Quantifying Ionospheric Effects on Global 21-cm Observations

Emma Shen,1* Dominic Anstey,1 Eloy de Lera Acedo,1,2 Anastasia Fialkov,2,3 Will Handley1,2
1 Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0HE, United Kingdom
2 Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, United Kingdom
3 Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom

ABSTRACT
We modelled the two major layer of Earth’s ionosphere, the F-layer and the D-layer, by a simplified spatial model with temporal variance to study the chromatic ionospheric effects on global 21-cm observations. From the analyses, we found that the magnitude of the ionospheric disruptions due to ionospheric refraction and absorption can be greater than the expected global 21-cm signal, and the variation of its magnitude can differ, depending on the ionospheric conditions. Within the parameter space adopted in the model, the shape of the global 21-cm signal is distorted after propagating through the ionosphere, while its amplitude is weakened. It is observed that the ionospheric effects do not cancel out over time, and thus should be accounted for in the foreground calibration at each timestep to account for the chromaticity introduced by the ionosphere.

Key words: methods: data analysis – atmospheric effects – cosmic dawn, reionisation, first stars, global 21-cm experiments

1 INTRODUCTION
A direct measurement of neutral hydrogen (HI) by measuring the 21-cm hyperfine transition line can provide a way to trace the intergalactic medium over time. However, the strong galactic foreground compared with the extremely weak signal is one of the main limitations of current 21-cm experiments, and in some cases their effects can be up to five orders of magnitude greater than the cosmological signal. In 2018, Bowman et al. (2018) reported to have found an absorption profile in the form of a flattened Gaussian centred at 78 MHz in the integrated spectrum by fitting it with the foreground model and a model for the 21-cm signal simultaneously. However, concerns have been raised for the unphysical parameters implied by their data, and the non-uniqueness of their solution (Hills et al. 2018; Singh & Subrahmanyan 2019; Bevins et al. 2020; Sims & Pober 2020).

The strong galactic foreground compared with the extremely weak signal is one of the main difficulties in achieving a detection of the 21-cm signal. Other astronomical radio sources emitting in the frequency band 30-240 MHz together are ~ 4-5 orders of magnitude larger than the 21-cm signal in brightness. These foreground sources include galactic synchrotron radiation, thermal free-free emission from galactic dust, and emissions from extragalactic point sources. Thus, precise understanding of frequency-dependent or chromatic effects of instrumental gain, instrumental noise contribution, antenna beam shape and ionospheric effects, coupled with spatially and spectrally varying foregrounds is critical to measuring the global 21-cm signal.

These effects are particularly challenging for detecting the global 21-cm signal because its relevant band lies in the lower frequency range; galactic foreground brightness temperature, with galactic synchrotron radiation being the major constituent, increases with decreasing frequency as a power law $T_{\text{sync}} \propto \nu^{-\alpha}$ with a spectral index of $\alpha \sim -2.54$ (Bowman et al. 2018). These foregrounds are considered spectrally smooth. It seems plausible that the signal could be extracted simply by fitting a power law. However, the inevitable chromaticity introduced by the antenna negates this assumption. Moreover, the increased fractional bandwidth in the lower frequency band leads to an increased variation of antenna beams across the measurement bandwidth, resulting in larger chromatic effects. Neither chromatic antenna beam nor the sky temperature is smooth as a function of frequency; after integration, the convolution would break the assumed smoothness.

The ionosphere further complicates chromatic mixing. The ionosphere is a magnetised plasma situated at the upper atmosphere spanning from ~ 50-600 km above the Earth’s surface. This ionised layer can corrupt a propagating signal. The most significant propagation effects relevant to the sky-averaged spectra are absorption and refraction. These effects increase with decreasing frequency, which scale approximately as $\nu^{-2}$, and are 2-3 orders of magnitude larger than the global 21-cm signal (Vedantham et al. 2014; Datta et al. 2016). This poses an especially large problem to experiments that adopt a wide field of view. In a global experiment, in order to achieve an achromatic antenna beam, it is necessary to adopt a wide field of view, as narrow beams often come with the penalty of chromatic effects. Also, a narrow-beam global experiment would be subject to cosmic variance (Muñoz & Cyr-Racine 2020). The global 21-cm signal could be measured directly by subtracting the modelled foreground contribution from the antenna temperature were our prior knowledge on the foregrounds and chromatic mixing accurate enough, without which we are left to make simple assumptions on those properties.

This work aims to address the chromatic effects due to the signal propagation through the temporally varying ionosphere, to which a satisfying solution has yet to be proposed. It analyses the chromatic ionospheric effects on global 21-cm observations by simulating the two major ionospheric layers, the F-layer and the D-layer, by a sim-
plified spatial model with temporal variance. The analysis focuses on the chromatic corruptions introduced by the ionosphere after applying a chromaticity correction factor. The model incorporates actual ionospheric data during nighttime, the time of day favourable to observations for Radio Experiment for the Analysis of Cosmic Hydrogen (REACH) (de Lera Acedo 2019).

2 IONOSPHERIC MODEL

The model adopted in this work is based on Vedantham et al. (2014); in their paper, it is shown that chromatic refraction and absorption on incoming rays due to the ionosphere would introduce a significant distortion to the foregrounds. Likewise, the model adopted in this paper considers only ionospheric refraction and absorption. Dynamic effects such as scintillation induced by ionospheric turbulence (Crane 1977), refraction resulting from large scale traveling ion-disturbances (Bougeret 1981), etc. are assumed to cancel out over time. Effects due to magnetic fields are smaller than day to day changes in electron density and thus are also ignored. Furthermore, Faraday rotation is assumed to be unpolarised on scales comparable to the antenna beam (Vedantham et al. 2014). The ionospheric layers are assumed to be homogeneous, and have no spatial variation across the azimuth.

