Assessment of Ionospheric Activity Tolerances for Epoch of Reionization Science with the Murchison Widefield Array

Cathryn M. Trott[1,2] , C. H. Jordan[1,2] , S. G. Murray[1,2] , B. Pindor[2,3], D. A. Mitchell[2,4] , R. B. Wayth[1,2] , J. Line[2,3],
B. McKinley[1,2], A. Beardsley[5] , J. Bowman[5] , F. Briggs[6], B. J. Hazelton[7] , J. Hewitt[8], D. Jacobs[5] , M. F. Morales[7] ,
J. C. Pober[9], S. Sethi[10], U. Shankar[10], R. Subrahmanyan[10] , M. Tegmark[11], S. J. Tingay[1,2] , R. L. Webster[2,3] , and
J. S. B. Wyithe[2,3]

[1] International Centre for Radio Astronomy Research—Curtin University 1 Turner Avenue, Bentley, WA 6102, Australia; cathryn.trott@.curtin.edu.au
[2] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[3] The University of Melbourne, School of Physics, Parkville, VIC 3010, Australia
[4] CSIRO Astronomy and Space Science (CASS), P.O. Box 76, Epping, NSW 1710, Australia
[5] Arizona State University, AZ, USA
[6] The Australian National University, Australia
[7] University of Washington, Department of Physics, Seattle, WA 98195, USA
[8] Massachusetts Institute of Technology, Cambridge, MA, USA
[9] Brown University, Department of Physics, Providence, RI 02912, USA
[10] Raman Research Institute, Bangalore, India
[11] MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

Structure imprinted in foreground extragalactic point sources by ionospheric refraction has the potential to contaminate Epoch of Reionization (EoR) power spectra of the 21 cm emission line of neutral hydrogen. The alteration of the spatial and spectral structure of foreground measurements due to total electron content gradients in the ionosphere creates a departure from the expected sky signal. We present a general framework for understanding the signatures of ionospheric behavior in the 2D neutral hydrogen power spectrum measured by a low-frequency radio interferometer. Two primary classes of ionospheric behavior are considered, corresponding to dominant modes observed in Murchison Widefield Array (MWA) EoR data, namely, anisotropic structured wave behavior and isotropic turbulence. Analytic predictions for power spectrum bias due to this contamination are computed and compared with simulations. We then apply the ionospheric metric described in Jordan et al. to study the impact of ionospheric structure on MWA data, by dividing MWA EoR data sets into classes with good and poor ionospheric conditions, using sets of matched 30-minute observations from 2014 September. The results are compared with the analytic and simulated predictions, demonstrating the observed bias in the power spectrum when the ionosphere is active (displays coherent structures or isotropic turbulence). The analysis demonstrates that unless ionospheric activity can be quantified and corrected, active data should not be included in EoR analysis in order to avoid systematic biases in cosmological power spectra. When data are corrected with a model formed from the calibration information, bias reduces below the expected 21 cm signal level. Data are considered "quiet" when the median measured source position offsets are less than $10''$–$15''$.

Key words: dark ages, reionization, first stars – instrumentation: interferometers – methods: statistical

1. Introduction

Detection and characterization of the neutral hydrogen signal from the early universe provides one of the best probes for exploring cosmology and source astrophysics at these epochs (Furlanetto et al. 2006). Because it is an emission-line signal emitted by the neutral intergalactic medium (IGM), study of its spatial and redshift structure and its amplitude encodes information about the growth of structure and evolution of the first ionizing sources of radiation. It therefore offers the potential to connect cosmic microwave background (CMB) observations from the last scattering surface with low- and intermediate-redshift observations of active galactic nuclei (AGNs) and galaxies, to obtain a complete evolutionary history of the universe.

Measurements of the integrated optical depth to electrons (Planck Collaboration et al. 2016) and of the Lyα Forest along sight lines to high-redshift quasars (Fan et al. 2006; Mortlock 2016) constrain the redshift range for the Epoch of Reionization (EoR), where the averaged neutral fraction of the IGM

reduces from 95% to 5%, in the range of $z = 5.5$–10, placing observations of the rest-frame 1420 MHz neutral hydrogen emission line within reach of low-frequency instruments ($\nu = 130$–220 MHz). The current low-frequency telescopes that are attempting to detect the EoR signal at lower redshift include the Murchison Widefield Array (MWA;[12] Bowman et al. 2013; Tingay et al. 2013; Jacobs et al. 2016), the Precision Array for Probing the Epoch of Reionization (PAPER;[13] Parsons et al. 2010), the Low Frequency Array (LOFAR;[14] van Haarlem et al. 2013; Patil et al. 2016), and the Long Wavelength Array (LWA;[15] Ellingson et al. 2009). These will soon be complemented by second-generation experiments such as HERA (DeBoer et al. 2017) and the Square Kilometre Array Low Telescope (SKA-Low; Koopmans et al. 2015). While current experiments aim to detect the EoR signal and

---

[12] http://www.mwatelescope.org
[13] http://eor.berkeley.edu
[14] http://www.lofar.org
[15] http://lwa.unm.edu

provide some broad evolutionary parameters, future experiments aim to explore this era, yielding a wealth of cosmological and astrophysical information.

Despite the substantial current efforts to detect the EoR signal from neutral hydrogen, a spatial fluctuation detection has yet eluded the community (unlike the global signal, potentially detected in the Cosmic Dawn by Bowman et al. 2018). While the signal weakness (tens of mK brightness temperature fluctuations) relative to measurement noise requires hundreds to thousands of hours of observations, the principal barriers are the complexity of the instruments and bright foreground (extragalactic and Galactic) emission. These foreground contaminants, observed through complex and chromatic radio interferometers, yield contamination in the signal across the parameter space of interest (angular and line-of-sight spatial scales). Careful instrument calibration and foreground treatment are crucial for accessing the weak, high-redshift signal (Barry et al. 2016; Patil et al. 2016; Trott & Wayth 2016). Any unmodeled structure to the foregrounds therefore presents a challenge, altering the signal processing that needs to be undertaken to perform the experiment, and having the potential to imprint residual structure that mimics the cosmological signal. One such source of unmodeled structure is refraction of the incoming foreground wave fronts by plasma nonuniformity in Earth's upper atmosphere, principally the ionosphere (Loi et al. 2016; Mevius et al. 2016; Jordan et al. 2017; de Gasperin et al. 2018). Gradients in the plasma across the pierce points to extragalactic sources yield source position offsets, while curvature induces flux density changes (scintillation) as the wave front acquires concavity or convexity (Vedantham & Koopmans 2016). Tasse et al. (2013), Prasad et al. (2014), Martí-Vidal et al. (2010), Wijnholds et al. (2010), Cornwell (2016), and Hurley-Walker & Hancock (2018) study the impact of the ionosphere on calibration and offer algorithms for correcting the direction-dependent effects. Most recently, Rioja et al. (2018) have suggested the LEAP algorithm to provide post-imaging corrections to the data, with application to the extended array configuration of MWA Phase II and SKA. Long-baseline arrays (e.g., LOFAR) have observed source decoherence (Mevius et al. 2016), and these have been the basis for much of the early work in this area (van der Tol et al. 2007). While Vedantham & Koopmans (2016) discussed the implications for EoR of scintillation noise on interferometric visibilities, we study the impact of residual noise power bias from refractive positional offsets of foreground point sources, which hinder our view of the weak cosmological signal.

The paper is structured as follows. In Section 2, we present an overview of the MWA EoR experiment, the data collected for this work, and the quantitative analysis described in Jordan et al. (2017). The analytic models for the ionosphere are then described in Section 3, and Section 4 provides details of the simulations used to support the analysis. Finally, in Section 5, we apply the models to the distribution of ionospheric conditions observed at the MRO and characterized by Jordan et al. (2017) and predict the impact on MWA EoR experiments. Throughout, vectors are denoted with an overhead arrow, matrices are uppercase and in script, and estimates of a quantity have a hat. Fourier transforms are denoted with script $\mathcal{F}$.

2. Methods

2.1. The MWA EoR Experiment

The MWA (Tingay et al. 2013) is a low-frequency aperture array interferometer, located in the Western Australian desert on the same site as the future SKA-Low and is an SKA precursor instrument. Its science goals are shared with many other low-frequency instruments, namely, early universe studies through redshifted neutral hydrogen, pulsar searches and studies, fast and slow transients, and low-frequency sky catalogs (Bowman et al. 2013; Wayth et al. 2015; Hurley-Walker et al. 2017). In Phase I (2013–2016), relevant for this work, it comprised 128 tiles of 16 dual-polarization crossed dipole antennas, spread over a diameter of ∼3 km.

The MWA EoR experiment (Jacobs et al. 2016) observes two primary fields (EOR0 and EOR1) in cold parts of the radio sky, with a total observing time exceeding 1500 hr, with the principal goal of statistical detection of the EoR and potential for estimation of the slope and amplitude of the spherically averaged power spectrum. It operates over two primary frequency bands: 138–167 MHz and 167–197 MHz.

2.2. Ionospheric Conditions at the MRO

Beardsley et al. (2016) have published the deepest upper limits on the EoR power spectrum in the redshift range $z = 6.5$–7.5 for the MWA EoR experiments, with inclusion of 32 hr of high-quality EoR0 data in the high band (167–197 MHz), derived from a raw sample of 86 hr. Much of the data that were unused in the analysis come from observations with poor calibration, attributable to periods of substantial ionospheric activity (Jordan et al. 2017). At small k, they found that the data were systematics dominated, and they attributed this excess power to foreground contamination. These published data, which included some remaining ionospheric activity (identified more recently with new techniques), were processed through the two independent analysis pipelines that are described in Jacobs et al. (2016). This dual processing ensured consistency and reliability of results, and the consequent processing through the MWA Real Time calibration System (RTS; Mitchell et al. 2008) resulted in calibration log files being available for the final data set. Jordan et al. (2017) analyzed these calibration files as a function of time and sky location, deriving ionospheric characteristics across ∼1000 point sources each 8 s. The analysis of these data resulted in the development of a metric to quantitatively describe the degree of ionospheric activity and identification and classification of four distinct ionospheric types. The metric has two independent components, which capture the two primary classes of activity observed: isotropic turbulent-like increases in the source position variations compared with thermal noise, and aniso-tropic structured waves of plasma density with a spatially periodic structure (see Loi et al. 2016, who discovered these density ducts in MWA data sets). The relevant components of the metric for these quantify the median absolute source offset, relative to a catalog position, and the anisotropy of position offset vectors (quantified via a principal component analysis of the ensemble source vectors), respectively. See Jordan et al. (2017) for relevant details.

