Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **520,** 2094–2108 (2023) https://doi.org/10.1093/mnras/stad191
Advance Access publication 2023 January 19

Towards 21-cm intensity mapping at $z = 2.28$ with uGMRT using the tapered gridded estimator – II. Cross-polarization power spectrum

Kh. Md. Asif Elahi [1]★ Somnath Bharadwaj,[1]★ Abhik Ghosh [2] Srijita Pal,[1] Sk. Saiyad Ali,[3] Samir Choudhuri,[4] Arnab Chakraborty,[5] Abhirup Datta,[6] Nirupam Roy,[7] Madhurima Choudhury [6,8] and Prasun Dutta[9]

[1]*Department of Physics and Centre for Theoretical Studies, IIT Kharagpur, Kharagpur 721302, West Bengal, India*
[2]*Department of Physics, Banwarilal Bhalotia College, Asansol 713303, West Bengal, India*
[3]*Department of Physics, Jadavpur University, Kolkata 700032, West Bengal, India*
[4]*Centre for Strings, Gravitation and Cosmology, Department of Physics, Indian Institute of Technology Madras, Chennai 600036, Tamil Nadu, India*
[5]*Department of Physics and McGill Space Institute, McGill University, Montreal, QC H3A 2T8, Canada*
[6]*Discipline of Astronomy, Astrophysics and Space Engineering, Indian Institute of Technology Indore, Indore 453552, Madhya Pradesh, India*
[7]*Department of Physics, Indian Institute of Science, Bangalore 560012, Karnataka, India*
[8]*ARCO (Astrophysics Research Center), Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel*
[9]*Department of Physics, IIT (BHU), Varanasi 221005, Uttar Pradesh, India*

Accepted 2023 January 15. Received 2023 January 6; in original form 2022 October 23

ABSTRACT

Neutral hydrogen (H I) 21-cm intensity mapping (IM) offers an efficient technique for mapping the large-scale structures in the Universe. We introduce the 'Cross' Tapered Gridded Estimator (Cross TGE), which cross-correlates two cross-polarizations (RR and LL) to estimate the multifrequency angular power spectrum $C_\ell(\Delta\nu)$. We expect this to mitigate several effects like noise bias, calibration errors, etc., which affect the 'Total' TGE that combines the two polarizations. Here, we apply the Cross TGE on 24.4-MHz-bandwidth uGMRT (upgraded Giant Metrewave Radio Telescope) Band 3 data centred at 432.8 MHz aiming H I IM at $z = 2.28$. The measured $C_\ell(\Delta\nu)$ is modelled to yield maximum likelihood estimates of the foregrounds and the spherical power spectrum $P(k)$ in several k bins. Considering the mean squared brightness temperature fluctuations, we report a 2σ upper limit $\Delta^2_{\rm UL}(k) \leq (58.67)^2\,{\rm mK}^2$ at $k = 0.804\,{\rm Mpc}^{-1}$, which is a factor of 5.2 improvement on our previous estimate based on the Total TGE. Assuming that the H I traces the underlying matter distribution, we have modelled $C_\ell(\Delta\nu)$ to simultaneously estimate the foregrounds and $[\Omega_{\rm H\,I} b_{\rm H\,I}]$, where $\Omega_{\rm H\,I}$ and $b_{\rm H\,I}$ are the H I density and linear bias parameters, respectively. We obtain a best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2 = 7.51 \times 10^{-4} \pm 1.47 \times 10^{-3}$ that is consistent with noise. Although the 2σ upper limit $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.061$ is ∼50 times larger than the expected value, this is a considerable improvement over earlier works at this redshift.

Key words: methods: data analysis – methods: statistical – techniques: interferometric – diffuse radiation – large-scale structure of Universe.

1 INTRODUCTION

The 21-cm line emission due to the hyperfine 'spin flip' transition of the neutral hydrogen (H I) atom is a unique observational probe to the high-redshift Universe. It can be used to study the large-scale structures in the Universe, probe the ionization state of the intergalactic medium, constrain the dark energy equation of state, put independent limits on various cosmological parameters, and quantify non-Gaussianity (Bharadwaj & Ali 2005; Loeb & Wyithe 2008; Mao et al. 2008; Wyithe, Loeb & Geil 2008; Bharadwaj, Sethi & Saini 2009; Visbal, Loeb & Wyithe 2009; Bagla, Khandai & Datta 2010; Morales & Wyithe 2010; Ansari et al. 2012; Hazra & Guha Sarkar 2012; Pritchard & Loeb 2012; Battye et al. 2013; Bull et al. 2015b; Long et al. 2022). At post-epoch of reionization (post-EoR; $z < 6$), the vast majority of H I is resided in discrete high-density clouds having column densities larger than 2×10^{20} atoms cm^{-2} (Wolfe et al. 1995; Ho, Bird & Garnett 2021). Instead of resolving these faint, discrete objects individually, the H I intensity mapping (IM) approach integrates the 21-cm emission over the large volumes of observation (Bharadwaj, Nath & Sethi 2001; Bharadwaj & Sethi 2001; Bharadwaj & Srikant 2004), enabling a full three-dimensional map of the H I distribution.

Several low-redshift ($z < 1$) single dish experiments (e.g. Pen et al. 2009a; Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013; Anderson et al. 2018; Wolz et al. 2021; Cunnington et al. 2023) have cross-correlated IM signal with optical galaxy redshift surveys (e.g. DEEP2; Newman et al. 2013) to constrain the H I distribution. Recently, cross-correlating with the eBOSS (Dawson et al. 2016) galaxy catalogues, the CHIME[1] (CHIME Collaboration

★ E-mail: asifelahi999@gmail.com (KMAE); somnath@phy.iitkgp.ac.in (SB)

[1]https://chime-experiment.ca/en/

2022a) interferometer has detected the 21-cm signal in the redshift range $0.78 < z < 1.43$ (CHIME Collaboration 2022b). However, an autocorrelation (i.e. not in cross-correlation with other probes) detection of the redshifted 21-cm signal is yet to be made. One of the primary science goals of the ongoing and upcoming IM experiments, such as BINGO[2] (Wuensche 2019), HIRAX[3] (Newburgh et al. 2016), MeerKAT[4] (Kennedy & Bull 2021), and the Tianlai project[5] (Chen 2012), is to measure the baryon acoustic oscillation (BAO) in the post-EoR 21-cm power spectrum (PS). Along with BAO, the next-generation IM surveys with the SKA[6] (Bull et al. 2015a) and recently upgraded OWFA[7] (Subrahmanya, Manoharan & Chengalur 2017) hold the promise to extract an ample amount of cosmological information through the 21-cm PS.

As a way forward in this direction, a few upper limits on the amplitude of the redshifted H I signal have been placed using the Giant Metrewave Radio Telescope (GMRT[8]; Swarup et al. 1991). Ghosh et al. (2011a, b) have used 610-MHz ($z = 1.32$) GMRT data to constrain $[\Omega_{\rm H\,I} b_{\rm H\,I}] < 0.11$ at 3σ level, where $\Omega_{\rm H\,I}$ is the comoving H I mass density in units of the present critical density and $b_{\rm H\,I}$ is the H I bias parameter. More recently, using the upgraded GMRT (uGMRT; Gupta et al. 2017) data, Chakraborty et al. (2021, hereafter Ch21) have put multiredshift constraints on the amplitude of 21-cm PS $\Delta^2_{\rm UL}(k) \leq (58.87)^2$, $(61.49)^2$, $(60.89)^2$, and $(105.85)^2\,{\rm mK}^2$ at $k = 1\,{\rm Mpc}^{-1}$. These values translate to the upper limits of $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.09$, 0.11, 0.12, and 0.24 at $z = 1.96$, 2.19, 2.62, and 3.58, respectively.

The biggest challenge to a high-redshift 21-cm IM experiment is perhaps the foregrounds (FGs) that are 4–5 orders of magnitude brighter than the predicted signal (e.g. Shaver et al. 1999; Di Matteo et al. 2002; Santos, Cooray & Knox 2005; Ali, Bharadwaj & Chengalur 2008; Ali & Bharadwaj 2014). The diffuse galactic synchrotron emission (DGSE) from our Galaxy and the extragalactic point sources (EPSs) – which are the diffused emission from the external galaxies – are the most dominant FG components considering the post-EoR observations (Haslam et al. 1981, 1982; Reich & Reich 1988; Condon 1989; Cress et al. 1996; Wilman et al. 2003; Blake, Ferreira & Borrill 2004; Owen & Morrison 2008; Singal et al. 2010; Condon et al. 2012; Randall et al. 2012; Zheng et al. 2017). The wide-field FGs (mainly the EPS) are very challenging to deal with even with the existing FG removal (Jelić et al. 2008; Bowman, Morales & Hewitt 2009; Paciga et al. 2011; Chapman et al. 2012; Trott, Wayth & Tingay 2012; Trott et al. 2016; Mertens, Ghosh & Koopmans 2018) and 'FG avoidance' (Datta, Bowman & Carilli 2010; Vedantham, Udaya Shankar & Subrahmanyan 2012; Thyagarajan et al. 2013; Pober et al. 2013, 2014; Liu, Parsons & Trott 2014a, b; Dillon et al. 2014, 2015) techniques.

The multifrequency angular power spectrum (MAPS; Zaldarriaga, Furlanetto & Hernquist 2004; Santos et al. 2005; Datta, Choudhury & Bharadwaj 2007) $C_\ell(\nu_a, \nu_b)$, which characterizes the second-order statistics of the sky signal jointly as a function of the angular multipole ℓ and frequencies ν, is a promising statistics to quantify the 21-cm signal (Mondal, Bharadwaj & Datta 2018; Mondal et al. 2019) and distinguish it from the FGs (Liu & Tegmark 2012; Trott et al. 2022). In this work, we have used the MAPS $C_\ell(\Delta\nu)$ that only depends on the frequency separation $\Delta\nu = |\nu_a - \nu_b|$. This is adequate when the statistical properties of the 21-cm signal do not vary significantly across the frequency bandwidth under consideration. FGs and the 21-cm signal are expected to show contrasting behaviour in $C_\ell(\Delta\nu)$. Being spectrally smooth, FGs are expected to show little or no variation with $\Delta\nu$ compared to the 21-cm signal that is expected to decorrelate with increasing $\Delta\nu$ (Bharadwaj & Sethi 2001; Bharadwaj & Ali 2005; Santos et al. 2005; Ali et al. 2008; Ali & Bharadwaj 2014). However, the wide-field point sources introduce oscillatory patterns along $\Delta\nu$ in the estimated $C_\ell(\Delta\nu)$ due to the inherent frequency response of the radio interferometers (Ghosh et al. 2011a, b). These oscillations, whose frequency increases at larger ℓ due to baseline migration, also manifest themselves as the 'FG wedge' (Datta et al. 2010; Morales et al. 2012) structure in the estimated cylindrical PS $P(k_\perp, k_\parallel)$ (Pal et al. 2022, hereafter, Paper I). The frequency structures in $C_\ell(\Delta\nu)$, or the FG wedge, jeopardize faithful FG removal and recovery of the 21-cm signal from the measured visibility data. Additionally, considering FG avoidance, various instrumental systematics, such as gain variations, primary beams (PBs), polarization leakage, calibration errors, missing channels flagged due to radio frequency interference (RFI), etc., extend the FG wedge to much higher $k_\parallel$ values, consequently reducing the otherwise FG-free '21-cm window' (TW) (Bowman et al. 2009; Pober et al. 2016; Thyagarajan et al. 2016).

The Tapered Gridded Estimator (TGE; Choudhuri et al. 2014, 2016a, b) is a visibility-based 21-cm PS estimator that allows us to taper the sky response to suppress the wide-field FG contributions arising from the side-lobe or periphery of the PB pattern. Additionally, it reduces computational load by using gridded visibilities, and internally subtracts out the positive-definite noise bias to produce unbiased estimates of the measured quantities. The TGE has been used to characterize the angular PS C_ℓ of the FGs at EoR frequencies (Choudhuri et al. 2017, 2020) as well as post-EoR frequencies (Chakraborty et al. 2019a, b; Mazumder et al. 2020). Bharadwaj et al. (2018) further developed upon this to introduce a MAPS-based TGE that first estimates the MAPS, and, from it, the PS, effectively dealing with the missing frequency channels in the visibility data while preserving all the qualities mentioned earlier. Pal et al. (2021) have used the MAPS-based TGE to estimate the MAPS and PS of the redshifted H I signal from EoR using an 8-MHz GMRT data set observed at 153 MHz.

In this work, we consider an observation of 25 h over four nights from the ELAIS-N1 field using a 200-MHz bandwidth at Band 3 (300–500 MHz) of uGMRT. These data were first introduced in Chakraborty et al. (2019a), and in a follow-up work, Chakraborty et al. (2019b) have presented the flagging, calibration, imaging, and point source subtraction from these data and also used the two-dimensional (2D) TGE to study the angular and spectral variation of $C_\ell(\nu)$ for the DGSE. Ch21 have conducted a multiredshift analysis of these data using a delay spectrum approach to estimate the PS of the 21-cm IM signal. In this approach, the missing frequency channels (flagged due to RFI) introduce ringing artefacts in the delay space, which can cause additional FG leakage and corrupt the estimated PS. The one-dimensional (1D) CLEAN (Parsons & Backer 2009) and the least-squares spectral analysis (LSSA; Trott 2016) are the two commonly used techniques that allow one to compensate for the missing frequency channels. Many recently developed algorithms, such as DAYENU filter (Ewall-Wice et al. 2021), Gaussian process regression (Mertens et al. 2020; Trott et al. 2020; Kern & Liu 2021) and Gaussian constrained realizations (Kennedy et al. 2022), have also aimed for an accurate recovery of the 21-cm PS from RFI-

[2]https://bingotelescope.org/
[3]https://hirax.ukzn.ac.za/
[4]https://www.sarao.ac.za/science/meerkat/
[5]http://tianlai.bao.ac.cn/
[6]https://www.skatelescope.org/
[7]http://rac.ncra.tifr.res.in/ort.html
[8]http://www.gmrt.ncra.tifr.res.in/

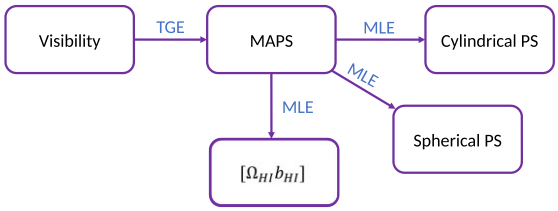

Figure 1. Flowchart of the paper, highlighting the main steps considered for the 21-cm IM from the calibrated visibilities.

contaminated data. Chakraborty, Datta & Mazumder (2022) have recently compared the 1D CLEAN and LSSA with simulated and actual visibility data to check which of the methods work better.