In this paper, we simulate the ionosphere with time-varying parameters using actual data for the F-layer (Fig. 1). The data is collected from Lowell GIRO Data Center at station Louisvale, South Africa (https://ulcar.uml.edu/DIDBase/), whose geographical location is close to where REACH is situated (the Karoo, South Africa). There is currently no available data required to simulate the D-layer, so the parameters are altered based on the F-layer data. The details will be elaborated in section 5.3. Moreover, the model is simulated with an actual beam whose gain pattern is different at each frequency. We investigate the change in the detected global 21-cm signal and foregrounds when subjected to the ionosphere. Also, chromaticity correction factors are applied when analysing the data.

Figure 1. Ionospheric data retrieved from Lowell GIRO Data Center at station Louisvale, South Africa. Each colour represents a different night. The upper panel shows the maximum electron density in the F-layer overnight. The lower panel shows the upper F-layer boundary in solid lines and the lower F-layer boundary in dashed lines.

Figure 2. Not-to-scale depiction of ionospheric refraction and absorption of homogeneous ionospheric layers (Vedantham et al. 2014).

2.1 Refractive Index

The refractive index \( \eta \) of the ionosphere is given by the generalised Appleton-Hartree equation (Shkarofsky 1961), in which \( \eta \) is described by electron density, magnetic field and wave propagation path. Because the change in refraction index due to the Earth’s magnetic field is insignificant (Vedantham et al. 2014), the magnetic field term is left out to simplify the complex calculation. The simplified Appleton-Hartree equation without the magnetic field term is:

\[
\eta^2 = 1 - \left( \frac{\nu_p}{\nu} \right)^2, \quad (1)
\]

where \( \nu_p \) is the electron plasma frequency, and \( \nu_c \) is the electron collision frequency. The electron plasma frequency can be determined by electron density \( N_e \) by the equation

\[
\nu_p = \frac{1}{2\pi} \sqrt{\frac{N_e e^2}{m_e \epsilon_0}}, \quad (2)
\]

where \( e \) is the electron charge, \( m_e \) is the effective mass of the electron, and \( \epsilon_0 \) is the permittivity of free space. The electric field of a plane wave traveling in a homogeneous ionospheric layer is given by

\[
E(\Delta s) = E_0 \exp \left( -i \frac{2\pi \nu \Delta s}{c} \right), \quad (3)
\]

where \( c \) is the speed of light, \( \Delta s \) is the distance measured along the propagation path, and \( E_0 \) is the initial electric field. The real part of the refractive index depends mostly on the electron density; it introduces a change in the phase velocity, resulting in refraction. On the other hand, the imaginary part depends mostly on the electron collision rate, which is determined by the electron density, gas density and temperature. Being imaginary, it serves to dampen the wave amplitude exponentially, resulting in absorption.

The ionospheric model in this work is composed of two distinct parts (see Fig. 2): one is the F-layer, accounting for refraction, and the other is the D-layer, accounting for absorption. The F-layer is characterised by low atmospheric gas density and high electron density, while the D-layer by high atmospheric gas density and low electron density. Thus, the F-layer is modelled with the real and frequency-dependent \( \eta_F \),

\[
\eta_F = \left( 1 - \left( \frac{\nu_p}{\nu} \right)^2 \right)^{1/2}, \quad \nu_c = 0, \quad (4)
\]
and the D-layer is modelled with the imaginary and also frequency-dependent \( \eta_D \).

\[
\eta_D \approx \frac{1}{2} \frac{v_e^2}{v^2 + v_e^2} \frac{1}{v_2}.
\]  

(5)

The parameters used to model the D-layer are drawn from a typical ionospheric condition (Mitra 1959; Evans & Hagfors 1968).

### 2.2 F-layer Refraction

Two different equations used to describe electron density distribution within the F-layer have been tested in the model: the first one assumes a homogeneous F-layer shell spanning between 200 km and 400 km from the Earth’s surface. The other one assumes an inhomogeneous layer at the same height, in which the electron density \( N_e \) varies parabolically with height, its peak lying in the middle of the layer (Bailey 1948). The parabolic electron density as a function of height \( h \) is then

\[
N_e(h) = N_0 \left[ 1 - \left( \frac{h - h_m}{d} \right)^2 \right],
\]

(6)

where \( N_0 \) is the maximum electron density at mean height \( h_m \), and \( d \) is the semi-thickness of the layer. It, if modelled with the same maximum electron density, would, imaginably, result in a lesser degree of refraction than the homogeneous model.

Due to the curvature of the Earth, the angle of the incoming ray is subject to deviation \( \delta \theta \) while passing through the boundaries of the F-layer, whose total deviation can be calculated by Snell’s law. For the parabolic layer, the total deviation experienced by an incoming ray of frequency \( \nu \) can be written as

\[
\delta \theta(\nu, \theta) = \left( \frac{v_p}{\nu} \right)^2 \frac{2}{d^2} \left[ a \sin \theta \int_{h_m-d}^{h_m+d} \frac{(h - h_m)dh}{\eta_F^2 ((h + a)^2 \eta_F^2 - a^2 \sin^2 \theta)^2} \right],
\]

(7)

where \( a \) is the Earth’s radius. For the homogeneous model, this integral can be simplified as a discretised sum, considering only the contribution from the two surfaces on which \( N_e \) suffers a change. By Snell’s law, any form of electron distribution can be easily calculated. Fig. 3 shows the deviation angle of incoming rays of different incidence angles as a function of frequency, calculated by equation (7), as is done in Vedantham et al. (2014).

To not further complicate the matter at hand, for the rest of this paper, the simpler homogeneous F-layer is adopted. The parameters used to model the F-layer are changed with time (every fifteen minutes) according to the data retrieved from Lowell GIRO Data Center.

