Visibility-based angular power spectrum estimation in low-frequency radio interferometric observations

Choudhuri, Samir; Bharadwaj, Somnath; Ghosh, Abhik; Ali, Sk. Saiyad

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stu2027

Publication date:
2014

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **445,** 4351–4365 (2014) doi:10.1093/mnras/stu2027

Visibility-based angular power spectrum estimation in low-frequency radio interferometric observations

Samir Choudhuri,[1] Somnath Bharadwaj,[1] Abhik Ghosh[2] and Sk. Saiyad Ali[3]★

[1]*Department of Physics, and Centre for Theoretical Studies, IIT Kharagpur, Kharagpur 721302, India*
[2]*Kapteyn Astronomical Institute, PO Box 800, NL-9700 AV Groningen, the Netherlands*
[3]*Department of Physics, Jadavpur University, Kolkata 700032, India*

Accepted 2014 September 26. Received 2014 September 23; in original form 2014 July 1

ABSTRACT

We present two estimators to quantify the angular power spectrum of the sky signal directly from the visibilities measured in radio interferometric observations. This is relevant for both the foregrounds and the cosmological 21-cm signal buried therein. The discussion here is restricted to the Galactic synchrotron radiation, the most dominant foreground component after point source removal. Our theoretical analysis is validated using simulations at 150 MHz, mainly for GMRT and also briefly for Low-Frequency Array. The Bare Estimator uses pair-wise correlations of the measured visibilities, while the Tapered Gridded Estimator uses the visibilities after gridding in the *uv* plane. The former is very precise, but computationally expensive for large data. The latter has a lower precision, but takes less computation time which is proportional to the data volume. The latter also allows tapering of the sky response leading to sidelobe suppression, an useful ingredient for foreground removal. Both estimators avoid the positive bias that arises due to the system noise. We consider amplitude and phase errors of the gain, and the w-term as possible sources of errors. We find that the estimated angular power spectrum is exponentially sensitive to the variance of the phase errors but insensitive to amplitude errors. The statistical uncertainties of the estimators are affected by both amplitude and phase errors. The w-term does not have a significant effect at the angular scales of our interest. We propose the Tapered Gridded Estimator as an effective tool to observationally quantify both foregrounds and the cosmological 21-cm signal.

Key words: methods: data analysis – methods: statistical – techniques: interferometric – diffuse radiation.

1 INTRODUCTION

Observations of the redshifted 21-cm radiation from the large-scale distribution of neutral hydrogen (H I) is one of the most promising probes to study the high-redshift Universe (recent reviews: Morales & Wyithe 2010; Mellema et al. 2013). This radiation appears as a very faint, diffuse background radiation in all low-frequency radio observations below 1420 MHz. At these frequencies, the sky signal is largely dominated by different foregrounds which are four to five orders of magnitude stronger than the redshifted 21-cm signal (Ali, Bharadwaj & Chengalur 2008; Bernardi et al. 2009; Ghosh et al. 2012; Pober et al. 2013). Foreground removal is possibly the most serious challenge for detecting the cosmological 21-cm signal. Various methodologies have been explored for foreground subtraction and for detecting the underlying 21-cm signal (Bowman, Morales & Hewitt 2009; Jelić et al. 2010; Ghosh et al. 2011b; Paciga et al. 2011, 2013; Petrovic & Oh 2011; Chapman et al. 2012, 2013; Cho, Lazarian & Timbie 2012; Liu & Tegmark 2012; Mao 2012; Trott, Wayth & Tingay 2012; Vedantham, Udaya Shankar & Subrahmanyan 2012; Jacobs, Bowman & Aguirre 2013; Thyagarajan et al. 2013; Dillon et al. 2014; Liu, Parsons & Trott 2014a,b; Parsons et al. 2014; Shaw et al. 2014a).

The Galactic synchrotron emission is expected to be the most dominant foreground at angular scale >10 arcmin after point source subtraction at 10–20 mJy level (Bernardi et al. 2009; Ghosh et al. 2012). A precise characterization and a detailed understanding of the Galactic synchrotron emission is needed to reliably remove foregrounds in 21-cm experiments. The study of the Galactic synchrotron emission is interesting in its own right. This will shed light on the cosmic ray electron distribution, the strength and structure of the Galactic magnetic field, and the magnetic turbulence (Waelkens, Schekochihin & Enßlin 2009; Lazarian & Pogosyan 2012; Iacobelli et al. 2013).

Bernardi et al. (2009) and Ghosh et al. (2012) have, respectively, analysed 150 MHz WSRT and GMRT observations where they find

★ E-mail: saiyad@phys.jdvu.ac.in

that the measured angular power spectrum can be well fitted with a power law ($C_\ell \propto \ell^{-\beta}$, $\beta = 2.2 \pm 0.3$ for WSRT and $\beta = 2.34 \pm 0.28$ for GMRT) up to $\ell \leq 900$. At relatively higher frequencies, Giardino et al. (2001) and Giardino et al. (2002) have analysed the fluctuations in the Galactic synchrotron radiation using the 2.3 GHz Rhodes Survey and the 2.4 GHz Parkes radio continuum and polarization survey, where they find a slope $\beta = 2.43 \pm 0.01$ ($2 \leq \ell \leq 100$) and $\beta = 2.37 \pm 0.21$ ($40 \leq \ell \leq 250$), respectively. At several tens of GHz, Bennett et al. (2003) have determined the angular power spectrum of the Galactic synchrotron radiation using the *Wilkinson Microwave Anisotropy Probe* data where they find a scaling $C_\ell \sim \ell^{-2}$ within $\ell \leq 200$. The structure of the Galactic synchrotron emission is not well quantified at the frequencies and angular scales relevant for detecting the cosmological 21-cm signal, and there is considerable scope for further work in this direction.

Radio interferometric observations measure the complex visibility. The measurement is done directly in Fourier space which makes interferometers ideal instruments for measuring the angular power spectrum of the sky signal. The visibility-based power spectrum estimator formalism has been extensively used for analysing CMB data from interferometers (Hobson, Lasenby & Jones 1995; White et al. 1999; Hobson & Maisinger 2002; Myers et al. 2003). A visibility-based estimator has also been successfully employed to study the power spectrum of the H I in the interstellar medium of several nearby galaxies (e.g. Begum, Chengalur & Bhardwaj 2006; Dutta et al. 2009). A direct visibility-based approach has been proposed for quantifying the power spectrum of the cosmological 21-cm signal expected at the GMRT (Bharadwaj & Sethi 2001; Bharadwaj & Pandey 2003; Bharadwaj & Ali 2005) and recently for the ORT (Ali & Bharadwaj 2013). Visibility-based power spectrum estimators have been used to analyse GMRT data in the context of H I observations (Ali et al. 2008; Ghosh et al. 2011a,b, 2012; Paciga et al. 2011). A recent paper (Paul et al. 2014) has proposed visibility correlations to detect the EoR signal using drift scan observations with the MWA.

It is possible to estimate the angular power spectrum of the sky signal from the synthesized radio image (e.g. Bernardi et al. 2009, 2010; Iacobelli et al. 2013). The noise properties of the visibilities are better understood than those of the image pixels. The noise in the different visibilities is uncorrelated, whereas the noise in the image pixels may be correlated depending on the baseline *uv* coverage. The visibility-based power spectrum estimators have the added advantage that they avoid possible imaging artefacts due to the dirty beam, etc. (Trott et al. 2011).

In this paper, we consider two estimators which use the measured visibilities to quantify the angular power spectrum of the sky signal. The Bare Estimator, which has been utilized in Ali et al. (2008) and Ghosh et al. (2011a), directly uses pairwise correlations of the measured visibilities. The Tapered Gridded Estimator, which has been utilized in Ghosh et al. (2011b, 2012), uses the visibilities after gridding on a rectangular grid in the *uv* plane. The latter incorporates the feature that it allows a tapering of the sky response and thereby suppresses the sidelobes of the telescope's primary beam. Earlier work (Ghosh et al. 2011b) has shown this to be a useful ingredient in foreground removal for detecting the cosmological 21-cm signal. In this paper, we have carried out a somewhat detailed investigation in order to place these two estimators on sound theoretical footing. The theoretical predictions are substantiated using simulations. As a test bed for the estimators, we consider a situation where the point sources have been identified and subtracted out so that the residual visibilities are dominated by the Galactic synchrotron radiation. We investigate how well the estimators are able to recover the angular power spectrum of the input model used to simulate the Galactic synchrotron emission at 150 MHz. We have also analysed the effects of gain errors and the w-term. Most of our simulations are for the GMRT, but we also briefly consider simulations for Low-Frequency Array (LOFAR). The estimators considered here can be generalized to the multifrequency angular power spectrum (MAPS; Datta, Roy Choudhury & Bharadwaj 2007), which can be used to quantify the cosmological 21-cm signal. We plan to investigate this in a future study.

A brief outline of the paper follows. In Section 2, we establish the relation between the visibility correlation and the angular power spectrum. In Section 3, we describe the simulations which we have used to validate the theoretical results of this paper. In Sections 4 and 5, we consider the Bare and the Tapered Gridded Estimators, respectively. The theoretical analysis and the results from the simulations are all presented in these two sections. Section 6 presents a brief comparison between the two estimators, and in Sections 7 and 8 we consider the effect of gain errors and the w-term, respectively. Much of the analysis of the previous sections is in the context of the GMRT. In Section 9, we apply the estimators to simulated LOFAR data and present the results. We present discussion and conclusions in Section 10.

2 VISIBILITY CORRELATIONS AND THE ANGULAR POWER SPECTRUM

In this section, we discuss the relation between the two-visibility correlation and the angular power spectrum of the specific intensity $I(\boldsymbol{\theta}, \nu)$ or equivalently the brightness temperature $T(\boldsymbol{\theta}, \nu)$ distribution on the sky under the flat-sky approximation. Here, $\boldsymbol{\theta}$ is a two-dimensional (2D) vector on the plane of the sky with origin at the centre of the field of view (FoV). It is useful to decompose the specific intensity as $I(\boldsymbol{\theta}, \nu) = \bar{I}(\nu) + \delta I(\boldsymbol{\theta}, \nu)$, where the first term $\bar{I}(\nu)$ is a uniform background brightness and the second term $\delta I(\boldsymbol{\theta}, \nu)$ is the angular fluctuation in the specific intensity. We assume that $\delta I(\boldsymbol{\theta}, \nu)$ is a particular realization of a statistically homogeneous and isotropic Gaussian random process on the sky. In radio interferometric observations, the fundamental observable quantity is a set of complex visibilities $\mathcal{V}(\boldsymbol{U}, \nu)$ which are sensitive to only the angular fluctuations in the sky signal. The baseline $\boldsymbol{U}$ quantifies the antenna pair separation $\boldsymbol{d}$ projected on the plane perpendicular to the line of sight in units of the observing wavelength λ. The measured visibilities are a sum of two contributions $\mathcal{V}(\boldsymbol{U}, \nu) = \mathcal{S}(\boldsymbol{U}, \nu) + \mathcal{N}(\boldsymbol{U}, \nu)$, the sky signal and system noise, respectively. We assume that the signal and the noise are both uncorrelated Gaussian random variables with zero mean. The visibility contribution $\mathcal{S}(\boldsymbol{U}, \nu)$ from the sky signal records the Fourier transform of the product of the primary beam pattern $\mathcal{A}(\boldsymbol{\theta}, \nu)$ and $\delta I(\boldsymbol{\theta}, \nu)$. The primary beam pattern $\mathcal{A}(\boldsymbol{\theta}, \nu)$ quantifies how the individual antenna responds to signals from different directions in the sky. Using the convolution theorem, we then have

$$\mathcal{S}(\boldsymbol{U}, \nu) = \int \mathrm{d}^2 U' \, \tilde{a}\left(\boldsymbol{U} - \boldsymbol{U}', \nu\right) \, \Delta\tilde{I}(\boldsymbol{U}', \nu), \tag{1}$$

where $\Delta\tilde{I}(\boldsymbol{U}, \nu)$ and $\tilde{a}(\boldsymbol{U}, \nu)$ are the Fourier transforms of $\delta I(\boldsymbol{\theta}, \nu)$ and $\mathcal{A}(\boldsymbol{\theta}, \nu)$, respectively. Typically, the term arising from the uniform specific intensity distribution $\bar{I}(\nu)\tilde{a}(\boldsymbol{U}, \nu)$ makes no contribution to the measured visibilities, and we have dropped this. We refer to $\tilde{a}(\boldsymbol{U}, \nu)$ as the aperture power pattern. The individual antenna response $\mathcal{A}(\boldsymbol{\theta}, \nu)$ for any telescope is usually quite complicated depending on the telescope aperture, the reflector, and the feed (Christiansen & Hogbom 1969). It is beyond the scope of this paper

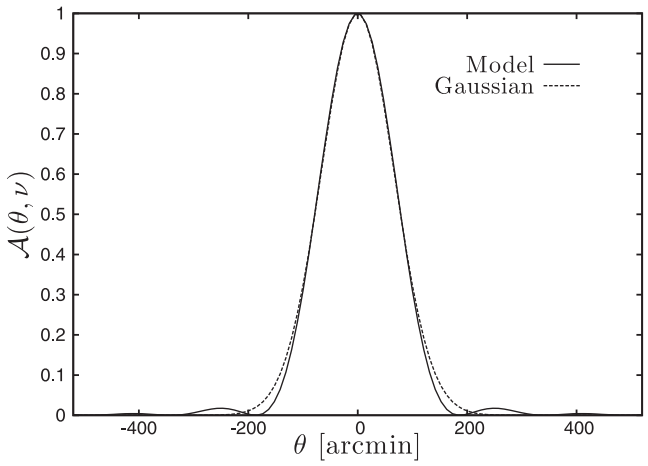

Figure 1. The solid curve shows the 150 MHz GMRT primary beam pattern $\mathcal{A}(\boldsymbol{\theta},\ \nu)$ predicted by equation (2), and the dashed curve shows Gaussian approximation (equation 4) with the same $\theta_{\rm FWHM}$.