In this work, we combine these two structural models to form a generic and tunable model for ionospheric structure, akin to that developed by Intema (2009, 2014) and van der Tol et al. (2007),

which also summed a turbulent and trigonometric component. The ionospheric phase screen derivative at position l, m and frequency ν is parameterized as

$$\phi(l, m, \nu; k_l, k_m, r_0, n) = A(\nu)\phi_{ani}(l, m, \nu; k_l, k_m) \\ + B(\nu)\phi_{turb}(l, m, \nu; r_0, n), \quad (1)$$

where k_l, k_m parameterize the spatial wavenumbers of the anisotropic component and r_0, n parameterize the characteristic scale and power-law exponent of the turbulent component. For the purposes of demonstrating the four classes of ionospheric conditions, the overall frequency-dependent amplitude of these components is controlled by $A(\nu)$ and $B(\nu)$, with a frequency-squared dependence demonstrated in the data (Jordan et al. 2017).

The anisotropic component is modeled explicitly by the observed *variation* of source number density compared with a stochastic (unclustered) field, such that

$$\Delta l(l, m, \nu; k_l) \propto \sin(2\pi k_l l), \quad (2)$$

where the shift in the apparent source position, Δl, is related to the gradient of the total electron content (TEC) screen (divergence of the source number density):

$$\nabla\phi_{ani}(l, m, \nu; k_l) \propto \Delta l. \quad (3)$$

The fields used by the MWA EoR experiment do not demonstrate any clustering for sources with $S > 100$ mJy (the confusion limit).

The turbulent component exhibits power-law behavior in Fourier space and is defined by its phase structure function (which is equivalent to the sum of the two-point correlation function and the variance—correlation at zero lag),

$$D_{\phi\phi}(r) = \langle[\phi_{turb}(r + p) - \phi_{turb}(p)]^2\rangle_p = \left(\frac{r}{r_0}\right)^n, \quad (4)$$

where p characterizes the spatial separation of points on the celestial sphere.

The four primary classes, as well as their relative occupancy in the MWA EoR data, identified by Jordan et al. (2017) are as follows:

1. Type I: quiet ionosphere; A, B small (74%).
2. Type II: turbulent ionosphere; A small, B large (15%).
3. Type III: structured, anisotropic ionosphere; A moderate, B small (3%).
4. Type IV: structured and turbulent ionosphere; A, B large (8%).

Thus, the bulk of the data were acquired during quiet or isotropic turbulent modes. These data are all using zenith-pointed primary beams, where the phase center HA < 1 hr. We now employ these types to form analytic models for EoR data and propagate into the 2D power spectrum. We then use the relative occupancy of data of each type to predict the impact on current and future EoR experiments at the MRO. For active modes, we assess the two qualitatively different modes in these active types: anisotropic, structured modes, and isotropic turbulence.

2.3. Power Spectrum

We aim to derive the expected power spectrum of the extragalactic point sources due to ionospheric disturbance of the wave front. The power spectrum of neutral hydrogen brightness temperature fluctuations is defined as

$$P_{21}(k = |\boldsymbol{k}|) = \frac{\delta(\boldsymbol{k} - \boldsymbol{k}')}{\Omega} \langle\tilde{T}(\boldsymbol{k})\tilde{T}(\boldsymbol{k}')\rangle \text{ mK}^2 h^{-3} \text{ Mpc}^3, \quad (5)$$

where Ω denotes the observation volume and T is the 21 cm brightness temperature field, and the delta function encodes the expectation of signal isotropy. In the flat-sky approximation, where the Fourier transform of the sky brightness can be equated with the measured interferometric visibilities, we can use the distribution of visibilities to estimate the power spectrum of the point-source foregrounds:

$$P(k) \propto \langle V^*(k)V(k)\rangle = \mathrm{Var}(V(k)) + \langle V(k)\rangle\langle V^*(k)\rangle, \quad (6)$$

where we have used the identity for the variance of a quantity in terms of its first and second moments. To understand the effect of the ionosphere on the power spectrum, we therefore only are required to determine the expected value and variance of the measured visibilities under active and quiet ionospheric conditions, where the former is studied for both turbulent and structured activity.

3. Ionospheric Models

Motivated by the knowledge that foreground power and structure will remain a primary impediment to detecting and characterizing the EoR cosmological signal, we are interested in understanding the impact of ionospheric activity on residual foreground signals. For this work, we employ a frequency-dependent model of point-source foregrounds (radio and star-forming galaxies at $z < 6$, but primarily below $z = 2$), to represent the residual sky model.

3.1. Foreground Point-source Models

Turbulent component: The isotropic component of the model is used to represent a Kolmogorov turbulence in the plasma. We follow the formalism of Murray et al. (2017), who developed a generic model for the statistical signature of clustering of extragalactic point sources in the 2D power spectrum. In addition to the intrinsic cosmological clustering of AGNs and star-forming galaxies, they extend the formalism to derive the power spectrum of apparent source overdensities due to a Kolmogorov turbulence in the ionosphere. This derivation is based on the weak-lensing formalism of the spatial distortion of a field according to a TEC gradient. The full derivation is presented in the Appendix, and the main results are reproduced here.

We denote the (scalar) TEC field by $\psi(\boldsymbol{l})$ and note that, in the weak-field regime with which we are here concerned, the first-order effect of the ionosphere is to impose a shift on the observed angular position of incoming sight lines. This shift can be written as

$$\Delta \boldsymbol{l} = K(\nu)\nabla\psi(\boldsymbol{l}), \quad (7)$$

where $K \propto \nu^{-2}$ is a scaling constant. The observed power spectrum (see Lewis & Challinor 2006, Equation (4.12)) is

$$P_{turb} \approx (1 - K(\nu)u^2 R^\psi)P_0(u) + K(\nu)P_u^\psi \star P_0(u), \quad (8)$$

with $\star$ indicating convolution, u encoding spatial wavenumber, and

$$R^\psi = \pi \int u^3 P_u^\psi \, du. \tag{9}$$

The second term in Equation (8) is subdominant, and we ignore it herein. We suppose that the power spectrum of ψ is Kolmogorov-like but has some scale above which the power is truncated. In practice, setting this scale above the resolution limit of an instrument serves to eliminate sensitivity to this choice. Thus, we consider

$$P_u^\psi = \begin{cases} \left(\dfrac{u}{u_0}\right)^\kappa & u \leqslant u_{\max} \\ 0 & u > u_{\max}, \end{cases} \tag{10}$$

where u_0 sets the characteristic scale of the turbulence. In this case, we have

$$R^\psi = \frac{u_{\max}^{4+\kappa}}{(4+\kappa)u_0^\kappa}. \tag{11}$$

Thus, the expected angular power spectrum due to Kolmogorov-like turbulence modifies the unperturbed power according to

$$P_{\text{turb}} = \left(1 - K(\nu)u^2 \frac{u_{\max}^{4+\kappa}}{(4+\kappa)u_0^\kappa}\right)P_u^S. \tag{12}$$

The last step is to include the line-of-sight contribution, where the amplitude of the turbulence follows a frequency-squared dependence, $K_u \propto \nu^{-2}$. For the prediction, we are assuming a perfect instrument, whereby the Fourier plane is completely filled and the "wedge"-like signature of smooth foregrounds (due to incomplete interferometric sampling) is replaced by a "brick"-like feature where power occupies the $\eta = 0$ mode and small η modes. If the ionospheric dependencies on frequency and spatial scale are assumed to be independent, we can approximately decouple the angular and line-of-sight Fourier transforms, yielding

$$P_{\text{turb}} = P_{u\eta}^S W_0(\eta) - K_0 W_2(\eta) u^2 \frac{u_{\max}^{4+\kappa}}{(4+\kappa)u_0^\kappa}P_{u\eta}^S, \tag{13}$$

where

$$W_2(\eta) = |\mathcal{F}(W(\nu)\nu^{\gamma-2})|^2 \tag{14}$$

and

$$W_0(\eta) = |\mathcal{F}(W(\nu)\nu^\gamma)|^2 \tag{15}$$

are the Fourier transforms of the window function multiplied by the decaying spectral behavior of the ionosphere and foreground sources (index $\gamma \sim -0.7$) and the bandwidth window function multiplied by the spectral behavior of the foregrounds alone, respectively. This term encodes the difference in line-of-sight spatial mode behavior for differing power-law dependencies on the frequency. In practice, for observational volumes with sufficient fractional bandwidth, the additional frequency dependence of the ionosphere will yield an additional leakage of foreground power from the brick into the EoR Window. This effect is shown in Section 5.

Structured component: In Trott et al. (2016), we present a framework for representing the statistical 2D power spectral signature of unclustered, Poisson-distributed point-source foregrounds within an interferometric instrument with a frequency-dependent primary beam and chromatic Fourier sampling. To extend this simple model, where the sources are assumed to be distributed randomly on the sky, to one where sources are apparently clustered into periodic anisotropic tubes via ionospheric refraction, a term with sky position dependence is added to the source number density model, characterized by amplitude ϵ, relative to the intrinsic case. Without loss of generality for a generic instrument, we consider the ionospheric perturbations to be aligned with y-axis, removing functional dependence on the m-component (north–south direction cosine):

$$\frac{dN}{dS}(l, m; \nu) = \alpha(1 + \epsilon(\nu)\sin[2\pi k_l l])\left(\frac{S_{\text{Jy}}}{S_0}\right)^{-\beta} \text{Jy}^{-1}\,\text{sr}^{-1}, \tag{16}$$

where

$$\epsilon(\nu) = \epsilon_0 \left(\frac{\nu}{\nu_0}\right)^{-2}. \tag{17}$$

We assume a 2D frequency-dependent primary beam with a Gaussian shape. This is broadly representative of the beam shape for a square dipole-based tile (such as the MWA's 4×4 dipole arrangement), allows decoupling of the l and m planes, and has an analytic Fourier transform:

$$B(l, n; \nu) = \exp\left(-(l^2 + m^2)/2\sigma_\nu^2\right). \tag{18}$$

Propagation of the source model through the frequency-dependent instrument model (primary beam, with characteristic size σ_ν) yields an expression for the variance and expected value at angular Fourier modes (u, v) and frequency ν, given by

$$\begin{aligned} \text{Var}(V(u, v)) = {} & \frac{\alpha S_{\max}^{3-\beta}}{3-\beta}\pi\sigma_\nu^2 \\ & \times \left(1 + \frac{\epsilon(\nu)^2}{2} - \frac{\epsilon(\nu)^2}{2}\exp(-4\pi^2 k_l^2 \sigma_\nu^2)\right), \end{aligned} \tag{19}$$

$$\langle V(u, v)\rangle = \frac{\alpha S_{\max}^{2-\beta}}{2-\beta}2\pi\sigma_\nu^2 \exp(-2\pi^2 v^2 \sigma_\nu^2)\, G, \tag{20}$$

where

$$\begin{aligned} G = {} & \exp(-2\pi^2 u^2 \sigma_\nu^2) \\ & + \frac{\epsilon(\nu)}{2}(\exp(-2\pi^2\sigma^2(u-k_l)^2) \\ & - \exp(-2\pi^2\sigma^2(u+k_l)^2)). \end{aligned} \tag{21}$$

These expressions can then be combined to form the expected power spectrum due to extragalactic point-source foregrounds distorted coherently by the ionosphere. The equivalent quiet ionosphere expressions can be recovered by setting $\epsilon = 0$.