The deviation angle will result in the change of direction for the incoming rays: it is shown in Vedantham et al. (2014) that a chromatic increase in sky area of a few per cent, a disruption \( \sim 2-3 \) orders of magnitude higher than the global signal, is introduced by the F-layer refraction. To model this effect, a \textit{lensing} effect is applied to the modified beam, and in so doing, the beam integrates to larger than unity over the entire sky. It also means different regions of the sky will be weighted differently according to their respective deviation angle. Some radiation close to the horizon coming from below may be able to acquire a new zenith angle smaller than \( 90^\circ \) after refraction, and hence could be captured by the antenna. The antenna temperature is expected to increase by \( \sim 1-10 \) K due to refraction. The effective antenna beam

\[
\hat{B}(\nu, \theta, \phi) = B(\nu, \theta, \theta + \delta \theta, \phi),
\]

(8)

due to the F-layer refraction can thus be written.

### 3 D-LAYER ABSORPTION

The model assumes a homogeneous D-layer shell spanning between 60 km and 90 km from the Earth’s surface with a constant electron density \( N_e = 2.5 \times 10^8 \) m\(^{-3} \). The high atmospheric gas densities at these heights yield large electron collision frequencies that can result in significant absorption. The electron collision frequency is assumed to be 10 MHz (Nicolet 1953). By equation (3), the amplitude of the electric field of a plane wave traveling in the D-layer is

\[
|E(\Delta s)| = E_0 \exp \left( -\frac{2\pi \nu \Delta s}{c \eta_D} \right).
\]

(9)

Since the intensity of an electromagnetic wave is proportional to the square of its amplitude, the loss factor that describes the loss of signal intensity can be written as

\[
\mathcal{L}(\nu, \theta) = \exp \left( -\frac{4\pi \nu \Delta s}{c \eta_D} \right).
\]

(10)

The effective beam subject to absorption in the D-layer is then

\[
\hat{B}(\nu, \theta, \phi) = B(\nu, \theta, \phi) L(\nu, \theta).
\]

(11)

Fig. 4 shows the loss factor of incoming rays of different incidence angles as a function of frequency calculated by equation (10). While the total distance propagation path can be integrated numerically by Snell’s law, it can also be approximated as

\[
\Delta s \approx h_D \left( 1 + \frac{h_D}{a} \right) \left( \cos^2 \theta + \frac{2h_D}{a} \right)^{-\frac{1}{2}},
\]

(12)

where \( h_D \) is the width of the D-layer, and \( h_D \) is the mean D-layer height. The absorption level in the relevant frequency range 50-200 MHz suggested by our model can be as large as 15%.
Combining the effects of both the F and the D-layer, the effective gain at the chosen reference frequency, and it will remove the chromaticity function is calculated and applied at each time step before it is corrected by the chromatic ionospheric correction, and then added back afterwards:

$$\hat{T}_A(\nu) = T_{\text{CMB}} + \frac{1}{C(\nu)} \int_0^{2\pi} \int_0^{\pi/2} d\phi \int_0^\nu d\nu \sin \theta [T_\nu(\nu, \theta, \phi) - T_{\text{CMB}}] B(\nu, \theta, \phi).$$

(18)

When these assumptions hold, the integrated antenna temperature at each frequency will be readjusted towards the corresponding beam gain at the chosen reference frequency, and it will remove the chromatic effect introduced by the originally chromatic beam. In order to see how well this equation could apply to the chromatic ionospheric effects, the modified equation is also tested:

$$C(\nu) = \int_0^{2\pi} \int_0^{\pi/2} d\phi \int_0^\nu d\nu \sin \theta [T_\nu(\nu, \theta, \phi) - T_{\text{CMB}}] B(\nu, \theta, \phi).$$

(17)

The difference between the integrated antenna temperature corrected by equation (16) and (17) exceeds 0.3 K at lower frequencies, which is significant enough to corrupt a signal as weak as the global 21-cm line. This difference is also apparently chromatic.

A similar modification should apply also to the antenna temperature integration: it should be subtracted before it is corrected by the chromatic correction, and then added back afterwards:

$$\hat{T}_A(\nu) = T_{\text{CMB}} + \frac{1}{C(\nu)} \int_0^{2\pi} \int_0^{\pi/2} d\phi \int_0^\nu d\nu \sin \theta [T_\nu(\nu, \theta, \phi) - T_{\text{CMB}}] B(\nu, \theta, \phi).$$

(18)

4.3 Data Structure and Gridding

The total antenna temperature is integrated in the ring pixel scheme adopted in the HEALpix coordinate instead of the conventional spherical coordinates suggested in equation (14). In our model, there are
12 \times 512^2 \ (N_{\text{side}} = 512) \ pixels \ in \ total, \ corresponding \ to \ a \ resolution \ of \ \sim 0.002 \ \text{in} \ \text{radian}, \ \text{or} \ \sim 6.87 \ \text{minutes} \ of \ \text{arc}. \ \text{High} \ \text{resolution} \ \text{is} \ \text{required} \ \text{for} \ \text{this} \ \text{model}, \ \text{as} \ \text{the} \text{-F-layer} \ \text{deviation} \ \text{can} \ \text{be} \ \text{very} \ \text{small} \ \text{at} \ \text{small} \ \text{zenith} \ \text{angle} \ \text{and} \ \text{at} \ \text{higher} \ \text{frequencies}, \ \text{as} \ \text{shown} \ \text{in} \ \text{Fig.} \ 3. 

While \ the \ sky \ map \ is \ stored \ in \ the \ HEALPix \ structure, \ the \ log-
\text{spiral} \ beam \ is \ designed \ in \ spherical \ poloidal \ coordinates, \ and \ thus \ needs \ to \ be \ transformed \ and \ remapped \ into \ the \ HEALPix \ structure \ by \ interpolation \ before \ the \ pixel-wise \ multiplication \ of \ the \ sky \ map \ and \ the \ effective \ antenna \ beam \ \text{equation} \ (20). \ \text{Before} \ \text{being} \ \text{subjected} \ \text{to} \ \text{the} \ \text{ionospheric} \ \text{expansion} \ \text{imposed} \ \text{by} \ \text{the} \ \text{F-layer}, \ \text{the} \ \text{beam} \ \text{is} \ \text{renormalised} \ \text{to} \ \text{make} \ \text{sure} \ \text{the} \ \text{beam} \ \text{integrates} \ \text{to} \ \text{unity} \ \text{over} \ \text{the} \ \text{entire} \ \text{sky} \ \text{at} \ \text{every} \ \text{frequency}.