The TGE is capable of recovering the 21-cm signal even when 80 per cent data from randomly selected frequency channels are flagged (Bharadwaj et al. 2018). The TGE first correlates the visibility data across frequency channels to estimate $C_\ell(\Delta\nu)$, and estimates the PS from $C_\ell(\Delta\nu)$. Even if there are a substantial number of missing frequency channels in the visibility data, it is possible that there are no missing frequency separations $\Delta\nu$ in the estimated $C_\ell(\Delta\nu)$. The entire procedure uses only the available frequency channels to estimate the PS; it is not essential to make up for any missing frequency channels. In Paper I, we have used the TGE on 24.4-MHz-bandwidth data at 432.8 MHz ($z = 2.28$) from the same observation where 55 per cent of the data were flagged. Paper I further used FG avoidance approach to constrain the mean squared brightness temperature fluctuations of the redshifted H I signal with a 2σ upper limit of $\Delta^2_{\rm UL}(k) \leq (133.97)^2\,{\rm mK}^2$ at $k = 0.347\,{\rm Mpc}^{-1}$, which corresponds to an upper limit $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.23$ at $z = 2.28$. The quoted upper limit was found to be ∼7 times larger than what Ch21 found at a close redshift of $z = 2.19$ ($\nu_{\rm c} = 445$ MHz).

This work considers the same data as used in Paper I, with two key differences introduced in the analysis technique. First, we define the 'Cross' TGE for MAPS that cross-correlates the two mutually orthogonal (Cross) polarization states (RR and LL) of the visibilities. We expect this cross-correlation approach to mitigate a number of issues, such as noise bias, calibration errors, etc., which affect the 'Total' TGE (used in Paper I) where the two polarizations are combined. Further, it is expected that this approach will also reduce contributions from polarization-dependent FGs and systematics. Secondly, we have introduced a novel maximum likelihood estimator (MLE) that estimates the spherical PS $P(k)$ of the 21-cm signal directly from the estimated $C_\ell(\Delta\nu)$ without explicitly referring to the cylindrical PS $P(k_\perp, k_\parallel)$. The MLE we present utilizes the statistical isotropy of the 21-cm signal that differentiates it from the FGs. The MLE is expected to be robust to outliers (Huber 1981), and is optimal as we use inverse noise covariance weightage in the likelihood. Apart from estimating $P(k)$, we have also used the MLE on the full data set (or a subset) to constrain the single parameter $[\Omega_{\rm H\,I} b_{\rm H\,I}]$, thus maximizing the signal-to-noise ratio (SNR). A simplified flowchart of our present work is presented in Fig. 1.

We have arranged the paper in the following way. We first summarize the observations and preliminary processing of the data in Section 2. Next, we present the formalism for the TGE along with the estimated MAPS in Section 3, and the formalism for cylindrical PS estimation along with the estimated $P(k_\perp, k_\parallel)$ in Section 4. In Sections 5 and 6, we present the MLE for obtaining the spherical PS $P(k)$ and $[\Omega_{\rm H\,I} b_{\rm H\,I}]$, respectively, and also present the corresponding results. We have summarized our findings in Section 7.

Same as Paper I, we have used a ΛCDM cosmology with $\Omega_{\rm m} = 0.309$, $h = 0.67$, $n_{\rm s} = 0.965$, and $\Omega_{\rm b}h^2 = 0.0224$ (Planck Collaboration VI 2020).

2 DATA DESCRIPTION

We have observed the ELAIS-N1 field during May 2017 for 25 h over four nights using a 200-MHz bandwidth at the Band 3 (300–500 MHz) of uGMRT with a frequency resolution ($\Delta\nu_{\rm c}$) of 24.4 kHz and an integration time of 2 s. The detailed description of the data, along with flagging, calibration, imaging, and point source subtraction, is presented in Chakraborty et al. (2019b). We have used the resulting flagged, calibrated, point source-subtracted visibility data for the entire analysis presented here. Note that polarization calibration is not performed on these data.

The subset of the above data that have been analysed here is the same as that in Paper I, with the difference that we have restricted the baselines to a smaller range $|\,\mathbf{U}\,| < 1000\lambda$ where the baseline coverage is found to be denser and nearly uniform (fig. 1 of Paper I). The data cover a 24.4-MHz bandwidth with a central frequency $\nu_{\rm c} = 432.84$ MHz.

Considering the visibility data that we have analysed here, $\mathcal{V}^x_i(\nu_a)$ refers to a visibility measured at the baseline $\mathbf{U}_i$, frequency ν_a, and polarization x. The present data contain two circularly polarized states RR and LL.

3 THE TGE FOR MAPS

The MAPS $C_\ell(\nu_a, \nu_b)$ quantifies the statistical properties of the sky signal jointly as a function of the angular multipoles and frequencies. The brightness temperature fluctuations in the sky are decomposed in terms of spherical harmonics $Y^{\rm m}_\ell(\hat{\mathbf{n}})$ as

$$\delta T_{\rm b}(\hat{\mathbf{n}},\ \nu) = \sum_{\ell,m} a_{\ell{\rm m}}(\nu)\, Y^{\rm m}_\ell(\hat{\mathbf{n}}). \tag{1}$$

We use this to define the MAPS as (Zaldarriaga et al. 2004; Santos et al. 2005; Datta et al. 2007)

$$C_\ell(\nu_a, \nu_b) = \left\langle a_{\ell{\rm m}}(\nu_a)\, a^*_{\ell{\rm m}}(\nu_b)\right\rangle. \tag{2}$$

Here, $\langle ... \rangle$ denotes an ensemble average over different statistically independent realizations of the random field $\delta T_{\rm b}(\hat{\mathbf{n}},\ \nu)$.

The TGE uses the measured visibilities to estimate MAPS. We note that the present analysis does not incorporate baseline migration. Considering a fixed antenna pair, the baseline $\mathbf{U}$ is held fixed at the value corresponding to the central frequency $\nu_{\rm c}$. The details of the visibility-based TGE are given in Paper I (also Bharadwaj et al. 2018 and Pal et al. 2021). Here, we briefly summarize the salient features of the formalism, and extend it to consider the polarization. We introduce a rectangular grid in the uv-plane and calculate $\mathcal{V}^x_{cg}(\nu_a)$ the convolved-gridded visibility for every grid point $\mathbf{U}_g$ using

$$\mathcal{V}^x_{cg}(\nu_a) = \sum_i \tilde{w}(\mathbf{U}_g - \mathbf{U}_i)\, \mathcal{V}^x_i(\nu_a)\, F^x_i(\nu_a). \tag{3}$$

Here, $F^x_i(\nu_a)$ is 0 if the visibility is flagged and 1 otherwise, and $\tilde{w}(\mathbf{U})$ is the Fourier transform of a suitably chosen window function $\mathcal{W}(\theta)$ that is introduced to taper the PB of the telescope far away from the phase centre.

The main lobe of the PB of any telescope with a circular aperture can be approximated as $\mathcal{A}(\theta) = e^{-\theta^2/\theta_0^2}$, where $\theta_0 \sim 0.6 \times \theta_{\rm FWHM}$, $\theta_{\rm FWHM}$ is the full width at half-maxima of $\mathcal{A}(\theta)$ (Bharadwaj & Sethi 2001; Choudhuri et al. 2014). Here, we have used a Gaussian window function $\mathcal{W}(\theta) = {\rm e}^{-\theta^2/[f\theta_0]^2}$ where the tapering parameter 'f' controls

the degree to which the PB pattern is tapered. Here, $f > 1$ provides minimal tapering, and $f < 1$ highly suppresses the sky response away from the phase centre. We had considered different values of f in the range $0.6 \leq f \leq 5$ in Paper I where we had found that it is possible to reduce oscillations (along frequency) in MAPS by reducing the value of f. However, this improvement was found to saturate around $f = 0.6$ that provides the best results. Based on this, we have used $f = 0.6$ for the entire analysis presented here.

Here, we assume that the 21-cm signal is unpolarized, and we treat the two polarizations (RR and LL) as independent measurements of the same 21-cm signal. In Paper I, we have combined the two polarizations using

$$\mathcal{V}_{cg}(\nu_a) = \mathcal{V}_{cg}^{\rm RR}(\nu_a) + \mathcal{V}_{cg}^{\rm LL}(\nu_a), \tag{4}$$

and used this to define the TGE for MAPS

$$\hat{E}_g(\nu_a,\, \nu_b) = M_g^{-1}(\nu_a,\, \nu_b)\mathcal{R}e\Big[\mathcal{V}_{cg}(\nu_a)\mathcal{V}_{cg}^{*}(\nu_b) - \sum_{x,\,i} F_i^x(\nu_a)F_i^x(\nu_b) \mid \tilde{w}(\mathbf{U}_g - \mathbf{U}_i) \mid^2 \mathcal{V}_i^x(\nu_a)\mathcal{V}_i^{*x}(\nu_b)\Big], \tag{5}$$

where $\mathcal{R}e[..]$ implies the real part of the expression within the brackets [..] and $M_g(\nu_a,\, \nu_b)$ is a normalization factor. In principle, it is adequate to consider the correlation $\mathcal{V}_{cg}(\nu_a)\mathcal{V}_{cg}^{*}(\nu_b)$ in order to estimate $C_\ell(\nu_a,\, \nu_b)$, except for the fact that we have an additive noise bias when $\nu_a = \nu_b$. The second term in the square brackets, which subtracts out the correlation of a visibility with itself, is introduced to remove the noise bias.

Instead of combining the two polarizations (equation 4), in this work we have used the correlation of the two cross-polarizations $\mathcal{V}_{cg}^{\rm RR}$ and $\mathcal{V}_{cg}^{\rm LL}$ to estimate $C_\ell(\nu_a,\, \nu_b)$. The cross-polarization correlation TGE for MAPS is defined as

$$\hat{E}_g(\nu_a,\, \nu_b) = M_g^{-1}(\nu_a,\, \nu_b)\mathcal{R}e \times \Big[\mathcal{V}_{cg}^{\rm RR}(\nu_a)\mathcal{V}_{cg}^{*\rm LL}(\nu_b) + \mathcal{V}_{cg}^{\rm LL}(\nu_a)\mathcal{V}_{cg}^{*\rm RR}(\nu_b)\Big]. \tag{6}$$

Since the noises in the two polarizations are uncorrelated, equation (6) has the advantage that it is not necessary to account for any noise bias in the cross-polarization estimator. We may also expect some further advantages if the calibration errors, FGs, and other systematics in the two polarizations are partially uncorrelated. We note that the estimator in equation (5) contains both the self-polarization correlations (RR × RR and LL × LL) and the cross-polarization correlations (RR × LL). In the subsequent discussion, we refer to the TGE in equations (5) and (6) as 'Total' and 'Cross', respectively. We have validated (Appendix A) the Cross TGE using simulations that incorporate the same flagging, frequency, and baseline coverage of the actual data. The validation of the Total estimator is given in Paper I.

We now discuss how we have determined the normalization factor $M_g^{-1}(\nu_a,\, \nu_b)$ for the Cross estimator. We first simulate multiple realizations of $[\delta T_{\rm b}(\hat{\mathbf{n}},\, \nu)]_{\rm uMAPS}$ the sky signal corresponding to a Gaussian random field having a unit MAPS [uMAPS; $C_\ell(\nu_a,\, \nu_b) = 1$]. We use this sky signal to simulate the corresponding visibilities $[\mathcal{V}_i^x(\nu_a)]_{\rm uMAPS}$ at the baselines, frequency channels, and polarizations identical to the data. These simulations incorporate both baseline migration and the frequency dependence of the telescope's PB. The flagging of the actual data $F_i^x(\nu_a)$ has been applied to the simulated visibilities $[\mathcal{V}_i^x(\nu_a)]_{\rm uMAPS}$ and used to obtain

$$M_g(\nu_a,\, \nu_b) = \mathcal{R}e\Big[\mathcal{V}_{cg}^{\rm RR}(\nu_a)\mathcal{V}_{cg}^{*\rm LL}(\nu_b) + \mathcal{V}_{cg}^{\rm LL}(\nu_a)\mathcal{V}_{cg}^{*\rm RR}(\nu_b)\Big]_{\rm uMAPS}. \tag{7}$$

We have averaged over multiple realizations of the simulated uMAPS to reduce the statistical uncertainties in the estimated M_g. For the subsequent analysis, we have used 50 realizations of uMAPS to estimate M_g. Note that M_g for the Total estimator (equation 5) is different from that given by equation (7), and the relevant equation for the Total estimator is presented in Paper I.

Both the estimators (equations 5 and 6) give unbiased estimate of the MAPS, i.e. $\langle \hat{E}_g \rangle = C_{\ell_g}$ at the grid point $\mathbf{U}_g$ that corresponds to an angular multipole $\ell_g = 2\pi \mid \mathbf{U}_g \mid$. Incorporating the fact that the statistics of the 21-cm signal is isotropic on the plane of the sky, we combine $\hat{E}_g$ at different grid points $\mathbf{U}_g$ within annular bins in the uv-plane. The bin-averaged TGE is defined as

$$\hat{E}_a = \frac{\sum_g w_g \hat{E}_g}{\sum_g w_g}, \tag{8}$$

where the sum is over all the grid points $\mathbf{U}_g$ in the $a^{\rm th}$ ℓ bin and w_g's are the corresponding weights. Here, we have used $w_g = M_g$, which implies that the weight is proportional to the baseline density in that particular grid point. The ensemble average of $\hat{E}_a$ gives an unbiased estimate of the bin-averaged MAPS $\bar{C}_{\bar{\ell}_a}$ at the bin-averaged angular multipole $\bar{\ell}_a = \frac{\sum_g w_g \ell_g}{\sum_g w_g}$. We subsequently use C_{ℓ_a} and ℓ_a to denote the bin-averaged values $\bar{C}_{\bar{\ell}_a}$ and $\bar{\ell}_a$, respectively.

The post-reionization 21-cm signal evolves relatively gradually with z (e.g. Sarkar, Bharadwaj & Anathpindika 2016), and it is quite reasonable to assume this to be ergodic (statistically homogeneous) along the line-of-sight direction for the 24.4-MHz bandwidth, which corresponds to the redshift interval $\Delta z = 0.19\,(z = 2.19\text{–}2.38)$, considered here. Instead of considering the entire covariance $C_\ell(\nu_a,\, \nu_b)$, it is now adequate to consider $C_\ell(\Delta\nu)$ that is a function of the frequency separations $\Delta\nu = |\nu_a - \nu_b|$. However, we note that $C_\ell(\nu_a,\, \nu_b)$ is a more accurate statistics (see e.g. Mondal et al. 2018, 2022) for a wide-band and high-redshift data (such as Trott et al. 2020).

