Bayesian constraints on the global 21-cm signal from the Cosmic Dawn

G. Bernardi,[1,2,3]★ J. T. L. Zwart,[4,5] D. Price,[3] L. J. Greenhill,[3] A. Mesinger,[6]
J. Dowell,[7] T. Eftekhari,[3] S. W. Ellingson,[8] J. Kocz[9] and F. Schinzel[7]

[1]*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands 7405, South Africa*
[2]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*
[3]*Harvard–Smithsonian Center for Astrophysics, Garden Street 60, Cambridge, MA 02138, USA*
[4]*Department of Physics & Astronomy, University of the Western Cape, Private Bag X17, Bellville, Cape Town 7535, South Africa*
[5]*Astrophysics, Cosmology & Gravity Centre, Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa*
[6]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[7]*Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA*
[8]*Bradley Department of Electrical & Computer Engineering, Virginia Tech, Blacksburg, VA 24061, USA*
[9]*Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91104, USA*

ABSTRACT

The birth of the first luminous sources and the ensuing epoch of reionization are best studied via the redshifted 21-cm emission line, the signature of the first two imprinting the last. In this work, we present a fully Bayesian method, HIBAYES, for extracting the faint, global (sky-averaged) 21-cm signal from the much brighter foreground emission. We show that a simplified (but plausible) Gaussian model of the 21-cm emission from the Cosmic Dawn epoch ($15 \lesssim z \lesssim 30$), parametrized by an amplitude $A_{H I}$, a frequency peak $\nu_{H I}$ and a width $\sigma_{H I}$, can be extracted even in the presence of a structured foreground frequency spectrum (parametrized as a seventh-order polynomial), provided sufficient signal-to-noise (400 h of observation with a single dipole). We apply our method to an early, 19-min-long observation from the Large aperture Experiment to detect the Dark Ages, constraining the 21-cm signal amplitude and width to be $-890 < A_{H I} < 0$ mK and $\sigma_{H I} > 6.5$ MHz (corresponding to $\Delta z > 1.9$ at redshift $z \simeq 20$) respectively at the 95-per cent confidence level in the range $13.2 < z < 27.4$ ($100 > \nu > 50$ MHz).

Key words: methods: data analysis – methods: statistical – cosmology: observations – dark ages, reionization, first stars – diffuse radiation.

1 INTRODUCTION

Observations of the highly redshifted 21-cm emission are considered the most powerful probe of the birth of the first luminous sources and the consequent epoch of reionization (for recent reviews, see McQuinn 2015; Furlanetto 2016). The 21-cm emission precisely traces and times the evolution of the average Hydrogen neutral fraction and the growth of the H II regions around ionizing sources throughout reionization (e.g. Ciardi, Stoehr & White 2003; Mellema et al. 2006; McQuinn et al. 2007; Lidz et al. 2008).

Prior to reionization, during the so-called Cosmic Dawn, the 21-cm signal marks the Lyα coupling and the X-ray heating eras, respectively. The Lyα coupling occurs with the birth of the first luminous sources that are expected to be highly effective in coupling the spin temperature to the intergalactic medium (IGM) temperature, generating 21-cm emission via the Wouthuysen–Field effect

(Wouthuysen 1952; Field 1959). Most models anticipate this transition to occur around $z \simeq 25$–30, when the IGM is colder than the cosmic microwave background (CMB), causing a 21-cm signal in absorption against the CMB. The 21-cm emission is here sensitive to the nature of the first luminous sources as well as the details of the formation of the first galaxies in the first minihaloes (e.g. Ciardi & Madau 2003; Furlanetto 2006; Fialkov et al. 2013). In particular, the redshift and amplitude of the peak in the 21-cm emission strongly depend on whether the dominant contribution to the Lyα coupling comes from atomically cooled galaxies or minihaloes, and how much the Lyman–Werner background suppresses star formation (e.g. Haiman, Abel & Rees 2000; Ricotti, Gnedin & Shull 2001; Fialkov et al. 2013).

As star formation progresses, X-rays are generated in the first galaxies by either early black holes or the diffuse, hot interstellar medium. Although other sources of energy injection due to dark matter annihilation (Valdés et al. 2007, 2013) and shocks from fluid motions may be present in the early Universe (McQuinn & O'Leary 2012), X-ray emission is commonly believed to be the

---

★ E-mail: giannibernardi75@gmail.com

most significant source of IGM heating that would, eventually, drive its temperature above the CMB temperature (Pritchard & Furlanetto 2007; Mesinger, Ferrara & Spiegel 2013; Pacucci et al. 2014; Tanaka, O'Leary & Perna 2016). The relative timing of this process is, however, very uncertain due to the essentially unknown properties of the first galaxies (Mesinger, Greig & Sobacchi 2016). In particular, most models assume that the IGM is heated well above the CMB temperature by the onset of reionization. However, if the first galaxies show a hard X-ray spectrum or their X-ray efficiency (commonly parametrized as the number of X-ray photons produced per stellar baryon) is low, then heating becomes inefficient and reionization begins when the IGM is still colder than the CMB (the 'cold reionization' scenario; Fialkov, Barkana & Visbal 2014; Mesinger, Ewall-Wice & Hewitt 2014). Such a scenario would also impact the subsequent morphology of reionization (Iliev et al. 2012; Ewall-Wice et al. 2016a).

This theoretical landscape is still completely unconstrained by observations, but first upper limits to the 21-cm fluctuations in the $12 < z < 18$ range are starting to appear at approximately three orders of magnitude higher than the expected signal (Ewall-Wice et al. 2016b). There is also initial evidence of heating prior to reionization provided by recent 21-cm power spectrum upper limits at $z = 8.4$ that constrain the IGM to be warmer than 8 K (Ali et al. 2015; Pober et al. 2015; Greig, Mesinger & Pober 2016).

While current experiments targeting 21-cm fluctuations are well placed to constrain the reionization process statistically, only the upcoming interferometric arrays like the Hydrogen Epoch of Reionization Array (Pober et al. 2014; DeBoer et al. 2016) and the Square Kilometre Array (Koopmans et al. 2015) will have sufficient sensitivity and frequency coverage to probe the Lyα and X-ray heating epochs (Mesinger et al. 2015; Ewall-Wice et al. 2016a). Therefore, increased attention has recently been devoted to observations targeting the global (sky-averaged) 21-cm emission (e.g. Pritchard & Loeb 2010; Morandi & Barkana 2012; Liu & Parsons 2016), including novel ways to use interferometric arrays to probe the global 21-cm signal (McKinley et al. 2013; Presley, Liu & Parsons 2015; Singh et al. 2015; Vedantham et al. 2015). Albeit challenged by the same requirements of accurate subtraction of bright foreground emission and control over systematic effects that affect its sibling 21-cm fluctuation observations, the global 21-cm signal may represent an alternative, relatively inexpensive way to achieve the millikelvin sensitivity needed to access the pre-reionization epoch. Instruments like the Experiment to Detect the Global Epoch-of-Reionization Signature (EDGES; Bowman, Rogers & Hewitt 2008; Bowman & Rogers 2010), the Large aperture Experiment to detect the Dark Ages (LEDA; Greenhill & Bernardi 2012; Bernardi et al. 2015; Kocz et al. 2015; Price et al., in preparation), SCI-HI (Voytek et al. 2014) and the *Dark Ages Radio Explorer* (Mirocha, Harker & Burns 2015; Harker et al. 2016) are (or will be) targeting such an epoch.