to consider the actual single antenna response of any particular telescope. We make the simplifying assumption that the telescope has a uniformly illuminated circular aperture of diameter D whereby we have the primary beam pattern (Fig. 1)

$$\mathcal{A}(\boldsymbol{\theta},\ \nu) = \left[\left(\frac{2\lambda}{\pi\theta D}\right) J_1\left(\frac{\pi\theta D}{\lambda}\right)\right]^2, \tag{2}$$

where J_1 is the Bessel function of the first kind of order 1, the primary beam pattern is normalized to unity at the pointing centre [$\mathcal{A}(0) = 1$], and the aperture power pattern is

$$\tilde{a}(\boldsymbol{U},\nu) = \frac{8\lambda^4}{\pi^2 D^4}\left[\left(\frac{D}{\lambda}\right)^2 \cos^{-1}\left(\frac{\lambda U}{D}\right) - U\sqrt{\left(\frac{D}{\lambda}\right)^2 - U^2}\right]. \tag{3}$$

We note that $\tilde{a}(\boldsymbol{U},\nu)$ in equation (3) peaks at $U = 0$, declines monotonically with increasing U, and is zero for $U \geq D/\lambda$. The primary beam pattern (Fig. 1) is well approximated by a circular Gaussian function

$$\mathcal{A}_{\rm G}(\boldsymbol{\theta},\nu) = \exp\left[-\theta^2/\theta_0^2\right] \tag{4}$$

of the same full width at half-maxima (FWHM) as equation (2). The parameter θ_0 here is related to the FWHM $\theta_{\rm FWHM}$ of the primary beam pattern $\mathcal{A}(\boldsymbol{\theta},\nu)$ (equation 2) as $\theta_0 = 0.6\theta_{\rm FWHM}$, and

$$\tilde{a}_{\rm G}(\boldsymbol{U},\nu) = \frac{1}{\pi U_0^2}\,{\rm e}^{-U^2/U_0^2}, \tag{5}$$

where $U_0 = (\pi\theta_0)^{-1} = 0.53/\theta_{\rm FWHM}$. While the Gaussian $\tilde{a}_{\rm G}(\boldsymbol{U},\nu)$ (equation 5) provides a good approximation to $\tilde{a}(\boldsymbol{U},\nu)$ (equation 3) particularly in the vicinity of $U = 0$, there is however a significant difference in that $\tilde{a}(\boldsymbol{U},\nu)$ has a compact support and is exactly zero for all $U \geq D/\lambda$, whereas $\tilde{a}_{\rm G}(\boldsymbol{U},\nu)$, though it has an extremely small value for $U \geq D/\lambda$, does not become zero anywhere. In practice, it is extremely difficult to experimentally determine the full primary beam pattern $\mathcal{A}(\boldsymbol{\theta},\ \nu)$ for a telescope. However, the value of $\theta_{\rm FWHM}$ is typically well determined. This has motivated the Gaussian approximation to be used extensively for both theoretical predictions (Bharadwaj & Sethi 2001; Bharadwaj & Ali 2005) and analysing observational data (Ali et al. 2008; Ghosh et al. 2012). The close match between $\mathcal{A}(\boldsymbol{\theta},\ \nu)$ (equation 2) and $\mathcal{A}_{\rm G}(\boldsymbol{\theta},\ \nu)$ (equation 4) indicates that we may also expect the Gaussian approximation to provide a good fit to the telescope's actual primary beam pattern, particularly within the main lobe. This, to some extent, justifies the use of the Gaussian approximation in the earlier works. The Gaussian approximation simplifies the calculations rendering them amenable to analytic treatment, and we use it on several occasions as indicated later in this paper. For much of the investigations presented in this paper we have considered $D = 45$ m and $\lambda = 2$ m which corresponds to GMRT 150 MHz observations. We have also considered $D = 30.75$ m and $\lambda = 2$ m which corresponds to LOFAR 150 MHz observations. For both these telescopes, Table 1 summarizes the values of some of the relevant parameters. Note that these values correspond to the idealized telescope model discussed above, and they are somewhat different from the values actually measured for the respective telescopes. For example, the GMRT primary beam pattern has $\theta_{\rm FWHM} = 186$ arcmin, whereas we have used $\theta_{\rm FWHM} = 157$ arcmin based on our idealized model. We discuss the observational consequence of this ∼16 per cent difference later in Section 10 of this paper. For the rest of this paper, we focus on the GMRT, except in Section 9 where we shift our attention to LOFAR. Our entire analysis is based on the idealized telescope model described above and the relevant parameters are listed in Table 1 for both these telescopes.

In the flat-sky approximation, the statistical properties of the background intensity fluctuations $\delta I(\boldsymbol{\theta},\ \nu)$ can be quantified through the 2D power spectrum $P(U,\ \nu)$ defined as

$$\langle \Delta\tilde{I}(\boldsymbol{U},\nu)\Delta\tilde{I}^*(\boldsymbol{U}',\nu)\rangle = \delta_{\rm D}^2(\boldsymbol{U}-\boldsymbol{U}')\,P(U,\nu), \tag{6}$$

where $\delta_{\rm D}^2(\boldsymbol{U}-\boldsymbol{U}')$ is a 2D Dirac delta function. The angular brackets $\langle\cdots\rangle$ here denote an ensemble average over different realizations of the stochastic intensity fluctuations on the sky. We also assume that the $P(U,\ \nu)$ depends only on the magnitude $U = |\boldsymbol{U}|$, i.e. the fluctuations are statistically isotropic. We note that $P(U,\ \nu)$ is related to $C_\ell(\nu)$ the angular power spectrum of the brightness temperature fluctuations through (Ali et al. 2008)

$$C_\ell(\nu) = \left(\frac{\partial B}{\partial T}\right)^{-2} P(\ell/2\pi,\ \nu)\,, \tag{7}$$

where the angular multipole ℓ corresponds to $U = \ell/2\pi$, B is the Planck function and $(\partial B/\partial T) = 2k_{\rm B}/\lambda^2$ in the Raleigh–Jeans limit which is valid at the frequencies of our interest. We will drop the ν dependence henceforth as the rest of the calculations are done at a fixed frequency $\nu = 150$ MHz.

Table 1. This shows some relevant parameters for the primary beam pattern calculated using the idealized telescope model (equations 2 and 3), and the Gaussian approximation (equations 4 and 5). The parameter σ_0 is defined in equation (10).

150 MHz	D	$\theta_{\rm FWHM}$ $1.03\lambda/D$	θ_0 $0.6\theta_{\rm FWHM}$	U_0 $0.53/\theta_{\rm FWHM}$	σ_0 $0.76/\theta_{\rm FWHM}$
GMRT	45 m	157 arcmin	95 arcmin	11.54	16.6
LOFAR	30.75 m	230 arcmin	139 arcmin	7.88	11.33

We now consider the two-visibility correlation which is defined as

$$V_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U}) = \langle \mathcal{V}(\boldsymbol{U})\mathcal{V}^*(\boldsymbol{U}+\Delta\boldsymbol{U})\rangle\,, \tag{8}$$

and which has the contribution

$$S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U}) = \int \mathrm{d}^2U'\, \tilde{a}(\boldsymbol{U}-\boldsymbol{U}')\, \tilde{a}^*(\boldsymbol{U}+\Delta\boldsymbol{U}-\boldsymbol{U}')\, P(U') \tag{9}$$

from the sky signal.

The visibilities at the baselines $\boldsymbol{U}$ and $\boldsymbol{U}+\Delta\boldsymbol{U}$ are correlated only if there is a significant overlap between $\tilde{a}(\boldsymbol{U}-\boldsymbol{U}')$ and $\tilde{a}^*(\boldsymbol{U}+\Delta\boldsymbol{U}-\boldsymbol{U}')$. The correlation $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U})$ is strongest when $|\Delta\boldsymbol{U}| = 0$, declines rapidly with increasing $|\Delta\boldsymbol{U}|$, and is zero for $|\Delta\boldsymbol{U}| \geq 2D/\lambda$. The correlation $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U})$ depends on both, the magnitude of $\Delta\boldsymbol{U}$ as well as the angle between $\Delta\boldsymbol{U}$ and $\boldsymbol{U}$, and an earlier work (Bharadwaj & Pandey 2003) has studied this in detail for the predicted post-reionization cosmological 21-cm signal. In this work, we have considered a power-law power spectrum $P(U) = AU^{-\beta}$ for different values of β in the range 1.5–3.5, and we have used equation (9) to study the $\Delta\boldsymbol{U}$ dependence of $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U})$. We find that the $\Delta\boldsymbol{U}$ dependence is isotropic to a great extent, and it can be well modelled using a Gaussian (Fig. 2) as

$$S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U}) = \exp\left[-\left(\frac{|\,\Delta\boldsymbol{U}\,|}{\sigma_0}\right)^2\right] S_2(U), \tag{10}$$

where $\sigma_0 = 0.76/\theta_{\mathrm{FWHM}}$ (Table 1) and $S_2(U) \equiv S_2(\boldsymbol{U}, \boldsymbol{U})$. While the approximation in equation (10) matches the result of equation (9) quite well for small $\Delta\boldsymbol{U}$, the approximation breaks down when $|\,\Delta\boldsymbol{U}\,| > 2D/\lambda$, where $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U}) = 0$ contrary to the prediction of equation (10). This discrepancy, however, does not significantly affect the estimators (defined later) because the value of $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U})$ predicted by equation (10) is extremely small for $|\,\Delta\boldsymbol{U}\,| > 2D/\lambda$.

A further simplification is possible for $U \gg U_0$ where it is possible to approximate $S_2(U)$ which is calculated using equation (9) by

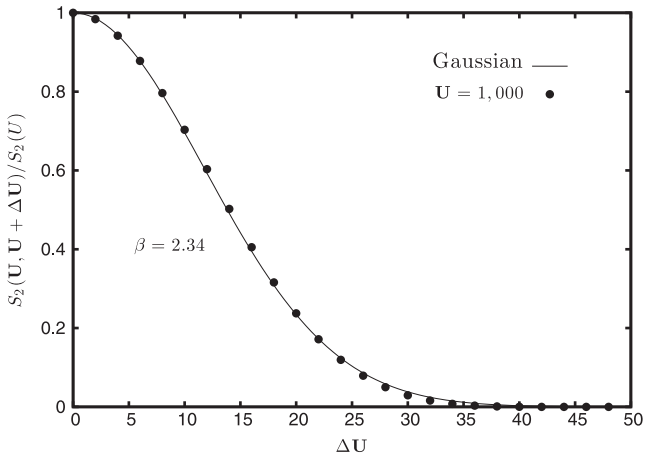

Figure 2. This figure shows how the sky signal contribution to the two-visibility correlation varies with $\Delta\boldsymbol{U}$ for a fixed value $U = 1000$. The points show the results from equation (9) for $P(U) = AU^{-2.34}$, and the solid line shows the Gaussian fit given in equation (10).

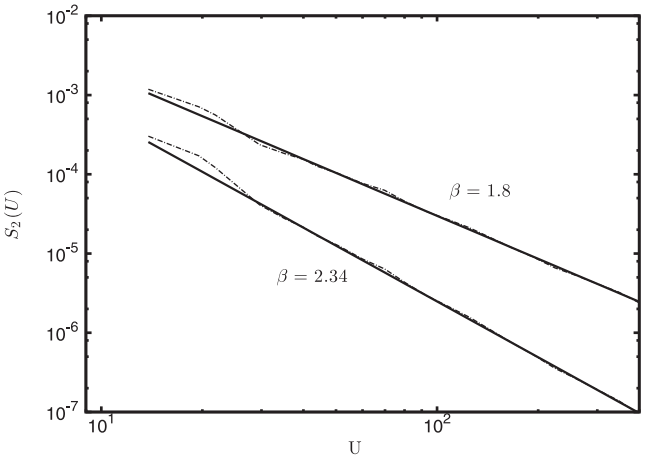

Figure 3. This shows the sky signal contribution to the visibility correlation ($S_2(U)$) for two different power spectra with slopes $\beta = 1.8$ and 2.34, respectively. The dash–dotted curves show the result of the convolution in equation (9) with $\Delta\boldsymbol{U} = 0$, whereas the solid curves show the result of approximating this with equation (11). We see that the approximation of equation (11) matches the convolution reasonably well at large baselines $U \geq 4U_0 \sim 45$.

assuming that the value of $P(U')$ does not change much within the width of the function $|\tilde{a}\left(\boldsymbol{U}-\boldsymbol{U}'\right)|^2$. We then obtain

$$S_2(U) = \left[\int \mathrm{d}^2U' \;|\,\tilde{a}(\boldsymbol{U}-\boldsymbol{U}')\,|^2\right] P(U)\,. \tag{11}$$

The integral in the square brackets has a constant value $\frac{\pi\theta_0^2}{2}$ in the Gaussian approximation which yields the value 1.19×10^{-3}, whereas we have 1.15×10^{-3} if we use equation (3) and numerically evaluate the integral in the square brackets. We see that the Gaussian approximation is adequate for the integral in equation (11), and we adopt the value $\pi\theta_0^2/2$ for the entire subsequent analysis. We have calculated $S_2(U)$ (Fig. 3) using the convolution in equation (9), and compared this with the approximation in equation (11). We find that the approximation in equation (11) matches quite well with the convolution (equation 9) for baselines $U \geq 4U_0 \sim 45$. Throughout the subsequent analysis, we have restricted the baselines to this range, and we have used equation (11) to evaluate $S_2(U)$, the sky signal contribution to the visibility correlation.