We have divided the $\mathbf{U}$ range $U \leq 1000\lambda$ into 10 bins of equal linear spacing, and evaluated $C_\ell(\Delta\nu)$ using the Cross estimator (equation 6). The estimated $C_\ell(\Delta\nu)$ are rather noisy at large $\Delta\nu$, and following Paper I we have only used $\Delta\nu \leq 12.2$ MHz for the subsequent analysis. The ℓ bins used here are somewhat different from those in Paper I; however, the ℓ values roughly match for the first six bins. Compared to Paper I that has used the Total estimator, we find that the $C_\ell(\Delta\nu)$ values obtained here have a different vertical offset that corresponds to a difference in the $\Delta\nu$-independent DC component of $C_\ell(\Delta\nu)$. This DC component of $C_\ell(\Delta\nu)$ only affects the lowest line-of-sight mode $k_\parallel = 0$ in $P(k_\perp, k_\parallel)$ the cylindrical PS. The $k_\parallel = 0$ mode is usually FG dominated, and we do not use it to constrain the 21-cm signal. Fig. 2 shows the DC-subtracted $C_\ell(\Delta\nu)$ obtained using both the Cross and Total estimators for the first six ℓ bins. The grey shaded regions show the 3σ errors for the Cross $C_\ell(\Delta\nu)$. These errors were estimated using simulations as described in Paper I, and also later in this paper.

We expect the estimated $C_\ell(\Delta\nu)$ (both Cross and Total) to be dominated by various FG components, mainly the DGSE and the radiation from unsubtracted EPSs. It has been reported in earlier studies that the measured C_ℓ is dominated by the DGSE at larger angular scales and by the residual point sources at smaller angular scales (Bernardi et al. 2009; Ghosh et al. 2012; Choudhuri et al. 2017; Chakraborty et al. 2019b). While both of these are expected to have intrinsically smooth frequency spectra, various observational effects introduce frequency-dependent structures in the estimated $C_\ell(\Delta\nu)$. For example, baseline migration, bandpass calibration errors, and polarization leakage introduce oscillations along $\Delta\nu$ in $C_\ell(\Delta\nu)$.

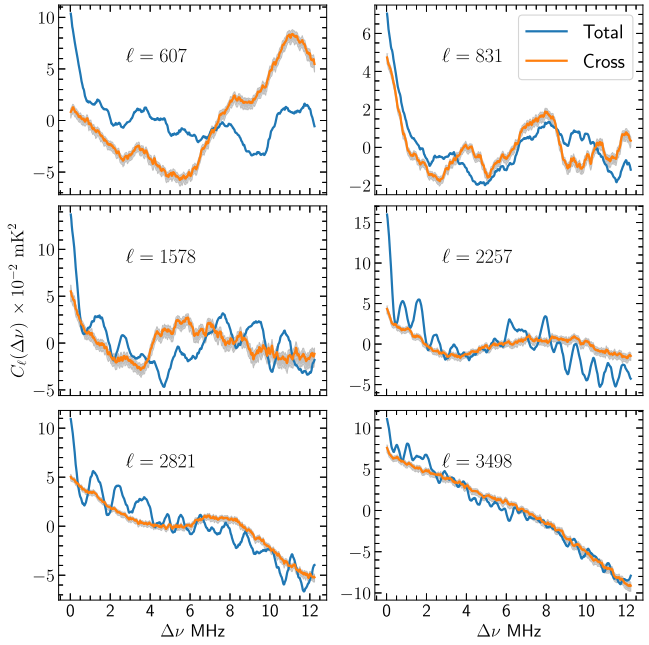

Figure 2. A comparison of mean-subtracted Total (blue) and Cross (orange) MAPS $C_\ell(\Delta\nu)$ for different ℓ values. The grey shaded regions show the 3σ error bars.

Considering the different panels of Fig. 2, we see that the two different estimates of $C_\ell(\Delta\nu)$ have comparable values once the DC is subtracted out. Considering the $\Delta\nu$ dependence, for many of the ℓ bins we find very similar slowly varying patterns in both the estimates. The degree of similarity appears to increase as we go to the larger ℓ bins. The Total $C_\ell(\Delta\nu)$ decorrelates very sharply when $\Delta\nu$ is increased from 0 to 1 MHz, and it also exhibits rapid oscillations at larger $\Delta\nu$. These rapid variations are considerably diminished in the Cross $C_\ell(\Delta\nu)$ that exhibits a much smoother $\Delta\nu$ dependence.

In the lowest ℓ bin, the Cross and Total $C_\ell(\Delta\nu)$ are found to differ in their $\Delta\nu$ dependence. We expect $C_\ell(\Delta\nu)$ in this bin to be DGSE dominated (Chakraborty et al. 2019b). The differences between the two estimates of $C_\ell(\Delta\nu)$ may arise due to polarized structure in the DGSE (Pen et al. 2009b). Further, these differences may also arise from differences in the instrumental calibration of the two polarizations, instrumental polarization leakage due to asymmetry of the PB response, and leakage from polarized point sources (Asad et al. 2015; Van Eck et al. 2018; Kumar et al. 2022). Faraday rotation in the magnetized plasma causes a phase difference between the left and right circularly polarized components (Smirnov 2011), and this also can contribute to the difference in the Cross and Total $C_\ell(\Delta\nu)$. The DGSE contribution decreases as we move to larger ℓ. In Fig. 2, we see that differences between the Total and Cross estimates of $C_\ell(\Delta\nu)$ go down as we move to larger ℓ bins. This supports the picture where a part of the difference between the Total and Cross estimates may be attributed to the DGSE. However, this does not explain why the sharp decorrelation around $\Delta\nu = 0$ and the rapid oscillations are mitigated for the Cross estimator. This possibly has to do with gain calibrations errors and other systematics that could be uncorrelated for the two polarizations.

4 THE CYLINDRICAL PS

Under the flat-sky approximation, $P(k_\perp, k_\parallel)$ the cylindrical PS of the 21-cm brightness temperature fluctuations $\delta T_{\rm b}(\hat{\mathbf{n}}, \nu)$ is related to the MAPS $C_\ell(\Delta\nu)$ through a Fourier transform along the Line of Sight (LoS) (Datta et al. 2007):

$$C_\ell(\Delta\nu) = \frac{1}{\pi r^2}\int_0^\infty {\rm d}k_\parallel \cos(k_\parallel r' \Delta\nu) P(k_\perp, k_\parallel), \tag{9}$$

where $k_\parallel$ and $k_\perp = \ell/r$ are the parallel and perpendicular to the LoS components of $\mathbf{k}$, respectively. The comoving distance r and its derivative with respect to frequency $r' = {\rm d}r/{\rm d}\nu$, which are evaluated at the reference frequency $\nu_{\rm c} = 432.8$ MHz ($z = 2.28$), have values 5703 Mpc and 9.85 Mpc MHz^{-1}, respectively.

We use an MLE to estimate the PS $P(k_{\perp a}, k_{\parallel m})$ from the measured $C_{\ell_a}(\Delta\nu_n)$, where $n, m \in [0, N_E - 1]$ and N_E is the number of frequency separations used in the PS estimation. In matrix notation,

$$C_{\ell_a}(\Delta\nu_n) = \sum_m \mathbf{A}_{nm} P(k_{\perp a}, k_{\parallel m}) + [{\rm Noise}]_n, \tag{10}$$

where $\mathbf{A}_{nm}$ are the components of the $N_E \times N_E$ Hermitian matrix $\mathbf{A}$ containing the coefficients of the Fourier transform, and $[{\rm Noise}]_n$ is an additive noise associated with each estimated $C_{\ell_a}(\Delta\nu_n)$. The maximum likelihood estimate of $P(k_{\perp a}, k_{\parallel m})$ is given by

$$P(k_{\perp a}, k_{\parallel m}) = \sum_n \left[\left(\mathbf{A}^\dagger \mathbf{N}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^\dagger\mathbf{N}^{-1}\right]_{mn} \mathcal{W}(\Delta\nu_n)\, C_{\ell_a}(\Delta\nu_n), \tag{11}$$

where $\mathbf{N}$ is the noise covariance matrix and '$\dagger$' denotes the Hermitian conjugate. Note that we have applied a Blackman–Nuttall (Nuttall 1981) window function $\mathcal{W}(\Delta\nu_n)$, normalized at $\Delta\nu = 0$, to reduce the ringing artefacts (ripples), which otherwise appears in the PS due to the discontinuity in $C_\ell(\Delta\nu)$ at the band edges.

We have estimated the noise covariance matrix $\mathbf{N}$ through multiple realizations (50 in this work) of 'noise-only' simulations of the measured complex visibilities. The random noise is assumed to follow a Gaussian distribution with zero mean and standard deviation $\sigma_N = 0.43$ Jy that is estimated from the real (or imaginary) part of the measured visibility data. Note that this assumption implies $\mathbf{N}$ to be diagonal. Further, we have also used these simulations to estimate $\delta P_N(k_\perp, k_\parallel)$ the system noise contribution to the uncertainty in the estimated PS. An analysis of the noise statistics for $P(k_\perp, k_\parallel)$, presented later in this section, leads us to believe that the actual noise level for the data is approximately 4.77 times larger than that obtained from these system noise-only simulations, and we have accounted for this by scaling up all the noise predictions by this factor.

Fig. 3 shows both the Total and Cross $P(k_\perp, k_\parallel)$ as a function of $k_\parallel$, where the different panels correspond to fixed values of $k_\perp$ that are in direct correspondence to the panels in Fig. 2. The grey shaded regions show the 3σ error bars for the Cross $P(k_\perp, k_\parallel)$ and the black solid lines show the theoretically predicted FG wedge boundary $[k_\parallel]_H = (r/r'\nu_{\rm c})k_\perp$ that corresponds to the FG contribution from a source located at the horizon. For each value of $k_\perp$, we have visually inspected the Cross $P(k_\perp, k_\parallel)$ and identified the region that is relatively free of FG contamination. We refer to this region as the TW whose boundary is demarcated by the green dashed line. In the subsequent discussion, we refer to the $(k_\perp, k_\parallel)$ modes complementary to the TW (i.e. from $k_\parallel = 0$ to the green dashed line) as the FG modes. We further refer to the region within $[k_\parallel]_H$ and the green dashed line as the buffer.

Considering the Total PS, as noted in Paper I, the amplitude of the PS starts with a high value ($\sim 10^9$ mK^2 Mpc^3) at $k_\parallel = 0$ and falls with increasing $k_\parallel$ and nearly flattens out at $k_\parallel \sim 0.2$–0.8 Mpc^{-1}. The amplitude then rises slightly in a few $k_\parallel$ bins just beyond $[k_\parallel]_H$ and then again falls to $\sim 10^5$ mK^2 Mpc^3 at $k_\parallel > 1$–2 Mpc^{-1} where it oscillate between positive and negative values that are comparable

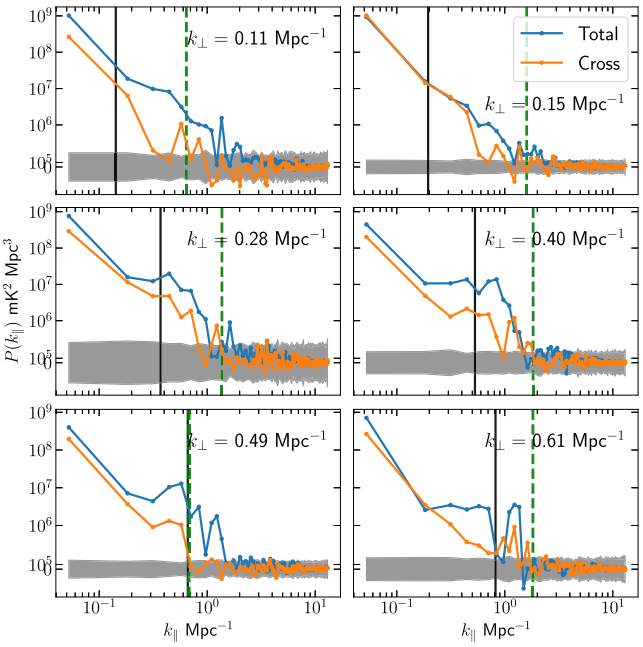

Figure 3. The Total (blue) and Cross (orange) cylindrical PS $P(k_\perp, k_\parallel)$ as a function of $k_\parallel$ for different values of $k_\perp$. The grey shaded regions show the 3σ error bars for the Cross PS. The vertical lines show the $[k_\parallel]_H$ (black solid) and the TW boundary (green dashed) for the respective $k_\perp$ bins of the Cross $P(k_\perp, k_\parallel)$.

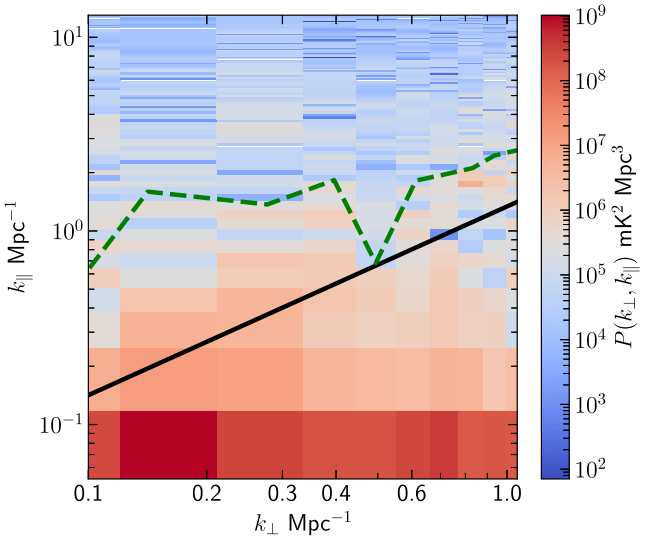

Figure 4. The Cross cylindrical power spectra $|P(k_\perp, k_\parallel)|$. Here, the black solid and green dashed lines denote $[k_\parallel]_H$ and the TW boundary, respectively. The region above the green dashed line was identified as being relatively free of FG contamination, and used for spherical binning.

with the noise. The Cross PS also shows a similar feature in all the $k_\perp$ bins but with a comparatively lower amplitude throughout the entire $k_\parallel$ range. This is particularly noticeable in the near-flat region where the Cross PS has an amplitude that is an order of magnitude smaller. Further, it reaches the noise level at a relatively lower $k_\parallel$ ($k_\parallel \sim 0.8$–$1\,\mathrm{Mpc}^{-1}$) as compared to the total PS.

The features in $P(k_\perp, k_\parallel)$ (Fig. 3) are directly related to the features seen in the MAPS $C_\ell(\Delta\nu)$ (Fig. 2). A sharp variation in $C_\ell(\Delta\nu)$ yields a smooth variation in $P(k_\perp, k_\parallel)$, and this is why the Total PS shows a wider flat region compared to the Cross PS, and it also reaches the noise level slower than the Cross PS. This feature is markedly visible in the last four $k_\perp$ bins that correspond to the last four ℓ bins of Fig. 2. In these bins, the Total MAPS decorrelates faster than the Cross MAPS, and so the Total PS goes to noise level much slower than the Cross PS.