In this paper, we present a Bayesian foreground separation method and show that it can extract the 21-cm signal from the Cosmic Dawn even in the presence of non-spectrally smooth foreground emission parametrized through high-order polynomials in frequency. We apply the algorithm to early LEDA data to derive upper limits on the global 21-cm signal in the $13.2 < z < 27.4$ range.

The paper is organized as follows. In Section 2, we describe the Bayesian method and its application to simulated data. In Section 3, we apply it to LEDA data; our results are discussed in Section 4.

2 BAYESIAN FRAMEWORK AND SIMULATIONS

Bayesian Monte Carlo sampling (for example using Markov chains; MCMC) has become a standard method for exploring a likelihood surface and reconstructing the posterior distribution in order to extract cosmological parameters from CMB observations and, recently, also in the 21-cm field (Harker et al. 2012; Greig & Mesinger 2015). Bayes' theorem indeed relates the posterior probability distribution $\mathcal{P}(\Theta|D, \mathcal{H})$ of a set of parameters Θ given the data D and a model $\mathcal{H}$, which includes the hypothesis and any related assumptions, to the likelihood $\mathcal{L}(D|\Theta, \mathcal{H})$ as

$$\mathcal{P}(\Theta|D, \mathcal{H}) = \frac{\mathcal{L}(D|\Theta, \mathcal{H})\,\Pi(\Theta|\mathcal{H})}{\mathcal{Z}(D|\mathcal{H})}, \tag{1}$$

where the priors $\Pi(\Theta|\mathcal{H})$ encode existing knowledge of parameter values and the evidence $\mathcal{Z}(D|\mathcal{H})$ is the integral of the likelihood $\mathcal{L}(D|\Theta, \mathcal{H})$ over the prior space, allowing not only normalization of the posterior but also model selection via its inherent ability to quantify Occam's razor (e.g. Mackay 2003; Liddle, Mukherjee & Parkinson 2006; Trotta 2008; Parkinson & Liddle 2013).

We implemented an algorithm for extracting the global 21-cm signal following Harker et al. (2012), who assume Gaussian measurement noise and hence write the likelihood $\mathcal{L}_j$ of measuring the observed sky temperature $T_{\rm ant}(\nu_j)$ at a single frequency ν_j as

$$\mathcal{L}_j\left(T_{\rm ant}(\nu_j)|\Theta\right) = \frac{1}{\sqrt{2\pi\sigma^2(\nu_j)}} e^{-\frac{[T_{\rm ant}(\nu_j) - T_m(\nu_j, \Theta)]^2}{2\sigma^2(\nu_j)}}, \tag{2}$$

where $T_m(\nu_j, \Theta)$ is the model spectrum and $\sigma(\nu_j)$ is the standard deviation of the frequency-dependent instrumental noise,

$$\sigma(\nu_j) = \frac{T_{\rm ant}(\nu_j)}{\sqrt{\Delta\nu\Delta t}},$$

where $\Delta\nu$ is the channel width and Δt is the total integration time. Assuming that $T_{\rm ant}(\nu)$ is measured at M discrete frequency channels and that the noise is uncorrelated between frequency channels, the (log-)likelihood for the full frequency spectrum becomes

$$\ln \mathcal{L}(T_{\rm ant}|\Theta) = \sum_{j=1}^{M} \ln \mathcal{L}_j\left(T_{\rm ant}(\nu_j)|\Theta\right). \tag{3}$$

The sky model at each frequency channel ν_j is the sum of the foreground T_f and the 21-cm signal $T_{\rm H\,I}$:

$$T_m(\nu_j) = T_f(\nu_j) + T_{\rm H\,I}(\nu_j). \tag{4}$$

Single-dipole observations measure the integrated Galactic foreground spectrum averaged over the whole sky, losing information about its spatial structure and how to separate foregrounds from the 21-cm global signal is still a very active debate in the community. Liu et al. (2013) and Switzer & Liu (2014) suggest taking advantage of the spatial structure of the Galactic foreground in order to improve its separation from the spatially constant 21-cm global signal. The most commonly adopted approach is to simply leverage the different spectral behaviour of foregrounds and the 21-cm signal, parametrizing the foreground spectrum through a principal component analysis (e.g. Vedantham et al. 2014) or a log-polynomial (e.g. Bowman & Rogers 2010; Pritchard & Loeb 2010; Harker et al. 2012; Voytek et al. 2014; Bernardi et al. 2015; Presley et al. 2015).
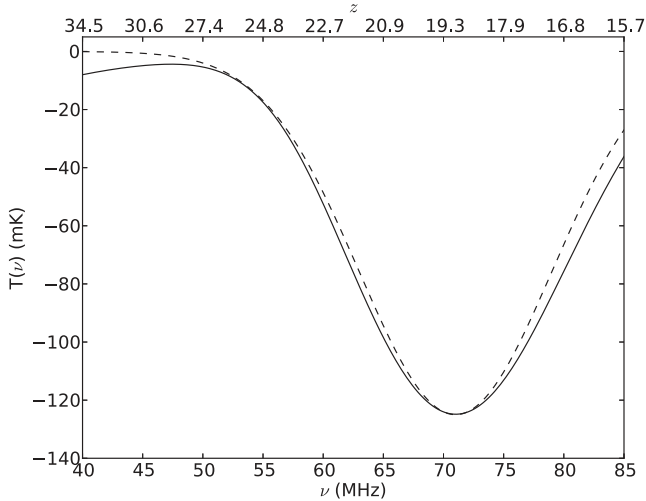
Figure 1. Comparison between a 21-cm empirical Gaussian model (dashed line) used in our current analysis and a physical model derived from ARES (solid line). The physical model is taken from Mirocha et al. (2015) and is defined by four parameters: the minimum virial temperature for star-forming haloes T_{min}, the efficiencies of Lyα and X-ray photon production, ξ_{LW} and ξ_X, respectively, and the IGM ionization efficiency ξ_{ion}. We set $T_{min} = 10^4$ K, $\xi_{LW} = 969$, $\xi_X = 0.02$ and $\xi_{ion} = 40$, respectively (see Mirocha et al. 2015, for details), and this can be considered as a reference model. The Gaussian model parameters are $A_{HI} = -125$ mK, $\nu_{HI} = 71$ MHz and $\sigma_{HI} = 8$ MHz, similar to the model chosen for our simulations. The agreement between the two profiles is at the 10–20-per cent level across most of the LEDA band.

In this paper, we therefore model the foreground emission as an Nth-order log-polynomial:

$$\log_{10} T_f(\nu_j) = \sum_{n=0}^{N} p_n \left[\log_{10} \left(\frac{\nu_j}{\nu_0} \right) \right]^n, \tag{5}$$

where we have adopted the convention $\nu_0 = 60$ MHz.

The choice of the polynomial order is critical in order to correctly model the foreground spectrum. Although earlier works showed that the foreground spectrum can be well described by very few components in frequency (e.g. de Oliveira-Costa et al. 2008; Pritchard & Loeb 2010), more recent simulations suggest that most of the frequency structure present in the observed sky arises from the coupling between the sky and the antenna beam pattern (Bernardi et al. 2015; Mozdzen et al. 2016).