We finally have the approximate relation between the sky signal contribution to the two-visibility correlation and the angular power spectrum

$$S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U}) = \frac{\pi\theta_0^2}{2}\left(\frac{\partial B}{\partial T}\right)^2 \exp\left[-\left(\frac{\Delta U}{\sigma_0}\right)^2\right] C_\ell, \tag{12}$$

where $\ell = 2\pi U$. We thus see that the visibilities at two different baselines $\boldsymbol{U}$ and $\boldsymbol{U}+\Delta\boldsymbol{U}$ are correlated only if the separation is small ($|\Delta U| \leq \sigma_0$), and there is negligible correlation if the separation is beyond a disc of radius σ_0. Further, the visibility correlation $S_2(\boldsymbol{U}, \boldsymbol{U}+\Delta\boldsymbol{U})$ gives a direct estimate of the angular power spectrum C_ℓ at the angular multipole $\ell = 2\pi U$. In addition to the sky signal $\mathcal{S}(\boldsymbol{U})$, each visibility also contains a system noise contribution $\mathcal{N}(\boldsymbol{U})$. For each visibility measurement, the real and imaginary parts of $\mathcal{N}(\boldsymbol{U})$ are both random variables of zero mean and rms σ_n. Further, the noise in any two different visibilities is uncorrelated. We can then write the total visibility correlation as

$$V_{2ij} \equiv \langle \mathcal{V}_i \mathcal{V}_j^* \rangle = V_0\, \mathrm{e}^{-|\Delta\boldsymbol{U}_{ij}|^2/\sigma_0^2}\, C_{\ell_i} + \delta_{ij} 2\sigma_n^2, \tag{13}$$

where $[\mathcal{V}_i, \mathcal{V}_j] \equiv [\mathcal{V}(\boldsymbol{U}_i), \mathcal{V}(\boldsymbol{U}_j)]$, $V_0 = \frac{\pi\theta_0^2}{2}\left(\frac{\partial B}{\partial T}\right)^2$, $\Delta\boldsymbol{U}_{ij} = \boldsymbol{U}_i - \boldsymbol{U}_j$, and the Kronecker delta δ_{ij} is non-zero only if we correlate a visibility with itself. Equation (13) relates the two-visibility correlation V_{2ij} to C_{ℓ_i} the angular power spectrum of the sky signal at the angular multipole $\ell_i = 2\pi U_i$ and σ_n^2 the mean square system noise, and we use this extensively in connection with the estimators that we consider in the subsequent sections.

3 SIMULATING THE SKY SIGNAL

We have used simulations of radio interferometric observations to validate the angular power spectrum estimators that we introduce in subsequent sections of this paper. In this section, we first describe the simulations of the sky signal, and then describe how these were used to simulate the expected visibilities. For the sky model, we assume that all point sources with flux above a sufficiently low threshold have been identified and removed from the data so that the 150 MHz radio sky is dominated by the diffuse Galactic synchrotron radiation.

The slope β of the angular power spectrum of diffuse Galactic synchrotron emission is within the range 1.5–3 as found by all the previous measurements at frequencies 0.15–94 GHz (e.g. La Porta et al. 2008; Bernardi et al. 2009). For the purpose of this paper, we assume that the fluctuations in the diffuse Galactic synchrotron radiation are a statistically homogeneous and isotropic Gaussian random field whose statistical properties are completely specified by the angular power spectrum. Further, we assume that the angular power spectrum of brightness temperature fluctuations is well described by a single power law over the entire range of angular scales of our interest. In this work, we have adapted the angular power spectrum

$$C_\ell^M = A_{150} \times \left(\frac{1000}{\ell}\right)^\beta, \tag{14}$$

where $A_{150} = 513\,\mathrm{mK}^2$ and $\beta = 2.34$ from Ghosh et al. (2012). This is the input model for all our simulations.

We have considered a $5.^\circ8 \times 5.^\circ8$ FoV for the GMRT simulations. This has been represented using a 2048×2048 grid with an angular resolution of $\sim$10.2 arcsec. We have first generated the Fourier components of the brightness temperature fluctuations on the grid using

$$\Delta\tilde{T}(\boldsymbol{U}) = \sqrt{\frac{\Omega\, C_\ell^M}{2}}[x(\boldsymbol{U}) + iy(\boldsymbol{U})], \tag{15}$$

where Ω is the total solid angle of the simulation, and $x(\boldsymbol{U})$ and $y(\boldsymbol{U})$ are independent Gaussian random variables with zero mean and unit variance. We then use a Fourier transform to generate the brightness temperature fluctuations $\delta T(\boldsymbol{\theta})$ or equivalently the specific intensity fluctuations $\delta I(\boldsymbol{\theta})$ on the grid. Fig. 4 shows one realization of the brightness temperature fluctuations generated using the procedure outlined above. We have generated 20 different independent realizations of the sky by considering different sets of random numbers in equation (15).

To simulate GMRT observations, we consider 8 h observations targeted on a field located at $+60^\circ$ Dec. for which the *uv* tracks for baselines within $|u|, |v| \leq 1000$ are shown in Fig. 5. We assume 16 s integration time for each sampled visibility data which gives us 217 457 visibility points. To calculate the visibilities, we have multiplied the simulated $\delta I(\boldsymbol{\theta})$ with the primary beam pattern $\mathcal{A}(\boldsymbol{\theta})$ (equation 2) and evaluated the Fourier transform of the product for each sampled baseline $\boldsymbol{U}$ on the *uv* track. In addition to the sky signal, each measured visibility will also have a system noise contribution. We have included this by adding independent Gaussian random noise contributions to both the real and imaginary parts of each visibility. This noise is predicted to have an rms of $\sigma_n = 1.03$ Jy for a single polarization at the GMRT.[1]

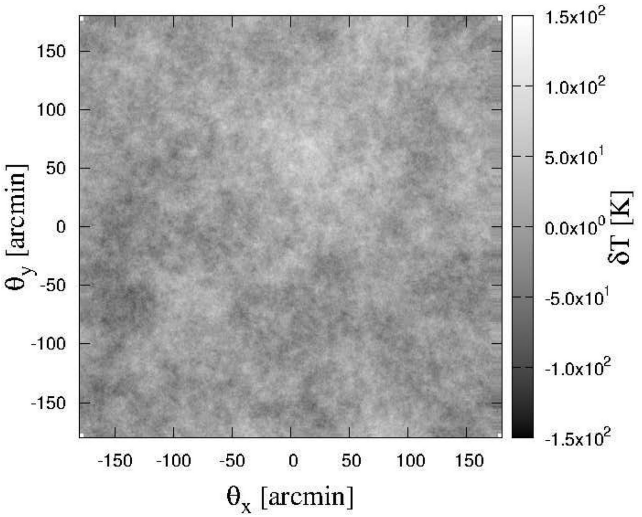

Figure 4. This shows a single realization of the simulated 150 MHz radio sky under the assumption that the bright point sources have been removed so that it is dominated by the diffuse Galactic synchrotron radiation. We have simulated a $5.^\circ8 \times 5.^\circ8$ FoV with $\sim$10.2 arcsec resolution.

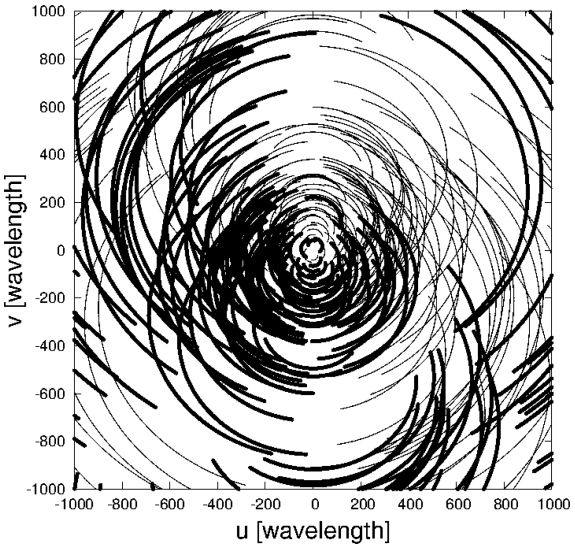

Figure 5. This shows the *uv* coverage for 8 h GMRT 150 MHz observations centred on a field at a declination of $\delta = +60^\circ$. Only baselines with $|u|, |v| \leq 1000$ have been shown. Note that u and v are antenna separations measured in units of the observing wavelength, and hence they are dimensionless.

It is clearly visible in Fig. 5 that the GMRT has a rather sparse *uv* coverage. The fact that we have data for only a limited number of the Fourier modes is expected to play an important role. This is particularly important for the cosmic variance which crucially depends on the number of independent Fourier modes. In order to assess the impact of the sparse *uv* coverage, we have also considered a

[1] http://www.gmrt.ncra.tifr.res.in

situation where exactly the same number of visibility measurements (217 457) are randomly distributed within the region $|u|, |v| \leq 1000$ on the *uv* plane.

In the subsequent sections of this paper, we have analysed 20 independent realizations of the sky signal, with visibilities points that correspond to the *uv* tracks shown in Fig. 5. We refer to this ensemble of 20 simulated data sets as 'GMRT'. We have also considered a random baseline distribution and calculated the visibilities for the same 20 realizations of the sky signal, and we refer to this as 'Random'. Finally, we have also carried out simulations for LOFAR which has a more uniform *uv* coverage as compared to the GMRT. These simulations are separately discussed in Section 9.

Finally, we note that the simulated baselines lying in the lower half of the *uv* plane (e.g. Fig. 5.) are all folded to the upper half using the property $\mathcal{V}(\boldsymbol{U}) = \mathcal{V}^*(-\boldsymbol{U})$. The simulated baseline distribution that we finally use for analysis is entirely restricted to the upper half of the *uv* plane.

4 THE BARE ESTIMATOR

The Bare Estimator directly uses the individual visibilities to estimate the angular power spectrum. Each measured visibility corresponds to a Fourier mode of the sky signal, and the visibility squared $| \mathcal{V}\mathcal{V}^* |$ straight away gives the angular power spectrum. This simple estimator, however, has a severe drawback because the noise contribution $2\sigma_n^2$ is usually much larger than the sky signal $V_0\, \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}\, C_\ell$ in equation (13). Any estimator that includes the correlation of a visibility with itself suffers from a very large positive noise bias. It is, in principle, possible to model the constant noise bias and subtract it out. This however is extremely difficult in practice because small calibration errors (discussed later in Section 7) would introduce fluctuations in the noise bias resulting in residuals that could exceed the sky signal. It is therefore desirable to avoid the noise bias by considering estimators which do not include the contribution from the correlation of a visibility with itself.

The Bare Estimator $\hat{E}_{\rm B}(a)$ is defined as

$$\hat{E}_{\rm B}(a) = \frac{\sum_{i,j}\, w_{ij}\, \mathcal{V}_i\, \mathcal{V}_j^*}{\sum_{i,j}\, w_{ij}\, V_0 \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}}\,, \tag{16}$$

where we have assumed that the baselines have been divided into bins such that all the baselines U in the range $U_1 \leq U < U_2$ are in bin 1, those in the range $U_2 \leq U < U_3$ are in bin 2 etc., and $\hat{E}_{\rm B}(a)$ refers to a particular bin a. The sum i, j is over all pairs of visibilities $\mathcal{V}_i, \mathcal{V}_j$ with baselines $\boldsymbol{U}_i, \boldsymbol{U}_j$ in bin a. We have restricted the sum to pairs within $| \boldsymbol{U}_i - \boldsymbol{U}_j | \leq \sigma_0$ as the pairs with larger separations do not contribute much to the estimator. The weight $w_{ij} = (1 - \delta_{ij}) K_{ij}$ is chosen such that it is zero when we correlate a visibility with itself, thereby avoiding the positive noise bias.

We now show that $\hat{E}_{\rm B}(a)$ gives an unbiased estimate of the angular power spectrum C_ℓ for bin a. The expectation value of the estimator can be expressed using equation (13) as

$$\langle \hat{E}_{\rm B}(a) \rangle = \frac{\sum_{i,j}\, w_{ij}\, V_{2ij}}{\sum_{i,j}\, w_{ij} V_0 \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}} = \frac{\sum_{i,j}\, w_{ij}\, \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2} C_{\ell_i}}{\sum_{i,j}\, w_{ij} \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}} \tag{17}$$

which can be written as

$$\langle \hat{E}_{\rm B}(a) \rangle = \bar{C}_{\bar{\ell}_a}, \tag{18}$$

where $\bar{C}_{\bar{\ell}_a}$ is the average angular power spectrum at

$$\bar{\ell}_a = \frac{\sum_{i,j}\, w_{ij}\, \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2} \ell_i}{\sum_{i,j}\, w_{ij} \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}} \tag{19}$$

which is the effective angular multipole for bin a.