The high values of the PS beyond $[k_\parallel]_H$ are related to the rapid oscillations we see in the MAPS. The amplitude of the PS depends on the oscillation amplitude in MAPS, whereas the oscillation period points to the $k_\parallel$ mode where the power corresponding to the oscillation arises. The oscillations with larger amplitudes and small periods thus show up as the spikes in the PS at the larger $k_\parallel$ modes. We have seen in Fig. 2 that the oscillation amplitude in the Cross MAPS is much smaller than that in the Total MAPS and this is why the amplitude of the spikes is much smaller in the Cross PS.

We have shown the Cross PS heatmap $|P(k_\perp, k_\parallel)|$ in Fig. 4. The black solid line and the green dashed line denote the wedge and the buffer boundary, respectively. We see that most of the power lies inside the wedge boundary where the PS varies $\sim 10^7$–$10^9\,\mathrm{mK}^2\,\mathrm{Mpc}^3$. There is considerable FG leakage in the buffer region where the PS varies $\sim 10^5$–$10^7\,\mathrm{mK}^2\,\mathrm{Mpc}^3$. The buffer boundary is chosen by inspecting the 1D slices (Fig. 3) in each $k_\perp$ bin. Considering the first $k_\perp$ bin, we have chosen a buffer of $\sim 0.5\,\mathrm{Mpc}^{-1}$. We have chosen a relatively larger buffer (0.8–$1.2\,\mathrm{Mpc}^{-1}$) in the subsequent bins that show additional leakage barring the fifth bin ($k_\perp = 0.49\,\mathrm{Mpc}^{-1}$) that looks clean beyond the wedge boundary.

It is necessary to ensure that the power in the TW is either strictly positive or consistent with noise. To ensure that our PS estimates are free from negative systematics, we study the quantity X, which is the ratio between the estimated cylindrical PS $P(k_\perp, k_\parallel)$ and the statistical fluctuation $\delta P_N(k_\perp, k_\parallel)$ expected due to the system noise:

$$X = \frac{P(k_\perp, k_\parallel)}{\delta P_N(k_\perp, k_\parallel)}. \tag{12}$$

We note that $\delta P_N(k_\perp, k_\parallel)$ in equation (12) does not include the factor of 4.77 that was mentioned earlier. We expect X to have a symmetric distribution with zero mean and unit standard deviation if the values of $P(k_\perp, k_\parallel)$ are entirely due to the system noise contribution.

Fig. 5 shows the histogram of X. We see that bulk of the data points (99.63 per cent) in the histogram lie in the central $|X| \leq 30$ region, which we delimit by the vertical black dashed lines. The probability density function (PDF) is mostly symmetric in the central region with a positive mean $\mu = 0.61$ and a standard deviation $\sigma_{\rm Est} = 4.77$. We do not see negative outlier values of X beyond the central region. The standard deviation $\sigma_{\rm Est} > 1$ suggests that the statistical fluctuation in $P(k_\perp, k_\parallel)$ is underestimated by the system noise-only simulations. This excess is possibly due to artefacts from calibration errors, inaccurate point source subtraction, and RFI. As mentioned earlier, we have scaled up all the error estimates by a factor of 4.77 to account for this.

We have seen in Paper I that a t-distribution adequately describes the X statistics of the Total PS near the central region, but it fails to fit the positive tail. Here also, we find the t-distribution (orange dashed line) to underfit the tail of the histogram. However, a Lorentzian distribution (green dashed line) seems to represent the statistics better. Also note that the positive tail is quite shorter for the Cross PS as compared to the Total PS (Paper I). This shorter positive tail suggests that we have a cleaner TW region with less FG contamination for the Cross PS in comparison to the Total PS. We have used all the $(k_\perp, k_\parallel)$ modes in the TW for spherical binning, which we describe in Section 5.

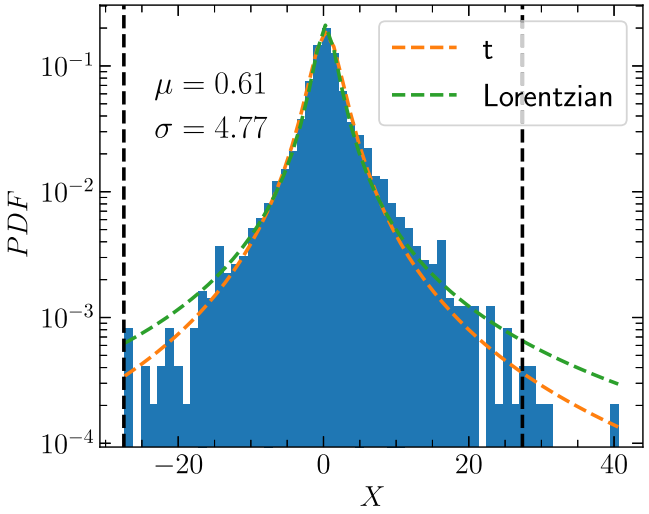

Figure 5. The histogram of the variable $X = \frac{P(k_\perp, k_\parallel)}{\delta P_N(k_\perp, k_\parallel)}$ is shown. The orange and green dashed curves show the fit with t and Lorentzian distributions, respectively.

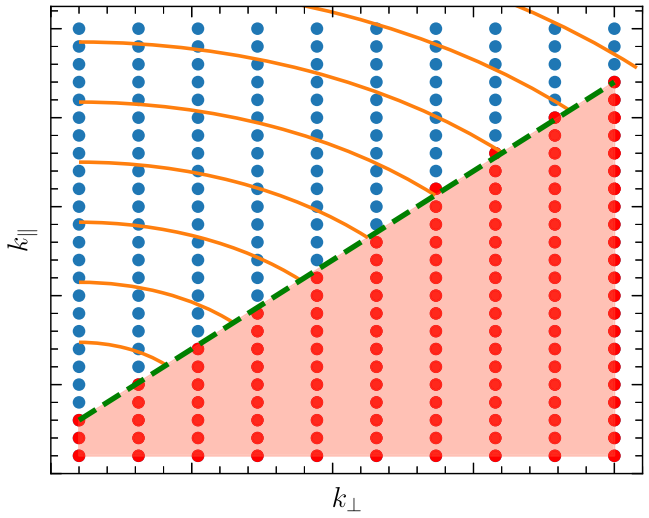

Figure 6. A schematic diagram to explain the spherically binning MLE. Each filled circle in the diagram represents a $(k_\perp, k_\parallel)$ mode. The FG modes are shown in red, whereas the modes in the TW are shown in blue. The green dashed curve demarcates the boundary of the TW that is divided into equally spaced logarithmic bins shown by the orange spherical arcs.

5 THE SPHERICAL PS

In this section, we utilize an important feature that distinguishes the redshifted 21-cm signal from the FGs. This arises from the fact that the 21-cm signal traces out the three-dimensional distribution of a cosmological density field. Like all cosmological density fields, we expect the 21-cm signal also to be statistically isotropic in three-dimensional space; i.e. its clustering properties depend only on the length of the spatial separation irrespective of the orientation with respect to the plane of the sky and the LoS directions. The allows us to quantify the 21-cm signal using the spherical PS $P_T(k)$, where $k = \sqrt{k_\perp^2 + k_\parallel^2}$. We note that this isotropy is broken by redshift space distortion (Bharadwaj & Ali 2005). While it is also possible to include this effect in our analysis, we have chosen to ignore it for this work. The 21-cm MAPS $[C_\ell(\Delta\nu)]_T$, which is related to $P_T(k)$ through equation (9), is expected to encode this isotropy through its ℓ and $\Delta\nu$ dependence. This distinguishes $[C_\ell(\Delta\nu)]_T$ from the other sources for which frequency separation $\Delta\nu$ does not correspond to a spatial separation.

Here, we have modelled the measured $C_\ell(\Delta\nu)$ as

$$C_{\ell_a}(\Delta\nu_n) = \left[C_{\ell_a}(\Delta\nu_n)\right]_{\rm FG} + \left[C_{\ell_a}(\Delta\nu_n)\right]_T + \left[C_{\ell_a}(\Delta\nu_n)\right]_R \quad (13)$$

and used this to estimate the spatially isotropic component $[C_{\ell_a}(\Delta\nu_n)]_T$. As mentioned earlier, $C_{\ell_a}(\Delta\nu_n)$ is dominated by $[C_{\ell_a}(\Delta\nu_n)]_{\rm FG}$ the FG contribution. The FGs are expected to have a smooth frequency dependence, and the $\Delta\nu$ dependence arises mainly due to instrumental effects like baseline migration (Morales et al. 2012; Hazelton, Morales & Sullivan 2013). Considering Fig. 4, we have identified a region of $(k_\perp, k_\parallel)$ plane where the modes are FG dominated (FG modes). Further, the complementary region, referred to as the TW, was identified as being relatively free of FG contamination. Here, we have assumed that $[C_{\ell_a}(\Delta\nu_n)]_{\rm FG}$ can be entirely quantified in terms of the FG modes as

$$\left[C_{\ell_a}(\Delta\nu_n)\right]_{\rm FG} = \sum_m A_{nm} \left[P(k_{\perp a}, k_{\parallel m})\right]_{\rm FG} \quad (14)$$

and we have excluded these modes for estimating the 21-cm signal. The modes within this region are schematically represented by the red points in Fig. 6, where the green dashed line denotes the boundary of the TW.

We have used only the TW modes to estimate the 21-cm signal. Here, we have assumed spatial isotropy and divided the TW into spherical bins (labelled $i = 1, ..., {\rm NBin}$) that are shown schematically in Fig. 6. We use $[P(k_i)]_T$ to denote the value of the spherical PS corresponding to the i-th bin. We have modelled the 21-cm signal as

$$\left[C_{\ell_a}(\Delta\nu_n)\right]_T = \sum_i B_i(a, n) \left[P(k_i)\right]_T \quad (15)$$

with $B_i(a, n) = \sum_m A_{nm}$, where this sum is over the $(k_{\perp a}, k_{\parallel m})$ modes that are within the i-th bin. Note that we have dropped the subscript 'T' in $[P(k_i)]_T$, and simply denote it as $P(k)$ when there is no ambiguity.

Considering equation (13), $[C_\ell(\Delta\nu)]_R$ refers to the residual MAPS, i.e. the component of $C_\ell(\Delta\nu)$ that is not included in the FGs or the isotropic 21-cm signal. Noise, systematics, and FG leakage are possible factors that contribute to $[C_\ell(\Delta\nu)]_R$ (Kumar, Dutta & Roy 2020). Ideally, we expect $[C_\ell(\Delta\nu)]_R$ to be consistent with our noise estimates, and we define chi-square (χ^2) as

$$\chi^2 = \sum_{a,n,m} \left[C_{\ell_a}(\Delta\nu_n)\right]_R \mathbf{N}^{-1}_{nm} \left[C_{\ell_a}(\Delta\nu_m)\right]_R, \quad (16)$$

where $\mathbf{N}$ is the noise covariance matrix introduced in equation (11). Our model for the measured $C_\ell(\Delta\nu)$ now has $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and $[P(k_i)]_T$ as parameters. We have maximized the likelihood $\mathcal{L} \propto \exp(-\chi^2/2)$ with respect to the parameters in order to determine the best-fitting parameter values. We have also used this likelihood analysis to obtain error estimates for the best-fitting parameter values.

Here, we have used a total 5000 measured data points $C_{\ell_a}(\Delta\nu_n)$ corresponding to 10 ℓ bins and 500 frequency separations $\Delta\nu$ to obtain maximum likelihood estimates for a total 664 parameters, of which 656 are the FG modes $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and the remaining 8 are the $[P(k_i)]_T$ corresponding to the 8 spherical k bins that span $0.804\,{\rm Mpc}^{-1} < k < 11.892\,{\rm Mpc}^{-1}$. We find that the goodness-of-fit parameter (reduced χ^2) has a value of 1.21, which indicates that our model provides an adequate fit for the measured $C_{\ell_a}(\Delta\nu_n)$ and the residual $[C_{\ell_a}(\Delta\nu_n)]_R$ is roughly consistent with noise. We have used the best-fitting $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and $[P(k_i)]_T$ in equations (14) and (15) to recover $[C_{\ell_a}(\Delta\nu_n)]_{\rm FG}$ and $[C_{\ell_a}(\Delta\nu_n)]_T$, respectively. The

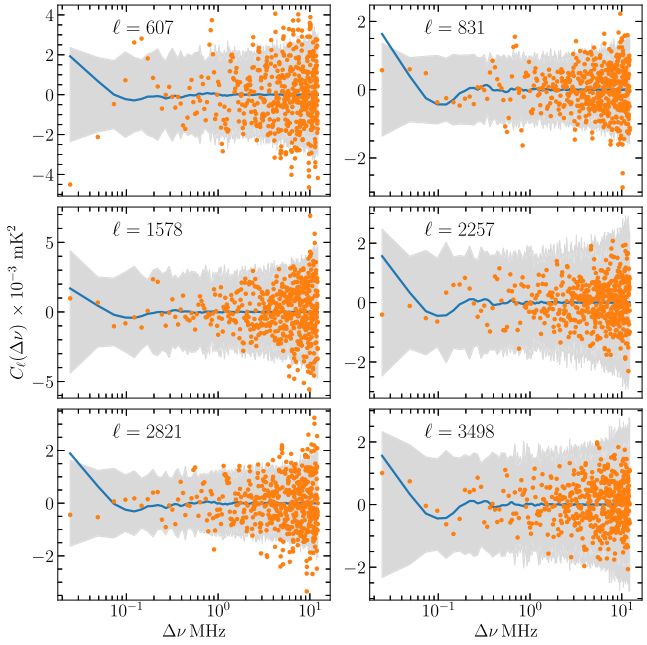

Figure 7. For the first 6 ℓ bins, this shows $[C_\ell(\Delta\nu)]_T$ (solid blue lines) and $[C_\ell(\Delta\nu)]_R$ (orange dots) corresponding to the best-fitting $[P(k_i)]_T$. The grey shaded regions show 2σ error bars for the measured $C_\ell(\Delta\nu)$.

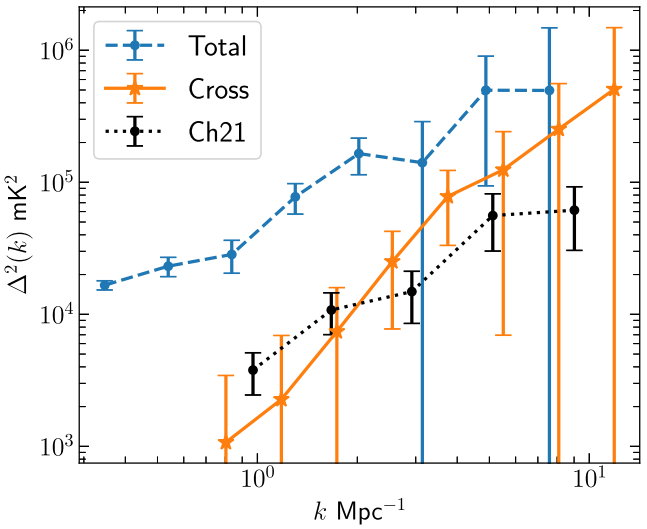

Figure 8. The mean squared brightness temperature fluctuations $\Delta^2(k)$ along with 2σ error bars. The orange asterisks show the results from this work (Cross), while the blue dashed and black dotted lines show the results from Paper I (Total) and Ch21, respectively.