Our implementation is focused on the pre-reionization, Cosmic Dawn signal at $15 \lesssim z \lesssim 30$, where the IGM is expected to be colder then the CMB. The 21-cm signal can be modelled as a Gaussian absorption profile (Bernardi et al. 2015; Presley et al. 2015):

$$T_{HI}(\nu_j) = A_{HI} e^{-\frac{(\nu_j - \nu_{HI})^2}{2\sigma_{HI}^2}}, \tag{6}$$

where A_{HI}, ν_{HI} and σ_{HI} are the amplitude, peak position and standard deviation of the 21-cm spectrum. We investigated how well this empirical model reproduces a physical 21-cm spectrum by using the publicly available code ARES[1] (Mirocha et al. 2012; Mirocha 2014; Mirocha et al. 2015). Fig. 1 shows that a Gaussian profile closely resembles the physical reference model defined in Mirocha et al. (2015) across most of the considered observing band. Deviations between the two models start to become noticeable at high

redshift when collisional coupling drives the 21-cm signal negative with respect to the Gaussian model. Although we plan to incorporate physical models in future analyses, the Gaussian profile is sufficiently accurate for the purpose of testing our signal-extraction method and applying it to establish first-order upper limits on the 21-cm signal (Section 3).

In order to efficiently explore the posterior probability distribution, we use the sampler MULTINEST (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009); crucially, it is an efficient calculator of the Bayesian evidence (with the posterior samples coming as a by-product) in relatively-low-dimensionality parameter spaces such as ours, and it robustly uncovers any degeneracies, skirts, wings or multimodalities in the posterior. We use an MPI-enabled PYTHON wrapper for MULTINEST (Buchner et al. 2014) that allows a full model fit to be evaluated in just a few minutes on a typical desktop machine. We have released a PYTHON implementation of our software, HIBAYES,[2] that incorporates the models described here, although the inclusion of different models is straightforward and will be the goal of future work.

We tested the signal extraction on a simulated case where we considered an $N = 7$ polynomial foreground model, representing the level of corruption of the intrinsic sky spectrum due to the primary beam for a simulated LEDA case (Bernardi et al. 2015). Such an assumption may be representative of other experiments – or considered a somewhat pessimistic case. We adopted the 21-cm model labelled as 'A' in Bernardi et al. (2015), which has an amplitude $A_{HI} = -100$ mK, a peak frequency $\nu_{HI} = 67$ MHz, a width $\sigma_{HI} = 5$ MHz and similar to the fiducial model of Pritchard & Loeb (2010) and Mirocha et al. (2015) plotted in Fig. 1. We considered a 400-h integration time with a 1-MHz channel width and a dual-polarization dipole. We also assumed the total bandwidth to span the 40–89 MHz range. These assumptions, although tuned to the LEDA case, can generally represent the observing specifications of any ground-based 21-cm global experiment that targets the pre-reionization era, with the 89-MHz cutoff being due to the radio frequency interference (RFI) caused by the radio FM band.

We assumed uniform priors on all the parameters and, in order (solely) to reduce the computing load, we set conservative priors on the 21-cm signal to be $-400 < A_{HI} < 0$ mK, $40 < \nu_{HI} < 89$ MHz and $0 < \sigma_{HI} < 35$ MHz. Whereas the priors on the peak position and width are essentially due to the observational constraints, the amplitude prior can be theoretically motivated by assuming an extreme (and somewhat unlikely) model with no gas heating occurring in the redshift range of interest.

The peak amplitude of the 21-cm signal may be estimated analytically from the expression for the 21-cm brightness temperature (e.g. Mesinger et al. 2015):

$$A_{HI} \approx 27 x_{HI} \left(1 - \frac{T_\gamma}{T_s} \right) (1 + \delta) \left(\frac{H}{dv_r/dr + H} \right)$$
$$\sqrt{\frac{1+z}{10} \frac{0.15}{\Omega_M h^2}} \left(\frac{\Omega_b h^2}{0.023} \right) \left(\frac{1 - Y_p}{0.75} \right) \text{ mK}. \tag{7}$$

For the global-signal case, we can ignore density fluctuations (i.e. set $\delta = 0$), peculiar velocities (i.e. set $dv_r/dr = 0$) and safely assume the helium fraction $Y_p = 0.25$. Assuming that the IGM is fully neutral

[1] https://bitbucket.org/mirochaj/ares

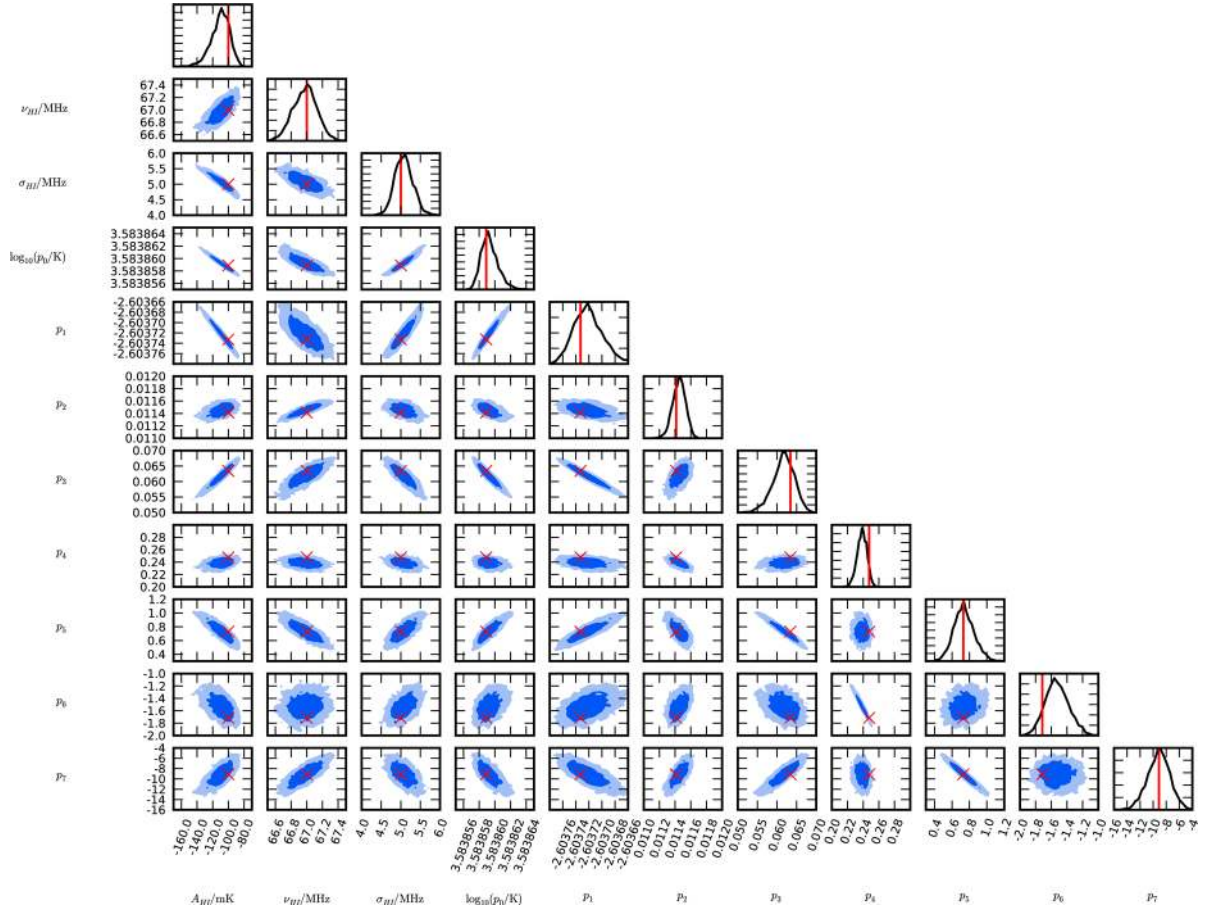[2] http://github.com/ska-sa/hibayes (Zwart, Price & Bernardi 2016).