We note that it is possible to express equation (17) using matrix notation as

$$\langle \hat{E}_{\rm B}(a) \rangle = \frac{\mathrm{Tr}(\boldsymbol{w}\boldsymbol{V}_2)}{\mathrm{Tr}(\boldsymbol{w}\boldsymbol{I}_2)}, \tag{20}$$

where we have the matrices $\boldsymbol{w} \equiv w_{ij}$, $\boldsymbol{V}_2 \equiv V_{2ij}$, $\boldsymbol{I}_2 = V_0 \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}$ and $\mathrm{Tr}(\boldsymbol{A})$ denotes the trace of a matrix $\boldsymbol{A}$.

We next evaluate $\sigma^2_{\hat{E}_{\rm B}}(a)$ the variance of $\hat{E}_{\rm B}(a)$. This gives δC_{ℓ_a} which is an estimate of the error in the angular power spectrum measured from the data. We have

$$[\delta C_{\ell_a}]^2 \equiv \sigma^2_{\hat{E}_{\rm B}}(a) = \left\langle \hat{E}^2_{\rm B}(a) \right\rangle - \left\langle \hat{E}_{\rm B}(a) \right\rangle^2 \tag{21}$$

which can be simplified to

$$\sigma^2_{\hat{E}_{\rm B}}(a) = \frac{\sum_{i,j,k,l}\, w_{ij}\, w_{kl}\, V_{2il}\, V_{2kj}}{[\mathrm{Tr}(\boldsymbol{w}\boldsymbol{I}_2)]^2} = \frac{\mathrm{Tr}(\boldsymbol{w}\boldsymbol{V}_2\boldsymbol{w}\boldsymbol{V}_2)}{[\mathrm{Tr}(\boldsymbol{w}\boldsymbol{I}_2)]^2} \tag{22}$$

under the assumptions that $\boldsymbol{w}$ is symmetric and the measured visibilities are Gaussian random variables.

The system noise only appears in the diagonal elements of the visibility correlation matrix $\boldsymbol{V}_2$, whereas the sky signal contributes to both the diagonal and the off-diagonal elements. Further, the diagonal elements of the weight matrix $\boldsymbol{w}$ are all zero. Consequently, the trace $\mathrm{Tr}(\boldsymbol{w}\boldsymbol{V}_2)$ in equation (18) does not pick up any contribution from the diagonal elements of $\boldsymbol{V}_2$, and the expectation value of the estimator is not affected by the system noise. The variance $\sigma^2_{\hat{E}_{\rm B}}(a)$ however has contributions from both diagonal and off-diagonal elements of $\boldsymbol{V}_2$. The diagonal elements are dominated by the system noise, whereas the off-diagonal elements contribute to the cosmic variance.

The weights w_{ij} should, in principle, be chosen so as to maximize the signal-to-noise ratio $\mathrm{SNR} = \langle \hat{E}_{\rm B}(a) \rangle / \sigma_{E_{\rm B}}(a)$. The optimal weights depend on the baseline distribution and $V_0 C_\ell / \sigma_n^2$, the relative amplitude of the signal to the noise in the individual visibilities. Here, we have made the simplifying assumption that all the visibility pairs contribute equally to $\sigma^2_{\hat{E}_{\rm B}}(a)$. Each visibility pair is assigned the weight $w_{ij} = (1 - \delta_{ij}) \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}$ which is proportional to its contribution to $\langle \hat{E}_{\rm B}(a) \rangle$.

To test the Bare Estimator, we have used it to estimate C_ℓ from the simulated GMRT and Random data. For this analysis, the visibilities with baselines U in the range $40 \leq U \leq 1000$ were divided in 20 equally spaced logarithmic bins. Fig. 6 shows the mean and the rms variation of $\ell(\ell + 1) C_\ell / 2\pi$ measured from the 20 independent realizations of the data. We find that the angular power spectrum estimated from the simulated GMRT data is in good agreement with the model (equation 14) that was used to simulate the data. We next test the predicted error estimate δC_ℓ given by equation (22). To do this, we have evaluated $\sigma^2_{\hat{E}_{\rm B}}(a)$ by explicitly carrying out the sum $\sum_{ijkl}$ where the indices each runs over all the baselines in bin a. For $\boldsymbol{V}_2$ (equation 13), we have used the mean C_ℓ estimated from the 20 realizations and the value of σ_n that was used for the system noise in the simulation. We find that δC_ℓ predicted by the analytic error estimate (equation 22) is in reasonably good agreement with the rms obtained from the 20 independent realizations of the data. The results for the Random data are very similar to those for GMRT, and we have not shown these separately here.

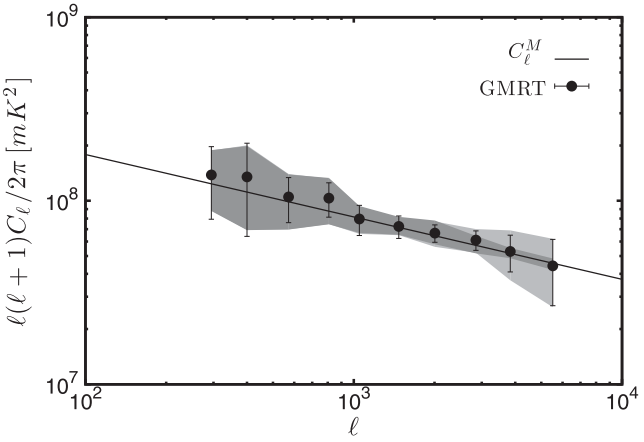

Figure 6. This shows C_ℓ multiplied with $\ell(\ell+1)/2\pi$, plotted as a function of ℓ. The solid line shows the input model (equation 14) used for the simulations, and the points show the values recovered by the Bare Estimator (equation 16). The points show the mean and the light shaded region shows the 1σ variation measured from 20 realizations of the GMRT simulations. The dark shaded region shows the cosmic variance which has been calculated by setting the system noise $\sigma_n = 0$ in the simulation, and the error bars show 1σ error bars predicted using equation (22). The errors are dominated by the cosmic variance at $\ell \leq 2500$ where the dark and faint shaded regions coincide. We see that the Bare Estimator correctly recovers the input model, and the predicted error bars are consistent with the errors measured from the simulations.

In conclusion of this section, we find that the Bare Estimator (equation 16) is able to successfully extract the angular power spectrum directly from the measured visibilities. We further show that (equation 22) provides a reasonably good estimate of the statistical errors for the measured angular power spectrum. The errors depend on the choice of the weights w_{ij}, the baseline distribution, the magnitude of the signal and the system noise. In Fig. 6, we see that the error decreases with increasing ℓ until $\ell \sim 2500$ beyond which the error increases again. We find that this feature does not change significantly between the GMRT and the Random simulations. Based on this, we conclude that this behaviour of the error is largely determined by the relative contributions from the signal whose magnitude falls with ℓ and the system noise which has been assumed to be constant across all baselines. The errors at $\ell \leq 2500$ are cosmic variance dominated, whereas the errors are dominated by the system noise at larger ℓ.

5 THE TAPERED GRIDDED ESTIMATOR

The telescope primary beam is usually not very well quantified at large angles where we have the frequency-dependent pattern of nulls and sidelobes (Fig. 1). Point sources located near the nulls and the sidelobes are a problem for estimating the angular power spectrum of the diffuse background radiation. Further, point sources located far away from the pointing centre, particularly those located near the nulls, introduce ripples along the frequency direction in the MAPS. This poses a severe problem for separating the foregrounds from the cosmological 21-cm signal. As pointed out in Ghosh et al. (2011b), it is possible to avoid these problems by tapering the sky response through a frequency-independent window function $\mathcal{W}(\theta)$. In this work, we choose a Gaussian $\mathcal{W}(\theta) = \mathrm{e}^{-\theta^2/\theta_\mathrm{w}^2}$ such that $\theta_\mathrm{w} = f\theta_0$ with $f \leq 1$ so that the window function cuts off the sky response well before the first null. This tapering is achieved by convolving the measured visibilities

$$\mathcal{V}_\mathrm{c}(\boldsymbol{U}) = \tilde{w}(\boldsymbol{U}) \otimes \mathcal{V}(\boldsymbol{U}) \tag{23}$$

where $\tilde{w}(\boldsymbol{U}) = \pi\theta_\mathrm{w}^2 \mathrm{e}^{-\pi^2 U^2 \theta_\mathrm{w}^2}$ is the Fourier transform of $\mathcal{W}(\theta)$. The convolved visibilities $\mathcal{V}_\mathrm{c}(\boldsymbol{U})$ are the Fourier transform of the product $\mathcal{W}(\theta)\,\mathcal{A}(\theta)\,\delta I(\boldsymbol{\theta})$ whose sky response can be well controlled through the window function $\mathcal{W}(\theta)$.

Current radio interferometers are expected to produce considerably large volumes of visibility data in observations spanning many frequency channels and large observing times. Given the potentially large computational requirement, it is useful to compress the visibility data by gridding it. We choose a rectangular grid in the *uv* plane and consider the convolved visibilities

$$\mathcal{V}_{cg} = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)\,\mathcal{V}_i \tag{24}$$

where $\boldsymbol{U}_g$ refers to the different grid points and $\mathcal{V}_i$ refers to the measured visibilities. We now focus our attention on $\mathcal{S}_{cg} = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)\,\mathcal{S}_i$ which is the sky signal contribution to $\mathcal{V}_{cg}$. This can be written as

$$\mathcal{S}_{cg} = \int \mathrm{d}^2U\ \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}) B(\boldsymbol{U}) \mathcal{S}(\boldsymbol{U}), \tag{25}$$

where $B(\boldsymbol{U}) = \sum_i \delta_\mathrm{D}^2(\boldsymbol{U} - \boldsymbol{U}_i)$ is the baseline sampling function of the measured visibilities and $\delta_\mathrm{D}^2(\boldsymbol{U})$ is the 2D Dirac delta function. The integral in equation (25) is dominated by the contribution from within a disc of radius $\sim(\pi\theta_\mathrm{w})^{-1}$ centred around $\boldsymbol{U}_g$. Assuming that the sampling function $B(\boldsymbol{U})$ is nearly uniform within this disc, we can replace $B(\boldsymbol{U})$ in equation (25) by its average value

$$\bar{B}(\boldsymbol{U}_g) = \left[\frac{\int \mathrm{d}^2U\ \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}) B(\boldsymbol{U})}{\int \mathrm{d}^2U\ \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U})}\right] \tag{26}$$

evaluated at the grid point $\boldsymbol{U}_g$. We then have the approximate equation

$$\mathcal{S}_{cg} = \bar{B}(\boldsymbol{U}_g) \int \mathrm{d}^2U\ \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}) \mathcal{S}(\boldsymbol{U})\,. \tag{27}$$

Considering equation (26) for $\bar{B}(\boldsymbol{U}_g)$, the denominator has value $\mathcal{W}(0) = 1$ whereby $\bar{B}(\boldsymbol{U}_g) = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)$ and we have

$$\mathcal{S}_{cg} = \left[\sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)\right] \int \mathrm{d}^2U\ \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}) \mathcal{S}(\boldsymbol{U})\,. \tag{28}$$

We note that equation (28) holds only if we have a uniform and sufficiently dense baseline distribution in the vicinity of the grid point $\boldsymbol{U}_g$. This breaks down if we have a patchy and sparse baseline distribution, and it is then necessary to use

$$\mathcal{S}_{cg} = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mathcal{S}(\boldsymbol{U}_i)\,. \tag{29}$$

In such a situation, it is necessary to take the exact patchy *uv* distribution into account, and it is difficult to make generic analytic predictions. Here, we have assumed a uniform baseline distribution, and we have used equation (28) extensively in the subsequent calculations.

The integral in equation (28) is the Fourier transform of the product $\mathcal{W}(\theta)\,\mathcal{A}(\theta)\,\delta I(\boldsymbol{\theta}) \equiv \mathcal{A}_\mathcal{W}(\theta)\,\delta I(\boldsymbol{\theta})$. We may think of $\mathcal{A}_\mathcal{W}(\theta)$ as a modified primary beam pattern which has a new θ_FWHM which is a factor $f/\sqrt{1+f^2}$ smaller than θ_FWHM given in Table 1 and whose sidelobes are strongly suppressed. We can approximate the modified primary beam pattern as a Gaussian $\mathcal{A}_\mathcal{W}(\theta) = \mathrm{e}^{-\theta^2/\theta_1^2}$ with $\theta_1 = f(1+f^2)^{-1/2}\theta_0$. Using this, we can generalize

equation (13) to calculate the correlation of the gridded visibilities $V_{c2gg'} = \langle \mathcal{V}_{cg} \mathcal{V}^*_{cg'} \rangle$. The crucial point is that we have to replace V_0 and σ_0 in equation (13) with $V_1 = \frac{\pi\theta_1^2}{2} \left(\frac{\partial B}{\partial T}\right)^2$ and $\sigma_1 = f^{-1}\sqrt{1+f^2}\sigma_0$ in order to account for the modified primary beam pattern $\mathcal{A}_{\mathcal{W}}(\theta)$. We then have

$$V_{c2gg'} = K_{1g} K^*_{1g'} V_1 \mathrm{e}^{-|\Delta \boldsymbol{U}_{gg'}|^2/\sigma_1^2} C_{\ell_g} + 2\sigma_n^2 K_{2gg'}, \tag{30}$$

where $\ell_g = 2\pi U_g$, $K_{1g} = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)$, $K_{2gg'} = \sum_i \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)\tilde{w}^*(\boldsymbol{U}_{g'} - \boldsymbol{U}_i)$, and $\Delta \boldsymbol{U}_{gg'} = \boldsymbol{U}_g - \boldsymbol{U}_{g'}$.