As with the turbulent component, the frequency dependence of the anisotropic component can be approximated by applying the Fourier transform of power laws with different indices to the quiet and active components. If the point-source foregrounds are considered to be spectrally flat (a reasonable assumption over a small fractional bandwidth compared with the inverse-square dependence of the ionospheric component),

then the spectral channels are completely correlated, and the $\eta = 0$ term contains all of the power (up to spectral leakage due to a finite bandwidth, which causes some spillage of power into small η modes, yielding the brick-like feature):[16]

$$\langle V(k, \eta)^* V(k, \eta) \rangle = W_0 \mathrm{Var}(V_o(k, \eta_0)$$
$$+ W_2 V_i(k, \eta_0)) + \langle \sqrt{W_0} V_o(k, \eta_0) + \sqrt{W_2} V_i(k, \eta_0) \rangle^2, \quad (22)$$

where V_o and V_i are the unperturbed ("original") and ionospheric components of the expression, respectively, and $\eta_0 = \eta(0)$ is the lowest spatial line-of-sight mode (the DC term). The power spectrum is then

$$P_{\mathrm{ani}} = \frac{1}{\Omega} \langle V(k, \eta)^* V(k, \eta) \rangle. \quad (23)$$

Note that this model corresponds to a "snapshot" observation of the ionosphere, during which the phase of the ionospheric structures has not evolved. In general, Jordan et al. (2017) observe anisotropic structures that are persistent in structure (amplitude and wavelength) but evolve in phase as they rotate relative to the celestial sphere (fixed in the geocentric frame, above the MRO site). For extended observations of the same field, this phase evolution will decohere sources, destroying power as the sources change apparent position. This can be quantified by modifying the expected visibility with an integral over an evolving phase term:

$$\langle V(u, v) \rangle = \int_{t=0}^{T} dt \frac{\alpha S_{\mathrm{max}}^{2-\beta}}{2 - \beta} 2\pi \sigma_\nu^2 \exp(-2\pi^2 v^2 \sigma_\nu^2) \, G(t), \quad (24)$$

with

$$G(t) = \exp(-2\pi^2 u^2 \sigma_\nu^2)$$
$$+ \frac{\epsilon(\nu)}{2} (\exp(i\phi(t)) \exp(-2\pi^2 \sigma^2 (u - k_l)^2)$$
$$- \exp(-i\phi(t)) \exp(-2\pi^2 \sigma^2 (u + k_l)^2)). \quad (25)$$

This term acts to decohere the visibility phases, destroying power at the characteristic scale of the structures, with maximal decoherence when the phase evolves by π over the coherently summed observation. For the data sets in question, the phase evolutions observed are $\Delta\phi \lesssim 0.5$.

Full model: With these two analytic expressions, Equations (13) and (22), for the contributions of the two components of the observed ionospheric activity, the expected impact due to each and due to a weighted combination of the two can be explored.

3.2. Estimating Ionospheric Parameters from Calibration Solutions

As already described, the current MWA EoR algorithms use the calibration and peeling information for the brightest apparent point sources in the field to form an assessment of ionospheric activity in each observation (through the

---

[16] Expanding the variance and squared expectation value, and propagating the variance through the Fourier transform, we obtain

$$\langle V(k, \eta)^* V(k, \eta) \rangle = \mathcal{F}^\dagger \langle V(k, \nu)^* V(k, \nu) \rangle \mathcal{F}$$
$$= \mathcal{F}^\dagger \mathrm{Var}(V_i + V_o) \mathcal{F} + \langle \mathcal{F}V_i + \mathcal{F}V_o \rangle^2$$
$$= \mathcal{F}^\dagger \mathrm{Var}(V_i) \mathcal{F} + \mathcal{F}^\dagger \mathrm{Var}(V_o) \mathcal{F} + \langle \mathcal{F}V_i \rangle^2 + \langle \mathcal{F}V_o \rangle^2 + 2\langle \mathcal{F}V_i \rangle \langle \mathcal{F}V_o \rangle$$
$$= W_0 \mathrm{Var}(V_o(k, \eta_0) + W_2 V_i(k, \eta_0)) + \langle \sqrt{W_0} V_o(k, \eta_0) + \sqrt{W_2} V_i(k, \eta_0) \rangle^2.$$

---

ionospheric metric of Jordan et al. 2017). They do not then use this information to reconstruct a TEC model to apply to the residual data for correction. Instead, the contaminated data are discarded from the analysis.

However, in principle, these records of apparent source positional shifts can be used to estimate the parameters of an ionospheric model and apply these to "de-warp" data in the image plane, allowing the observations to be included in EoR analysis. Such algorithms have, and are, being developed for LOFAR (SPAM; Intema 2009; Prasad et al. 2014), MWA (Hurley-Walker et al. 2017; Rioja et al. 2018), and SKA (internal SKA Resolution Teams and Tasse et al. 2013; Trott 2015; Cornwell 2016; Rioja et al. 2018). Such an approach has not yet been required by the MWA EoR projects, where the baselines are relatively short, the data set is large, and only a small fraction were found to contain ionospheric activity. However, for future precision 21 cm science with long-baseline arrays (e.g., SKA), this will become an imperative.

As such, we estimate the precision with which the parameters of an ionospheric model could theoretically be estimated with the information available in the MWA calibration solutions, using a Fisher Analysis and the Cramer–Rao Bound (CRB; Kay 1998). We apply this specifically to the MWA case but show scalings for a general interferometer.

Our measurement set is the residual visibility set for each baseline, where the model is fitted to the difference between the expected (catalog) and measured positions of sources in the field:

$$V_{\mathrm{mod}}(\boldsymbol{u}, ; \nu) = V_{\mathrm{meas}}(\boldsymbol{u}; \nu) - V_{\mathrm{catal}}(\boldsymbol{u}; \nu), \quad (26)$$

where

$$V_{\mathrm{meas}}(\boldsymbol{u}; \nu) = \sum_{i=0}^{N_p} S_i B_i(\boldsymbol{l} + \Delta\boldsymbol{l}) \exp(-2\pi i(u(\boldsymbol{l} + \Delta\boldsymbol{l})))$$
$$+ \sum_{i=N_p}^{N_T} S_i B_i(\boldsymbol{l} + \Delta\boldsymbol{l}) \exp(-2\pi i(\boldsymbol{u} \cdot (\boldsymbol{l} + \Delta\boldsymbol{l}))), \quad (27)$$

$$V_{\mathrm{catal}}(\boldsymbol{u}; \nu) = \sum_{i=0}^{N_p} S_i B_i(\boldsymbol{l}) \exp(-2\pi i(\boldsymbol{u} \cdot \boldsymbol{l})). \quad (28)$$

Here, the measured data include the N_p sources that are used for calibration and source peeling, and these data yield and yield ionospheric model information, and the remaining sky of $N_T - N_p$ unmodeled sources. The latter are fundamentally confused sources and will be treated as an additional noise term in the Fisher Analysis according to a statistical description of their contribution to visibility noise. Also, $\Delta\boldsymbol{l}$ contains the vector apparent positional offsets of each source due to the ionospheric phase screen, and this variable encodes the ionospheric parameters to be fitted:

$$\Delta\boldsymbol{l}(\nu) = \epsilon(\nu) \sin(2\pi(k_l l + k_l m))$$
$$= \epsilon_0 \sin(2\pi(k_l l + k_l m)) \left(\frac{\nu}{\nu_0}\right)^{-2}, \quad (29)$$

with $\vec{\theta} = [\epsilon_0, k_l, k_m]$ as the vector of parameters to be estimated for the CRB in the case of duct-like behavior. Preempting the results of future sections, we restrict our attention to this ionospheric type because we find it to have impact on the EoR

experiments and because it provides a deterministic model for the Fisher Analysis.