Figure 2. Posterior probability distribution, marginalized into one and two dimensions, for the $N = 7$ foreground and the 21-cm models fitted to the simulated data. The dark and light shaded regions indicate the 68- and 95-per cent confidence regions. The simulated parameter values are indicated in red. The marginalized probability distributions are plotted in the [0, 1] range.

during this epoch (i.e. $x_{H_I} = 1$), equation (7) becomes

$$A_{H_I} \approx 27 \left(1 - \frac{T_\gamma}{T_s}\right) \sqrt{\frac{1+z}{10} \frac{0.15}{\Omega_M} \frac{\Omega_b h}{0.023}} \text{ mK}, \quad (8)$$

where T_γ is the CMB temperature, T_s is the spin temperature, $\Omega_M = 0.315$ is the matter density, $\Omega_b = 0.049$ is the baryon density and $h \equiv H/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.673$ is the normalized Hubble parameter (Planck Collaboration 2015).

Assuming no gas heating, the gas temperature T_K can be calculated from thermal decoupling (where $T_\gamma = T_K$) following the $(1 + z)^2$ adiabatic cooling. Also assuming that the Lyα emission from the first luminous sources is very effective in completely coupling the spin temperature T_s to the gas temperature, we can write

$$T_s = T_K = T_{\gamma,0}(1 + z_d)\left[\frac{1+z}{1+z_d}\right]^2, \quad (9)$$

with $T_{\gamma,0} = 2.73$ K the CMB temperature at the present time and $z_d \approx 200$ is the redshift of the thermal decoupling between the IGM and the CMB. Substituting everything into equation (8), we obtain

$$A_{H_I} \approx 27 \left(1 - \frac{1+z_d}{1+z}\right) \sqrt{\frac{(1+z)}{10} \frac{0.15}{\Omega_M} \frac{\Omega_b h}{0.023}} \text{ mK}, \quad (10)$$

which gives $A_{H_I} \approx -380$ mK at $z = 15.7$, corresponding to the lowest redshift of the considered observing band.

Results of the HIBAYES fit to the simulated data are shown in Figs 2 and 3. Most of the parameters are well recovered to within the 68-per cent contours, although some of the best-fitting values are marginally offset from their true values. Correlations between some parameters are apparent, although the one-dimensional marginalized distributions are fairly smooth for all the parameters, with no evidence for multimodality. Most of the foreground parameters are very tightly constrained, as are the 21-cm peak frequency and width. The 21-cm amplitude shows the largest relative errors – at the 12-per cent level – and noticeable anti-correlation with the foreground amplitude and slope. Such anti-correlation can be explained by the degeneracy between these parameters at the peak frequency ν_{H_I}: if the foreground amplitude is overestimated (underestimated), the 21-cm amplitude will be underestimated (overestimated) or the foreground slope will be steeper (flatter), in order to preserve the same observed spectrum value. We also note that there are correlations between the higher order polynomial coefficients, although they can be disentangled given the high level of sensitivity simulated here. The results presented here are in agreement with the Fisher matrix estimates from Bernardi et al. (2015) and show that, even in the presence of spectrally unsmooth foreground emission requiring high-order polynomials to be modelled, our method is able to extract the 21-cm signal provided sufficient signal-to-noise ratio. Our results are also broadly consistent with the MCMC analysis presented by Harker et al. (2012) and Harker (2015), although they used a fairly distinct frequency band and 21-cm model from ours.
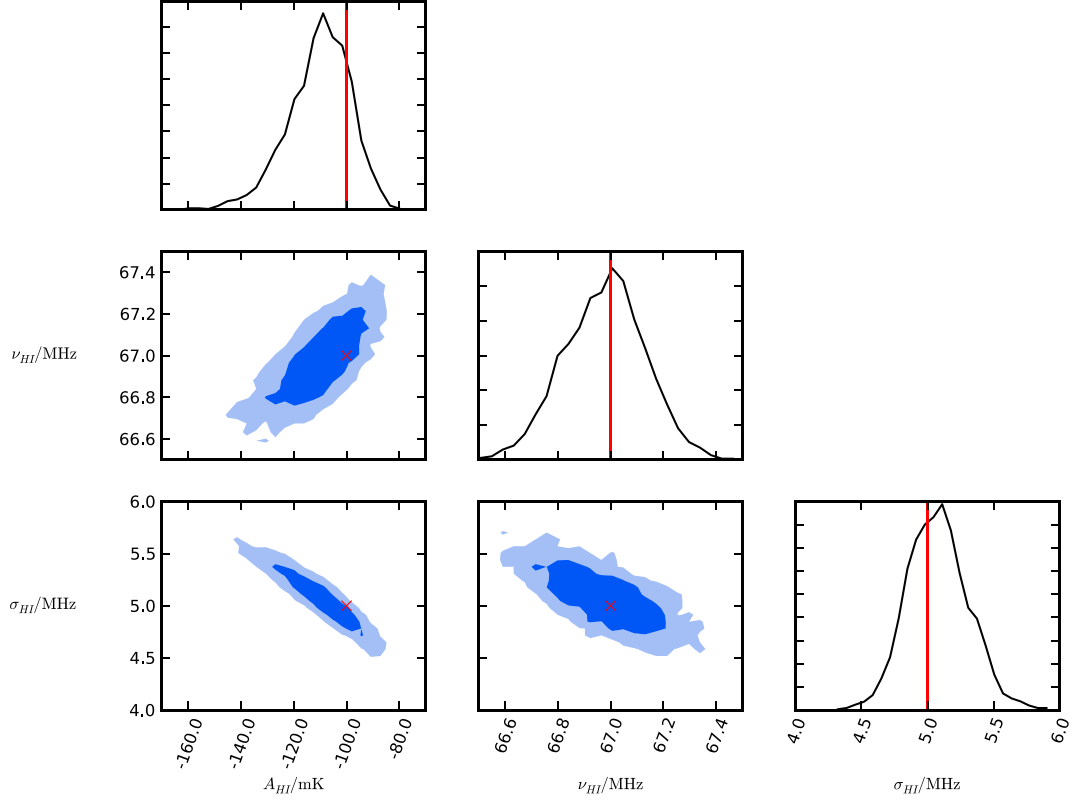
Figure 3. Zoom-in on the 21-cm parameters from Fig. 2. The dark and light shaded regions indicate the 68- and 95-per cent confidence regions. The simulated parameter values are indicated in red. The marginalized probability distributions are plotted in the [0, 1] range.

3 ANALYSIS OF LEDA DATA

We next applied our method to preliminary LEDA data. The LEDA instrument is described in detail in upcoming papers (Schinzel et al., in preparation; Price et al., in preparation); here, we briefly describe the system along with the observations and data-reduction approach.

LEDA is a sub-instrument of the Long Wavelength Array at the Owens Valley Radio Observatory (LWA-OVRO; Hallinan et al., in preparation). LWA-OVRO is primarily an all-sky imaging radio interferometer, designed to operate in the frequency range 10–88 MHz. It consists of a 'core' of 251 dual-polarization dipole-type antennas within a 200-m diameter, plus an additional 5 'outrigger' antennas, located a few hundred metres from the core, customized for LEDA. In addition, 32 'expansion' antennas are quasi-randomly distributed up to 1500 m from he core.