We now define the estimator $\hat{E}_g$ for the angular power spectrum at a single grid point g as

$$\hat{E}_g = \frac{(\mathcal{V}_{cg}\mathcal{V}^*_{cg} - \sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 \mid \mathcal{V}_i \mid^2)}{(\mid K_{1g} \mid^2 V_1 - K_{2gg} V_0)}. \tag{31}$$

Using equations (30) and (13), respectively, to evaluate the expectation values

$$\left\langle \mathcal{V}_{cg}\mathcal{V}^*_{cg} \right\rangle = \mid K_{1g} \mid^2 V_1 C_{\ell_g} + 2\sigma_n^2 K_{2gg} \tag{32}$$

and

$$\sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 \langle \mid \mathcal{V}_i \mid^2 \rangle = V_0 \sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 C_{\ell_i} + 2\sigma_n^2 K_{2gg}, \tag{33}$$

we see that the system noise contributions to these two terms are exactly equal and it exactly cancels out in $\langle \hat{E}_g \rangle$. Further, assuming that $\sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 C_{\ell_i} \approx C_{\ell_g} K_{2gg}$, we have

$$\langle \hat{E}_g \rangle = C_{\ell_g}. \tag{34}$$

We see that $\hat{E}_g$ defined in equation (31) gives an unbiased estimate of the angular power spectrum C_ℓ avoiding the positive noise bias caused by the system noise.

The terms K_{1g} and K_{2gg} in equation (31) are both proportional to N_g the number of visibilities that contribute to the grid point g. For large N_g, it is reasonable to assume that $|K_{1g}|^2 \gg K_{2gg}$ and we thereby simplify equation (31) to obtain

$$\hat{E}_g = \frac{(\mathcal{V}_{cg}\mathcal{V}^*_{cg} - \sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 \mid \mathcal{V}_i \mid^2)}{\mid K_{1g} \mid^2 V_1} \tag{35}$$

for the estimator.

We use this to define the binned Tapered Gridded Estimator

$$\hat{E}_{\mathrm{G}}(a) = \frac{\sum_g w_g \hat{E}_g}{\sum_g w_g}, \tag{36}$$

where w_g refers to the weight assigned to the contribution from any particular grid point. This has an expectation value

$$\langle \hat{E}_{\mathrm{G}}(a) \rangle = \frac{\sum_g w_g C_{\ell_g}}{\sum_g w_g} \tag{37}$$

which can be written as

$$\langle \hat{E}_{\mathrm{G}}(a) \rangle = \bar{C}_{\bar{\ell}_a}, \tag{38}$$

where $\bar{C}_{\bar{\ell}_a}$ is the average angular power spectrum at

$$\bar{\ell}_a = \frac{\sum_g w_g \ell_g}{\sum_g w_g} \tag{39}$$

which is the effective angular multipole for bin a.

We next calculate the variance of $\hat{E}_{\mathrm{G}}(a)$ defined as

$$[\delta C_{\ell_a}]^2 \equiv \sigma^2_{E_{\mathrm{G}}}(a) = \langle \hat{E}^2_{\mathrm{G}}(a) \rangle - \langle \hat{E}_{\mathrm{G}}(a) \rangle^2. \tag{40}$$

Explicitly using equation (35) yields a rather unwieldy expression which is not very useful for making analytic predictions for the variance. The first term in the numerator of equation (35) which is of order N_g^2 makes a much larger contribution to the variance than the second term $\sum_i \mid \tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i) \mid^2 \mid \mathcal{V}_i \mid^2$ which is of order N_g. In our analysis, we make the simplifying assumption that we can drop the second term which yields

$$\sigma^2_{E_{\mathrm{G}}}(a) = \frac{\sum_{gg'} w_g w'_g \mid K^{-1}_{1g} K^{*-1}_{1g'} V_{c2gg'} \mid^2}{V_1^2 [\sum_g w_g]^2}. \tag{41}$$

We further approximate $K_{2gg'} = \mathrm{e}^{-|\Delta \boldsymbol{U}_{gg'}|^2/\sigma_1^2} K_{2gg}$ which allows us to write the variance as

$$\sigma^2_{E_{\mathrm{G}}}(a) = \frac{\sum_{gg'} w_g w'_g \mathrm{e}^{-2|\Delta \boldsymbol{U}_{gg'}|^2/\sigma_1^2} \mid C_{\ell_g} + \frac{2K_{2gg'}\sigma_n^2}{K_{1g}K^*_{1g'}V_1} \mid^2}{[\sum_g w_g]^2} \tag{42}$$

using equation (30).

We have applied the Tapered Gridded Estimator to the simulated GMRT and Random data. The 20 realizations were used to calculate the mean and the variance of the estimated C_ℓ. We have considered the values $f = 1.0$, 0.8, 0.65, and 0.4 for the tapering window, and have also tried two different weight schemes $w_g = 1$ and K^2_{1g}, respectively. The former assigns equal weight to every grid point that has same data, this is expected to minimize the cosmic variance. The latter scheme assigns a larger weight to grid points which have a denser visibility sampling relative to the grid points with sparser sampling. This is expected to minimize the system noise contribution. The grid spacing ΔU in the uv plane is chosen based on two considerations. A very small value of ΔU results in a very large number of grid points which do not contain independent signal contributions. This also unnecessarily increases the computation time. In contrast, a large value of ΔU implies that the signal in many visibilities is very poorly represented in the gridded data, resulting in a loss of signal. We have chosen a grid spacing $\Delta u = \sqrt{\ln 2}/(2\pi\theta_{\mathrm{w}})$ which corresponds to one-fourth of the FWHM of $\tilde{w}(\boldsymbol{U})$ as an optimum value. For any fixed grid position $\boldsymbol{U}_g$, we have restricted the contribution to baselines $\boldsymbol{U}_i$ within $\mid \boldsymbol{U}_g - \boldsymbol{U}_i \mid \leq 6\Delta U$. The weight function $\tilde{w}(\boldsymbol{U}_g - \boldsymbol{U}_i)$ falls considerably and we do not expect a significant contribution from the visibilities beyond this baseline separation. The tapering also modifies the smallest baseline where the approximation of equation (11) is valid, and the grid points $\boldsymbol{U}_g$ in the range $U_{\min} = \sqrt{1+f^2} f^{-1} 40$ to 1000 were binned into 10 equally spaced logarithmic bins for this analysis.

Fig. 7 shows the results for $f = 0.8$ and $w_g = |K_{1g}|^2$. We see that for both GMRT and Random the estimated C_ℓ are roughly within the 1σ region of the input model angular power spectrum C^M_ℓ. For GMRT, however, the estimated C_ℓ values all appear to be somewhat in excess of C^M_ℓ indicating that we have an overestimate of the angular power spectrum relative to C^M_ℓ. In comparison, the C_ℓ values are in better agreement with C^M_ℓ for the Random simulation. For both GMRT and Random, the error estimates predicted by equation (42) are in good agreement with the rms fluctuation estimated from the 20 realizations. We note that the rms fluctuation of C_ℓ is more for GMRT in comparison to Random.

The Tapered Gridded Estimator is expected to give an unbiased estimate of C_ℓ provided we have a uniform and sufficiently dense baseline distribution. We test this using the Random simulations which have a uniform baseline distribution. In such a situation, we expect the deviation $C_\ell - C^M_\ell$ to arise purely from statistical fluctuations. The deviation is expected to have values around $\sigma/\sqrt{N_{\mathrm{r}}}$ and converge to 0 as N_{r}, the number of realizations, is increased. For

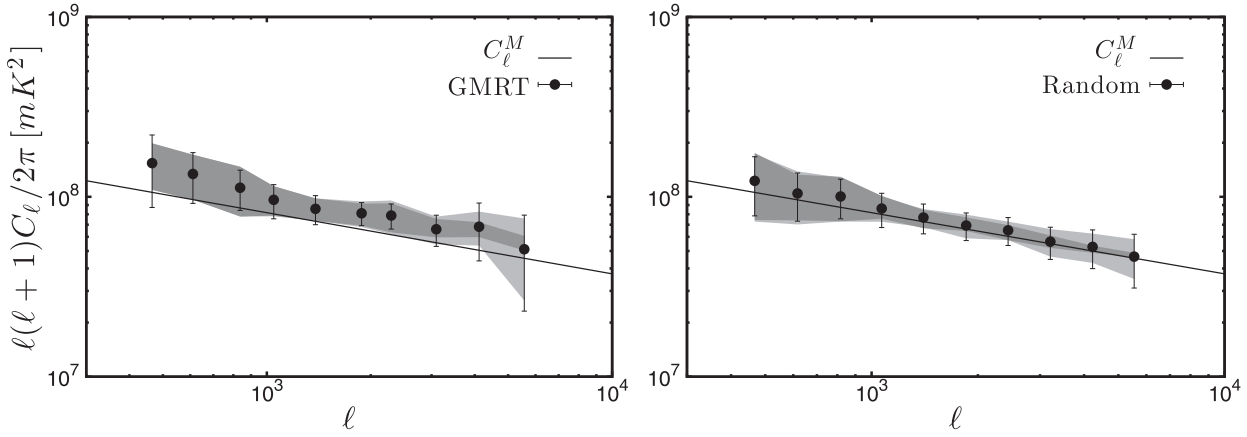

Figure 7. Same as Fig. 6, but for the Tapered Gridded Estimator.

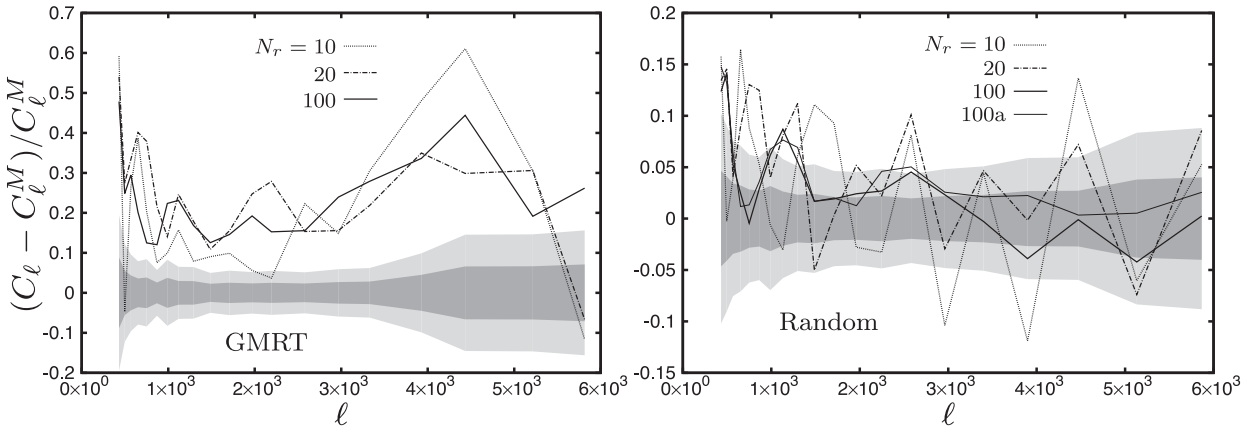

Figure 8. The different curves show the fractional deviation $(C_\ell - C_\ell^M)/C_\ell^M$ for the different numbers of realizations (N_r) shown in the figure. The curve 100a corresponds to $N_r = 100$ with 869 828 baselines, which is four times the number of baselines in the other simulations. The two shaded region show $\sigma/(\sqrt{N_r}\, C_\ell^M)$ for $N_r = 20$ and 100, respectively. We have used $f = 0.8$ and $w_g = |K_{1g}|^2$, with 20 equally spaced logarithmic bins in ℓ.

this purpose, we have studied (Fig. 8) how the fractional deviation $(C_\ell - C_\ell^M)/C_\ell^M$ varies if we increase the number of realizations from $N_r = 10$ to 100. We find that it is more convenient to use 20 equally spaced logarithmic bins in ℓ to highlight the convergence of the fractional deviation with increasing N_r. Note that we have used 10 bins (as mentioned earlier) everywhere except in Fig. 8. For the Random simulation (right-hand panel), we find that as the number of realizations is increased the convergence of the fractional deviation to 0 is clearly visible for $\ell \geq 1.2 \times 10^3$ ($U \geq 200$). Further, the fractional deviation is also found to be consistent with $\sigma/(\sqrt{20}\, C_\ell^M)$ and $\sigma/(10\, C_\ell^M)$ expected for $N_r = 20$ and 100, respectively. At smaller baselines, however, the behaviour is not so clear. The approximation equation (11) for the convolution and the approximation for the primary beam pattern each introduce around 2–5 per cent errors in the estimated C_ℓ at small baselines. Further, for a uniform baseline distribution, the bins at the smallest ℓ values contain fewer baselines and also fewer grid points, and are susceptible to larger fluctuations. The discrete *uv* sampling due to the finite number of baselines is also expected to introduce some errors at all values of ℓ. To test this effect, we have considered a situation where $N_r = 100$ and the total number of baselines is increase to 869 828 which is a factor of 4 larger compared to the other simulations. We find that for $\ell \geq 3 \times 10^3$ the fractional deviation falls from ∼5 per cent to ∼2 per cent when the baseline density is increased, this difference is not seen at smaller baselines. In summary, the tests clearly show that for a uniform baseline distribution the estimator is unbiased for $\ell \geq 1.2 \times 10^3$. In contrast, for the GMRT (left-hand panel) the fractional deviation does not converge to 0 as N_r is increased. We see that C_ℓ is overestimated at all values of ℓ. As mentioned earlier, the GMRT has a patchy *uv* coverage for which equation (27), which assumes a uniform baseline distribution, breaks down. The overestimate is a consequence of GMRT's patchy *uv* coverage, and is not inherent to the Tapered Gridded Estimator. The rms fluctuations also are larger for GMRT in comparison to the Random simulations (Fig. 7). This too is a consequence of GMRT's patchy *uv* coverage.