Table 1. The mean squared brightness temperature fluctuations $\Delta^2(k)$ and the corresponding statistical error predictions σ for different k bins. The 2σ upper limits $\Delta^2_{\rm UL}(k) = \Delta^2(k) + 2\sigma$ and corresponding $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL}$ values are also provided.

k (Mpc^{-1})	$\Delta^2(k)$ (mK)^2	1σ (mK)^2	SNR	$\Delta^2_{\rm UL}(k)$ (mK)^2	$[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL}$
0.804	$(32.75)^2$	$(34.42)^2$	0.905	$(58.67)^2$	0.072
1.181	$(47.64)^2$	$(48.19)^2$	0.977	$(83.15)^2$	0.089
1.736	$(86.05)^2$	$(65.31)^2$	1.736	$(126.24)^2$	0.121
2.551	$(158.47)^2$	$(93.18)^2$	2.892	$(206.11)^2$	0.177
3.748	$(279.47)^2$	$(149.74)^2$	3.483	$(350.64)^2$	0.273
5.507	$(352.76)^2$	$(242.38)^2$	2.118	$(491.87)^2$	0.350
8.093	$(502.73)^2$	$(391.31)^2$	1.651	$(747.65)^2$	0.490
11.892	$(712.14)^2$	$(698.77)^2$	1.039	$(1218.07)^2$	0.589

measured $C_\ell(\Delta\nu)$ is FG dominated, and we find that the recovered $[C_{\ell_a}(\Delta\nu_n)]_{\rm FG}$ closely matched the measured $C_\ell(\Delta\nu)$ shown earlier in Fig. 2. The recovered $[C_{\ell_a}(\Delta\nu_n)]_T$ are shown in the different panels of Fig. 7. The residuals $[C_\ell(\Delta\nu)]_R$ and the 2σ errors due to noise are shown using the orange dots and the grey shaded regions, respectively. We see that, in all the ℓ bins shown here, $[C_{\ell_a}(\Delta\nu_n)]_T$ varies within 0–0.002 mK^2 and the values lie within the predicted 2σ noise levels. Further, in all cases the recovered $[C_{\ell_a}(\Delta\nu_n)]_T$ has maximum value at $\Delta\nu = 0$, the value decreases with increasing $\Delta\nu$, and is close to 0 for large $\Delta\nu$. We also notice some oscillatory features in $[C_{\ell_a}(\Delta\nu_n)]_T$ that reflect the fact that some of the $(k_{\perp a}, k_{\parallel m})$ modes (in the FG region) were excluded when calculating $[C_{\ell_a}(\Delta\nu_n)]_T$. Although there are some outliers at large $\Delta\nu$ (possibly due to larger cosmic variance), the residual $[C_\ell(\Delta\nu)]_R$ are found to be largely consistent with the $0 \pm 2\sigma$ noise levels.

We have used the best-fitting $[P(k)]_T$ values to calculate the mean squared brightness temperature $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$ shown by the orange asterisks in Fig. 8 along with the corresponding 2σ error bars. The $\Delta^2(k)$ values, σ, and the SNR $[\Delta^2(k)/\sigma]$ are tabulated for different k bins in Table 1. We find that $\Delta^2(k) > 0$ for all the k bins. The $\Delta^2(k)$ values in the first three and the last 2 k bins are consistent with noise at the $0 \pm 2\sigma$ level, whereas it is slightly above 2σ in the sixth bin. We interpret $\Delta^2(k)$ estimated in these six bins as arising from noise. The values of $\Delta^2(k)$ in the fourth and fifth bins exceed $0 + 2\sigma$, but are within $0 + 3\sigma$ and $0 + 5\sigma$, respectively. The $\Delta^2(k)$ values estimated in these two bins may have a contribution from residual systematics or FG leakage.

Considering the $\Delta^2(k)$ values, we find that $\Delta^2(k)$ has the smallest value $(32.75)^2$ mK^2 at the lowest k bin where $k = 0.804$ Mpc^{-1}. The values of $\Delta^2(k)$ as well as the 1σ errors are found to increase with increasing k as a power-law k^n, where the exponent $n \sim 2.5$ for $\Delta^2(k)$ and $n \sim 2.4$ for $\sigma(k)$, respectively. We have used the estimated $\Delta^2(k)$ and the σ values to place 2σ upper limits $\Delta^2_{\rm UL}(k) = \Delta^2(k) + 2\sigma$ on the 21-cm brightness temperature fluctuations at different k values. The 2σ upper limits $\Delta^2_{\rm UL}(k)$ are also tabulated in Table 1. We find the tightest constraint on the upper limit to be $\Delta^2_{\rm UL}(k) \leq (58.67)^2$ mK^2 at $k = 0.804$ Mpc^{-1}.

Fig. 8 also shows (blue dashed line) the results from Paper I where we have used the Total TGE as against the Cross TGE used here. We find that the present analysis shows significant improvement over Paper I throughout the entire k range. Particularly near $k \sim 1$ Mpc^{-1}, the value of $\Delta^2(k)$ is nearly 20 times smaller in the present analysis as compared to Paper I, whereas this factor is around 4–6 for the larger k bins. Comparing the upper limits, we had $\Delta^2_{\rm UL}(k) \leq (133.97)^2$ mK^2 at $k = 0.347$ Mpc^{-1} in Paper I, which is tightened to $\Delta^2_{\rm UL}(k) \leq (58.67)^2$ mK^2 at $k = 0.804$ Mpc^{-1} in this work. Note that the lowest k bin here is somewhat larger than that in Paper I. We have also compared our findings with Ch21, who have conducted a multiredshift analysis of the same observational data after splitting it into four subbands, each of 8-MHz bandwidth. The black dashed line in Fig. 8 shows the $\Delta^2(k)$ values from their $z = 2.19$ subband that is the close to our analysis ($z = 2.28$). We find that the present upper limits are close to the findings of Ch21, who reported $\Delta^2_{\rm UL}(k) \leq (61.49)^2$ mK^2 at $k = 1$ Mpc^{-1} at the redshift $z = 2.19$. We note that the bandwidth of the data analysed here is larger than that used in Ch21. For a nearly one-to-one comparison with Ch21, we have repeated the analysis using the same 8-MHz bandwidth for which

the results are presented in Appendix C. We find that the results are very similar to those presented here.

The upper limits on the 21-cm brightness temperature fluctuations allow us to constrain the cosmological H I abundance parameter $[\Omega_{\rm H\,I} b_{\rm H\,I}]$. Here, $\Omega_{\rm H\,I}$ is the comoving H I mass density in units of the current critical density (Lanzetta, Wolfe & Turnshek 1995), and $b_{\rm H\,I}$ is the H I bias parameter. The assumption here is that the H I distribution traces the underlying matter distribution through $b_{\rm H\,I}$. This assumption allows us to express $P_T(\mathbf{k})$ in terms of $P_m^s(\mathbf{k})$ the underlying matter PS in redshift space. Here, we use equations (13) and (14) of Paper I (which has been taken from Bharadwaj & Ali 2005),

$$P_T(\mathbf{k}) = [\Omega_{\rm H\,I} b_{\rm H\,I}]^2\, \bar{T}^2\, P_m^s(\mathbf{k}) \tag{17}$$

with the mean brightness temperature $\bar{T}$

$$\bar{T}(z) = 133\,{\rm mK}\,(1+z)^2 \left(\frac{h}{0.7}\right)\left(\frac{H_0}{H(z)}\right) \tag{18}$$

and $P_m^s(\mathbf{k})$ is the underlying dark matter PS in redshift space for which we have used a fitting formula (Eisenstein & Hu 1998), ignoring the effect of redshift space distortion.

We have used the estimated $\Delta^2_{\rm UL}(k)$ to place the corresponding 2σ upper limits $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL}$ that are also tabulated in Table 1. We obtain the tightest constraint of $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.072$ from the smallest bin $k = 0.804\,{\rm Mpc}^{-1}$. This is a factor of 3 improvement over Paper I where we were able to constrain $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.23$ at $k = 0.347\,{\rm Mpc}^{-1}$. Ch21 reported $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.11$ at $k = 1\,{\rm Mpc}^{-1}$ that is close to the upper limit that we obtain here.

This maximum likelihood approach of estimating $P(k)$ is different from the usual spherical binning approach (e.g. Paper I). The MLE is robust in the presence of small numbers of somewhat larger outliers (Huber 1981), and is optimal as we use inverse noise covariance weightage in the likelihood. We have validated the MLE in Appendix A. We have also carried out a consistency check on the best-fitting solutions and the error estimates of MLE by sampling the posterior probability distributions of the parameters using a Markov chain Monte Carlo (MCMC) algorithm. The details of the MCMC analysis are presented in Appendix B.

6 CONSTRAINING

In this section, we consider the possibility of utilizing the entire set of measured $[C_{\ell_a}(\Delta\nu_n)]$ values to directly constrain a single parameter $[\Omega_{\rm H\,I} b_{\rm H\,I}]$, without involving an intermediate step of estimating the spherical PS $[P(k)]_T$. Here also, we have modelled the measured $C_{\ell_a}(\Delta\nu_n)$ using equation (13), and used equation (14) to model the FG component $\left[C_{\ell_a}(\Delta\nu_n)\right]_{\rm FG}$. Considering the 21-cm signal, we have used only the TW modes to model $\left[C_{\ell_a}(\Delta\nu_n)\right]_T$ using

$$\left[C_{\ell_a}(\Delta\nu_n)\right]_T = [\Omega_{\rm H\,I} b_{\rm H\,I}]^2\, \bar{T}^2 \sum_q A_{nq}\, P_m(k_{\perp a}, k_{\parallel q}), \tag{19}$$

where $P_m(\mathbf{k})$ is the dark matter PS (equation 17) ignoring the effect of redshift space distortion. The entire 21-cm signal is now quantified by a single parameter $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$. Here, we have used maximum likelihood to simultaneously estimate the best-fitting values of the amplitude of FG modes $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$.

We have considered the measured $[C_{\ell_a}(\Delta\nu_n)]$ from all the available ℓ bins to constrain $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$. The total number of parameters now becomes 657, 656 of which represent $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and 1 parameter for $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$. The best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ is found to be 3.57×10^{-3} with an associated 1σ uncertainty of 1.41×10^{-3}.

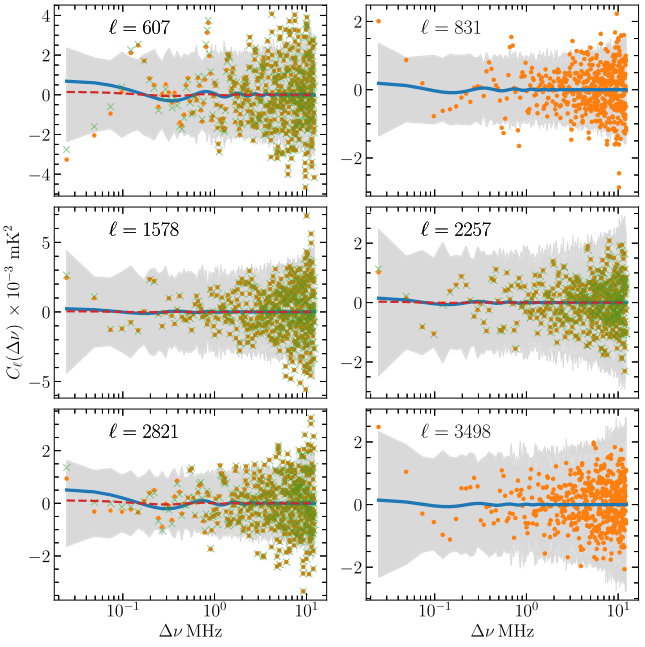

Figure 9. For the first 6 ℓ bins, the figure shows the recovered $[C_\ell(\Delta\nu)]_T$ (solid blue lines) and $[C_\ell(\Delta\nu)]_R$ (orange dots) corresponding to the best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ obtained from Set I (Table 2). The recovered $[C_\ell(\Delta\nu)]_T$ (red dashed lines) and $[C_\ell(\Delta\nu)]_R$ (green crosses) corresponding to the best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ obtained from Set II are also shown. The 2σ error in the measured $C_\ell(\Delta\nu)$ is shown by the grey shaded regions.

Table 2. $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$, its associated error due to noise, SNR, and the 2σ upper limits $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL}$ are shown.

Set	ℓ	$[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ $\times 10^{-4}$	Error $\times 10^{-3}$	SNR	$[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL}$
I	All	35.74	1.41	2.53	0.080
II	617, 1578, 2257, and 2821	7.51	1.47	0.51	0.061

The reduced χ^2 is found to be 1.21, which is the same as found in Section 5 where we estimated the spherical PS. The blue solid lines in the different panels of Fig. 9 show the recovered $[C_\ell(\Delta\nu)]_T$ corresponding to the best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$. The residuals $[C_\ell(\Delta\nu)]_R$ and the 2σ errors due to noise are shown using the orange dots and the grey shaded regions, respectively. We find that the recovered $[C_\ell(\Delta\nu)]_T$ as well as the residuals $[C_\ell(\Delta\nu)]_R$ are roughly consistent with the 2σ noise level. The 2σ upper limit on $[\Omega_{\rm H\,I} b_{\rm H\,I}]$ is found to be $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.080$. These results, hereafter referred to as 'Set I', are tabulated in Table 2.

We next consider the possibility of improving the constraints on $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ by using a subset of the measured $[C_{\ell_a}(\Delta\nu_n)]$. Here, we have repeated the analysis considering various combinations of ℓ bins to find that a particular set ($\ell = 617$, 1578, 2257, and 2821) provides better constraints. These results, hereafter referred to as 'Set II', are also tabulated in Table 2. We have modelled the 2000 available $[C_{\ell_a}(\Delta\nu_n)]$ values using 175 parameters for $[P(k_{\perp a}, k_{\parallel m})]_{\rm FG}$ and 1 parameter for $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$, and find the reduced χ^2 to be 1.58, which indicates an acceptable fit. The best-fitting value of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ is found to be $7.51 \times 10^{-4} \pm 1.47 \times 10^{-3}$ that can be attributed to noise. The dashed red lines in Fig. 9 show the recovered $[C_\ell(\Delta\nu)]_T$ whose values are found to be close to zero throughout the $\Delta\nu$ range.

The residuals $[C_\ell(\Delta\nu)]_R$ (green crosses) are found to be quite similar to those found for Set I. We conclude that we are able to separate out the FGs from the measured $C_\ell(\Delta\nu)$ in Set II. We place a 2σ upper limit of $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.061$ that is better than the upper limit obtained from Set I or from the spherical PS (Table 1).