Each LEDA outrigger antenna is equipped with a receiver board designed for precision radiometry. Here we present the total-power data taken from a single outrigger antenna, whereas future LEDA analyses will make use of data from all five outrigger antennas, supported by the analysis of cross-correlations with the core antennas in order to improve the instrument calibration by measuring the antenna primary beam (see the discussion in Bernardi et al. 2015) and ionospheric distortions.

Observations were made during the nights of 2016 February 11 and 12 over a 2-h period centred at LST = $10^h 30^m$ when the Galactic Centre, Cassiopeia A and Cygnus A were near or below the horizon. The antenna total-power data were digitized at a rate of 196.608 MHz, giving a 0–98.304 MHz bandwidth, covered by 4096 channels, each of them 24-kHz wide. Data were integrated over 1 s.

We calibrated spectra using a multi-stage approach, as follows. The first calibration stage was a modified version of the three-state switching calibration technique employed by EDGES (Rogers & Bowman 2012). The three-state switching removes the effect of time variations in the system gain $G(\nu, t)$ and receiver temperature $T_{rx}(\nu, t)$, and imposes an absolute temperature scale on the data. The LEDA outrigger antennas switch between the sky, and two calibration references – referred to as 'hot' and 'cold' – with different noise-equivalent temperatures $T_{hot}(\nu, t)$ and $T_{cold}(\nu, t)$, respectively. The power measured in each state is then given by

$$P_{ant}(\nu, t) = G(\nu, t)\, \Delta\nu\, k_B (T'_{ant}(\nu, t) + T_{rx}(\nu, t))$$
$$P_{hot}(\nu, t) = G(\nu, t)\, \Delta\nu\, k_B (T_{hot}(\nu, t) + T_{rx}(\nu, t))$$
$$P_{cold}(\nu, t) = G(\nu, t)\, \Delta\nu\, k_B (T_{cold}(\nu, t) + T_{rx}(\nu, t)), \tag{11}$$

where P_{ant}, P_{hot} and P_{cold} are the powers for the antenna, hot calibration reference and cold calibration reference states, respectively; T'_{ant} is the antenna noise-equivalent temperature; $\Delta\nu = 24$ kHz is the channel width and k_B is the Boltzmann constant. The first-stage calibrated antenna temperature T'_{ant} is recovered via

$$T'_{ant}(\nu, t) = (T_{hot} - T_{cold}) \frac{P_{ant} - P_{cold}}{P_{hot} - P_{cold}} + T_{cold}, \tag{12}$$

where T_{hot} and T_{cold} are measured before the observation and where we drop the explicit dependence on time and frequency on the r.h.s. of the equation for simplicity. As the receiver switched between the three states every 5 s, with the first second of data in each state blanked, the total on-sky time was eventually 1152 s ($\approx$19 min).

The second-stage calibrated antenna temperature T''_{ant} is obtained by correcting T'_{ant} for the reflection coefficient Γ:

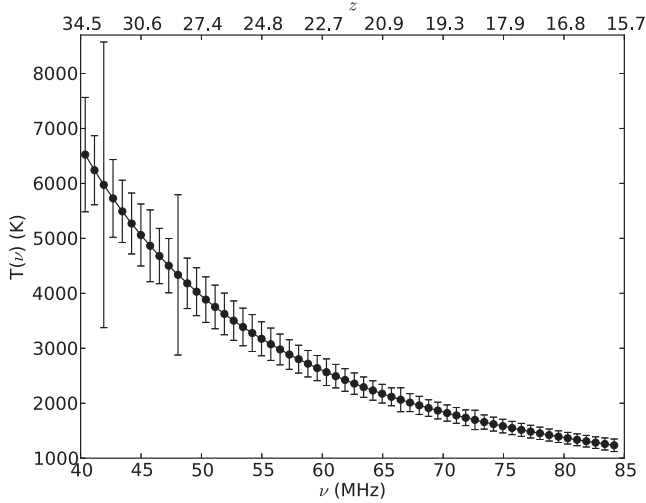$$T''_{ant}(\nu, t) = T'_{ant}(\nu, t)[1 - |\Gamma|^2(\nu)], \tag{13}$$

Figure 4. Measured sky spectrum for a 2-h observation ($\approx$19- min effective integration time, 2016 February 12 at $9^{\rm h}5 <$ LST $< 11^{\rm h}5$). Note that the error bars have been inflated by a factor of 1000 in order to make them visible.

where $\Gamma(\nu)$ measures the impedance mismatch between the receiver and the antenna and was determined using a vector network analyzer (Price et al., in preparation).

At this point, spectra were flagged for RFI using the SUMTHRESH-OLD algorithm (Offringa et al. 2010), then averaged in frequency to achieve a final resolution of 768 kHz. The 40–85 MHz band of interest was subsequently extracted, with a few MHz lost at the upper end of the bandwidth due to filter roll-off.

The final calibration stage was performed using a sky spectrum model $\hat{T}_{\rm ant}(\nu, t)$:

$$\hat{T}_{\rm ant}(\nu, t) = \frac{\int d\Omega B(\theta, \phi, \nu) T_{\rm sky}(\theta, \phi, \nu, t)}{\int d\Omega B(\theta, \phi, \nu)}, \tag{14}$$

where $B(\theta, \phi, \nu)$ is the antenna beam pattern and $T_{\rm sky}(\theta, \phi, \nu, t)$ is the model sky brightness distribution evaluated using the de Oliveira-Costa et al. (2008) global sky model.[3] We used the dipole beam model from Dowell et al. (2012). Dowell et al. (in preparation) have recently completed a sky survey covering the frequency range 35–80 MHz using the LWA, and these data may be used to improve the calibration in the future.

The calibrated antenna temperature $T_{\rm ant}(\nu, t)$ measured at time t is then obtained as

$$\begin{aligned} T_{\rm ant}(\nu, t) &= \frac{\int_T \hat{T}_{\rm ant}(\nu, t')\, dt'}{\int_T T''_{\rm ant}(\nu, t')\, dt'} T''_{\rm ant}(\nu, t) \\ &= \alpha(\nu) T''_{\rm ant}(\nu, t), \end{aligned} \tag{15}$$

where the average occurs over the full $T = 2$ h of observation. The calibration $\alpha(\nu)$ is calculated from observations on 2016 February 11, and then applied to the observations on February 12. The final calibrated sky spectrum $T_{\rm ant}(\nu)$ after averaging in time is shown in Fig. 4 and becomes the input for HIBAYES. The thermal noise can be estimated by propagating the uncertainties of equation (11):

$$\begin{aligned} \sigma^2(\nu) = {}& \left(\frac{\partial T'_{\rm ant}}{\partial P_{\rm ant}}\right)^2 (\Delta P_{\rm ant})^2 + \left(\frac{\partial T'_{\rm ant}}{\partial P_{\rm cold}}\right)^2 (\Delta P_{\rm cold})^2 \\ & + \left(\frac{\partial T'_{\rm ant}}{\partial P_{\rm hot}}\right)^2 (\Delta P_{\rm hot})^2. \end{aligned} \tag{16}$$