The CRB yields the minimum estimation precision for parameter estimation for an unbiased and optimal estimator (which may not exist) where all of the information is used efficiently to estimate model parameters. The Maximum Likelihood Estimator is asymptotically efficient (for large data sets and signal-to-noise ratios [S/Ns]) and reduces to a weighted least-squares estimator when the data are Gaussian distributed. For data that are distributed as a multivariate generalized Gaussian, the Fisher Information is encased in an $N_\theta \times N_\theta$ matrix $\mathcal{I}$, where N_θ is the number of unknown parameters. The ijth element of this matrix is given by

$$\mathcal{I}_{ij} = \left(\frac{\partial \boldsymbol{V}_{\mathrm{mod}}}{\partial \theta_i}\right)^\dagger \mathcal{C}^{-1}\left(\frac{\partial \boldsymbol{V}_{\mathrm{mod}}}{\partial \theta_j}\right), \tag{30}$$

where the sum over the data (baselines) is encoded in the vector model visibility derivatives and $\mathcal{C}$ is the data covariance matrix. This form is appropriate for the case where no unknown parameters are contained within the data covariance model. Before describing the form of the covariance, we extend the expression for the Fisher Information specifically for the current case where the parameters are encoded within the individual source position shifts:

$$\mathcal{I}_{ij} = \left(\frac{\partial \boldsymbol{V}_{\mathrm{mod}}}{\partial \Delta \boldsymbol{l}}\frac{\partial \Delta \boldsymbol{l}}{\partial \theta_i}\right)^\dagger \mathcal{C}^{-1}\left(\frac{\partial \boldsymbol{V}_{\mathrm{mod}}}{\partial \Delta \boldsymbol{l}}\frac{\partial \Delta \boldsymbol{l}}{\partial \theta_j}\right). \tag{31}$$

The data covariance matrix in its simplest form carries the information on the thermal (radiometric) noise uncertainty on each datum (visibility noise in Jy indexed by baseline and frequency):

$$\sigma_{\mathrm{therm}} = \frac{2k_B T_{\mathrm{sys}}(\nu)}{\sqrt{\Delta\nu\Delta t}} \text{ Jy}, \tag{32}$$

with $\Delta\nu$ and Δt describing the spectral and temporal resolution of each measurement, and with the system temperature (sky dominated at these frequencies) described by (Furlanetto et al. 2006)[17]

$$T_{\mathrm{sys}}(\nu) \simeq 180 \text{ K}\left(\frac{\nu_{\mathrm{MHz}}}{180}\right)^{-2.6}. \tag{33}$$

In addition to this measurement noise, the estimation of ionospheric parameters occurs within a data set that contains an unmodeled confused source background. For a single frequency, the point-source contribution to the data covariance matrix is uncorrelated between baselines and reduces to (cf. Equation (19))

$$\mathrm{Var}(V(u, v)) = \frac{\alpha S_{\mathrm{max}}^{3-\beta}}{3-\beta}\pi\sigma_\nu^2 \text{ Jy}^2, \tag{34}$$

with S_{max} corresponding to the peeling limit and σ_ν again representing the primary beam width. This is determined by the dominant image noise term (either the confusion level or

---

[17] A more accurate form is (Braun 2013)

$$T_{\mathrm{sys}}(\nu) \approx 20(408/\nu_{\mathrm{MHz}})^{2.75} + 2.73 + 288(0.005$$
$$+ 0.1314 \exp 8(\ln(\nu_{\mathrm{MHz}}/1000) - \ln(22.23)) \text{ K}.$$

Table 1
Observational Parameters for Estimating Ionospheric Parameters for the MWA EoR Experiment and Output Amplitude Signal-to-noise Ratio

Parameter	Value	Value
	150 MHz	100 MHz
S_{max}	200 mJy	300 mJy
N_p	1200	1600
N_a	128	128
Δt	120 s	120 s
$\Delta\nu$	30.72 MHz	30.72 MHz
$\frac{\epsilon_0}{\Delta\epsilon_0}$	200	130

radiometric noise). In the case of thermal noise dominance and considering an estimation threshold of F times the noise, the maximum peeling flux density is

$$S_{\mathrm{max}} = F\frac{\sigma_{\mathrm{therm}}}{\sqrt{N_a(N_a-1)/2}} \text{ Jy} \tag{35}$$

for N_a antennas. This can be used to estimate the number of sources that are peeled (and therefore used in the estimation) and their flux density distribution according to a broken power-law parameterization of the source number density (where the exponent of the flux density steepens from $\beta \simeq -1.59$ to -2.5 above 1 Jy, yielding fewer bright sources).

The full data covariance therefore reduces to a variance, and each diagonal entry is given by

$$\mathcal{C}_{ii} = \sigma_{\mathrm{therm}}^2 + \frac{\alpha S_{\mathrm{max}}^{3-\beta}}{3-\beta}\pi\sigma_\nu^2. \tag{36}$$

This general model is applied to the specific MWA case. Table 1 describes the relevant parameters. We implement this model and compute the CRB. Table 1 also contains the resulting S/N on the ionospheric amplitude, $\frac{\epsilon_0}{\Delta\epsilon_0}$. At both frequencies, the large number of sources available to be used for estimation yield high-precision estimates. Notably, the precision on the amplitude is factors of hundreds, corresponding to factors of 10^4 reduction in power bias. In the predictions for the MWA EoR experiment presented shortly in Section 5, we apply both the current approach (no attempt to correct residual data) and also a theoretical approach where the ionospheric measurements are used to produce an ionospheric model and residual data are corrected before formation of the power spectrum.

4. Results

4.1. Simulations

To test the appropriateness of the analytic models we have formed for the two ionospheric components, we perform simple simulations of perturbed point-source fields. Figure 1 displays the ratio (active divided by quiet), difference (active minus quiet), and fractional difference (difference power divided by unperturbed power) for a simple model with 20,000 point sources distributed across the sky and 40 sky realizations. Anisotropic ionospheric activity is imprinted by altering the probability of a source being located at a physical location, to match a sinusoidal distribution aligned with the m-axis (as in Equation (16)). The characteristic wavelength is
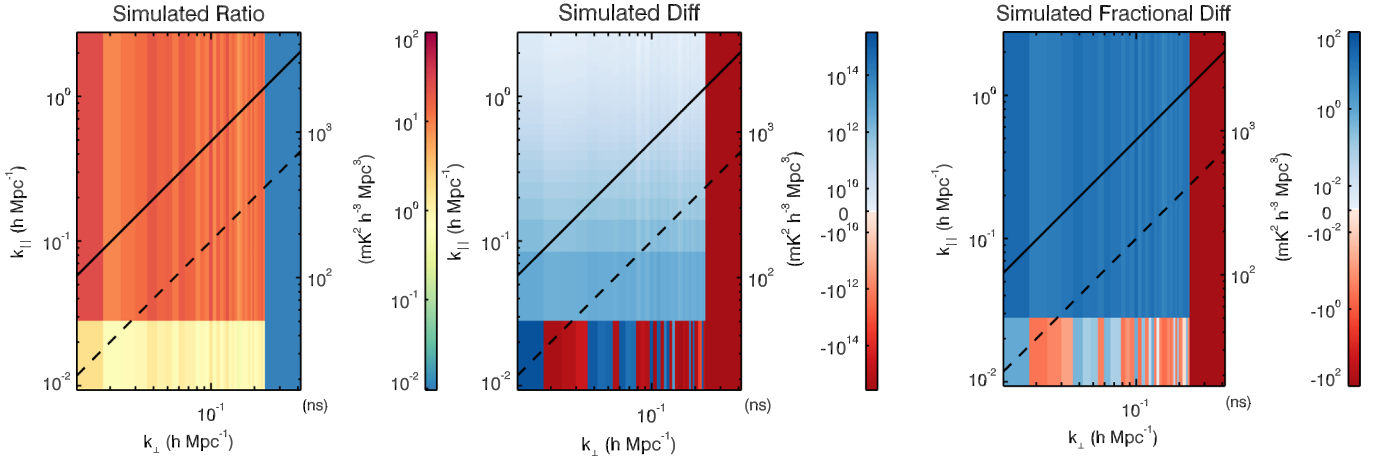
Figure 1. Simulated prediction for the ratio (active/quiet), difference (active–quiet), and fractional difference of the anisotropic ionosphere compared with the unperturbed ionosphere. The power is increased on all scales as the coherent structures concentrate the signal in the sky.
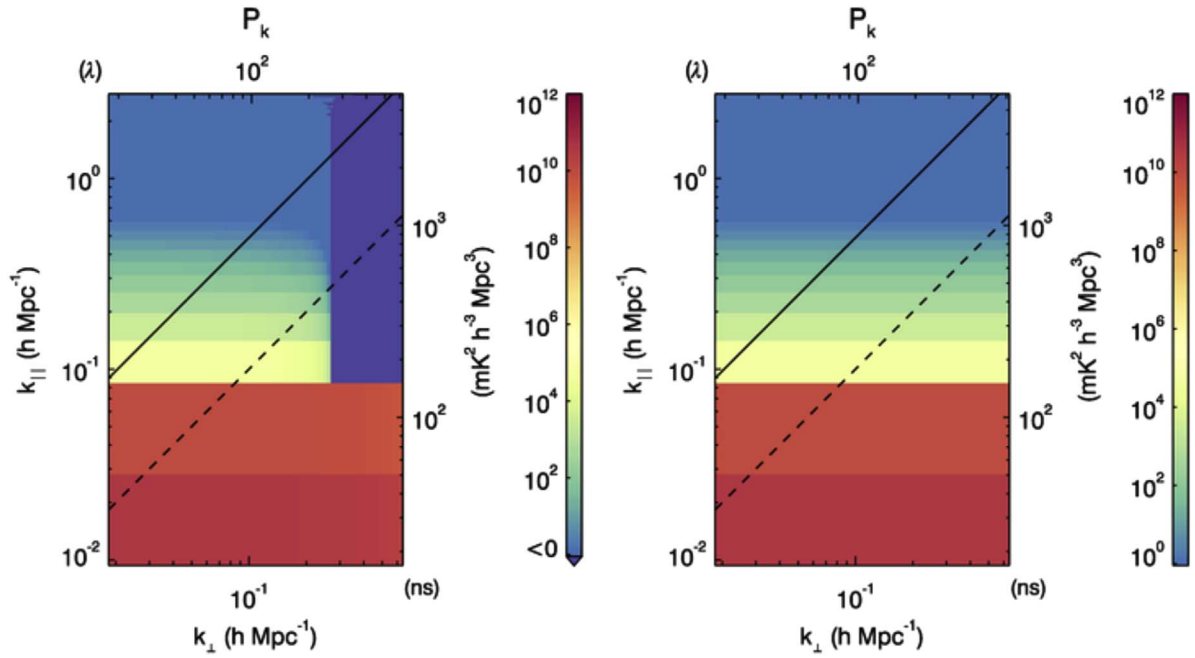