We now study how the estimator behaves for different values of f. Figs 9 and 10, respectively, show the relative deviation $(C_\ell - C_\ell^M)/C_\ell^M$ and the relative error σ/C_ℓ^M for different values of f with $w_g = |K_{1g}|^2$. Here, C_ℓ and σ refer to the mean and rms estimated from the 20 realizations. We find that the deviations are roughly within the 1σ errors for all the cases that we have considered. For GMRT, the deviation increases with decreasing f. This effect is only visible at low ℓ for Random. The error σ increases with f for both GMRT and Random. In all cases, the error is found to decrease until $\ell \sim 2000$ and then increase subsequently. As mentioned earlier for the Bare Estimator, we interpret this as a transition from cosmic-variance- to system-noise-dominated errors as ℓ is increased. The sky coverage of the modified primary beam $\mathcal{A}_{\mathcal{W}}(\theta)$ falls with a decrease in f. This explains the behaviour of the cosmic variance contribution which increases as f is reduced. We further see that the

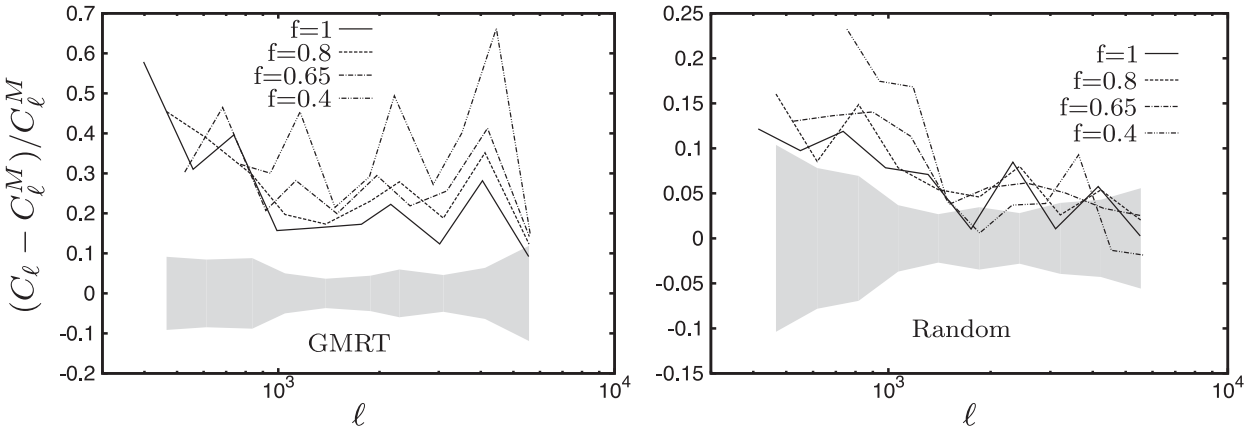

Figure 9. This shows how the fractional deviation varies with f for $w_g = |K_{1g}|^2$. The results are averaged over 20 realizations of the sky signal. For comparison, the shaded region shows $\sigma/(\sqrt{20}\, C_\ell^M)$ which is the expected statistical fluctuation for $f = 0.8$.

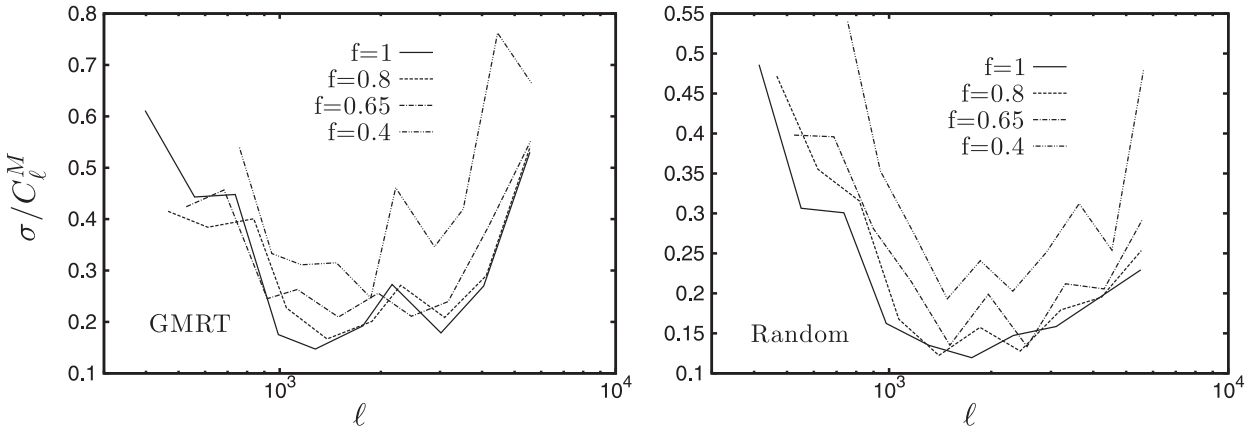

Figure 10. This figure shows the relative error (σ/C_ℓ^M) estimated from 20 realization of the simulation. Here, we have used $w_g = |K_{1g}|^2$ and the different f values shown in this figure.

system noise contribution also increases as f is reduced. This can be attributed to the term $V_1 = \frac{\pi\theta_1^2}{2}$ which appears in equation (42). This effectively increases the system noise contribution relative to C_ℓ as f is reduced.

We have studied the relative performance of the two weight schemes mentioned earlier. Fig. 11 shows the relative deviation and the relative error for both $w_g = 1$ and $|K_{1g}|^2$ for $f = 0.8$. As expected, the first scheme performs better in the cosmic-variance-dominated regime. The difference between the two weight schemes, however, is not very large in this regime. The second weight scheme performs significantly better in the system-noise-dominated region. In this region, the errors are nearly doubled if we use $w_g = 1$ instead of $|K_{1g}|^2$.

In summary, we have introduced a Gridded Estimator for the angular power spectrum where it is possible to avoid the positive noise bias which arises due to the contribution from the correlation of a visibility with itself. Further, the estimator allows the possibility to taper the sky response and thereby implement sidelobe suppression. We have used simulated visibility data to validate the estimator. We find that the estimator provides an unbiased estimate of C_ℓ for $\ell \geq 1.2 \times 10^3$ if we have a sufficiently dense, uniform baseline distribution. We also find that equation (42) provides a good analytic estimate of the errors in the measured C_ℓ. The estimator is found to be sensitive to the telescope's *uv* coverage, and we have somewhat of an overestimate for the GMRT which has a patchy *uv* coverage. This deviation, however, is roughly within the 1σ error bars and is not expected to be a serious issue. It is possible to carry out simulations with the actual observational *uv* coverage and use these to compensate for the overestimate. The new telescopes like LOFAR (discussed later) have a denser and more uniform *uv* coverage, and we do not expect this issue to be of concern there. The 1σ errors, we find, increase as the tapering is increased. The choice of f, however, is decided by issues related to point source removal not considered here. We find that the weight scheme $w_g = |K_{1g}|^2$ performs better than $w_g = 1$, and we use the former for the subsequent analysis.

6 A COMPARISON OF THE TWO ESTIMATORS

Comparing the Bare Estimator with the Tapered Gridded Estimator, we see (left-hand panel of Fig. 11) that the former is more successful in recovering the input sky model. The statistical errors also (right-hand panel of Fig. 11), we find, are somewhat smaller for the Bare Estimator. The Bare Estimator deals directly with the measured visibilities, and in a sense we expect it to outperform any other estimator which deals with gridded visibilities. What then is the motivation to consider a Gridded Estimator which is not able to recover the input model with as much accuracy as the Bare Estimator? The Bare Estimator deals directly with the

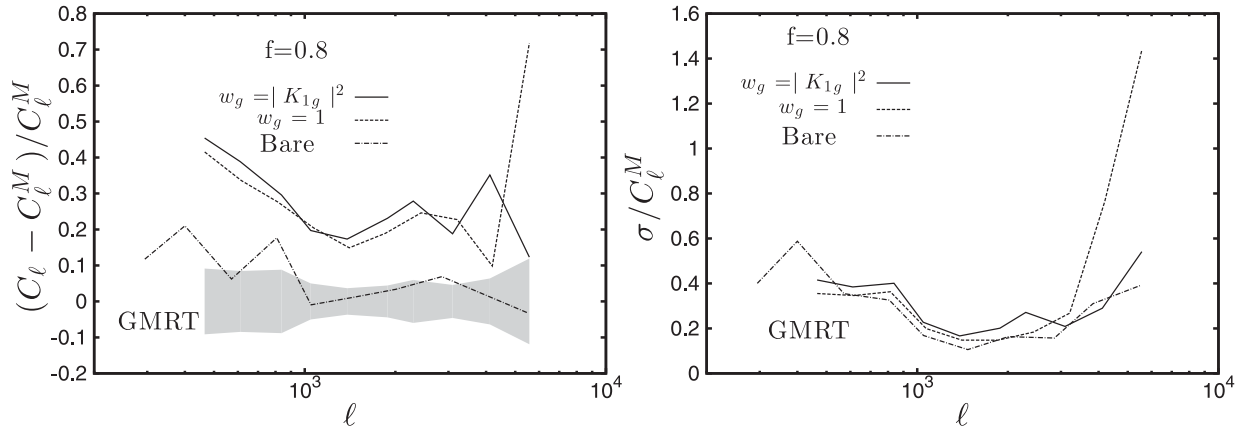

Figure 11. The left-hand (right-hand) panel shows the fractional deviation (error) for the two weight schemes $w_g = 1$ and $|K_{1g}|^2$, respectively, both with $f = 0.8$. The results for the Bare Estimator have also been shown for comparison. The results are based on 20 realizations of the sky signal. For comparison, the shaded region in the left-hand panel shows $\sigma/(\sqrt{20}\,C_\ell^M)$ which is the expected statistical fluctuation for $w_g = |K_{1g}|^2$.

visibilities and the computational time for the pairwise correlation in equation (16) scales proportional to N^2, where N is the total number of visibilities in the data. Further, the error calculation in equation (22) is expected to scale as N^4. In contrast, the computation time is expected to scale as N for Tapered Gridded Estimator. This N dependence arises in the process of gridding the visibilities, the correlation equation (35) and the error estimate equation (42) are both independent of N.

Fig. 12 show the computation time for the two estimators as the number of visibilities varied. We see that the computation time shows the expected N dependence for large values of $N(>1000)$. The Bare Estimator takes less computation time when N is small ($N \leq 10^4$). However, the computation time for the Bare Estimator and its error estimate are larger than that for the Tapered Gridded Estimator for $N \geq 10^5$. The Bare Estimator is extremely computation extensive for a large N and it is preferable to use the Gridded Estimator when $N \geq 10^5$. Based on this, we focus on the Tapered Gridded Estimator for most of the subsequent discussion.

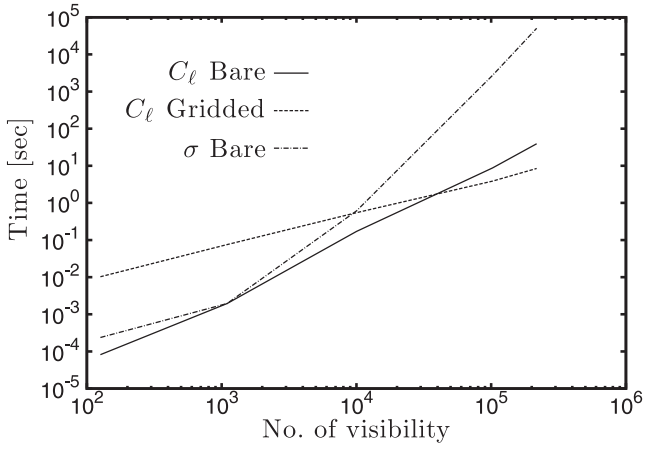

Figure 12. This shows how the computation time varies with the number of visibility data for the two different estimators. The computation time for analytically predicting the error (equation 22) for the Bare Estimator is also shown.

7 GAIN ERRORS

The measured visibilities have undetermined time varying gains which arise due to the atmosphere, receiver system, etc. The calibration procedure attempts to determine these gains and correct for them, but this generally leaves unknown residual gain errors in the data. Datta, Bhatnagar & Carilli (2009) and Datta, Bowman & Carilli (2010) have studied the impact of the residual gain errors on bright source subtraction and place a tolerance limit for detecting the reionization 21-cm signal. Recently, Shaw et al. (2014b) have discussed how amplifier gains effect on the detection of the 21-cm signal from post-reionization and reionization era, respectively. Here, we study the effect of gain errors on the estimators that we have defined earlier. For this work, we assume antenna-dependent gain errors whereby the calibrated visibilities can be written as

$$\mathcal{V}(\boldsymbol{U}_{ab}) = g_a g_b^* [\mathcal{S}(\boldsymbol{U}_{ab}) + \mathcal{N}(\boldsymbol{U}_{ab})], \tag{43}$$

where a, b refer to the two antennas corresponding to the baseline $\boldsymbol{U}_{ab}$, and $g_a = (1+\alpha_a)\mathrm{e}^{\mathrm{i}\phi_a}$ and $g_b = (1+\alpha_b)\mathrm{e}^{\mathrm{i}\phi_b}$ are the respective antenna gains. Here, α_as and the ϕ_as are, respectively, the amplitude and the phase errors of the individual antenna gains. We have assumed that both α_a and ϕ_a are Gaussian random variables of zero mean and variance σ_α^2 and σ_ϕ^2, respectively. The errors are assumed to be independent in different antennas and at different time instants.