---

[3] Our PYTHON-based implementation that includes observer-centred sky models is available at https://github.com/telegraphic/pygsm (Price 2016).
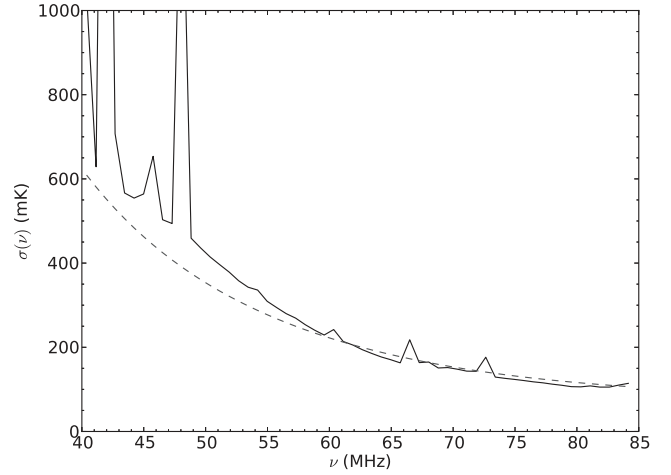
Figure 5. Comparison of the estimated thermal noise (dashed line) and the noise measured as the standard deviation of the observed data (solid line).
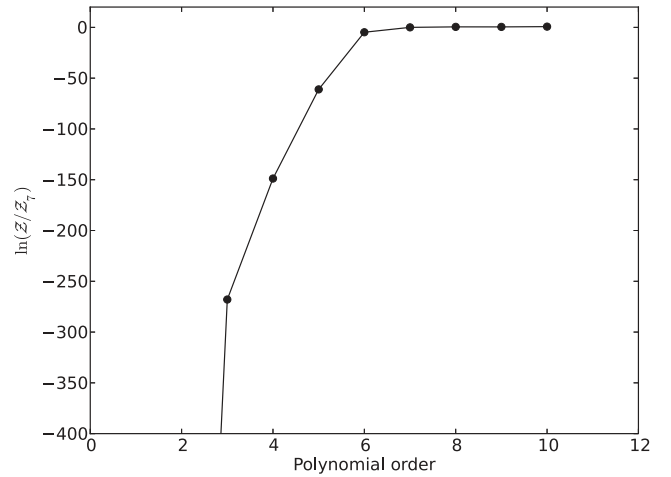

Figure 6. Bayesian evidence for foreground models fitted to the LEDA data, relative to the $N = 7$ polynomial foreground model, as a function of polynomial order N. The uncertainties on the evidence are of the same magnitude as the filled circle size.

Fig. 5 compares the estimated thermal noise with that derived as the standard deviation of the calibrated antenna temperature as a function of time for each frequency channel. The measured and the expected thermal noise levels are consistent above 55 MHz, whereas the measured noise is higher than expected at lower frequencies. The large spikes below 50 MHz correlate with known RFI sources, where a larger fraction of data are flagged, causing an effective decrease in integration time.

In the HIBAYES analysis, we first looked to confirm the foreground parametrization used in our simulations (Section 2). Following Harker (2015), we sought to establish the foreground model by fitting the data with increasing polynomial order assuming that the 21-cm signal is fairly described by equation (6). We found that the evidence (Fig. 6) increases sharply as a function of polynomial order until $N = 6$, after which it starts to flatten. According to the scale of Jeffreys (1939), the $N = 7$ model is still decisively (odds $>100 : 1$) preferred over the $N = 6$ model, whereas the $N = 8$ model is disfavoured (negative odds) over the $N = 7$ model. In practice, the evidence remains essentially flat as the polynomial increases beyond the $N = 7$ order and small (positive or negative) variations are likely due to sampling accuracy. We therefore fitted the data
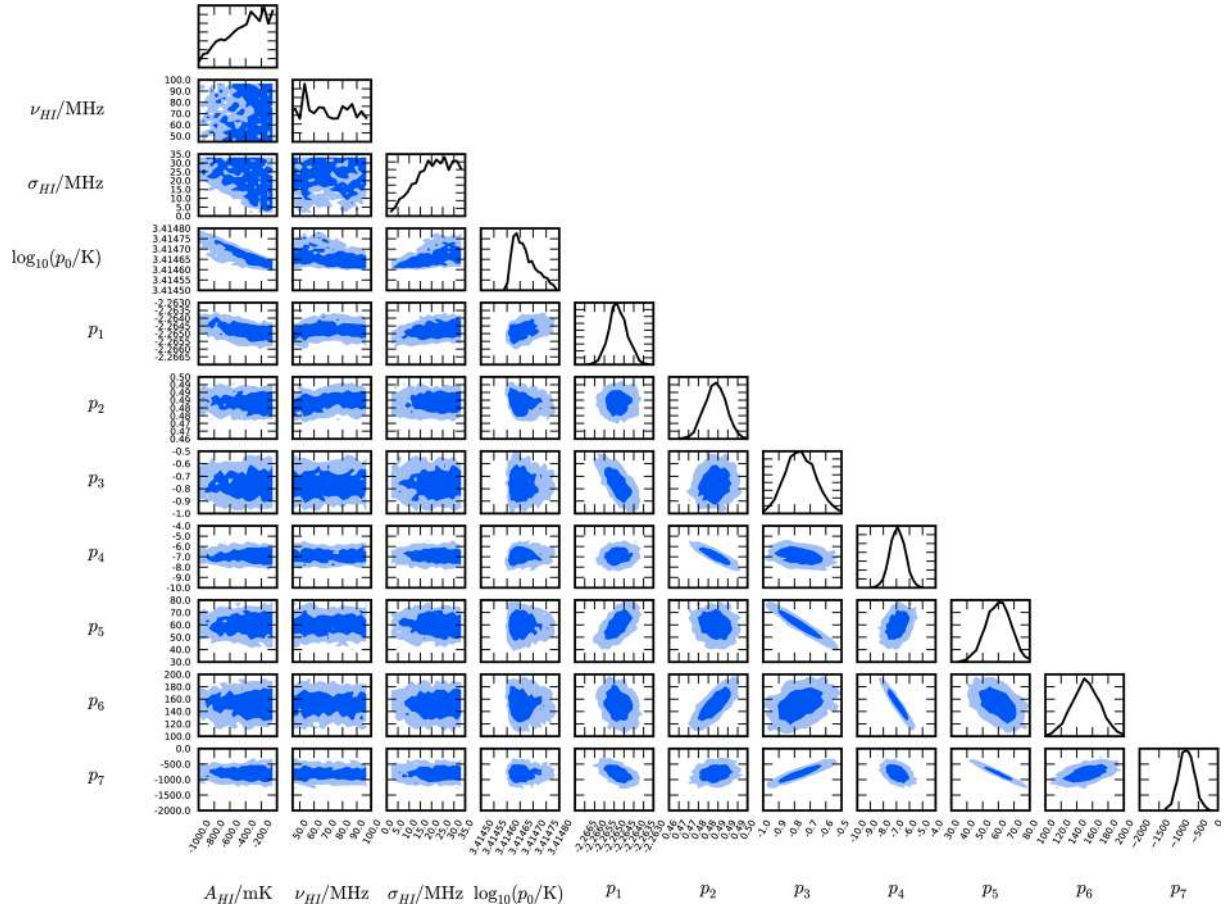
Figure 7. Posterior probability distribution, marginalized into one and two dimensions, for the $N = 7$ polynomial foreground and 21-cm models, fitted to the LEDA data. The dark and light shaded regions indicate the 68- and 95-per cent confidence regions. The marginalized probability distributions are plotted in the [0, 1] range.

using an evidence-motivated model that includes the 21-cm signal and an $N = 7$ polynomial foreground model, and we used the measured noise as a function of frequency. We emphasize here that the chosen $N = 7$ polynomial is not intended to represent the spectral structure of the intrinsic sky emission but rather the 'observed foregrounds', i.e. the convolution of the intrinsic sky emission with the instrumental response. It is also worth noticing that the evidence here favours a model that is in fair agreement with earlier LEDA simulations presented in Bernardi et al. (2015). In a future work, we will investigate the possibility of parametrizing the instrument and the sky emission separately and of using the evidence to indicate the best choice of sub-models, rather than assuming a combined parametrization as here.