4.4 Modelled Global 21-cm Signal

In the simulations presented in this work, a 21-cm signal, assumed to be a normal Gaussian:

\[ T_{21}(v) = A_0 e^{-\left(v-v_0\right)^2/2\sigma^2}, \tag{21} \]

is inserted to the sky map to test the effect the ionospheric disturbances have on the existing method to extract the signal. The modelled 21-cm signal is assumed to be constant across the sky with no spatial variation. Gaussian signals motivated by Bowman et al. (2018) as well as those of various amplitudes \( A_0 \), central frequencies \( \nu_0 \), and width \( \sigma \) are tested. The sky temperature with an inserted global 21-cm signal is given by

\[ T_{\text{sky}}(v, \theta, \phi) = T_1(v, \theta, \phi) + T_{21}(v), \tag{22} \]

where \( T_1(v, \theta, \phi) = T_1, \nu \) in equation (15). This is the sky temperature we use in the model when simulating with a global 21-cm signal.

4.5 Data Analysis

It is essential to know the change in the detected antenna temperature the ionospheric effects would introduce, for it is what affects the values we get from direct observations.

As the simulation considers a ionosphere that varies in time, it should be noted that the change in antenna temperature is dependent on the sky region at which the beam is pointing, as well as the ionospheric condition at the time of observation. It is speculated that the chromatic ionospheric disturbances would cancel out over time, so, in such cases, in order to make a fair comparison between different integration periods, the integrated antenna temperature needs to be scaled to a common ground.

A similar problem arises when it comes to the chromaticity correction mentioned earlier. Equation (16) shows that when this chromaticity correction is applied, it is corrected with respect to the beam gain \( B(v_0, \Omega) \) at the reference frequency \( v_0 \), where \( \Omega \) is the angular position of the sky area at which the beam is pointing, which means the final product of this correction will differ according to the reference frequency; if the beam gain \( B(v, \Omega) \) at \( v_1 \) is smaller than at \( v_2 \) by a factor of \( a \) at \( a(\Omega) = B(v_1, \Omega)/B(v_2, \Omega) \), with \( B(v, \Omega) d\Omega = 1 \) for all frequencies:

\[ C_{v_2}(v) = \frac{1}{A} \times \int \frac{T_1(v_1, \Omega) B(v, \Omega) d\Omega}{T_1(v_1, \Omega) B(v_1, \Omega) d\Omega} = \frac{C_{v_1}(v)}{A}, \tag{23} \]

where \( A \) is a constant.

As shown in the derivation, the difference is not a simple constant but a factor convolved with \( T_{\text{CMB}} \) and it is also dependant on the angular location. This correction is adopted in some of the calculations just to show a fairer comparison when a different reference frequency is applied, but a fair comparison might not be needed in practice since the raw, unscaled, value is perhaps what really matters in the experiment.

Besides the difference in magnitude, it is perhaps more critical to see how the chromatic profile is altered; it is the intermixed chromatic effect that confuses the signal measurement. As the foreground can be said to obey a power law \( T_{\text{sync}} \propto \nu^{-\alpha} \), the \( n \)th order log-polynomial fitting of the following form is adopted to fit the foregrounds:

\[ T_M(\nu) = T_{\text{CMB}} + 10^{\log_{10} a_i \log(\nu)/\alpha_i}. \tag{24} \]

In this paper, we adopt a 5th log-polynomial to fit the data, as including higher order terms would risk modelling out the cosmological signal. This technique has been proposed in the literature (Pritchard & Loeb 2008; Harker et al. 2012) and is widely used in analysing global signal. However, the merit of this kind of fitting works well under the assumption of a smooth foreground; it falls short when a chromatic beam is used, where the spatial variation in the sky temperature is coupled into the spectral structure, especially for a wide frequency band. necessarily imply a.

Since the simulations include a signal in the form of a Gaussian, the complete form of the model that is used to fit the simulated data is:

\[ T_M(\nu) = T_{\text{CMB}} + 10^{\log_{10} a_i \log(\nu)/\alpha_i} + A_0 e^{-\left(v-v_0\right)^2/2\sigma^2}, \tag{25} \]

with constraints on each parameter, the values of which depend on the signal added in the simulation. If our model for the foregrounds is accurate enough, the latter form (equation 25) should fit the simulated data better than the previous form (equation 24) that only takes the foregrounds into account. Both forms of fitting will be shown in section 5. Also, one must stress that the yielded residuals after fitting do not necessarily imply an ionospheric feature, as any distortion to a originally smooth foreground would introduce an undulation to the residual of a log-polynomial fit.

5 SIMULATION RESULTS

In this section, the plots and analyses of the simulated data generated by the aforementioned model will be shown. The effects coming from the two different layers will be looked into separately before moving on to the simulations where both layers are in effect.

5.1 Separate Effects

Fig. 5 shows that the F-layer alone increases the foreground temperature by \( \sim 10 \ \text{K} \) at 50 MHz, resulting in a \( \sim 0.2\% \) increment compared to the original antenna temperature, while the D-layer reduces the signal by \( \sim 800 \ \text{K} \) at 50 MHz, resulting in a \( \sim 16\% \) deficit compared to the original antenna temperature. The D-layer dominates the ionospheric effect in magnitude, about 15 times that of the F-layer. However, we are more interested in how it affects the shape, which, when unaffected by the ionosphere or other effects, should resemble the power law that describes the foregrounds dominated by synchrotron radiation.