Although we do not report a detection of the 21-cm signal, we compare the upper limit with the available observational and theoretical constraints on the parameters $\Omega_{\rm H\,I}$ and $b_{\rm H\,I}$ at various redshifts in the post-reionization Universe (Padmanabhan, Choudhury & Refregier 2015, and references therein). The H I spectral stacking analysis (e.g. Rhee et al. 2016) at $z < 0.5$, H I IM experiments in cross-correlation with galaxy surveys (e.g. Chang et al. 2010; Masui et al. 2013; CHIME Collaboration 2022b) at $z < 1.3$, and observations of DLAs and sub-DLAs from quasar spectra (e.g. Prochaska & Herbert-Fort 2004; Péroux et al. 2005; Kanekar et al. 2009; Prochaska & Wolfe 2009; Noterdaeme et al. 2012; Zafar et al. 2013) at $2 < z < 5.5$ estimate $\Omega_{\rm H\,I} \sim 10^{-3}$. On the other hand, various simulations (e.g. Bagla et al. 2010; Marín et al. 2010; Guha Sarkar et al. 2012; Sarkar et al. 2016) indicate that $1 \leq b_{\rm H\,I} \leq 2$ for the redshift we have considered here. These values of $\Omega_{\rm H\,I}$ and $b_{\rm H\,I}$ imply that our present upper limit $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.061$ is roughly 30–60 times larger than currently estimated values.

7 SUMMARY AND CONCLUSIONS

The 21-cm IM has long been recognized as a powerful technique for efficiently mapping the large-scale structures in the Universe out to high redshifts. Aiming a 21-cm IM at $z = 2.28$ in Paper I, we considered 24.4-MHz-bandwidth data that were taken from a four-night observation of the ELAIS-N1 field at the Band 3 (250–500 MHz) of uGMRT. We analysed the flagged, calibrated, and point source-subtracted visibility data (details in Chakraborty et al. 2019b) with the TGE, which allows us to taper the sky response to suppress the contribution from sources in the periphery of the telescope's field of view. In this work, we introduce a Cross TGE that grids the cross-polarizations (i.e. two mutually perpendicular polarizations) RR and LL of visibilities independently, and then correlates them to obtain $C_\ell(\Delta\nu)$. We expect this to mitigate several effects like noise bias, calibration errors, etc., which affect the Total $C_\ell(\Delta\nu)$ (used in Paper I) that combines the two polarizations.

We have compared the $C_\ell(\Delta\nu)$ estimated from the Total and Cross estimators in Section 3. We find that $C_\ell(\Delta\nu)$ from both the estimators have comparable values once the DC component is subtracted out (Fig. 2). Considering the $\Delta\nu$ dependence, we find that the Total $C_\ell(\Delta\nu)$ sharply decorrelates within $\Delta\nu < 1$ MHz and exhibits rapid oscillations at larger $\Delta\nu$. In comparison, the Cross $C_\ell(\Delta\nu)$ decorrelates smoothly, and also with a considerably smaller oscillation amplitude. A combination of these two effects is reflected in the cylindrical PS $P(k_\perp, k_\parallel)$ that is evaluated from $C_\ell(\Delta\nu)$ using equation (11). Compared to the Total PS, we find that the Cross PS reaches the expected noise level at comparatively smaller $k_\parallel$ modes (Fig. 3), considerably restricting the FG leakage. Further, the smaller oscillation amplitude in the Cross $C_\ell(\Delta\nu)$ results in significantly lower power in the Cross $P(k_\perp, k_\parallel)$ beyond the theoretically predicted wedge boundary $[k_\parallel]_H$.

The Cross $P(k_\perp, k_\parallel)$ heatmap (Fig. 4) shows that although the bulk of the FGs lie inside $[k_\parallel]_H$, there is a considerable amount of leakage beyond it. We have avoided these FG-dominated modes (FG modes) and selected the relatively FG-free TW region to put constraints (upper limits) on the cosmological 21-cm signal. We have checked the noise statistics of the PS in the TW region through the quantity X (defined in equation 12) that is expected to follow a standard normal distribution if the values of $P(k_\perp, k_\parallel)$ in the TW region are entirely due to the system noise. The PDF of X, in the central region ($|X| \leq 30$), is found to be mostly symmetric with a positive mean $\mu = 0.61$ and a standard deviation $\sigma_{\rm Est} = 4.77$ with no negative outlier values (Fig. 5). The absence of large negative outliers ensures that systematics, like discontinuities in the band edges, large phase errors, etc., are not affecting our PS results. The standard deviation $\sigma_{\rm Est} > 1$ suggests that the statistical fluctuations in $P(k_\perp, k_\parallel)$ are underestimated by the system noise-only simulations. To deal with this, we have scaled up our error estimates by the $\sigma_{\rm Est}$ factor. On a related note, we find that (similar to Paper I) a t-distribution adequately describes the central region of the PDF, but it fails to fit the positive tail. We show that a Lorentzian distribution represents the PDF better. We also notice the positive tail in the PDF to be more restricted (as compared to that of Total PS shown in Paper I) in the Cross PS analysis, suggesting a cleaner (less FG contaminated) TW region.

In Section 5, we have explored the fact that the 21-cm signal is isotropic in three-dimensional Fourier space, and hence its fluctuations can be entirely quantified with a spherical 21-cm PS $P(k)$. We introduced an MLE that estimates $P(k)$ directly from the measured $C_\ell(\Delta\nu)$ without explicitly estimating $P(k_\perp, k_\parallel)$. We model the measured $C_\ell(\Delta\nu)$ as a combination of FGs, 21-cm signal, and residual systematics (equation 13). We use the FG modes to model the FGs (equation 14), and the TW modes to model the 21-cm signal (equation 15). We incorporate the isotropy of the 21-cm signal by dividing the TW region into spherical bins where the amplitude of the 21-cm PS $P(k)$ is a constant. This approach is further illustrated with a schematic diagram in Fig. 6. We maximize the likelihood $\mathcal{L} \propto \exp(-\chi^2/2)$, where the χ^2 is defined through equation (16), to find the best-fitting values of the free model parameters, $P(k)$ and the amplitudes of the FG modes $[P(k_\perp, k_\parallel)]_{\rm FG}$. We have validated the MLE using simulations in Appendix A, and also presented an MCMC analysis (Appendix B) to show the consistency of the best-fitting MLE solutions and their error estimates. The MCMC analysis also shows that the parameters $P(k)$ are uncorrelated in different spherical k bins. The maximum likelihood estimation of the 21-cm PS is likely to be more robust (than spherical averaging) in the presence of outliers (Huber 1981). This framework is also more suitable for propagating any correlation between FG and 21-cm modes. Thus, this method provides a self-consistent way of determining unbiased error bars on the 21-cm PS modes. We expect that the error estimation will play an increasingly important role when IM experiments come close to making the first detections.

We have used the best-fitting values of $P(k)$ to recover the isotropic component $[C_\ell(\Delta\nu)]_T$ that is found to be largely consistent with the noise fluctuations at 2σ level (Fig. 7). The brightness temperature fluctuation, $\Delta^2(k)$, is also found to be consistent with the noise at 2σ level in most of the k bins. The $\Delta^2(k)$ values are found to be 6–20 times tighter than Paper I, and are comparable to the findings of Ch21 (Fig. C1). The tightest constraint on the upper limits $\Delta^2_{\rm UL}(k) \leq (58.67)^2\,{\rm mK}^2$ at $k = 0.804\,{\rm Mpc}^{-1}$ suggests $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.072$. These results are tabulated in Table 1. The upper limits are nearly 5.2 times better than our earlier results (Paper I), where we have reported $\Delta^2_{\rm UL}(k) \leq (133.97)^2\,{\rm mK}^2$ and $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.23$ at $k = 1.03\,{\rm Mpc}^{-1}$. Note that Ch21 results [$\Delta^2_{\rm UL}(k) \leq (61.49)^2\,{\rm mK}^2$ and $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.11$ at $k = 1\,{\rm Mpc}^{-1}$] are similar to our current upper limits.

Finally, we have also considered the possibility of using the entire set of $C_\ell(\Delta\nu)$ measurements to directly constrain a single parameter $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$, without involving the intermediate step of estimating $P(k)$. In this approach, we model the 21-cm signal using equation (19), and estimate $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$ using the MLE. We

find by combining four ℓ bins, as quoted in Table 2 (Set II), $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2 = 7.51 \times 10^{-4} \pm 1.47 \times 10^{-3}$ that is attributable to noise at 1σ. Although the 2σ upper limit $[\Omega_{\rm H\,I} b_{\rm H\,I}]_{\rm UL} \leq 0.061$ is ∼50 times larger than the expected value (see e.g. Padmanabhan et al. 2015), this is a considerable improvement over earlier IM works at this redshift.

Although the upper limit is a significant improvement over Paper I, a tighter constraint on the upper limit is expected if we can completely remove FGs from the data. Recently, Trott et al. (2022) have used a smooth FG filter DAYENU (Ewall-Wice et al. 2021) and estimated the MAPS $C_\ell(\nu_a, \nu_b)$ from high-redshift ($z = 6.2$–7.5) Murchison Widefield Array (Tingay et al. 2013) data. Furthermore, the full MAPS $C_\ell(\nu_a, \nu_b)$, which does not assume the 21-cm signal to be ergodic (Mondal et al. 2018), also provides possibilities for FG removal using eigendecomposition as presented in Liu & Tegmark (2012), and also discussed in Mondal et al. (2022). The idea is that the FGs, being featureless in frequency, can be accurately captured by means of only a few of the leading eigenmodes and this can be used to subtract out the FG contribution. We plan to consider these possibilities in future work.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for a careful reading of the manuscript and for the insightful comments and suggestions. We thank the staff of GMRT for making this observation possible. GMRT is run by National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research. AE thanks Sukhdeep Singh for valuable discussions. AG would like to acknowledge IUCAA, Pune for providing support through the associateship programme. SB would like to acknowledge funding provided under the MATRICS grant SERB/F/9805/2019-2020 of the Science & Engineering Research Board, a statutory body of Department of Science & Technology (DST), Government of India. A part of this work has used the Supercomputing facility 'PARAM Shakti' of IIT Kharagpur established under National Supercomputing Mission (NSM), Government of India and supported by Centre for Development of Advanced Computing (CDAC), Pune.

DATA AVAILABILITY

The data used are available upon reasonable request to the corresponding authors.

REFERENCES

Ali S. S., Bharadwaj S., 2014, J. Astrophys. Astron., 35, 157
Ali S. S., Bharadwaj S., Chengalur J. N., 2008, MNRAS, 385, 2166
Anderson C. J. et al., 2018, MNRAS, 476, 3382
Ansari R. et al., 2012, A&A, 540, A129
Asad K. M. B. et al., 2015, MNRAS, 451, 3709
Bagla J. S., Khandai N., Datta K. K., 2010, MNRAS, 407, 567
Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, MNRAS, 434, 1239
Bernardi G. et al., 2009, A&A, 500, 965
Bharadwaj S., Ali S. S., 2005, MNRAS, 356, 1519
Bharadwaj S., Sethi S. K., 2001, J. Astrophys. Astron., 22, 293
Bharadwaj S., Srikant P. S., 2004, J. Astrophys. Astron., 25, 67
Bharadwaj S., Nath B. B., Sethi S. K., 2001, J. Astrophys. Astron., 22, 21
Bharadwaj S., Sethi S. K., Saini T. D., 2009, Phys. Rev. D, 79, 083538
Bharadwaj S., Pal S., Choudhuri S., Dutta P., 2018, MNRAS, 483, 5694
Blake C., Ferreira P. G., Borrill J., 2004, MNRAS, 351, 923
Bowman J. D., Morales M. F., Hewitt J. N., 2009, ApJ, 695, 183
Bull P., Camera S., Raccanelli A., Blake C., Ferreira P., Santos M., Schwarz D. J., 2015a, Proc. Sci., Measuring Baryon Acoustic Oscillations with Future SKA Surveys. SISSA, Trieste, PoS#24
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015b, ApJ, 803, 21
Chakraborty A. et al., 2019a, MNRAS, 487, 4102
Chakraborty A. et al., 2019b, MNRAS, 490, 243
Chakraborty A. et al., 2021, ApJ, 907, L7
Chakraborty A., Datta A., Mazumder A., 2022, ApJ, 929, 104
Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Chapman E. et al., 2012, MNRAS, 423, 2518
Chen X., 2012, Int. J. Mod. Phys.: Conf. Ser., 12, 256
CHIME Collaboration, 2022a, ApJS, 261, 29
CHIME Collaboration, 2022b, preprint (arXiv:2202.01242)
Choudhuri S., Bharadwaj S., Ghosh A., Ali S. S., 2014, MNRAS, 445, 4351
Choudhuri S., Bharadwaj S., Roy N., Ghosh A., Ali S. S., 2016a, MNRAS, 459, 151
Choudhuri S., Bharadwaj S., Chatterjee S., Ali S. S., Roy N., Ghosh A., 2016b, MNRAS, 463, 4093
Choudhuri S., Bharadwaj S., Ali S. S., Roy N., Intema H. T., Ghosh A., 2017, MNRAS, 470, L11
Choudhuri S., Ghosh A., Roy N., Bharadwaj S., Intema H. T., Ali S. S., 2020, MNRAS, 494, 1936
Condon J. J., 1989, ApJ, 338, 13
Condon J. J. et al., 2012, ApJ, 758, 23
Cress C. M., Helfand D. J., Becker R. H., Gregg M. D., White R. L., 1996, ApJ, 473, 7
Cunnington S. et al., 2023, MNRAS, 518, 6262
Datta A., Bowman J. D., Carilli C. L., 2010, ApJ, 724, 526
Datta K. K., Choudhury T. R., Bharadwaj S., 2007, MNRAS, 378, 119
Dawson K. S. et al., 2016, AJ, 151, 44
Di Matteo T., Perna R., Abel T., Rees M. J., 2002, ApJ, 564, 576
Dillon J. S. et al., 2014, Phys. Rev. D, 89, 023002
Dillon J. S. et al., 2015, Phys. Rev. D, 91, 123011
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Ewall-Wice A. et al., 2021, MNRAS, 500, 5195
Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Ghosh A., Bharadwaj S., Ali S. S., Chengalur J. N., 2011a, MNRAS, 411, 2426
Ghosh A., Bharadwaj S., Ali S. S., Chengalur J. N., 2011b, MNRAS, 418, 2584
Ghosh A., Prasad J., Bharadwaj S., Ali S. S., Chengalur J. N., 2012, MNRAS, 426, 3295
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Guha Sarkar T., Mitra S., Majumdar S., Choudhury T. R., 2012, MNRAS, 421, 3570
Gupta Y. et al., 2017, Curr. Sci., 113, 707
Haslam C. G. T., Klein U., Salter C. J., Stoffel H., Wilson W. E., Cleary M. N., Cooke D. J., Thomasson P., 1981, A&A, 100, 209
Haslam C. G. T., Salter C. J., Stoffel H., Wilson W. E., 1982, A&AS, 47, 1
Hazelton B. J., Morales M. F., Sullivan I. S., 2013, ApJ, 770, 156
Hazra D. K., Guha Sarkar T., 2012, Phys. Rev. Lett., 109, 121301
Ho M.-F., Bird S., Garnett R., 2021, MNRAS, 507, 704
Högbom J. A., 1974, A&AS, 15, 417
Huber P. J., 1981, Robust Statistics. Wiley Series in Probability and Statistics. Wiley, New York
Jelić V. et al., 2008, MNRAS, 389, 1319
Kanekar N., Prochaska J. X., Ellison S. L., Chengalur J. N., 2009, MNRAS, 396, 385
Kennedy F., Bull P., 2021, MNRAS, 506, 2638
Kennedy F., Bull P., Wilensky M., Choudhuri S., 2022, preprint (arXiv:2211.05088)
Kern N. S., Liu A., 2021, MNRAS, 501, 1463
Kumar J., Dutta P., Roy N., 2020, MNRAS, 495, 3683
Kumar J., Dutta P., Choudhuri S., Roy N., 2022, MNRAS, 512, 186
Lanzetta K. M., Wolfe A. M., Turnshek D. A., 1995, ApJ, 440, 435
Liu A., Parsons A. R., Trott C. M., 2014a, Phys. Rev. D, 90, 023018