After having established the foreground model, we set the priors on the width of the 21-cm signal to be the same used for simulations as they encompass the full breadth of theoretical predictions. We set uniform priors on the 21-cm peak position to be $50 < \nu_{H_I} < 100$ MHz as such a range brackets both models with the most extreme star formation efficiency – which would shift the peak at low frequencies – and with the most extreme X-ray efficiency – which would shift the peak at high frequencies (Pritchard & Loeb 2010; Mirocha et al. 2015). We relaxed the constraints on the depth of the 21-cm peak amplitude that we used for the simulated case because we seek to derive data-driven upper limits even in the case of models that are disfavoured by theory. We still assumed that the

21-cm signal cannot be positive, i.e. that $T_s = T_k < T_\gamma$, which is accepted in any model for the redshift range considered here.

We ran different chains with decreasing lower bounds of the A_{H_I} prior. In the initial case, we set $-380 < A_{H_I} < 0$ mK and found that the whole prior range is within the 95-per cent confidence region. We found that the two-dimensional posterior distributions for both the 21-cm amplitude and the width showed monotonically decreasing profiles with increasing prior range until an area of the prior range is clearly excluded at a confidence level greater than 95 per cent for $-1000 < A_{H_I} < 0$ mK. The posterior probability distribution for this final run is displayed in Fig. 7.

The foreground parameters are very well constrained and all their marginalized, one-dimensional distributions are Gaussian-like, similar to the simulated case, with the exception of the foreground amplitude p_0, whose marginalized posterior is slightly asymmetric. The best-fitting foreground parameters are fairly different from the simulated case apart from the first two coefficients that indicate the foreground amplitude at 60 MHz and its power-law slope. This is not unexpected as higher order polynomials are most likely compensating for limitations in the instrumental calibration that were not included in the simulations (e.g. errors on the reflection coefficient or calibration load). It is interesting, however, that the best-fitting foreground coefficients show correlations between foreground parameters similar, albeit at a qualitative level, to the simulated case, for example between the p_3 and p_5, and p_5 and p_7, coefficients.

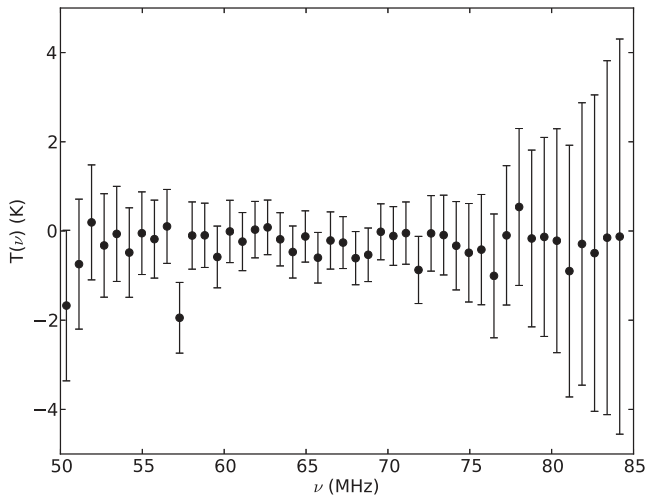**Figure 8.** Residual spectrum after subtraction of the best-fitting, maximum a posteriori foreground model. Error bars are plotted at the 2σ confidence level and include both the measured and the best-fitting parameter uncertainties. All the data points but one at 57.7 MHz are compatible with zero.

The 68- and 95-per cent confidence contours of the 21-cm parameters are fairly different and, essentially, identify upper limit regions, as expected given the noise levels. Whereas no constraints can be placed on the peak position $\nu_{\rm H\,I}$ within the prior range, bright, narrow 21-cm Gaussian profiles are disfavoured by the data. Quantitatively, we constrain $A_{\rm H\,I} > -890$ mK and $\sigma_{\rm H\,I} > 6.5$ MHz at the 95-per cent level in the $13.2 < z < 27.4$ range: this amplitude limit is only a factor of ≈ 2.5 away from constraining the extreme model with no heating described in Section 2.

In the absence of a detection, the root mean square (rms) of the residuals is a metric often used in observations – although less statistically rigorous than limits derived directly from the posterior probability distribution. Fig. 8 shows the residual spectrum after subtraction of the best-fitting, maximum a posteriori foreground model. We find a 470 mK rms residual over the whole redshift range, that, if considered a proxy for the 68-per cent confidence level, is approximately consistent with the HIBAYES constraints on the 21-cm peak amplitude.

4 SUMMARY AND CONCLUSIONS

We have presented a fully Bayesian algorithm for simultaneously fitting the global 21-cm signal in the presence of sky foregrounds. Our algorithm capitalizes on the Bayesian evidence's Occam's razor effect for model selection, with posterior probability distributions coming as a by-product.

We tested the method on simulated data and showed that, assuming a seventh-order polynomial foreground spectrum, the 21-cm global signal – parametrized as a Gaussian absorption profile – can be strongly constrained with a 400-h integration time for a LEDA-like observing setup. This result more quantitatively confirms the Fisher matrix analysis previously carried out by Bernardi et al. (2015). Although here we presented a specific application for the 21-cm signal from the Cosmic Dawn, the code can easily be extended to the full redshift range of interest for global-signal measurements. The code (Zwart et al. 2016) is publicly available at http://github.com/ska-sa/hibayes.

We applied the method to observations in order to derive upper limits on the 21-cm signal from the Cosmic Dawn. We showed that

the Bayesian evidence can guide the choice of foreground model (Harker 2015), with a maximum for a seventh-order polynomial in the present case. We emphasize that such a model reflects the combination of the intrinsic foregrounds and the spectral structure introduced by the instrument. In this respect, the evidence does not yet constrain the intrinsic foreground spectrum as suggested by Harker (2015), and future work will be dedicated to incorporating both the intrinsic sky and instrument models in the analysis and placing constraints on the intrinsic foreground spectrum.

The best-fitting foreground parameters are very well constrained; in particular, we derive a spectral index for the diffuse Galactic emission $\beta(\equiv p_1) = 2.27 \pm 0.04$. This value is consistent with early measurements of the Galactic radio background at 81.5 MHz by Bridle (1967), but is noticeably flatter than what was measured at 150 MHz by Rogers & Bowman (2008). This possible flattening of the spectral index may be good news for foreground subtraction for future 21-cm (global and interferometric) observations targeting the pre-reionization epoch.

Voytek et al. (2014) report the only other broad-band measurements at these frequencies. A direct comparison with their results is not straightforward as they do not report either rms residuals or direct upper limits on 21-cm parameters. We note, however, that our best-fitting spectral index is consistent with theirs, although their spectrum normalization is about 40 per cent greater than what we report here. Our measurements, however, are within 10 per cent of the carefully absolutely calibrated Galactic spectrum measured by EDGES at 150 MHz (Rogers & Bowman 2008) once it is scaled down to 60 MHz using the EDGES spectral index. We therefore believe our absolute flux density scale to be appropriate and its uncertainty to be negligible at the present level of sensitivity.