Figure 2. Analytic prediction for the 2D power spectrum of a turbulent ionosphere (left) and unperturbed ionosphere (right). The turbulent parameters set a scale of $1/u_0 = 2'$ and amplitude $K_0 = 5 \times 10^{-7}$.
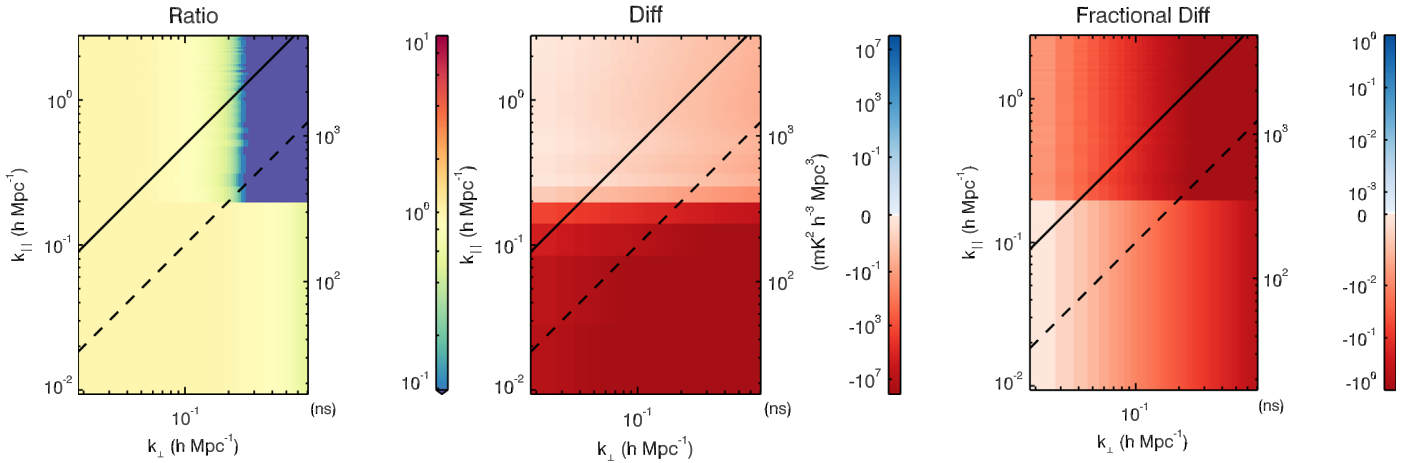

Figure 3. Analytic prediction for the ratio, difference, and fractional difference of the full turbulent ionosphere compared with the unperturbed ionosphere. The primary loss of power occurs on scales smaller than the characteristic turbulent scale.
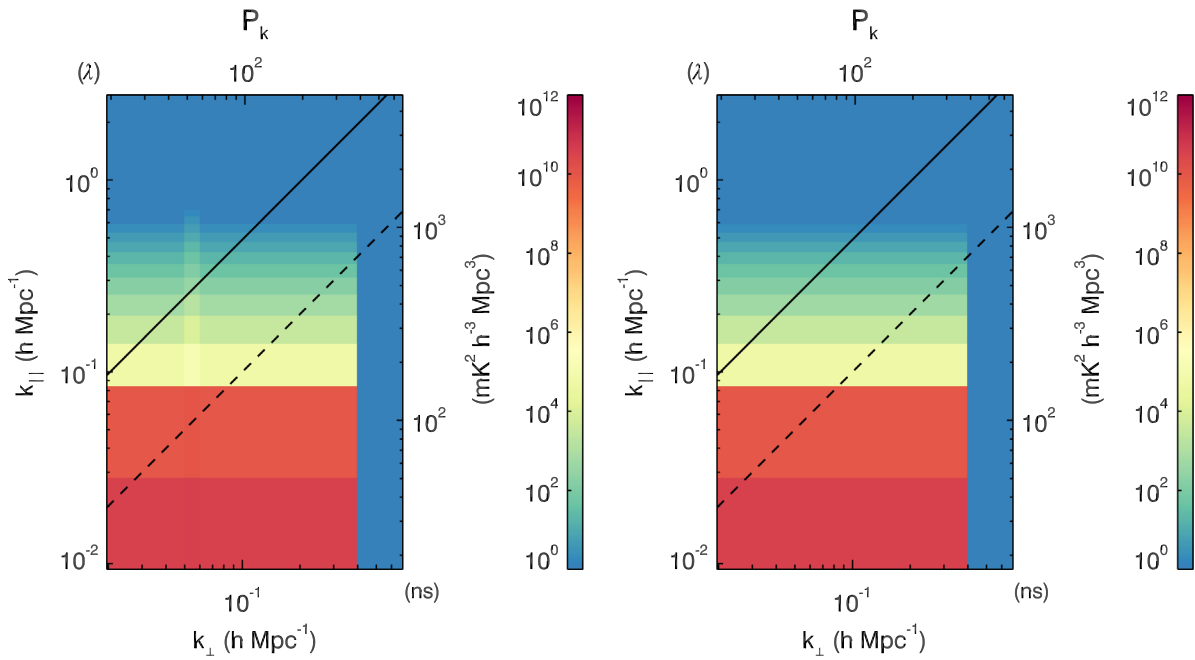
Figure 4. Analytic prediction for the power spectrum of a snapshot anisotropic ionospheric perturbation, characterized by coherent density variability of the foreground sources (left), compared with the unperturbed prediction (right). The characteristic scale is $\sim1^\circ.5$, with a 10% amplitude.
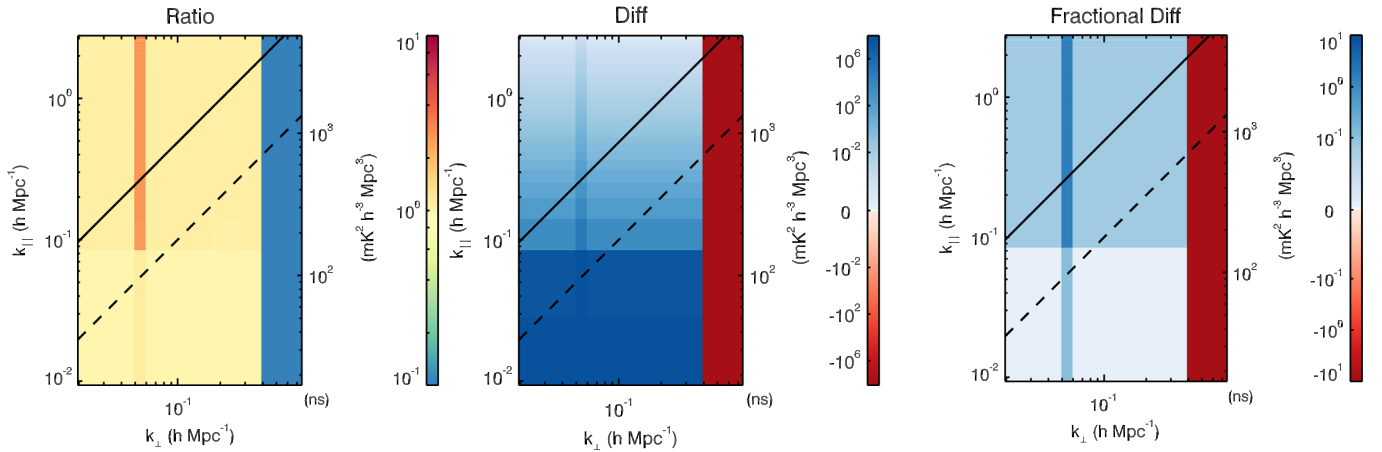
Figure 5. Analytic prediction for the ratio, difference, and fractional difference of the anisotropic ionosphere compared with the unperturbed ionosphere. The power is increased on all scales as the coherent structures concentrate the signal in the sky.

$k_l = 45$ ($1^\circ.5$), with a fractional amplitude of 10%. The power cutoff above $k_\perp = 0.4$ is due to the baseline sampling, not the ionosphere.

4.2. Analytic Predictions

The analytic predictions for the turbulent and structured components of the observed ionospheric activity can be compared with the unperturbed models. The turbulent component power spectra are displayed in Figures 2–3, showing the perturbed (left) and unperturbed (right) spectra for a characteristic scale of $1/u_0 = 2'$ and amplitude $K_0 = 5 \times 10^{-7}$, as well as their ratio and difference. This corresponds to an effective power cutoff at $1/u_{\rm cut} \sim 4'$, which is larger than the characteristic scale observed in the data ($1/u_{\rm cut} \sim 0.2'$) but allows us to explore how the turbulence manifests in the power spectrum.

The snapshot anisotropic component is similarly shown in Figures 4 and 5 with $k_l = 45$, corresponding to a scale of $\sim1^\circ.5$ and a 10% amplitude.

There are key differences between these two modes of activity. While the turbulent component is shown to destroy power on all scales, it is primarily scales smaller than the characteristic turbulent scale that are most affected. The fractional power loss on larger scales is very small and mostly corresponds to the differing spectral behavior. When considering data with thermal noise, the turbulent component is expected to have an impact only at smaller scales. The anisotropic mode, however, displays increased power, because the ducts are effectively concentrating signal on the sky, leading to less decoherence between (physically unassociated) foreground radio galaxies. This concentration occurs only in one dimension, spreading the increased power across all angular modes in the average of $k_\perp = \sqrt{u^2 + v^2}$.
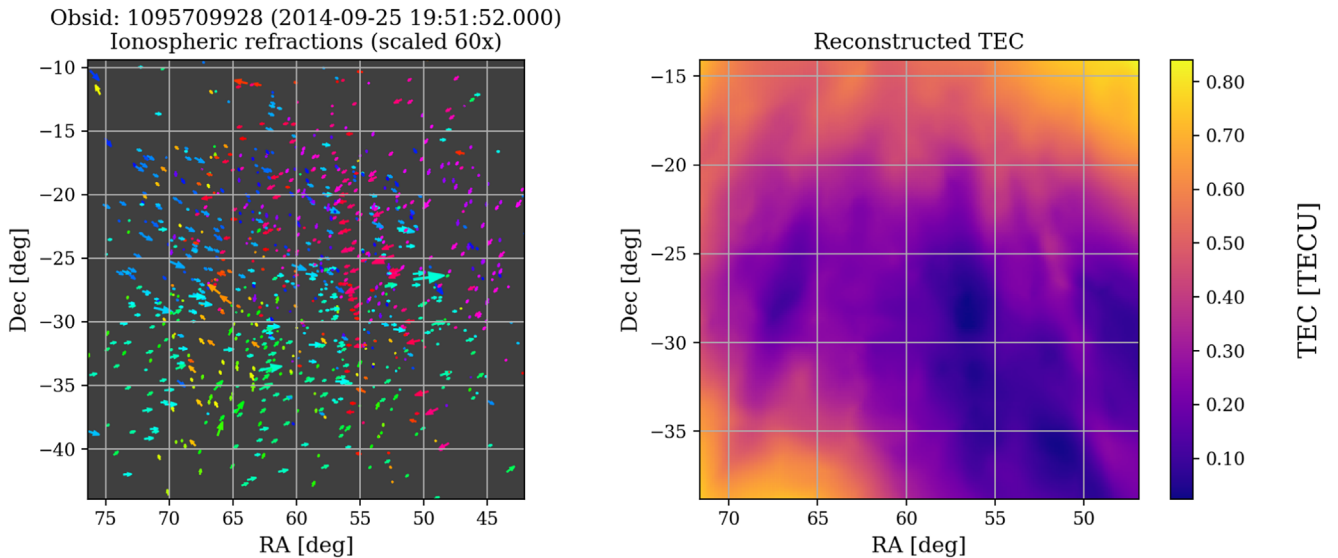
Figure 6. Example reconstruction of slant TEC field from the measured point-source deviations from catalog positions, for a 2-minute observation from 2014 September 25, displaying anisotropic duct-like behavior. Left: vectors of point-source positional offsets; right: reconstructed slant TEC scalar field (Jordan et al. 2017).