Liu A., Tegmark M., 2012, MNRAS, 419, 3491
Liu A., Parsons A. R., Trott C. M., 2014b, Phys. Rev. D, 90, 023019
Loeb A., Wyithe J. S. B., 2008, Phys. Rev. Lett., 100, 161301
Long H., Morales-Gutiérrez C., Montero-Camacho P., Hirata C. M., 2022, preprint (arXiv:2210.02385)
Mao Y., Tegmark M., McQuinn M., Zaldarriaga M., Zahn O., 2008, Phys. Rev. D, 78, 023529
Marín F. A., Gnedin N. Y., Seo H.-J., Vallinotto A., 2010, ApJ, 718, 972
Masui K. W. et al., 2013, ApJ, 763, L20
Mazumder A., Chakraborty A., Datta A., Choudhuri S., Roy N., Wadadekar Y., Ishwara-Chandra C. H., 2020, MNRAS, 495, 4071
Mertens F. G., Ghosh A., Koopmans L. V. E., 2018, MNRAS, 478, 3640
Mertens F. G. et al., 2020, MNRAS, 493, 1662
Mondal R., Bharadwaj S., Datta K. K., 2018, MNRAS, 474, 1390
Mondal R., Bharadwaj S., Iliev I. T., Datta K. K., Majumdar S., Shaw A. K., Sarkar A. K., 2019, MNRAS, 483, L109
Mondal R., Mellema G., Murray S. G., Greig B., 2022, MNRAS, 514, L31
Morales M. F., Hewitt J., 2004, ApJ, 615, 7
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, ApJ, 752, 137
Newburgh L. B. et al., 2016, in Hall H. J., Gilmozzi R., Marshall H. K., eds, Proc. SPIE Conf. Ser. Vol. 9906, Ground-Based and Airborne Telescopes VI. SPIE, Bellingham, p. 99065X
Newman J. A. et al., 2013, ApJS, 208, 5
Noterdaeme P. et al., 2012, A&A, 547, L1
Nuttall A. H., 1981, IEEE Trans. Acoust. Speech Signal Process., 29, 84
Owen F. N., Morrison G. E., 2008, AJ, 136, 1889
Paciga G. et al., 2011, MNRAS, 413, 1174
Padmanabhan H., Choudhury T. R., Refregier A., 2015, MNRAS, 447, 3745
Pal S., Bharadwaj S., Ghosh A., Choudhuri S., 2021, MNRAS, 501, 3378
Pal S. et al., 2022, MNRAS, 516, 2851
Parsons A. R., Backer D. C., 2009, AJ, 138, 219
Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, ApJ, 756, 165
Pen U.-L., Staveley-Smith L., Peterson J. B., Chang T.-C., 2009a, MNRAS, 394, L6
Pen U.-L., Chang T.-C., Hirata C. M., Peterson J. B., Roy J., Gupta Y., Odegova J., Sigurdson K., 2009b, MNRAS, 399, 181
Péroux C., Dessauges-Zavadsky M., D'Odorico S., Sun Kim T., McMahon R. G., 2005, MNRAS, 363, 479
Planck Collaboration VI, 2020, A&A, 641, A6
Pober J. C. et al., 2013, ApJ, 768, L36
Pober J. C. et al., 2014, ApJ, 782, 66
Pober J. C. et al., 2016, ApJ, 819, 8
Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
Prochaska J. X., Herbert-Fort S., 2004, PASP, 116, 622
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Randall K. E., Hopkins A. M., Norris R. P., Zinn P. C., Middelberg E., Mao M. Y., Sharp R. G., 2012, MNRAS, 421, 1644
Reich P., Reich W., 1988, A&AS, 74, 7
Rhee J., Lah P., Chengalur J. N., Briggs F. H., Colless M., 2016, MNRAS, 460, 2675
Roberts D. H., Lehar J., Dreher J. W., 1987, AJ, 93, 968
Santos M. G., Cooray A., Knox L., 2005, ApJ, 625, 575
Sarkar D., Bharadwaj S., Anathpindika S., 2016, MNRAS, 460, 4310
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Singal J., Stawarz Ł., Lawrence A., Petrosian V., 2010, MNRAS, 409, 1172
Smirnov O. M., 2011, A&A, 527, A107
Subrahmanya C. R., Manoharan P. K., Chengalur J. N., 2017, J. Astrophys. Astron., 38, 10
Swarup G., Ananthakrishnan S., Kapahi V. K., Rao A. P., Subrahmanya C. R., Kulkarni V. K., 1991, Curr. Sci., 60, 95
Switzer E. R. et al., 2013, MNRAS, 434, L46
Thyagarajan N. et al., 2013, ApJ, 776, 6
Thyagarajan N., Parsons A. R., DeBoer D. R., Bowman J. D., Ewall-Wice A. M., Neben A. R., Patra N., 2016, ApJ, 825, 9
Tingay S. J. et al., 2013, PASA, 30, e007
Trott C. M., 2016, MNRAS, 461, 126
Trott C. M., Wayth R. B., Tingay S. J., 2012, ApJ, 757, 101
Trott C. M. et al., 2016, ApJ, 818, 139
Trott C. M. et al., 2020, MNRAS, 493, 4711
Trott C. M., Mondal R., Mellema G., Murray S. G., Greig B., Line J. L. B., Barry N., Morales M. F., 2022, A&A, 666, A106
Van Eck C. L. et al., 2018, A&A, 613, A58
Vedantham H., Udaya Shankar N., Subrahmanyan R., 2012, ApJ, 745, 176
Visbal E., Loeb A., Wyithe S., 2009, J. Cosmol. Astropart. Phys., 2009, 030
Wilman R. J., Röttgering H. J. A., Overzier R. A., Jarvis M. J., 2003, MNRAS, 339, 695
Wolfe A. M., Lanzetta K. M., Foltz C. B., Chaffee F. H., 1995, ApJ, 454, 698
Wolz L. et al., 2021, MNRAS, 510, 3495
Wuensche C., 2019, J. Phys.: Conf. Ser., 1269, 012002
Wyithe J. S. B., Loeb A., Geil P. M., 2008, MNRAS, 383, 1195
Zafar T., Péroux C., Popping A., Milliard B., Deharveng J. M., Frank S., 2013, A&A, 556, A141
Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622
Zheng H. et al., 2017, MNRAS, 464, 3486

APPENDIX A: VALIDATION OF CROSS ESTIMATOR

In Paper I, we have validated the Total correlation TGE (equation 5) using simulated visibilities corresponding to a sky signal that is assumed to be a Gaussian random field with a PS $P^m(k)$

$$P^m(\mathbf{k}) = A \left(\frac{k}{k_0}\right)^n \mathrm{mK^2\,Mpc^3} \tag{A1}$$

having an arbitrarily chosen value of $A = 10$, $k_0 = 1\,\mathrm{Mpc}^{-1}$, and a power-law index $n = -2$. The simulated visibilities incorporated the same parameters (such as the baseline distribution, flagging etc.) of the data being used here. The details of the simulations can be found in Paper I. This work uses the Cross TGE (equation 6) that we validate here using the same simulated data.

We have applied the Cross estimator (equation 6) on the simulated visibilities, and analysed the simulated data identical to the actual data, to estimate $C_\ell(\Delta\nu)$. We have used 16 independent realizations of the simulation to estimate the mean $C_\ell(\Delta\nu)$ and the 2σ errors shown in the uppermost panel of Fig. A1 at three representative values of ℓ. We have also shown (solid lines) the analytical model predictions calculated using equation (9). We see that the $C_\ell(\Delta\nu)$ estimated from the simulations closely matches the analytical predictions, which are mostly within the shaded region showing the 2σ uncertainty.

The last two panels show the validation of the MLE as a PS estimator. The middle panel shows the estimated spherical PS $P(k)$ (blue filled circles) and 2σ error bars due to the cosmic variance. The input model $P^m(k)$ is shown with the magenta solid line. We see that $P(k)$ is in reasonably good agreement with $P^m(k)$ across the entire k range considered here. The bottom panel shows the fractional deviation $\delta = [P(k) - P^m(k)]/P^m(k)$ (data points) and the expected 2σ statistical fluctuations for the same (grey shaded region). We have $|\,\delta\,| \lesssim 10$ per cent in most of the k bins shown here. We see that the δ values are all consistent with the predicted 2σ errors. We have somewhat larger error bars at the smallest k bin. The convolution with the tapering window function (equation 3) is expected to become important at the small baselines (Choudhuri et al. 2014), and this possibly contributes to enhancing the statistical fluctuations in the small k-bins. A part of the deviations could also arise from the low baseline densities in some of the bins. The entire validation presented here used *exactly* the same $(k_\perp, k_\parallel)$ modes as those that have been used for the actual data. In summary, we have validated the Cross estimator

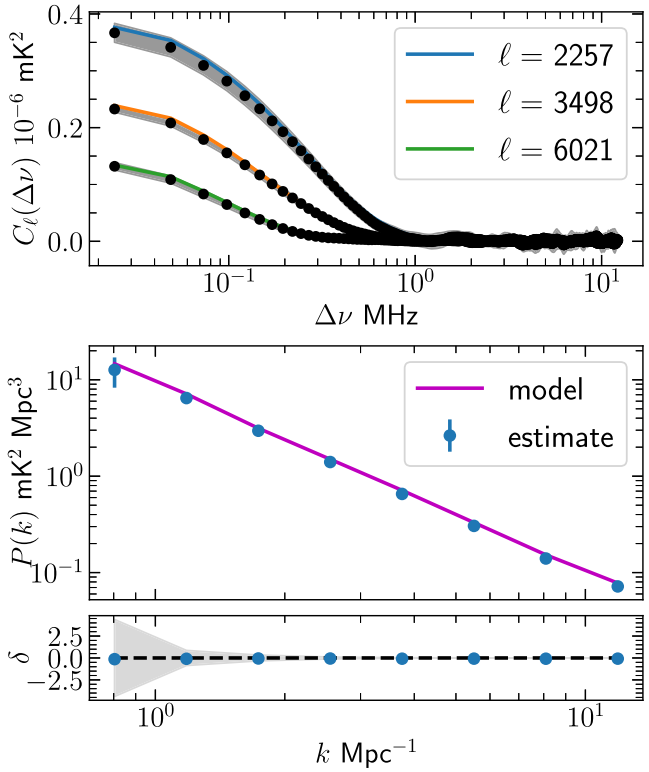

Figure A1. The uppermost panel shows the validation of Cross TGE (equation 6). The data points (black filled circles) show the mean $C_\ell(\Delta\nu)$ with 2σ errors (shaded region) estimated from 16 realizations of the simulated sky signal. The solid lines show the analytical predictions of $C_\ell(\Delta\nu)$ corresponding to the input model $P^m(k)$. The middle panel shows the estimated spherically binned PS $P(k)$ (blue filled circles) and 2σ error bars estimated using the MLE. The input model $P^m(k)$ is shown with the purple solid line. The bottom panel shows the fractional deviation $\delta = [P(k) - P^m(k)]/P^m(k)$ (data points) and the expected 2σ statistical fluctuations for the same (grey shaded region).

and find that it can recover the input model PS to an accuracy close to $\lesssim$10 per cent across the entire k range considered here.

APPENDIX B: MCMC ANALYSIS

An MCMC algorithm allows us to draw parameter samples that are consistent with the measured data. Using the parameter samples, we construct the probability distribution of the model parameters and also characterize the correlation between the parameters. Here, we use an MCMC to check the consistency of the best-fitting values and their error estimates that we have obtained using the MLE presented in Sections 5 and 6. For this purpose, we first consider the posterior probability distribution of the parameters $[P(k)]_T$. We keep $[P(k_\perp, k_\parallel)]_{\rm FG}$, the amplitudes of the FG modes, which are the other free parameters of our model, fixed at the maximum likelihood estimates. Since the FG and the TW modes are uncorrelated (assuming that the initial density fluctuations are Gaussian on large scales), keeping $[P(k_\perp, k_\parallel)]_{\rm FG}$ fixed does not change the posterior of $[P(k)]_T$. We have used a uniform prior $\mathcal{U}(-\infty, \infty)$ on $[P(k)]_T$; this allows $[P(k)]_T$ to have any possible real numbers with an equal probability. The prior, along with the likelihood defined through equation (16), yields the posterior ($\propto$ likelihood $\times$ prior) from which we draw samples using an MCMC.

Table B1. Same as Table 1 except that the mean value of $\Delta^2(k)$ and standard deviations σ are estimated using an MCMC.

k (Mpc^{-1})	$\Delta^2(k)$ (mK)^2	1σ (mK)^2	SNR	$\Delta^2_{\rm UL}(k)$ (mK)^2	$[\Omega_{\rm H\,I}b_{\rm H\,I}]_{\rm UL}$
0.804	$(32.71)^2$	$(34.33)^2$	0.908	$(58.55)^2$	0.072
1.181	$(47.60)^2$	$(48.17)^2$	0.976	$(83.11)^2$	0.089
1.736	$(86.16)^2$	$(65.19)^2$	1.747	$(126.19)^2$	0.121
2.551	$(158.63)^2$	$(93.12)^2$	2.902	$(206.17)^2$	0.177
3.748	$(279.53)^2$	$(149.72)^2$	3.486	$(350.67)^2$	0.273
5.507	$(353.38)^2$	$(242.35)^2$	2.126	$(492.29)^2$	0.350
8.093	$(501.87)^2$	$(390.10)^2$	1.655	$(745.81)^2$	0.488
11.892	$(711.87)^2$	$(696.66)^2$	1.044	$(1215.50)^2$	0.588

We have used the affine-invariant ensemble sampling (AIES; Goodman & Weare 2010) algorithm implemented in the PYTHON module EMCEE (Foreman-Mackey et al. 2013) to draw samples from the posterior probability distribution of $[P(k)]_T$. Fig. B1 shows the posterior probability distributions of $\Delta^2(k)$ that is obtained by scaling the MCMC samples of $[P(k)]_T$ with $k^3/2\pi^2$. The panels in the main diagonal show the 1D marginalized posterior probability distribution of $\Delta^2(k)$, whereas the off-diagonal panels show the 2D projections of the posterior probability distribution of each pair of the parameters. In each panel, the solid green lines show the mean value of $\Delta^2(k)$ derived from the MCMC samples. The maximum likelihood solutions (hereafter, MLE solutions), which are obtained from maximizing the likelihood (Section 5), are also shown (dashed red lines) along with the MCMC solutions. Note that all values quoted in the figure are in mK^2 units. Considering the diagonal panels, the dashed blue and the dashed–dot orange vertical lines demarcate the 1σ and 2σ levels, respectively. In the off-diagonal panels, the blue (dashed), orange (dashed–dot), and pink (dotted) contours show the 1σ, 2σ, and 3σ levels, respectively. For all the k values considered here, the MLE solutions are found to lie within the 1σ uncertainty intervals of the MCMC solutions. We also do not find any correlation among the parameters. The error estimates are also quite similar in both analyses. We have also computed the 2σ upper limits on $\Delta^2(k)$ and $[\Omega_{\rm H\,I}b_{\rm H\,I}]$ using the MCMC samples. The MCMC means, their uncertainties, and the upper limits are highlighted in Table B1.