Fig. 6 shows the residual after being fitted using equation (24) and (25). The fitted Gaussian signal is closer to the simulated signal in terms of amplitude and the central frequency for the case where only the F-layer is simulated. Moreover, the residual of the F-layer is weaker than that of the D-layer. It suggests that the D-layer has a greater effect on the overall chromatic variance of the integrated antenna temperature. In the presence of only the F-layer, a global
signal of amplitude $\sim 0.5$ K is still detectable by the simple log-polynomial fitting (equation 25). However, in other cases simulated with a different set of parameters, where the residual yielded by the fitting exceeds 0.2 K, the F-layer alone can introduce significant distortions to the foregrounds. Although the signal in this simulation has bigger amplitude, $\sim 0.5$ K, in reality, it is possible that the signal is as small as 0.1 K (Cohen et al. 2017). In such cases, a 0.2 K disturbance can still overwhelm the signal.

Fig. 6 also shows the residual of the data integrated over a longer time interval (five nights), and a longer integration time appears to reduce the residual in the D-layer case. However, it is most likely due to the lesser effect at other time-steps that the overall distortion averages out when integrated, and this result is therefore not sufficient enough to be taken as an evidence that the ionospheric effect cancels out over time. The discussion will be elaborated in section 5.3.

### 5.2 Combined Effect

To get a grasp of the range of the total ionospheric effect can have on the integrated antenna temperature within the ionospheric parameter space, the parameters including electron density $N_e$, electron collision frequency $\nu_c$, height, and thickness, twelve sets of randomised parameters within respective realistic ranges in a normal ionospheric weather were simulated. This result suggests that the integrated antenna temperature subjected to ionospheric distortions modelled within a given scope can have a 2500 K difference at 50 MHz and the range narrows as the frequency increases. The percentage of difference also decreases with growing frequency.

It is also critical to know how the inserted global 21-cm signal could be affected due to the ionosphere. Fig. 7 shows the change due to the ionosphere simulated using randomised parameters. The plots show that, for a signal centring at 78.1 MHz, the amplitude can be weakened by $\sim 15\%$, which is similar to the D-layer absorption rate shown in Fig. 5. The change in the Gaussian shape shows the chromatic distortions introduced by the ionosphere, whose effects increase with lowering frequency.

### 5.3 Different Integration Time Length

A case with a constant and another with a time-varying D-layer, both with a time-varying F-layer, are shown in this section. There is currently no available data required to simulate the D-layer. Nevertheless, we are still interested in investigating the behaviour of a time-varying D-layer. To achieve that, we instead change the D-layer parameters according to the actual time-varying F-layer data by scaling, so that each scale factor of the same parameter of the D-layer and the F-layer corresponds to the respective scale factor suggested in the theoretical values adopted in Vedantham et al. (2014) (the same values are also adopted in Fig. 3 and 4). We have also made sure that those values lie within the theoretical ranges. Since electron collision frequency $\nu_c$ is considered null when modelling the F-layer, a constant $\nu_c = 10$ MHz is adopted for the D-layer. This is only an attempt to study a varying D-layer, and may not be representative of an actual ionosphere.

Fig. 8 shows the residuals of the data integrated over different lengths of time. It would be convenient if the chromatic ionospheric effects cancel over a long integration time; in such a case, the ionosphere can be ignored simply by temporal integration. However, by the way the data is analysed and fitted in this work (equation 25), there is no clear sign showing the cancellation of chromatic ionospheric effects over time for both constant and time-varying D-layer; the residual is still comparable to the global 21-cm signal in magnitude even after integration over five nights. The absence of this property suggests that an alternative approach to remove this significant distortion to the foregrounds should be proposed to accurately measure the global 21-cm signal.

One can also observe that the residuals decrease by about an order of magnitude in the case of a time-varying D-layer. It is due to a milder ionospheric condition compared to the theoretical values used to simulate a constant D-layer. It suggests that the significance of the ionospheric effects depends greatly on the current ionospheric condition.

### 5.4 Pattern of the Residuals

When simulated without a global 21-cm signal, the pattern of the residual after log-polynomial fitting (equation 24) is highly consistent in its shape in several different cases. This pattern is shown in Fig. 9 to be dominated by the D-layer, with the F-layer having a minor effect. Fig. 9 shows the residuals after log-polynomial fitting to the data simulated with six sets of randomised D-layer parameters and without the F-layer. The frequencies at which the local maxima and minima of the residuals are situated are consistent in all cases. The amplitude fluctuation with frequency is also consistent. The physics underlying this pattern and its mathematical composition is still unclear.

This characteristic motivates us to modify the fitting equation (equation 25) by adding a term that describes this pattern, so it becomes:

$$T_i(\nu) = T_{CMB} + 10^{12} \alpha_i \log(\nu)^j + A_0 e^{-\nu^2/2\sigma^2} + B_0 R_D(\nu),$$

where $B_0$ is a constant and $R_D(\nu)$ is the pattern as a function of fre-
Figure 6. The plots show the residuals of the signal modelled with ionospheric effects, fitted by equation (24) and (25) using least-squares fit. Simulated with only the (top-left) D-layer, a snapshot, (top-right) F-layer, a snapshot, (bottom-left) D-layer, constant in time, integrated every 15 minutes over 5 nights, (bottom-right) F-layer, time-varying, integrated every 15 minutes over 5 nights. We show: black – residual of the least-squares fit of a 5th order log-polynomial and a Gaussian (equation 25), olive – residual of the least-squares fit of a 5th order log-polynomial only (equation 24), red – the Gaussian component of the combined least-squares fit (equation 25), and dotted-magenta – the 21-cm signal added in the simulation. Greater residuals and more ill-fitted signals are yielded in the D-layer cases, suggesting that the D-layer has a greater overall effect on the chromatic variance.

Figure 7. Change in the global 21-signal due to the ionosphere. The parameters used to model the ionosphere are randomised within the theoretical ranges for a normal ionospheric weather. The plot shows the inserted signal in the dotted green line and the ionospherically affected signal in the solid black line. In this case, the amplitude is weakened by 15%. One can also observe a chromatic distortion in the signal, that it suffers a greater change at lower frequencies.

