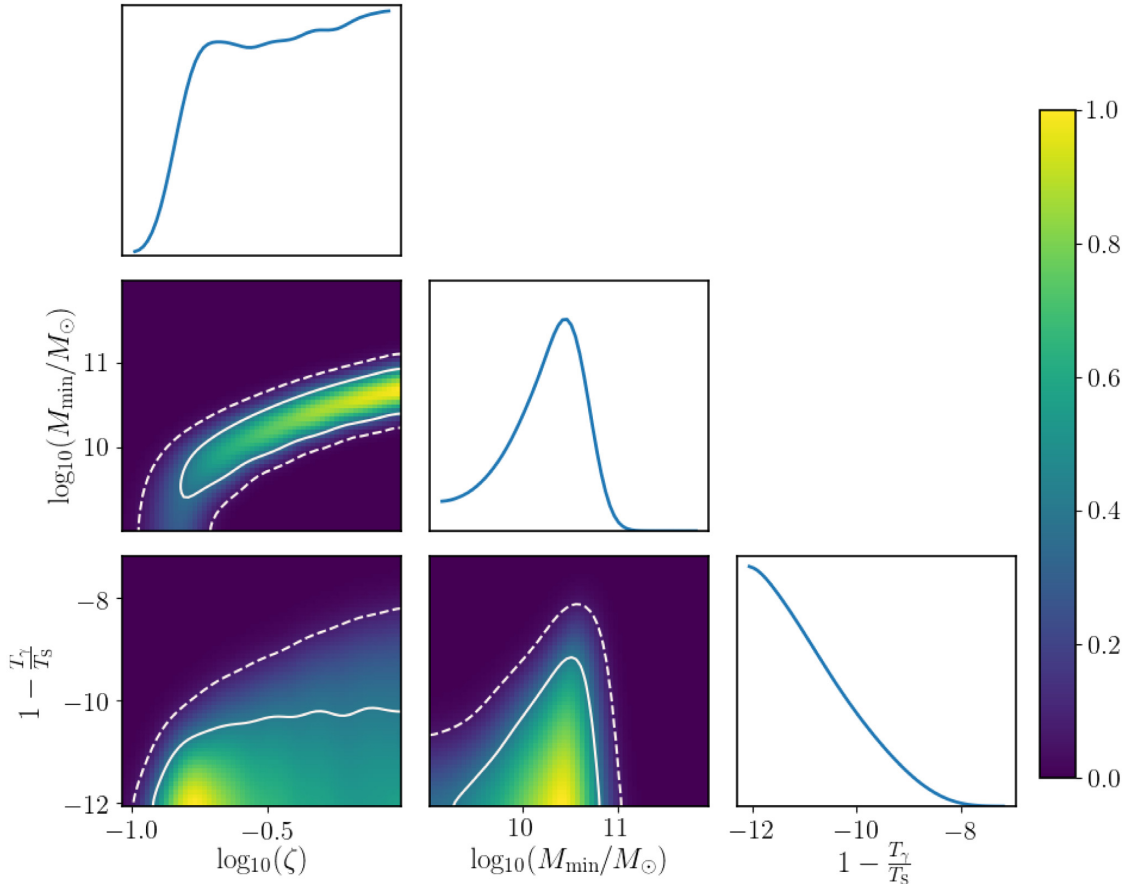


Figure C1. Similar to Fig. 6 but for a different source model. Here, we consider $\dot{N}_i \propto M_{\text{halo}}^\beta$ with $\beta = 1.2$.