The two-visibility correlation can be written as

$$\langle \mathcal{V}(\boldsymbol{U}_{ab})\mathcal{V}^*(\boldsymbol{U}_{cd})\rangle = \langle g_a g_b^* g_c^* g_d\rangle [S_2(\boldsymbol{U}_{ab}, \boldsymbol{U}_{cd}) + N_2(\boldsymbol{U}_{ab}, \boldsymbol{U}_{cd})], \tag{44}$$

where the product of the gains is to be averaged over different realizations of the gain errors α and ϕ. We now have three different possibilities which we discussed below.

Case I. The two visibilities $\mathcal{V}(\boldsymbol{U}_{ab})$ and $\mathcal{V}(\boldsymbol{U}_{cd})$ are at two different time instants or they have no antenna in common. In this situation, we have

$$\langle g_a g_b^* g_c^* g_d\rangle = \mathrm{e}^{-2\sigma_\phi^2}. \tag{45}$$

Case II. The two visibilities $\mathcal{V}(\boldsymbol{U}_{ab})$ and $\mathcal{V}(\boldsymbol{U}_{cd})$ are at the same time instant and have only one antenna in common. In this situation, we have

$$\langle g_a g_b^* g_c^* g_d\rangle = \left(1+\sigma_\alpha^2\right)\mathrm{e}^{-\sigma_\phi^2}. \tag{46}$$

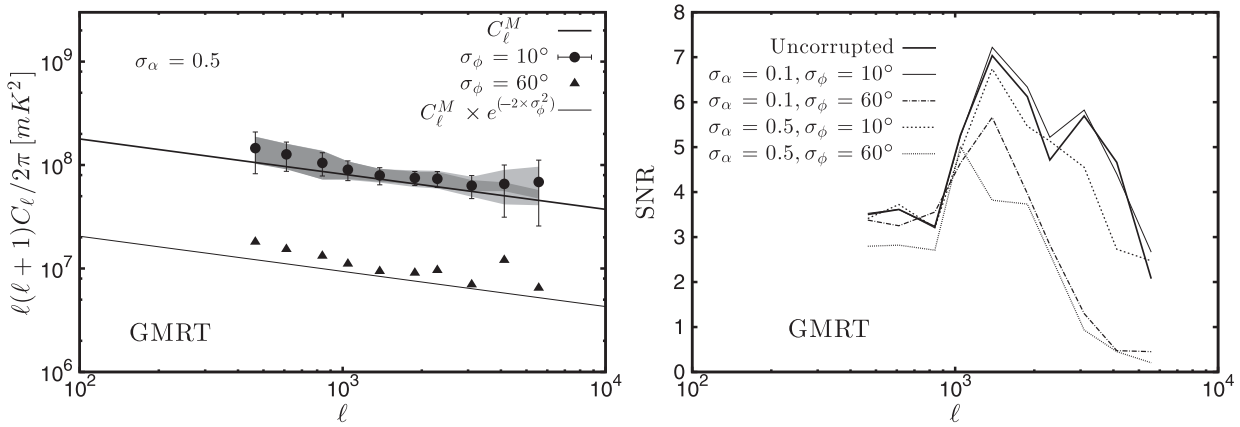

Figure 13. The left-hand panel shows the same as Fig. 6 for the Tapered Gridded Estimator using corrupted visibilities with the σ_α and σ_ϕ values shown in the figure. We have also shown $C_\ell^M \times e^{(-2\times\sigma_\phi^2)}$ with $\sigma_\phi = 60^\circ$ for comparison. The right-hand panel shows the SNR for different values of σ_α and σ_ϕ.

Case III. Both $\mathcal{V}(\boldsymbol{U}_{ab})$ and $\mathcal{V}(\boldsymbol{U}_{cd})$ referred the same measured visibility. In this situation, we have

$$\langle g_a g_b^* g_c^* g_d \rangle = \left(1+\sigma_\alpha^2\right)^2 . \tag{47}$$

The signal contribution to both the estimators defined earlier is dominated by Case I, whereas the noise is dominated by Case III. Based on this, it is possible to generalize equation (13) to obtain the approximate relation

$$V_{2ij} = \mathrm{e}^{-2\sigma_\phi^2} V_0\, \mathrm{e}^{-|\Delta \boldsymbol{U}_{ij}|^2/\sigma_0^2}\, C_{\ell_i} + \left(1+\sigma_\alpha^2\right)^2 \delta_{ij} 2\sigma_n^2 \tag{48}$$

which takes into account the effect of gain errors. It is also possible to generalize equation (30) for the gridded visibilities in a similar fashion. Using these to calculate the effect of gain errors on the estimators defined earlier, we have

$$\langle \hat{E}(a) \rangle = \mathrm{e}^{-2\sigma_\phi^2} \bar{C}_{\bar{\ell}_a} \tag{49}$$

for both the Bare and the Tapered Gridded Estimators. We see that both the estimators are unaffected by the error in the gain amplitude; however, the phase errors cause the expectation value of the estimator to decrease by a factor $\mathrm{e}^{-2\sigma_\phi^2}$. It is quite straightforward to generalize equations (22) and (41) to incorporate the effect of the gain errors in the variance of the Bare and the Tapered Gridded Estimators, respectively. The main effect is that the signal contribution is suppressed by a factor of $\mathrm{e}^{-2\sigma_\phi^2}$, whereas the system noise contribution is jacked up by a factor of $(1+\sigma_\alpha^2)^2$ (equation 48). We consequently expect the SNR to remain unchanged in the cosmic-variance-dominated regime at low ℓ, whereas we expect the SNR to fall in the system-noise-dominated regime (large ℓ). Further, we also expect the transition from the cosmic-variance- to the system-noise-dominated regime to shift to smaller ℓ values if the gain errors increase.

We have carried out simulations to test the effect of gain errors on the angular power spectrum estimators. For this, we have generated 20 different realizations of the random gain errors and used these to corrupt the simulated visibilities described in Section 3. The simulations were carried out for different values of σ_α and σ_ϕ. We have applied both the Bare and the Tapered Gridded Estimators on the corrupted visibilities. Both the estimators show very similar behaviour under gain errors, and we show the results for only the Tapered Gridded Estimator.

We have considered two values $\sigma_\alpha = 0.1$ and 0.5 which, respectively, correspond to 10 and 50 per cent errors in the gain amplitude. The left-hand panel of Fig. 13 shows the results for $\sigma_\alpha = 0.5$. We see that the expectation value of the estimator is unaffected by the errors in the gain amplitude. For the phase errors, we have considered the values $\sigma_\phi = 10^\circ$ and 60° for which $\mathrm{e}^{-2\sigma_\phi^2}$ have values 0.94 and 0.11, respectively. The left-hand panel of Fig. 13 shows that equation (49) provides a good description for the effect of the gain errors on the angular power spectrum estimator. We see that the net result of the phase errors is that the estimated angular power spectrum is reduced by a factor $\mathrm{e}^{-2\sigma_\phi^2}$ relative to the input model.

The right-hand panel of Fig. 13 shows the SNR for different values of σ_α and σ_ϕ. The rms fluctuation σ_{E_G} of the estimator is expected to depend exponentially as $\mathrm{e}^{-2\sigma_\phi^2}$ on the phase errors and have a $(1+\sigma_\alpha^2)^2$ dependence on the amplitude errors (equation 48). We find that the simulated SNR are more sensitive to the phase errors in comparison to the amplitude errors. The SNR is nearly invariant to gain errors in the cosmic-variance-dominated regime (low ℓ) where σ_{E_G} is reduced by the same factor $\mathrm{e}^{-2\sigma_\phi^2}$ as the expectation value of the estimator. However, the transition from the cosmic-variance-dominated to the system-noise-dominated regime (approximately the peak of the SNR curves) shifts to smaller ℓ if the gain errors are increased. The amplitude errors, we see reduces the SNR at large ℓ where the error is dominated by the system noise.

8 THE *w*-TERM

The entire analysis, until now, has been based on the assumption that the visibility contribution $\mathcal{S}(\boldsymbol{U})$ from the sky signal is the Fourier transform of the product of $\mathcal{A}(\boldsymbol{\theta})$ and $\delta I(\boldsymbol{\theta})$. This is only an approximate relation which is valid only if the FoV is sufficiently small. The actual relation is

$$\mathcal{S}(u,v,w) = \int \mathrm{d}l\mathrm{d}m \frac{\delta I(l,m)\mathcal{A}(l,m)}{\sqrt{1-l^2-m^2}} \mathrm{e}^{-2\pi\mathrm{i}[ul+vm+w(\sqrt{1-l^2-m^2}-1)]}, \tag{50}$$

where the w-term, which we have ignored until now, is the baseline component along the line of sight to the phase centre and l, m are the direction cosines corresponding to any point on the sky. In a situation where the primary beam pattern falls of within a small angle from the phase centre, it is adequate to treat the region of sky under observation as a 2D plane and use $(l, m) = (\theta_x, \theta_y)$. For example, the GMRT has an FWHM of 186 arcmin for which $\sqrt{1-l^2-m^2} \approx 0.997$. The term $\sqrt{1-l^2-m^2}$ which appears in

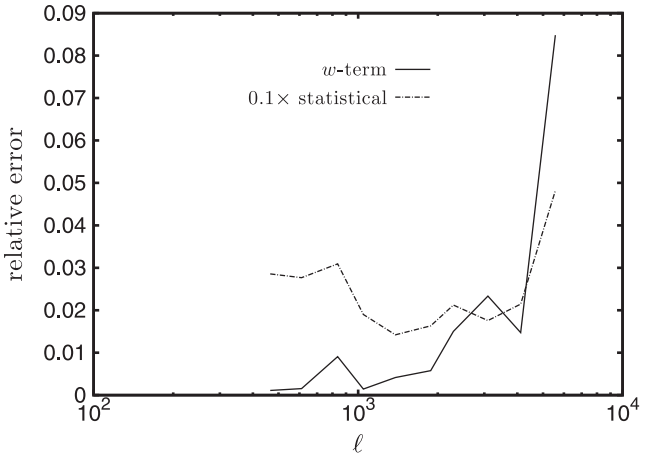

Figure 14. This shows the relative change in the estimated angular power spectrum using Tapered Gridded Estimator due to the w-term. For comparison, we have also shown $0.1 \times \delta C_\ell / C_\ell$ which corresponds to 10 per cent of the relative statistical error in C_ℓ.

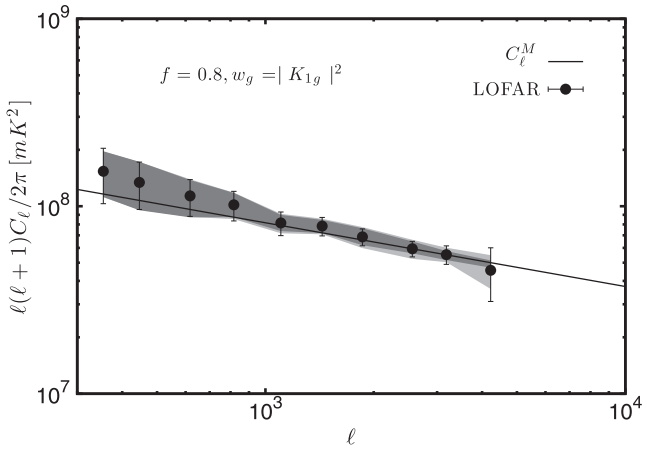

Figure 15. Same as Fig. 6 for the Tapered Gridded Estimator and the simulated LOFAR data.

the denominator of equation (50) incorporates the curvature of the sky. We see that this makes an insignificant contribution at the small angles of our interest, and hence may be ignored. The term $w(\sqrt{1-l^2-m^2}-1)$ which appears in the phase in equation (50) has a value $\sim 10^{-3} \times w$ for the angle mentioned earlier, and this is not necessarily small. The value of w depends on the telescope configuration and the observing direction, and may be quite large ($>10^3$). It is therefore necessary to assess the impact of the w-term on the angular power spectrum estimators defined earlier.

We have simulated GMRT visibilities using equation (50) keeping the w-term. The 20 realizations of the sky signal and the baseline tracks are the same as described in Section 3, and we have used the flat-sky approximation (i.e. we have dropped $\sqrt{1-l^2-m^2}$ from the denominator). We have applied both the Bare and the Tapered Gridded Estimator to these simulated visibility data. We show results for only the Tapered Gridded Estimator, the results are very similar for the Bare Estimator and we have not shown these separately. Fig. 14 shows the relative change in the estimated angular power spectrum if we include the w-term. We find that the change due to the w-term is less than 3 per cent for all values of ℓ barring the largest ℓ value where there is a 9 per cent change. The w-term has a larger effect at the large baselines which also correspond to a larger value of w. We find that the change caused by the w-term is less than 10 per cent of the statistical fluctuations for most values of ℓ. In summary, for angular power spectrum estimation, it is adequate to ignore the w-term at the angular scales of our interest for the GMRT.