4.3. Comparison with Data

Figure 6 displays data from a zenith-pointed observation from 2014 September 25 of the EoR1 field (R.A. = 4^h, decl. = $-30°$) corresponding to anisotropic duct-like behavior. The apparent source offsets are used to reconstruct a slant TEC field across the FOV, from which an ionospheric metric of activity can be derived and a spatial power spectrum formed. In these 30 minutes of contiguous data, the coherent structures are seen to move with respect to the geocentric frame, but they mostly maintain their position in the celestial frame, consistent with a slowly moving local ionospheric structure above the observatory. The visibilities are formed with respect to a phase center that is fixed in the celestial frame, corresponding to an ionospheric phase screen that is constant with time for these data. Therefore, we would expect minimal decoherence of visibilities that are coherently combined over the 30 minutes. The observed wavelength of the structures is $\sim5°–8°$, corresponding to $k_l \sim 10$. Figure 7 shows the ratio and difference for these data, compared with a matching set from the same week with quiet conditions. A comparison analytic prediction for these parameters is shown in Figure 8. The prediction is broadly consistent with the data: the overall power is increased, with the primary change occurring at large scales, consistent with the duct structures observed in the reconstructed scalar TEC fields (see Figure 6).

5. Tolerances

Despite individual types of ionospheric conditions being observed at the MRO, it is the integrated effect of these that will decohere otherwise coherent visibility data and destroy power on scales corresponding to that for the ionospheric activity. Jordan et al. (2017) characterized four qualitatively distinct types of activity and their relative occupancy of EoR data from the 2015 observing season, but each type will affect the power spectrum in a different way. For the Type II (turbulent) mode, occupying 15% of that data set, the typical median positional offset is $\sim0.2'$, corresponding to a scale of little relevance to EoR science ($k_\perp = 16 \, h \, \mathrm{Mpc}^{-1}$). Because the turbulent mode primarily destroys power on scales smaller

than the characteristic scale, this mode is not likely to be relevant for the EoR experiment.[18] The anisotropic modes, however, routinely display activity on degree scales, and these require further investigation into their impact.

The analytic predictions for bias in the power spectrum space must be compared with expectations for cosmological signal strength for a given experiment, redshift range, and spatial scale. This allows tolerances to be set for the degree of ionospheric activity acceptable in the data before it is predicted to bias the experimental result. This can be applied to both the uncorrected and corrected data sets.

To estimate these tolerances, we used simulated cosmological power spectra as a function of spatial scale, and at different redshifts, from 21cmFAST (Mesinger et al. 2011) and assuming the favored faint galaxy model for reionization. These simulated power spectra are used as reference EoR signals. Power spectra are extracted for redshifts $z = 8.5$ and 13.2, and we employ the spherically averaged dimensionless power spectrum:

$$\Delta^2(k) = \frac{P(k)k^3}{2\pi^2} \mathrm{mK}^2. \tag{37}$$

Given that the ionosphere is not of a persistent structure over the full course of a $\sim$1000 hr experiment, the timescale of power bias is relevant. Observations using the MWA at the MRO in EoR data, reported by Jordan et al. (2017), demonstrate that each 4 hr observing night typically shows persistent structure (with a few notable exceptions where activity changes rapidly owing to external factors). Therefore, we assume that the ionospheric activity coherence time is equivalent to this timescale. We can therefore use the observed occupancy of different ionospheric types at the MRO to assess the impact on the EoR power spectrum. The instrument model considers the station size and array layout for the MWA Phase I (relevant for comparing with 2014 data), to compute the

---

[18] A separate question, not addressed here, is the ability of the direction-dependent calibration and peeling of bright sources to correctly measure this and remove power cleanly. This is likely to be a much larger effect than residual power loss from turbulence shifting the positions of weaker sources.
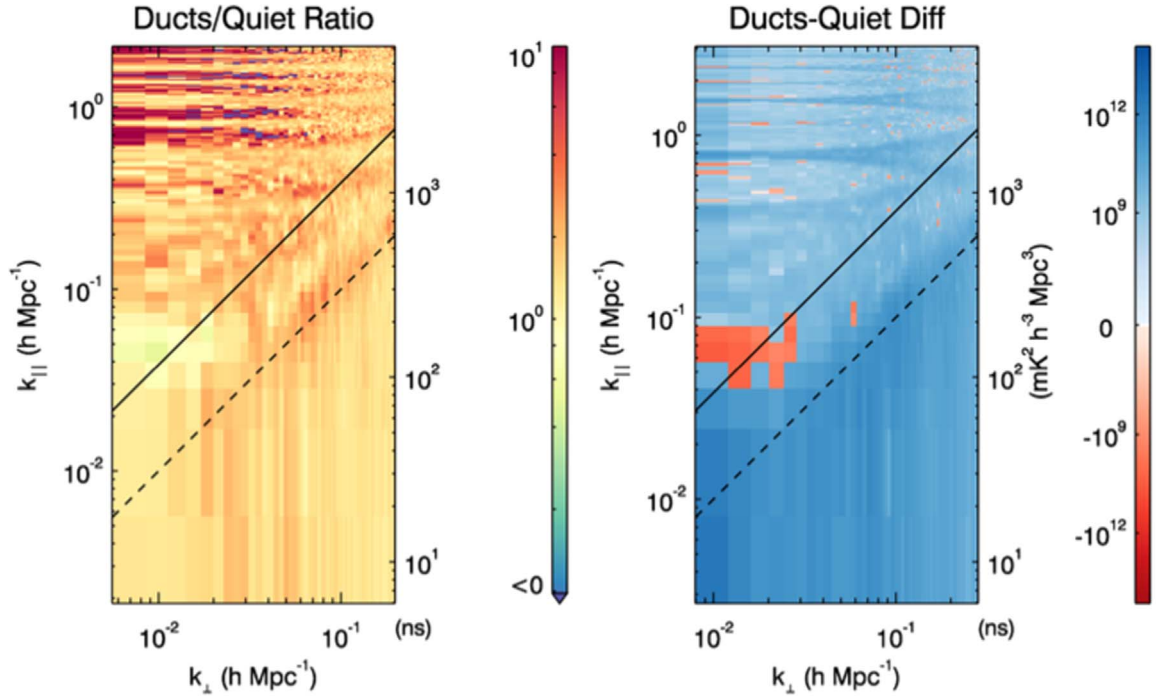
Figure 7. Ratio and difference of 30 minutes of zenith-pointed observations from 2014 September, during which the MWA observed Type IV (anisotropic structures) and Type I (quiet) conditions. Each data set is formed from data that are contiguous in time, with the same observation parameters, phase center, and a zenith-pointed beam.
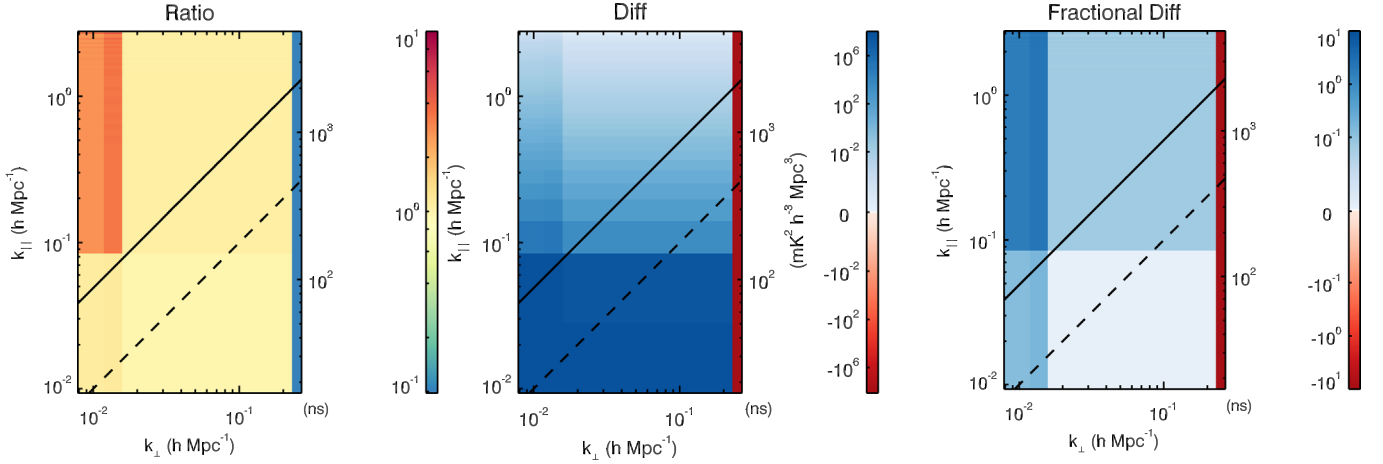

Figure 8. Analytic prediction for the ratio, difference, and fractional difference of the anisotropic ionosphere compared with the unperturbed ionosphere for structured ionospheric parameters consistent with that observed in the 2014 data.

relevant foreground signal in the data and the resolution of the *uv*-plane.