Next, we carry out an MCMC analysis to constrain $[\Omega_{\rm H\,I}b_{\rm H\,I}]$, which we have previously done using the MLE in Section 6. We have used the measured $C_\ell(\Delta\nu)$ values from the same $4\,\ell$ bins mentioned in the Set II of Table 2. We have kept the $[P(k_\perp, k_\parallel)]_{\rm FG}$ values fixed at the maximum likelihood estimates, and used a uniform prior $\mathcal{U}(-2, 2)$ on $[\Omega_{\rm H\,I}b_{\rm H\,I}]^2$. We have also checked that a broader range on the prior does not alter the posterior. Fig. B2 shows the resulting posterior probability distribution of the model parameter $[\Omega_{\rm H\,I}b_{\rm H\,I}]^2$. The vertical lines show the mean (solid green) and the associated 1σ (dashed blue) and 2σ (dashed–dot orange) uncertainties along with the best-fitting value obtained from MLE (red dashed). We find $[\Omega_{\rm H\,I}b_{\rm H\,I}]^2 = 7.50 \times 10^{-4} \pm 1.46 \times 10^{-3}$ that translates into a 2σ upper limit $[\Omega_{\rm H\,I}b_{\rm H\,I}]_{\rm UL} \leq 6.06 \times 10^{-2}$. These values are found to be close to the MLE solutions.

In both the MCMC runs, we have used 200 random walkers and initialized them to the vicinity of the parameters derived from the MLE to ensure a faster convergence. We first performed 300 burn-in steps before running the full chain of 10 000 steps. To check the convergence of the MCMC, we have considered the quantity τ_f, the integrated autocorrelation time, which gives an

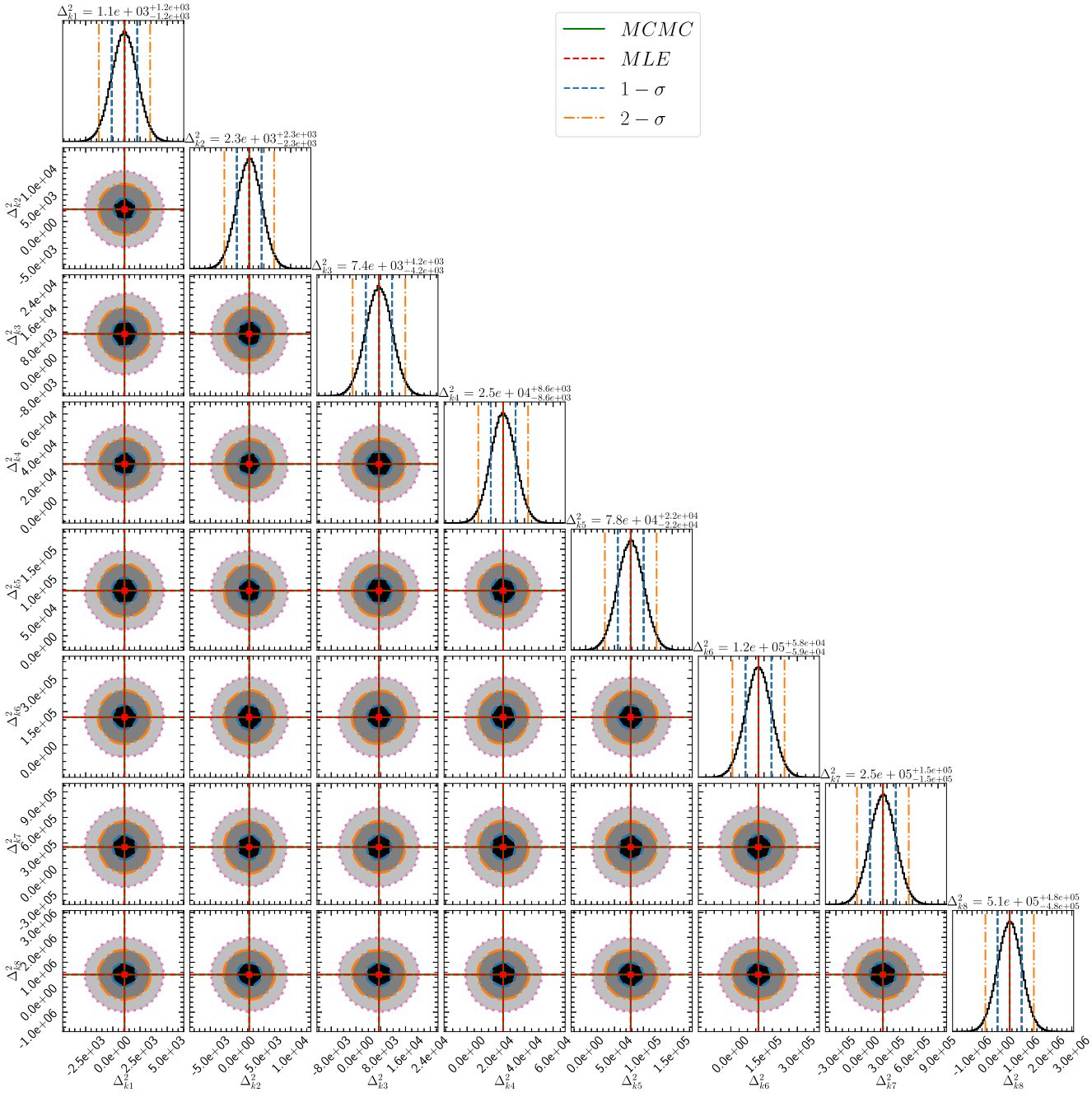

Figure B1. The diagonal shows the marginalized 1D posterior probability distribution of $\Delta^2(k)$ in mK^2. The vertical lines display the mean (solid green) and the associated 1σ (dashed blue) and 2σ (dashed–dot orange) standard deviations of the MCMC samples along with the best-fitting value from MLE (red dashed). Each panel of the off-diagonal shows the marginalized 2D projections of the posterior probability distribution of each pair of the parameters. The blue (dashed), orange (dashed–dot), and pink (dotted) contours show the 1σ, 2σ, and 3σ levels, respectively. The green (solid) lines and the red (dashed) lines show the MCMC median and the MLE solutions, respectively. The plot has made use of the PYTHON module CORNER (Foreman-Mackey 2016).

estimate of the number of steps required for the chains to converge (Goodman & Weare 2010). Here, we quote the mean autocorrelation time, which is the mean of the integrated autocorrelation time estimated for the chains corresponding to each parameter. The mean autocorrelation time is found to be ∼82 and ∼26 steps for the 8 and 1 parameter cases of $P(k)$ and $[\Omega_{\mathrm{H\,I}} b_{\mathrm{H\,I}}]^2$, respectively. We have conservatively chosen a significantly large number of steps ($> 50 \times \tau_f$) to reduce the sample variance. We have also checked the trace plots of the MCMC to ensure the convergence of AIES.

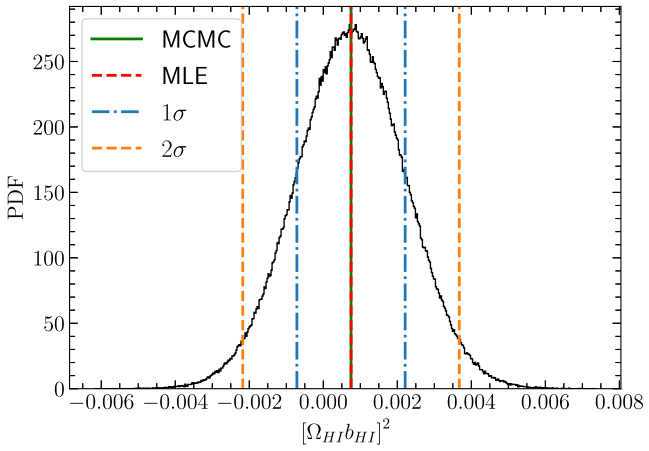

Figure B2. This figure shows the posterior probability distribution of $[\Omega_{\rm H\,I} b_{\rm H\,I}]^2$. The vertical lines show the mean (solid green) and the associated 1σ (dashed blue) and 2σ (dashed–dot orange) errors estimated from the MCMC samples, along with the best-fitting value obtained from the MLE (red dashed) in Section 6.

APPENDIX C: A COMPARISON

In this appendix, we present an analysis of the same 8-MHz-bandwidth data as analysed by Ch21 for a comparison between the two different estimators used in these two works. The data are drawn from the central frequency $\nu_{\rm c} = 445$ MHz that corresponds to the redshifted 21-cm signal from $z = 2.19$. In short, we have used the Cross TGE (equation 6) to estimate $C_\ell(\Delta\nu)$, and use the MLE (Section 5) to estimate the spherical PS $P(k)$ at several k bins. The blue solid line in Fig. C1 shows the $\Delta^2(k)$ values obtained from the present analysis, whereas the black dotted line shows the same as shown in the middle panel of fig. A1 of Ch21. We find that the $\Delta^2(k)$ values in the two analyses are comparable and consistent within the 2σ error bars. The error bars in the present analysis are found to be larger ($\sim\sqrt{2}$ times) than Ch21. These larger error bars in the present analysis can be attributed to the fact that we have used only the correlation between the cross-polarizations (RR × LL) and discarded the correlation of the self-polarizations (RR × RR and LL × LL). For comparison, Ch21 reported 2σ upper limits $\Delta^2_{\rm UL}(k) \leq (61.49)^2$ mK^2 at $k = 0.97$ Mpc^{-1}, whereas we find $\Delta^2_{\rm UL}(k) \leq (68.13)^2$ mK^2 at $k = 0.90$ Mpc^{-1}.

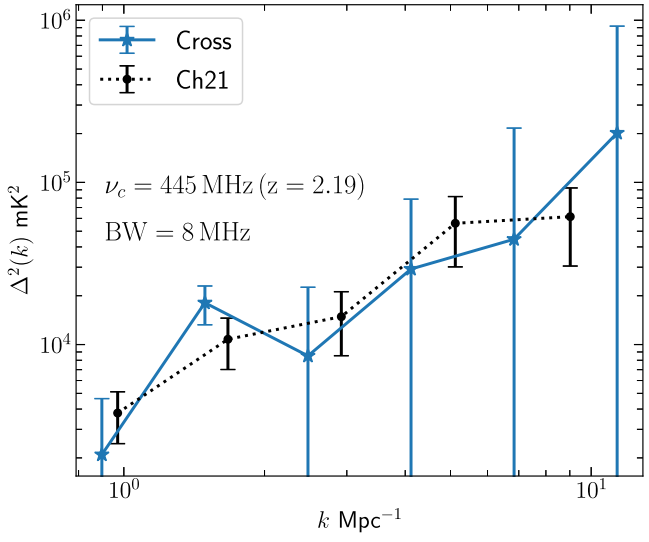

Figure C1. The mean squared brightness temperature fluctuations $\Delta^2(k)$ along with 2σ error bars from the analysis of an 8-MHz-bandwidth data at $\nu_{\rm c} = 445$ MHz ($z = 2.19$). The blue solid and black dotted lines show the $\Delta^2(k)$ values obtained from this work and Ch21, respectively.

A notable feature in TGE is that by tapering, it restricts the wide-angle point source contributions to comparably small $k_\parallel$ and thereby broadens the accessible TW region, which enables us to probe larger scales. We also found that (Fig. 3) the Cross TGE significantly reduces various polarization-dependent systematics in $C_\ell(\Delta\nu)$, allowing access to the smaller $k_\parallel$ modes. As the two estimators yield different (relatively) FG-free TW regions, we have not attempted to compare them on a one-to-one basis (i.e. same k value).

In addition to the cross-polarization PS described here, another significant distinction between the TGE and the techniques employed Ch21 is the treatment of the missing frequency channels that are flagged due to RFI. This issue is highlighted in Paper I (Section 5), which we briefly reiterate here. For each baseline, Ch21 have computed a Fourier transform of the measured visibilities along frequency to estimate the delay space visibilities (Morales & Hewitt 2004), which are then used (Parsons et al. 2012) to estimate the PS. The missing frequency channels introduce ringing artefacts in the Fourier transform and corrupt the estimated PS. Ch21 have used the 1D CLEAN introduced by Parsons & Backer (2009) to get uncorrupted delay space visibilities from RFI-contaminated data. This 1D CLEAN, which is adapted (Roberts, Lehar & Dreher 1987) from the 2D CLEAN deconvolution algorithm (Högbom 1974) used in aperture synthesis, performs a non-linear deconvolution in the delay space, equivalent to a least-squares interpolation in the frequency domain. In contrast, the TGE first correlates the visibility data across frequency channels to estimate $C_\ell(\Delta\nu)$. Despite having a substantial number of missing frequency channels in the visibility data (55 per cent here), there are no missing frequency separations $\Delta\nu$ in the estimated $C_\ell(\Delta\nu)$. The MLE (Section 5) is then used to estimate the spherical PS $P(k)$ from the $C_\ell(\Delta\nu)$. It is not essential to make up for any missing frequency channels because the entire procedure uses only the available frequency channels to estimate the PS. Bharadwaj et al. (2018) have used simulations to demonstrate that TGE can successfully recover the PS even when the data in 80 per cent randomly chosen frequency channels are flagged. For the present analysis, we have validated the estimator (Appendix A) using simulations where the flagging of the simulated data exactly matches that of the actual data. In addition to this direct validation of the estimator using simulated data, this comparison with Ch21, and the broadly consistent match of $\Delta^2(k)$ values from the two very distinct methods, makes our results assuredly more reliable.

This paper has been typeset from a TEX/LATEX file prepared by the author.