9 LOFAR

LOFAR, the Low-Frequency Array, is an innovative new radio telescope which operates at the lowest radio frequencies (10–240 MHz; var Haarlem et al. 2013). It consists of an interferometric array of dipole antenna stations distributed throughout the Netherlands and Europe. The individual stations perform the same basic functions as the dishes of a traditional interferometric radio telescope. Hence, the station beam which is analogous to the primary beam ultimately determines the FoV for a given observation. In the High Band Antennas (HBAs, 110–240 MHz), groups of dipole pairs are combined into HBA tiles and the station beam is formed from the combined signal from the tiles. The HBA tiles are sensitive to two orthogonal linear polarizations. Close to the phase centre, the LOFAR station beam can be well modelled with a circular Gaussian and the FWHM of the Gaussian varies approximately from 3.°0 to 5.°0 in the frequency range 115–185 MHz with $\theta_{\rm FWHM} = 3.°8$ at 150 MHz.

In this section, we consider the possibility of using LOFAR to estimate the angular power spectrum of the 150 MHz sky signal after point source subtraction. The LOFAR has a wider FoV compared to the GMRT and we have simulated an $\sim 8^\circ \times 8^\circ$ region of the sky with an angular resolution of 14 arcsec × 14 arcsec. Here, again we have generated 20 independent realizations of the sky signal. The simulations were carried out in exactly the same way as described in Section 3 using the LOFAR parameters given in Table 1. We have generated the LOFAR baseline distribution for the 62 antennas in the central core region for 8 h of observing time. Visibilities were generated with a time interval of 40 s and we obtain a total of 669 809 visibilities in the baseline range $30 \le \boldsymbol{U} \le 800$. We have included the w-term for calculating the LOFAR visibilities. The LOFAR has a denser uv coverage compared to the GMRT, and the simulated baseline range is nearly uniformly covered. We have used $\sigma_n = 2.2$ Jy (var Haarlem et al. 2013) for the system noise in the simulations. Given the large volume of data, we have only used the Tapered Gridded Estimator with $f = 0.8$ and $w_g = K_{1g}^2$.

Fig. 15 shows the angular power spectrum estimated from our simulations. We see that the estimated C_ℓ values are all within the 1σ region of the input model angular power spectrum C_ℓ^M. The estimated C_ℓ values, however, are somewhat in excess of C_ℓ^M at small ℓ (<1000). The fractional deviation $(C_\ell - C_\ell^M)/C_\ell^M$ is around $\sim$30 per cent at the smallest ℓ bin, and it is $\sim$15 per cent at $\ell \sim 800$. The excess is not seen at larger ℓ where the estimated values are in excellent agreement with C_ℓ^M. We also see that the error estimates predicted by equation (42) are in good agreement with the rms fluctuation estimated from the 20 realizations. The transition from cosmic-variance-dominated errors to system-noise-dominated errors occurs at $\ell \sim 2000$ similar to the GMRT. The LOFAR has considerably more baselines compared to the GMRT, and the errors in the estimated angular power spectrum are smaller for LOFAR in comparison to GMRT.

As mentioned earlier in the context of the GMRT, the excess in the estimated C_ℓ may be a consequence of patchy uv coverage at small baselines ($U < 160$). The average baseline density in the region $U < 160$ is several times larger than the average within $U < 800$; however, this does not guaranty that the former is less patchier than

the latter. Further, it is not possible to say anything definite from a visual inspection of the baseline distribution. The convolution with the primary beam pattern and the window function introduces an ∼8 per cent deviation between C_ℓ and C_ℓ^M at $U < 160$. The exact cause of the excess at small ℓ is at present not fully understood.

10 DISCUSSION AND CONCLUSIONS

In this paper, we have introduced two estimators for quantifying the angular power spectrum of the sky brightness temperature. Both of these estimators use the visibilities measured in radio interferometric observations. The Bare Estimator works directly with the measured visibilities, and uses pairwise visibility correlations to estimate the angular power spectrum. The Tapered Gridded Estimator uses the visibility data after gridding on a rectangular grid in the *uv* plane. Here, it is possible to taper the sky response so as to suppress the sidelobes and reduce the FoV. Earlier work (Ghosh et al. 2011b) shows tapering to be an important ingredient in foreground removal for detecting the cosmological 21-cm signal. We have investigated the properties of the estimators, and present analytic formulae for the expectation value (equations 18 and 38) and the variance (equations 22 and 42). The expectation value of both the estimators is free from the positive system noise bias which arises due to the correlation of a visibility with itself. The system noise affects only the variance.

We have carried out simulations to validate the estimators. The simulated sky signal assumes that the point sources have been removed and the residuals are dominated by the diffuse Galactic synchrotron radiation which is modelled as a homogeneous and isotropic Gaussian random field with a power-law angular power spectrum. We consider GMRT observations for most of the analysis. We find that the Bare Estimator is able to recover the input model to a good level of precision. The computation time is found to scale as N^2 with the number of visibility data. Further, the scaling is N^4 for the variance.

We find that the Tapered Gridded Estimator is able to recover the input model C_ℓ^M to a high level of precision provided the baselines have a uniform *uv* coverage. For the GMRT which has a patchy *uv* coverage, the C_ℓ estimated from the Tapered Gridded Estimator is largely within the 1σ errors from the input model C_ℓ^M. There is, however, indication that the angular power spectrum is overestimated to some extent. Comparing the results to a situation with a uniform random baseline distribution, we conclude that the overestimate is a consequence of GMRT's patchy *uv* coverage and is not inherent to the Tapered Gridded Estimator which is unbiased in the ideal situation of uniform *uv* coverage. It is possible to use simulations to quantify this overestimate and correct for this in a real observation. We do not anticipate this overestimate to be a very major obstacle for the Tapered Gridded Estimator. The computation time for this estimator and its variance both scale as N. Long observations spanning many frequency channels will produce large volumes of visibility data. The Bare Estimator is computationally very expensive for large N, and a Gridded Estimator is the only feasible alternative. Consequently, we have focused on the Tapered Gridded Estimator for much of the analysis in the later part of this paper.

Residual gain errors corrupt the measured visibilities, and this is a potential difficulty for estimating the angular power spectrum. We have analysed the effect of gain errors on the two estimators introduced in this paper. Our analysis, validated by simulations, shows that the expectation value of the estimators is unaffected by amplitude errors. The phase errors cause a decrement by the factor $e^{-2\sigma_\phi^2}$ in the expectation value. The statistical errors in the estimated C_ℓ are affected by both the amplitude and the phase errors; however, this is more sensitive to the phase errors relative to the amplitude errors. We have also investigated the effect of the w-term. We find that the w-term does not cause a very big change in the estimated C_ℓ at the scales of our interest here. Our analysis here shows that the residual phase errors can lead to the angular power spectrum being underestimated by a factor of $e^{-2\sigma_\phi^2}$ which has a value ∼0.1 for $\sigma_\phi = 60^\circ$. It is therefore imperative to independently quantify the magnitude of the residual phase errors for a correct estimate of the angular power spectrum.

In addition to GMRT, we have also applied the estimators to simulated LOFAR data. We find that the C_ℓ estimated using the Tapered Gridded Estimator is within the 1σ errors of the input model. There is, however, indication that there is some overestimation (15–30 per cent) at low ℓ ($<$1000). The exact cause of this excess at small ℓ is at present not fully understood.

The two estimators considered here both avoid the positive noise bias which arises due to the system noise contribution in the visibilities. This is achieved by not including the contribution from the correlation of a visibility with itself. As an alternative, one could consider an estimator which straight away squared the measured or the gridded visibilities. In this situation, it is necessary to separately identify the noise bias contribution and subtract it out. The noise bias contribution is expected to be independent of frequency and ℓ. It is, in principle, possible to identify a frequency- and ℓ-independent component and subtract it out. However, our analysis in this paper shows that the errors in the amplitude of the calibrated gains affect the noise bias. Frequency- and baseline-dependent gain errors would manifest themselves as the frequency and ℓ dependence of the noise bias. This is a major obstacle which is bypassed by our estimators.

The MAPS (Datta et al. 2007) jointly quantifies the angular and frequency dependence of the fluctuations in the sky signal. This can be estimated directly from the measured visibilities (e.g. Ali et al. 2008), and it can be used to detect the cosmological 21-cm signal (Ghosh et al. 2011b). In future work, we plan to generalize the analysis of this paper to the MAPS and address various issues, including point source removal, which are relevant for detecting the cosmological 21-cm signal.

ACKNOWLEDGEMENTS

SC would like to acknowledge UGC, India for providing financial support. SB would like to thank Ayesha Begum, Jayaram N. Chengalur, Prasun Dutta and Jasjeet S. Bagla for useful discussions. AG acknowledges the financial support from the European Research Council under ERC-Starting Grant FIRSTLIGHT - 258942. SSA would like to acknowledge C.T.S, I.I.T. Kharagpur for the use of its facilities and the support by DST, India, under Project No. SR/FTP/PS-088/2010. SSA would also like to thank the authorities of the IUCAA, Pune, India for providing the Visiting Associateship programme.

REFERENCES

Ali S. S., Bharadwaj S., 2014, J. Astrophys. Astron., 35, 157
Ali S. S., Bharadwaj S., Chengalur J. N., 2008, MNRAS, 385, 2166
Begum A., Chengalur J. N., Bhardwaj S., 2006, MNRAS, 372, L33
Bennett C. L. et al., 2003, ApJS, 148, 97
Bernardi G. et al., 2009, A&A, 500, 965
Bernardi G. et al., 2010, A&A, 522, A67

Bharadwaj S., Ali S. S., 2005, MNRAS, 356, 1519
Bharadwaj S., Pandey S. K., 2003, J. Astrophys. Astron., 24, 23
Bharadwaj S., Sethi S., 2001, J. Astrophys. Astron., 22, 293
Bowman J. D., Morales M. F., Hewitt J. N., 2009, ApJ, 695, 183
Chapman E. et al., 2012, MNRAS, 423, 2518
Chapman E. et al., 2013, MNRAS, 429, 165
Cho J., Lazarian A., Timbie P. T., 2012, ApJ, 749, 164
Christiansen W. N., Hogbom J. A., 1969, Radio Telescopes, 2nd edn. Cambridge Univ. Press, Cambridge
Datta K. K., Roy Choudhury T., Bharadwaj S., 2007, MNRAS, 378, 119
Datta A., Bhatnagar S., Carilli C. L., 2009, ApJ, 703, 1851
Datta A., Bowman J. D., Carilli C. L., 2010, ApJ, 724, 526
Dillon J. S. et al., 2014, Phys. Rev. D, 89, 023002
Dutta P., Begum A., Bharadwaj S., Chengalur J. N., 2009, MNRAS, 398, 887
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
Giardino G., Banday A. J., Fosalba P., Górski K. M., Jonas J. L., O'Mullane W., Tauber J., 2001, A&A, 371, 708
Giardino G., Banday A. J., Górski K. M., Bennett K., Jonas J. L., Tauber J., 2002, A&A, 387, 82
Hobson M. P., Maisinger K., 2002, MNRAS, 334, 569
Hobson M. P., Lasenby A. N., Jones M., 1995, MNRAS, 275, 863
Iacobelli M. et al., 2013, A&A, 558, A72
Jacobs D. C., Bowman J., Aguirre J. E., 2013, ApJ, 769, 5
Jelić V. et al., 2010, MNRAS, 409, 1647
La Porta L., Burigana C., Reich W., Reich P., 2008, A&A, 479, 641
Lazarian A., Pogosyan D., 2012, ApJ, 747, 5
Liu A., Tegmark M., 2012, MNRAS, 419, 3491
Liu A., Parsons A. R., Trott C. M., 2014a, Phys. Rev. D, 90, 023018
Liu A., Parsons A. R., Trott C. M., 2014b, Phys. Rev. D, 90, 023019
Mao X.-C., 2012, ApJ, 744, 29
Mellema G. et al., 2013, Exp. Astron., 36, 235
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Myers S. T. et al., 2003, ApJ, 591, 575
Paciga G. et al., 2011, MNRAS, 413, 1174
Paciga G. et al., 2013, MNRAS, 433, 639
Parsons A. R. et al., 2014, ApJ, 788, 106
Paul S. et al., 2014, ApJ, 793, 28
Petrovic N., Oh S. P., 2011, MNRAS, 413, 2103
Pober J. C. et al., 2013, ApJ, 768, L36
Shaw J. R., Sigurdson K., Pen U.-L., Stebbins A., Sitwell M., 2014a, ApJ, 781, 57
Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U.-L., 2014b, preprint (arXiv:1401.2095)
Thyagarajan N. et al., 2013, ApJ, 776, 6
Trott C. M., Wayth R. B., Macquart J.-P. R., Tingay S. J., 2011, ApJ, 731, 81
Trott C. M., Wayth R. B., Tingay S. J., 2012, ApJ, 757, 101
van Haarlem M. P. et al., 2013, A&A, 556, A2
Vedantham H., Udaya Shankar N., Subrahmanyan R., 2012, ApJ, 745, 176
Waelkens A. H., Schekochihin A. A., Enßlin T. A., 2009, MNRAS, 398, 1970
White M., Carlstrom J. E., Dragovan M., Holzapfel S. W. L., 1999, ApJ, 514, 12

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