We start by considering the uncorrected residual signal model, as is currently implemented in the MWA EoR pipelines. The analytic model described in Sections 3.1 and 4.2 is used to compute the impact of ionospheric activity for the MWA experiments at relevant redshifts ($z = 8.5$ and 13.2, corresponding to 150 and 100 MHz, respectively). As a metric for the realistic impact of including active data, we compute the bias in the slope of the 1D 21 cm power spectrum when residual power from activity is included. As described above, reference EoR signal power is extracted from simulated models for reionization (Mesinger et al. 2011). We consider the relative observed fractions of each type of ionospheric activity at the MRO to construct the model observed power spectrum if all data were

included, for example, allowing 8% of the data to contain anisotropic structures and 15% to contain turbulent modes. We consider a 1000 hr experiment with a bandwidth of 10 MHz. Figures 9 and 10 display the cosmological signal alone (green), cosmological signal plus ionospheric foreground bias (black), the thermal noise level (red), and linear fits to the logarithmic slope in the region where the S/N exceeds unity (orange, blue). The structure in the biased power is due to the cylindrical averaging over modes with differing levels of foreground contamination. Omitting these modes ("working outside of the wedge") leaves few *k*-modes available with S/N > 1.

Also shown in these figures is the corrected case (purple), where we have propagated the uncertainties on the ionospheric amplitude and angular scale factors computed in Section 3.2 to the residual (post-data-correction) power spectrum. The
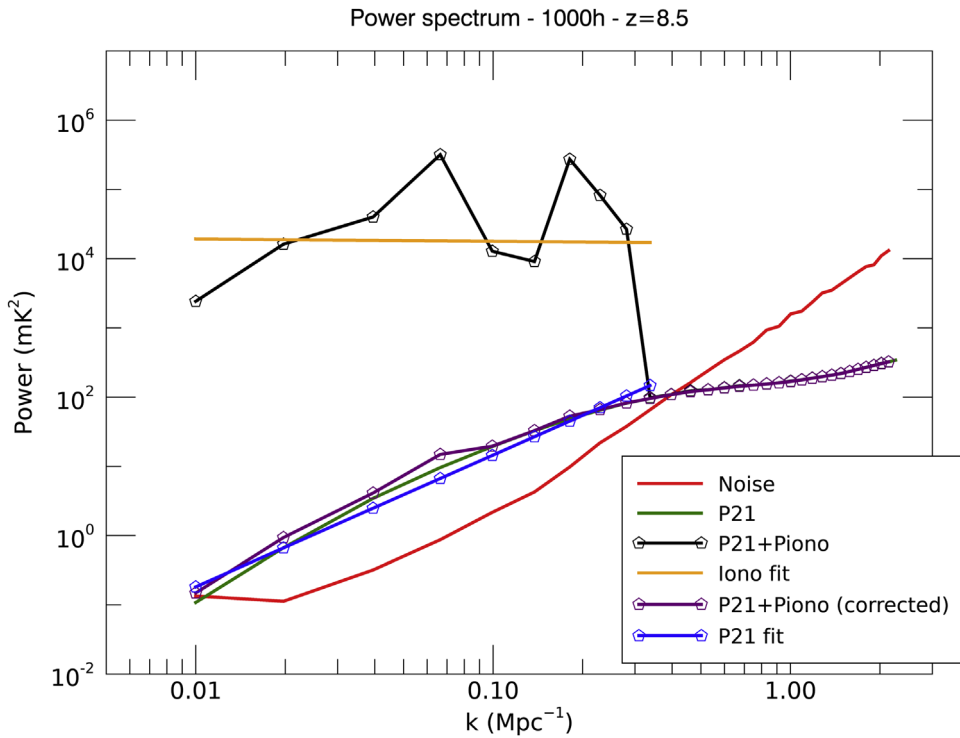
Figure 9. Spherically averaged dimensionless power spectrum for a simulated 21 cm EoR signal at $z = 8.5$ (150 MHz), when no ionospherically active data are included (green) and when 8% of data contain anisotropic activity (black = uncorrected; purple = corrected). Linear fits to the logarithmic slope in S/N > 1 regions are also displayed (orange, blue).
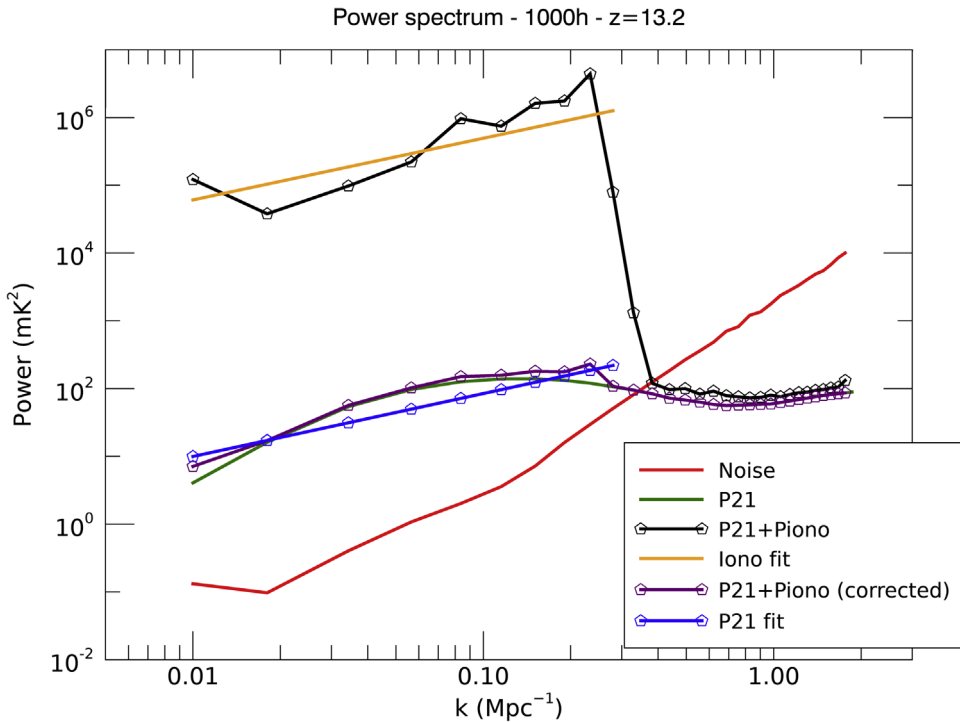

Figure 10. Spherically averaged dimensionless power spectrum for a simulated 21 cm EoR signal at $z = 13.2$ (100 MHz), when no ionospherically active data are included (green) and when 8% of data contain anisotropic activity (black = uncorrected; purple = corrected). Linear fits to the logarithmic slope in S/N > 1 regions are also displayed (orange, blue).

correction here is marginally adequate to be able to use these data for the EoR experiments.

At both redshifts, the ionospheric bias from including 8% of data from ionospherically active nights (assuming that the foregrounds are not treated with knowledge of the activity and

that no correction is applied) leads to systematic biases in the 21 cm parameter estimation. Note that both of these cases show extreme biases, implying that there are large fractions of k-space where the ionosphere makes the experiment impossible. However, these are for the extreme case where almost
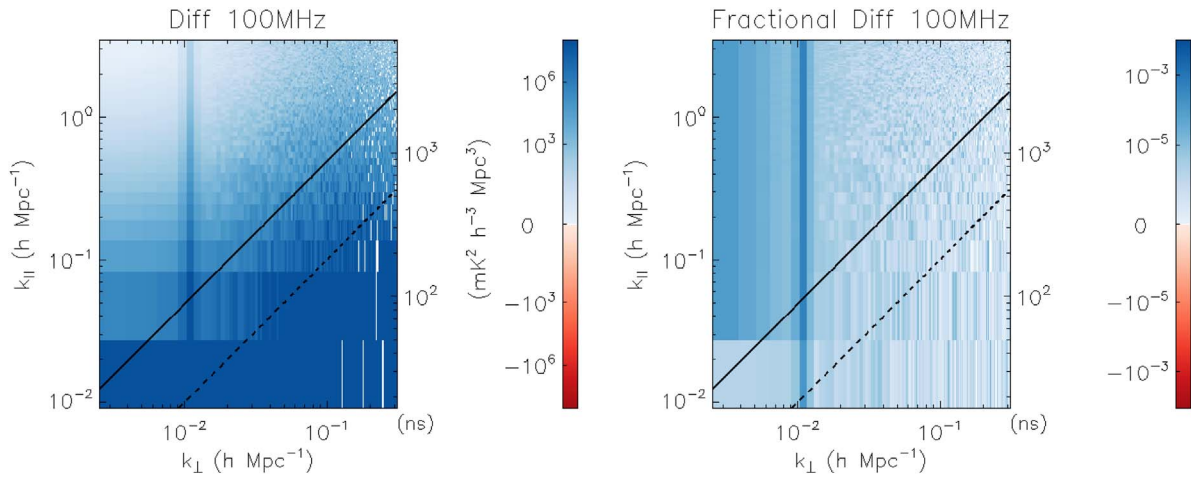
Figure 11. Difference (left) and fractional difference (right) for an MWA experiment at 100 MHz ($z = 13.2$), with anisotropic turbulence occurring at the amplitude observed in 2014 data.

10% of the data display coherent, structured behavior, and where no point sources have been correctly extracted. For a less extreme case where perhaps only 1% of the data contain structure and these data are included in the analysis, the bias will be a factor of 100 lower. In addition, if the brightest sources are well measured and subtracted cleanly, then the bias will again be reduced. Nonetheless, the potential for a large power bias implies that exclusion of data with active ionospheres is prudent for such a precision experiment.

Figure 11 further displays the difference and fractional difference for the uncorrected case at $z = 13.2$. Here the instrumental sampling has been included and the foreground wedge is present. Although there is a maximum of $10^6 \, \mathrm{mK^2 \, Mpc^3}$ difference in the 2D power spectrum, when averaged to 1D it corresponds to a factor of $10^4 \, \mathrm{mK^2}$ increase on EoR scales.

We therefore conclude that inclusion of ionospherically active data (where no ionospheric correction has been applied) will lead to systematic biases in the outputs of MWA EoR experiments.

When the information is used to estimate the ionospheric model parameters with an ideal estimator, and these are then used to correct the residual data, the bias is reduced by a factor of $\sim 10^4$, yielding acceptable levels for EoR experiments with the MWA.