Our analysis constrains the 21-cm signal amplitude and width to be $-890 < A_{\rm H\,I} < 0$ mK and $\sigma_{\rm H\,I} > 6.5$ MHz, respectively, at the 95-per cent confidence level in the $13.2 < z < 27.4$ ($100 > \nu > 50$ MHz) range. Note that the constraint on $\sigma_{\rm H\,I}$ corresponds to a redshift width $\Delta z \approx 1.9$ at redshift $z \simeq 20$. Our results are the tightest upper limits on the 21-cm signal from the Cosmic Dawn to date and are encouraging in terms of achieving a factor of a few improvement in the sensitivity necessary to start placing significant constraints on structure prior reionization and on the thermal history of the IGM and the related sources of heating.

REFERENCES

Ali Z. S. et al., 2015, ApJ, 809, 61
Bernardi G., McQuinn M., Greenhill L. J., 2015, ApJ, 799, 90
Bowman J. D., Rogers A. E. E., 2010, Nature, 468, 796

Bowman J. D., Rogers A. E. E., Hewitt J. N., 2008, ApJ, 676, 1
Bridle A. H., 1967, MNRAS, 136, 219
Buchner J. et al., 2014, A&A, 564, A125
Ciardi B., Madau P., 2003, ApJ, 596, 1
Ciardi B., Stoehr F., White S. D. M., 2003, MNRAS, 343, 1101
DeBoer D. R. et al., 2016, PASA, preprint (arXiv:1606.07473)
de Oliveira-Costa A., Tegmark M., Gaensler B. M., Jonas J., Landecker T. L., Reich P., 2008, MNRAS, 388, 247
Dowell J., Wood D., Stovall K., Ray P. S., Clarke T., Taylor G., 2012, J. Astron. Instrum., 01, 1250006
Ewall-Wice A. et al., 2016a, MNRAS, 460, 4320
Ewall-Wice A., Hewitt J., Mesinger A., Dillon J. S., Liu A., Pober J., 2016b, MNRAS, 458, 2710
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Fialkov A., Barkana R., Visbal E., Tseliakhovich D., Hirata C. M., 2013, MNRAS, 432, 2909
Fialkov A., Barkana R., Visbal E., 2014, Nature, 506, 197
Field G. B., 1959, ApJ, 129, 536
Furlanetto S. R., 2006, MNRAS, 371, 867
Furlanetto S. R., 2016, in Mesinger A., ed., Astrophysics and Space Science Library, Vol. 423, Understanding the Epoch of Cosmic Reionization. Springer Int. Publ., Switzerland, p. 247
Greenhill L. J., Bernardi G., 2012, preprint (arXiv:1201.1700)
Greig B., Mesinger A., 2015, MNRAS, 449, 4246
Greig B., Mesinger A., Pober J. C., 2016, MNRAS, 455, 4295
Haiman Z., Abel T., Rees M. J., 2000, ApJ, 534, 11
Harker G. J. A., 2015, MNRAS, 449, L21
Harker G. J. A., Pritchard J. R., Burns J. O., Bowman J. D., 2012, MNRAS, 419, 1070
Harker G. J. A., Mirocha J., Burns J. O., Pritchard J. R., 2016, MNRAS, 455, 3829
Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., Mao Y., Koda J., Ahn K., 2012, MNRAS, 423, 2222
Jeffreys H., 1939, Theory of Probability. Oxford Univ. Press, Oxford
Kocz J. et al., 2015, J. Astron. Instrum., 4, 1550003
Koopmans L. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 1, Available at http://pos.sissa.it/cgi-bin/reader/conf.cgi?
Liddle A. R., Mukherjee P., Parkinson D., 2006, Astron. Geophys., 47, 4.30
Lidz A., Zahn O., McQuinn M., Zaldarriaga M., Hernquist L., 2008, ApJ, 680, 962
Liu A., Parsons A. R., 2016, MNRAS, 457, 1864
Liu A., Pritchard J. R., Tegmark M., Loeb A., 2013, Phys. Rev. D, 87, 043002
Mackay D. J. C., 2003, Information Theory, Inference and Learning Algorithms. Cambridge Univ. Press, Cambridge, p. 640
McKinley B. et al., 2013, AJ, 145, 23
McQuinn M., 2015, ARA&A, preprint (astro-ph/1512.00086)
McQuinn M., O'Leary R. M., 2012, ApJ, 760, 3

McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043
Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006, MNRAS, 372, 679
Mesinger A., Ferrara A., Spiegel D. S., 2013, MNRAS, 431, 621
Mesinger A., Ewall-Wice A., Hewitt J., 2014, MNRAS, 439, 3262
Mesinger A., Ferrara A., Greig B., Iliev I., Mellema G., Pritchard J., Santos M., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 11, Available at http://pos.sissa.it/cgi-bin/reader/conf.cgi?
Mesinger A., Greig B., Sobacchi E., 2016, MNRAS, 459, 2342
Mirocha J., 2014, MNRAS, 443, 1211
Mirocha J., Skory S., Burns J. O., Wise J. H., 2012, ApJ, 756, 94
Mirocha J., Harker G. J. A., Burns J. O., 2015, ApJ, 813, 11
Morandi A., Barkana R., 2012, MNRAS, 424, 2551
Mozdzen T. J., Bowman J. D., Monsalve R. A., Rogers A. E. E., 2016, MNRAS, 455, 3890
Offringa A. R., de Bruyn A. G., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010, MNRAS, 405, 155
Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, MNRAS, 443, 678
Parkinson D., Liddle A. R., 2013, preprint (arXiv:1302.1721)
Planck Collaboration, 2015, preprint (astro-ph/1502.01589)
Pober J. C. et al., 2014, ApJ, 782, 66
Pober J. C. et al., 2015, ApJ, 809, 62
Presley M. E., Liu A., Parsons A. R., 2015, ApJ, 809, 18
Price D. C., 2016, Astrophysics Source Code Library, record ascl:1603.013
Pritchard J. R., Furlanetto S. R., 2007, MNRAS, 376, 1680
Pritchard J. R., Loeb A., 2010, Phys. Rev. D, 82, 023006
Ricotti M., Gnedin N. Y., Shull J. M., 2001, ApJ, 560, 580
Rogers A. E. E., Bowman J. D., 2008, AJ, 136, 641
Rogers A. E. E., Bowman J. D., 2012, Radio Sci., 47, RS0K06
Singh S., Subrahmanyan R., Udaya Shankar N., Raghunathan A., 2015, ApJ, 815, 88
Switzer E. R., Liu A., 2014, ApJ, 793, 102
Tanaka T. L., O'Leary R. M., Perna R., 2016, MNRAS, 455, 2619
Trotta R., 2008, Contemp. Phys., 49, 71
Valdés M., Ferrara A., Mapelli M., Ripamonti E., 2007, MNRAS, 377, 245
Valdés M., Evoli C., Mesinger A., Ferrara A., Yoshida N., 2013, MNRAS, 429, 1705
Vedantham H. K., Koopmans L. V. E., de Bruyn A. G., Wijnholds S. J., Ciardi B., Brentjens M. A., 2014, MNRAS, 437, 1056
Vedantham H. K. et al., 2015, MNRAS, 450, 2291
Voytek T. C., Natarajan A., Jauregui-Garcia J. M., Peterson J. B., Lopez-Cruz O., 2014, ApJ, 782, L9
Wouthuysen S. A., 1952, AJ, 57, 31
Zwart J. T. L., Price D., Bernardi G., 2016, Astrophysics Source Code Library, record ascl:1606.004

This paper has been typeset from a TeX/LaTeX file prepared by the author.