Table 2. The table shows the root mean square (RMS) of the residuals after fitting and the amplitude $|A_0|$, central frequency $\mu$, and width $\sigma$ (standard deviation) of the fitted Gaussian signal. The reference (in red) is the original input signal used in the model. The residual does not have a consistent decrease as the integration time period increases.

| Residual Fitted Gaussian Signal | RMS (K) | $|A_0|$ (K) | $\mu$ (MHz) | $\sigma$ (MHz) |
|--------------------------------|---------|------------|-------------|---------------|
| Reference                      | –       | 0.53       | 78.1        | 18.7          |
| time-varying D-layer           |         |            |             |               |
| snapshot                       | 3.92 × 10^{-2} | 0.21 | 81.3        | 7.38          |
| 1 hour                         | 1.41 × 10^{-2} | 0.22 | 80.7        | 9.86          |
| 4 hours                        | 3.40 × 10^{-2} | 0.38 | 79.6        | 15.6          |
| 6 hours                        | 2.45 × 10^{-2} | 0.30 | 79.7        | 13.5          |
| 43 hours                       | 3.01 × 10^{-2} | 0.27 | 79.1        | 12.2          |
| constant D-layer               |         |            |             |               |
| snapshot                       | 3.51 × 10^{-1} | 0.68 | 89.3        | 23.8          |
| 2 hour                         | 2.25 × 10^{-1} | 0.45 | 75.9        | 41.4          |
| 3 hours                        | 3.35 × 10^{-1} | 0.71 | 88.3        | 25.0          |
| 5 hours                        | 2.39 × 10^{-1} | 0.37 | 90.0        | 25.0          |
| 24 hours                       | 2.45 × 10^{-1} | 0.46 | 82.0        | 40.0          |

Figure 8. Residuals of the least-squares fit of a 5th order log-polynomial and a Gaussian (equation 25) to data integrated every 15 minutes over different lengths of time. The upper panel is simulated with a time-varying D-layer and F-layer, while the lower panel is simulated with a constant D-layer and a time-varying F-layer. Those integrated over a longer time do not yield a weaker residual or a significantly better fit, showing no clear sign that the chromatic ionospheric effects cancel over time in both cases.
Table 3. The table shows the root mean square (RMS) of the residuals after fitting and the amplitude $|A_0|$, central frequency $\mu$, and width $\sigma$ (standard deviation) of the fitted Gaussian signal. The reference (in red) is the original input signal used in the model. The same sets of data are fitted using two different equations, one is fitted with the additional D-layer pattern term. Some cases show improvements when the D-layer pattern is also being fitted while some yield a null fitted signal.

| Residual Gaussian Signal | Fitted without the D-Layer Pattern (equation 25) | Fitted with the D-Layer Pattern (equation 26) |
|--------------------------|------------------------------------------------|---------------------------------------------|
| Reference                | $0.53$ | $78.1$ | $18.7$ | $0.23 \times 10^{-1}$ | $0.00$ | $79.7$ | $41.6$ | $2.23 \times 10^{-1}$ | $0.00$ | $79.7$ | $41.6$ |
| snapshot 1               | $2.35 \times 10^{-1}$ | $0.45$ | $80.8$ | $39.6$ | $2.10 \times 10^{-2}$ | $0.45$ | $78.2$ | $19.7$ |
| snapshot 2               | $4.12 \times 10^{-1}$ | $0.99$ | $88.6$ | $25.6$ | $1.67 \times 10^{-2}$ | $0.20$ | $82.0$ | $11.5$ |
| snapshot 3               | $3.25 \times 10^{-1}$ | $0.99$ | $85.9$ | $27.0$ | $2.15 \times 10^{-1}$ | $0.00$ | $64.0$ | $38.6$ |
| 2 hours                  | $2.25 \times 10^{-1}$ | $0.45$ | $75.9$ | $41.4$ | $9.18 \times 10^{-2}$ | $0.99$ | $61.9$ | $7.22$ |
| 24 hours                 | $2.45 \times 10^{-1}$ | $0.46$ | $82.0$ | $40.0$ |                                                              |

Figure 9. Residuals (upper panel) and their Fourier components (lower panel), (purple) simulated with the D-layer, the F-layer, and a global 21-cm signal, (red) simulated with the D-layer, and the F-layer, (blue) simulated with the D-layer only. The frequencies at which the local maxima and minima of the residuals are situated are consistent in all cases. The amplitude fluctuation with frequency is also consistent. This pattern is dominated by the D-layer, with the F-layer having a minor effect.

counter the purpose of time integration, which is to let the chromatic ionospheric effect cancel itself out over time before signal extraction.

5.5 Modified Beam Chromaticity Correction and Spatially Varying Spectral Index

We are interested in knowing whether the modified chromaticity correction that bears a similar form to equation (17), used to correct the beam chromaticity, can also help mitigate the chromaticity imposed by the ionosphere. Technically, it is correcting for both the ionosphere and the beam, for the two are convolved for cases simulated with an ionosphere.

The difference between the data corrected by the two different equations is small, $\sim 0.1$ K at lower frequencies. Although the magnitude of this difference is still comparable to the brightness of a global 21-cm signal, the difference does not decrease with frequency like a power law, but seems to fluctuate about 0 K. The residuals are similar in both cases, with the one corrected by the modified equation having a slightly smaller residual (Fig. 10), suggesting that the modified equation has more or less improved the result. However, more cases need to be tested in order to ascertain the effectiveness of the modified correction equation.

We have also studied the residual of a case simulated with a uniform spectral index $\alpha = 2.5$ and another with realistic spectral indices that vary across the sky. The map with realistic spectral indices is built by interpolating the GSM and the Haslam map (Anstey et al. 2020). As the adopted chromaticity correction (equation 17) only works under the assumption that the spectral index is uniform, it is expected that the correction would fail to effectively correct for the coupled chromatic disruption caused by the antenna beam itself and the ionospheric effects when the spectral index varies spatially.