6. Discussion and Conclusions

We have presented analysis of the bias in the EoR power spectrum due to unaccounted-for residual point-source foreground contribution when ionospheric activity is present. This analysis pertains to the case where the activity is assessed, but residual data (after peeling of bright sources) are not corrected and retain the imprint of the activity, and the case where the residual data are corrected with a model formed from the peeling information. In the former case, aside from the bias in the power, combining active and inactive data coherently (averaging visibilities) will lead to decoherence of the signal and will have smaller effects on the 21 cm power spectrum.

Of the two primary types of ionospheric activity identified at the Murchison Radio-astronomy Observatory in MWA EoR data sets (Type I: isotropic small-scale turbulence; Type IV: turbulent, anisotropic structure) and in other MWA data sets (Loi et al. 2016; Hurley-Walker et al. 2017; Rioja et al. 2018),

the scale of the turbulence is such that it should not affect EoR power spectra (where power is principally on larger scales), while anisotropic behavior (observed to occur in 8% of observations) imprints residual foreground power that can affect parameter estimation. We therefore conclude that data with anisotropic ionospheric activity should be omitted from deep EoR observations, and data showing turbulent behavior should be monitored to ensure that the spatial scales are not of relevance to the EoR.

In future, with the construction of the low-frequency component of the SKA at the MRO site, expected ionospheric activity can be guided by these results. We expect the same distribution of ionospheric classes (on average) to be present for the SKA-Low experiments. For the SKA, however, the much longer baselines demand ionospheric correction to be performed. While the MWA's short baselines allow for a single phase screen model to be fitted across the array, longer baselines demand more sophisticated models, where individual sources are observed through different ionospheric conditions across the array. The array has been designed to allow for enough ionospheric pierce points to be available to calibrate the instrument, particularly at the challenging lower frequency of 50 MHz. The HERA array, in South Africa, as the other upcoming EoR array, shares a similar compact configuration to the MWA, and its ionospheric calibration requirements are likely to be similar to the MWA and less stringent than SKA-Low.

Appendix
Turbulence Power Spectrum Model

While the effect of high-amplitude features of the ionosphere such as TIDs may be individually modeled, it is expected that a temporally dynamic low-amplitude statistical field exists at all times. The form of this field is further expected to exhibit Kolmogorov-like turbulence, i.e., an isotropic angular power spectrum with power-law dependence on angular modes u, with slope $\kappa \sim -5/3$. In this appendix, we describe the statistical effect of this turbulent field on the observed EoR + foreground signal.

We will exclusively work in the regime in which the flat-sky approximation holds, i.e., in which angular scale u and angle l are related as Fourier duals. We furthermore take the Fourier convention most popular in radio interferometry, i.e., $\hat{f} = \int f \exp(-iu \cdot l)$.

We denote the (scalar) TEC field by $\psi(l)$ and note that in the weak-field regime with which we are here concerned the first-order effect of the ionosphere is to impose a shift on the observed angular position of incoming sight lines. This shift can be written as

$$\Delta l = K(\nu)\vec{\nabla}\psi(l), \tag{38}$$

where $K \propto \nu^{-2}$ is a scaling constant.

Our problem bears immediate similarities to that of weak lensing of the CMB. In both cases, a background angular signal is shifted by the gradient of a scalar field. Thus, to solve our problem, we follow the derivation of Lewis & Challinor (2006; see especially Section 4.1). Briefly, the observed field (we shall denote the field S to signify flux density, though it is interchangeable for temperature) is defined by

$$\tilde{S}(l) = S(l + \Delta l). \tag{39}$$

The right-hand side here affords a series expansion that is a good approximation on our scales of interest for a weak ψ. Truncating the series at quadratic order and taking the Fourier transform, while maintaining independence of the S and ψ fields, yields an observed power spectrum (see Lewis & Challinor 2006, Equation (4.12))

$$P_{\text{turb}} \approx (1 - K(\nu)u^2 R^\psi)P_u^S + K(\nu)P_u^\psi \star P_u^S, \tag{40}$$

with $\star$ indicating convolution and

$$R^\psi = \pi \int u^3 P_u^\psi \, du. \tag{41}$$

The final term in Equation (40) is subdominant owing to the weak TEC field. Indeed, when the field is strong and the convolution becomes important, the approximations made here themselves become invalid. We thus ignore the convolution for the remainder of this analysis.

The dominant effect of the turbulent ionosphere then is to destroy small-scale power. Specifically, the observed power falls to zero at a scale

$$u' = \sqrt{\frac{1}{K(\nu)R^\psi}}. \tag{42}$$

The cutoff is rather sharp (in log space), and we propose that the most effective way to treat the ionospheric distortion is as a sharp cut in EoR window at some k_{max}.

We suppose that the power spectrum of ψ is Kolmogorov-like but has some scale above which the power is truncated. In practice, setting this scale above the resolution limit of an instrument serves to eliminate sensitivity to this choice. Thus, we consider

$$P_u^\psi = \begin{cases} \left(\dfrac{u}{u_0}\right)^\kappa & u \leqslant u_{\text{max}} \\ 0 & u > u_{\text{max}}. \end{cases} \tag{43}$$

In this case, we have

$$R^\psi = \frac{u_{\text{max}}^{4+\kappa}}{(4+\kappa)u_0^\kappa}. \tag{44}$$

Suppose we wish to place a hard cut on the spectrum at u_f, when the power is diminished by a fraction f of its original amplitude. Further suppose that we set u_{max} to be the limiting resolution of our telescope at a given frequency. The natural question to ask then is how the scale u_f compares to u_{max}, by considering their ratio. If it is greater than unity, then we need not worry about the effects of the ionosphere, whereas if the opposite is true, then the ionosphere begins to encroach on the observed modes. This scale can be written as

$$u_f = \sqrt{\frac{f(4+\kappa)}{K(\nu)}\left(\frac{u_0}{u_{\text{max}}}\right)^\kappa \frac{1}{u_{\text{max}}^3}}. \tag{45}$$

ORCID iDs

Cathryn M. Trott ⓘ https://orcid.org/0000-0001-6324-1766
C. H. Jordan ⓘ https://orcid.org/0000-0002-1220-2940
S. G. Murray ⓘ https://orcid.org/0000-0003-3059-3823
D. A. Mitchell ⓘ https://orcid.org/0000-0002-1828-1969
R. B. Wayth ⓘ https://orcid.org/0000-0002-6995-4131
A. Beardsley ⓘ https://orcid.org/0000-0001-9428-8233
J. Bowman ⓘ https://orcid.org/0000-0002-8475-2036
B. J. Hazelton ⓘ https://orcid.org/0000-0001-7532-645X
D. Jacobs ⓘ https://orcid.org/0000-0002-0917-2269
M. F. Morales ⓘ https://orcid.org/0000-0001-7694-4030
R. Subrahmanyan ⓘ https://orcid.org/0000-0001-9913-900X
M. Tegmark ⓘ https://orcid.org/0000-0001-7670-7190
S. J. Tingay ⓘ https://orcid.org/0000-0002-8195-7562
R. L. Webster ⓘ https://orcid.org/0000-0002-5325-2709
J. S. B. Wyithe ⓘ https://orcid.org/0000-0001-7956-9758

References

Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., & Pober, J. C. 2016, MNRAS, 461, 3135
Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, ApJ, 833, 102
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. 2018, Natur, 555, 67
Braun, R. 2013, A&A, 551, A91
Cornwell, T. 2016, SDP Memo Ser., SKA-TEL-SDP-0000088
DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
de Gasperin, F., Mevius, M., Rafferty, D. A., Intema, H. T., & Fallows, R. A. 2018, A&A, 615, 179
Ellingson, S. W., Clarke, T. E., Cohen, A., et al. 2009, IEEEP, 97, 1421
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117

Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
Hurley-Walker, N., Callingham, J. R., Hancock, P. J., et al. 2017, MNRAS, 464, 1146
Hurley-Walker, N., & Hancock, P. J. 2018, arXiv:1808.08017
Intema, H. T. 2009, PhD thesis, Leiden Univ.
Intema, H. T. 2014, arXiv:1402.4889
Jacobs, D. C., Hazelton, B. J., Trott, C. M., et al. 2016, ApJ, 825, 114
Jordan, C. H., Murray, S., Trott, C. M., et al. 2017, MNRAS, 471, 3974
Kay, S. M. 1998, Fundamentals of Statistical Signal Processing: Detection Theory (Englewood Cliffs, NJ: Prentice-Hall)
Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14), ed. T. L. Bourke et al. (Trieste: POS), 001
Lewis, A., & Challinor, A. 2006, PhR, 429, 1
Loi, S. T., Cairns, I. H., Murphy, T., et al. 2016, JGRA, 121, 1569
Martí-Vidal, I., Guirado, J. C., Jiménez-Monferrer, S., & Marcaide, J. M. 2010, A&A, 517, A70
Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955
Mevius, M., van der Tol, S., Pandey, V. N., et al. 2016, RaSc, 51, 927
Mitchell, D. A., Greenhill, L. J., Wayth, R. B., et al. 2008, ISTSP, 2, 707

Mortlock, D. 2016, ASSL, 423, 187
Murray, S. G., Trott, C. M., & Jordan, C. H. 2017, ApJ, 845, 7
Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, AJ, 139, 1468
Patil, A. H., Yatawatta, S., Zaroubi, S., et al. 2016, MNRAS, 463, 4317
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Prasad, P., Wijnholds, S. J., Huizinga, F., & Wijers, R. A. M. J. 2014, A&A, 568, A48
Rioja, M. J., Dodson, R., & Franzen, T. M. O. 2018, MNRAS, 478, 2337
Tasse, C., van der Tol, S., van Zwieten, J., van Diepen, G., & Bhatnagar, S. 2013, A&A, 553, A105
Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, 7
Trott, C. M. 2015, SKA Memo Ser., 154
Trott, C. M., Pindor, B., Procopio, P., et al. 2016, ApJ, 818, 139
Trott, C. M., & Wayth, R. B. 2016, PASA, 33, e019
van der Tol, S., Jeffs, B. D., & van der Veen, A.-J. 2007, ITSP, 55, 4497
van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
Vedantham, H. K., & Koopmans, L. V. E. 2016, MNRAS, 458, 3099
Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015, PASA, 32, e025
Wijnholds, S., van der Tol, S., Nijboer, R., & van der Veen, A. J. 2010, ISPM, 27, 30