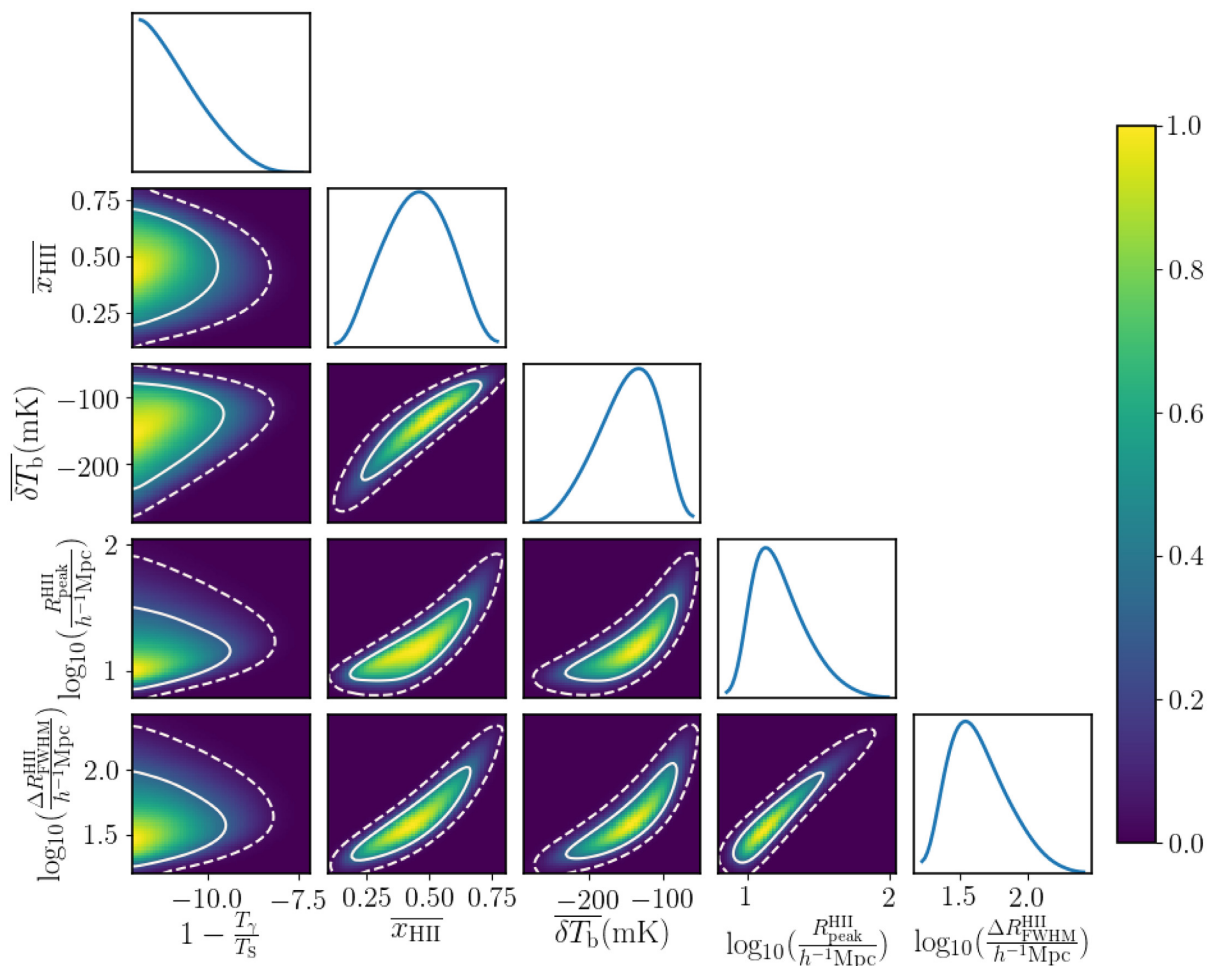


Figure C2. Similar to Fig. 7 but for a different source model. Here, we consider $\dot{N}_i \propto M_{\text{halo}}^\beta$ with $\beta = 1.2$.

obtained from the upper limit observation are indeed independent of the source model as expected as the LOFAR observations do not directly measure any source properties.

¹The Oskar Klein Centre, Department of Astronomy, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden

²Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel

³Department of Physics, Technion, Haifa 32000, Israel

⁴Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

⁵Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Straße 1, D-85748 Garching, Germany

⁶Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700AV Groningen, the Netherlands

⁷Astronomy Centre, Department of Physics and Astronomy, Penvensey II Building, University of Sussex, Brighton BN1 9QH, UK

⁸Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK

⁹School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

¹⁰Department of Physics, University of the Western Cape, Cape Town 7535, South Africa

¹¹SARAO, 2 Fir Street, Black River Park, Observatory, Capetown, South Africa

¹²Department of Physics, Banwarilal Bhalotia College, Asansol, West Bengal, India

¹³Ruder Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

¹⁴LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Université, F-75014 Paris, France

¹⁵Leiden Observatory, Leiden University, PO Box 9513, NL-2300RA Leiden, the Netherlands

¹⁶Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, N-0315 Oslo, Norway

¹⁷Independent University Bangladesh, Plot 16, Block B, Aftabuddin Ahmed Road, Bashundhara R/A, Dhaka, Bangladesh

¹⁸Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

¹⁹ASTRON, PO Box 2, NL-7990 AA Dwingeloo, the Netherlands

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.