5.6 Different Reference Frequency

We study the difference between two sets of data simulated with the same ionosphere, but whose beam/ionosphere chromaticity is corrected by equation (19) using different reference frequencies. When doing a magnitude comparison, they are also corrected by a factor calculated by equation (23) to compensate for the different reference temperature at different frequency caused by a chromatic beam.

The difference of the former case can be seen as null, which is expected, since the original correction factor derived by equation (17) should make the same correction irrespective of the reference frequency, as long as the spectral index is constant across the entire sky. The latter case does not have the same property, as the beam no longer integrates to unity but to different values at different frequencies after being stretched by the ionospheric refraction. The difference, $\sim 0.2$ K at lower frequencies is significant enough to overwhelm the global 21-cm signal. The residuals in Fig. 11 shows obvious changes in the residual when a different reference frequency is adopted. This is due to the intrinsic beam chromaticity being coupled with the chromatic ionospheric effects, that it no longer integrates to unity.

5.7 Different Gaussian Signals

Apart from the signal motivated by Bowman et al. (2018), two different global 21-cm signals of lower amplitudes ($\sim 0.1$ K) and centred at other frequencies are also tested in the simulation. The parameters of these two signals lie within the possible range of high-redshift parameters suggested in Cohen et al. (2017). So far only two have...
Table 4. The table shows the root mean square (RMS) of the residuals after fitting and the amplitude $|A_0|$, central frequency $\mu$, and width $\sigma$ (standard deviation) of the fitted Gaussian signal. The reference (in red) is the original input signal used in the model. The residual increases with growing reference frequency. The overall fit shows an improvement as the reference frequency decreases, with the exception of width, which is shown to suffer greater reduction.

| Residual Fitted Gaussian Signal | RMS (K) | $|A_0|$ (K) | $\mu$ (MHz) | $\sigma$ (MHz) |
|--------------------------------|---------|------------|-------------|----------------|
| Reference                      | -       | 0.53       | 78.1        | 18.7           |
| 90 MHz                         | 2.09 x 10^{-1} | 0.87       | 80.8        | 16.7           |
| 120 MHz                        | 2.96 x 10^{-1} | 0.99       | 83.5        | 24.2           |
| 150 MHz                        | 3.41 x 10^{-1} | 0.99       | 86.1        | 25.6           |
| 180 MHz                        | 3.59 x 10^{-1} | 0.99       | 86.9        | 26.2           |

Figure 11. Residuals of the least-squares fit of a 5th order log-polynomial and a Gaussian (equation 25) to data corrected using different reference frequencies: 90 MHz, 120 MHz, 150 MHz, and 180 MHz. The difference in the residuals can be observed. The chromatic beam no longer integrates to unity when coupled with the chromatic ionospheric effects, and thus introduces different distortions at different frequencies.

Figure 12. The panels show the residuals of the signal modelled with ionospheric effects, fitted by equation (24) and (25) using least-squares fit. The inserted Gaussian signals in the different panels: (top) $A_0 = 0.53$, $\mu = 78.1$ MHz, $\sigma = 18.7$ MHz (middle) $A_0 = 0.15$, $\mu = 100$ MHz, $\sigma = 10$ MHz (bottom) $A_0 = 0.2$, $\mu = 60$ MHz, $\sigma = 10$ MHz, integrated over 5 nights. Each panel consists of three different lines: black – residual of the least-squares fit of a 5th order log-polynomial and a Gaussian (equation 25), olive – residual of the least-squares fit of a 5th order log-polynomial only (equation 24), red – the Gaussian component of the combined least-squares fit (equation 25), and dotted-magenta – the 21-cm signal added in the simulation. The residuals can be as large as 0.6 K at the lower frequency range, similar to the case where the Gaussian signal is only a compensation for the deviation from the pattern at the less stable points at the extrema.

5.8 REACH Pipeline Analysis

We also tried to fit the simulated data using the current version of the REACH analysis pipeline, in which no feature to correct the ionospheric distortion has been added. In the pipeline, POLYCHORD, a state-of-the-art nested sampling algorithm for high-dimensional parameter spaces using Bayesian inference (Anstey et al. 2020), is adopted. It tries to fit the data using the log likelihood given by

$$
\log(L) = - \sum_{\nu} \frac{1}{2} \log(2\pi \sigma^2) + \frac{1}{2} \left( \frac{T_A(\nu) - T_1(\nu, \theta_f) - T_2(\nu, \theta_{sig})}{\sigma} \right)^2
$$

where $\sigma$ is a free parameter with a log uniform prior of 0.001-0.1 K. Fig. 13 shows the reconstructed Gaussian signal, in blue, and the residual, in red. It is apparent that the Gaussian signal cannot be recovered by the current version of the pipeline when the signal is contaminated by ionospheric effects, with a residual exceeding 4 K at lower frequencies, much higher than the intended signal detection at $\sim 0.5$ K.

6 CONCLUSIONS

The works shown in this paper can be summarised by the following points:

(i) The total chromatic ionospheric effect is dominated by the D-layer absorption both in magnitude and chromatic distortion. The F-layer has a lower, yet significant, contribution. The residuals yielded by the simple log-polynomial fit seem to share a similar pattern, the local extrema occurring at the same frequencies in multiple cases when simulated without a signal.

(ii) Irrespective of the extreme ionospheric conditions, foreground
temperature can differ within a $\sim 2500$ K range, depending on the ionospheric condition. The relevant parameters include electron density, electron collision frequency, height and width in both layers.

(iii) The global 21-cm signal modelled as a Gaussian can have its amplitude weakened and its shape distorted after being subjected to ionospheric interference based on the parameters used in this analysis.

(iv) Preliminary results show that the chromaticity introduced by the ionosphere does not cancel out over time for a constant or a time-varying D-layer, with a time-varying F-layer, when using our simplified spatial model. It also suggests that the significance of the ionospheric effects greatly depend on the current ionospheric condition.

(v) When simulated with a uniform spectral index, the beam chromaticity function and its modified counterpart have a limited use to correcting the antenna temperature subjected to chromatic ionospheric interference. Moreover, the modified equation yields different results when adopting different reference frequencies. The adopted chromaticity corrections have little to no use in the cases where spatially varying spectral indices are adopted.

(vi) As precise removal of the foregrounds is critical in extracting the faint global 21-cm signal, the ionospheric model should be included in data analysis pipelines to properly account for the chromaticity introduced by the ionosphere.

The ionospheric model adopted in this work makes strong assumptions on the complex ionosphere, excluding spatial azimuthal variations and the dynamic phenomena that might affect the result significantly. Based on our results, we conclude that better ionospheric models are essential to understanding the chromatic global 21-cm signal distortion in ground-based experiments.

A ionospheric model has not yet been included in the actual foreground model adopted in REACH. As a precise knowledge of the ionospheric condition is not currently achievable, the current ionospheric model can be improved by gaining further insights on the mathematical formulation of the ionospheric effects based on the observables as well as the simulations. With some final adjustments, it should eventually be deployed in the REACH analysis pipeline to enable a more accurate removal of the foregrounds.

**REFERENCES**

- Anstey D., de Lera Acedo E., Handley W., 2020, A General Bayesian Framework for Foreground Modelling and Chromaticity Correction for Global 21cm Experiments (arXiv:2010.09644)
- Bailey D. K., 1948, Terrestrial Magnetism and Atmospheric Electricity, 53, 41
- Bevins H. T. J., Handley W. J., Fialkov A., de Lera Acedo E., Greenhill L. J., Price D. C., 2020, arXiv e-prints, p. arXiv:2007.14970
- Bougeret J. L., 1981, Astronomy and Astrophysics, 96, 259
- Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
- Cohen A., Fialkov A., Barkana R., Lotem M., 2017, Monthly Notices of the Royal Astronomical Society, 472, 1915–1931
- Crane R. K., 1977, IEEE Proceedings, 65, 180
- Datta A., Bradley R., Burns J. O., Harker G., Komjathy A., Lazio T. J. W., 2016, Astrophysical Journal, 831, 6
- Evans J. V., Hagfors T., 1968, Radar astronomy. New York McGraw-Hill
- Harker G. J. A., Pritchard J. R., Burns J. O., Bowman J. O., 2012, Monthly Notices of the Royal Astronomical Society, 419, 1070
- Hills R., Kuikka R., Meier D. P., Puchwein E., 2018, Nature, 564, 332
- Mitra A. P., 1959, Journal of Geophysical Research (1896–1977), 64, 733
- Mozdzen T. J., Bowman J. D., Monsalve R. A., Rogers A. E. E., 2017, Monthly Notices of the Royal Astronomical Society, 464, 4995
- Muñoz J. B., Cyr-Racine F.-Y., 2020, arXiv e-prints, p. arXiv:2005.03664
- Nicolet M., 1953, Journal of Atmospheric and Terrestrial Physics, 3, 200
- Pritchard J. R., Loeb A., 2008, Phys. Rev. D, 78, 103511
- Shkarofsky I. P., 1961, Proceedings of the IRE, 49, 1857
- Sims P. H., Pober J. C., 2020, Monthly Notices of the Royal Astronomical Society, 492, 22
- Singh S., Subrahmanyan R., 2019, ApJ, 880, 26
- Vedantham H. K., Koopmans L. V. E., de Bruyn A. G., Wijnholds S. J., Ciardi B., Brentjens M. A., 2014, Monthly Notices of the Royal Astronomical Society, 437, 1056
- de Lera Acedo E., 2019, 2019 International Conference on Electromagnetics in Advanced Applications (ICEAA), pp 0626–0629

**ACKNOWLEDGEMENTS**

AF is supported by the Royal Society University Research Fellowship. WH was supported by a Gonville & Caius research fellowship. EdLa and DA are supported by STFC.

**DATA AVAILABILITY**

Two sets of simulation data (integrated antenna temperature) for reference are available at doi:10.5281/zenodo.4105855. The ionospheric data collected from Lowell GIRO Data Center are available at https://ulcar.uml.edu/DIDBase/.
APPENDIX A: DERIVATION OF THE EXTRA CHROMATICITY CORRECTION FACTOR

This is the full derivation of the extra chromaticity correction factor (equation 23 in section 4.5):

\[ C_{v_2}(v) = \frac{\int [T_i(v_2, \Omega) - T_{\text{CMB}}] B(v, \Omega) d\Omega}{\int [T_i(v_2, \Omega) - T_{\text{CMB}}] B(v, \Omega) d\Omega} \]

\[ = \frac{\int [T_i(v_1, \Omega) - T_{\text{CMB}}] (v_2/v_1)^{-a} B(v, \Omega) d\Omega}{\int [T_i(v_1, \Omega) - T_{\text{CMB}}] B(v, \Omega) d\Omega} \]

\[ = \frac{(v_2/v_1)^{-a} \int [T_i(v_1, \Omega) - T_{\text{CMB}}] a(\Omega) B(v, \Omega) d\Omega}{\int [T_i(v_1, \Omega) - T_{\text{CMB}}] a(\Omega) B(v, \Omega) d\Omega} \]

\[ \approx \frac{\int T_i(v_1, \Omega) a(\Omega) B(v, \Omega) d\Omega}{\int T_i(v_1, \Omega) a(\Omega) B(v_1, \Omega) d\Omega} \text{, when } T_{\text{CMB}} \text{ is small} \]

\[ = \frac{1}{A} \times \frac{\int T_i(v_1, \Omega) B(v, \Omega) d\Omega}{\int T_i(v_1, \Omega) B(v_1, \Omega) d\Omega} = \frac{C_{v_1}(v)}{A}. \]

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