

Establishing a relationship between the cosmological 21 cm power spectrum and interferometric closure phases

P. M. Keller^{1,*}, B. Nikolic^{2,*}, and N. Thyagarajan^{3,*}

¹ Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands

² Cavendish Astrophysics, University of Cambridge, Cambridge CB3 0HE, UK

³ Space & Astronomy, Commonwealth Scientific and Industrial Research Organisation (CSIRO), P. O. Box 1130, Bentley, WA 6102, Australia

Received 1 October 2025 / Accepted 21 November 2025

ABSTRACT

Context. Measurements of the cosmic 21 cm background need to achieve a high dynamic range to isolate it from bright foreground emissions. Instrumental calibration inaccuracies can compromise the spectral fidelity of the smooth foreground continuum, thereby limiting the dynamic range of the measurement and potentially precluding the detection of the cosmic line signal. In light of this calibration challenge, recent work has proposed using the calibration-independent closure phase to search for the spectral fluctuations of the cosmic 21 cm background signal. However, so far there has been only a heuristic understanding of the mapping between closure phases and the cosmological power spectrum of the background line signal.

Aims. This work aims to establish a more accurate mathematical relationship between closure phase measurements and the cosmological power spectrum of the background line signal.

Methods. Building on previous work, we treat the cosmic signal component as a perturbation to the closure phase and use a delay spectrum approach to estimate the power of the perturbing signal. We establish the relationship between this estimate and the cosmological power spectrum using standard Fourier transform techniques, and validate it using simulated observations from the Hydrogen Epoch of Reionization Array (HERA).

Results. We find that, statistically, the power spectrum estimate from closure phases is approximately equal to the true cosmological power spectrum convolved with a foreground-dependent window function, provided that the signal-to-foreground ratio is small. Compared with standard approaches, the foreground dependence of the window function results in an increased amount of mode-mixing and a more pronounced proliferation of foreground power along the line-of-sight dimension of the cylindrical power spectrum. These effects can be mitigated by flagging instances where the window function is broad. Crucial to gaining the necessary sensitivity, this mapping will allow us to average the measurements of closure triads of different shapes based on their imprint in cylindrical Fourier space.

Key words. methods: data analysis – techniques: interferometric – large-scale structure of Universe – dark ages, reionization, first stars

1. Introduction

The 21 cm line of neutral hydrogen (HI) is a rich probe of astrophysics and cosmology across a wide range of cosmic epochs (see reviews by Pritchard & Loeb 2012; Furlanetto 2016). At the redshifts of the dark ages ($z \gtrsim 30$), the HI 21 cm signal is expected to trace the density of baryonic matter and hence will be a pristine probe of early structure formation. Following cosmic dawn ($z \sim 30$), the thermal state of hydrogen coupled to the photons emitted by astrophysical sources, which culminated in the reionisation of the intergalactic medium (IGM) expected to have occurred at $z \sim 6 - 12$. The HI 21 cm signal from cosmic dawn and the epoch of cosmic reionisation (EoR) should be rich in information about the first generation of stars and black holes. In the post-reionisation Universe, neutral hydrogen resides predominantly in galaxies. The cosmic HI 21 cm signal at these epochs can thus be used to detect Baryon acoustic oscillations (BAO), which serve as a standard ruler for measuring the expansion history of the Universe. Such measurements are of particu-

lar interest at the redshifts where cosmic expansion started to be increasingly influenced by dark energy ($z \lesssim 3$).

Several current and future interferometric experiments aim to detect the spatial fluctuations of the cosmic HI 21 cm signal. Low-frequency experiments ($\sim 30-200$ MHz) targeting cosmic dawn and the EoR include HERA (DeBoer et al. 2017), the MWA (Tingay et al. 2013), LOFAR (van Haarlem et al. 2013), NenuFAR (Zarka et al. 2012), the OVRO-LWA (Eastwood et al. 2019), the GMRT EoR experiment (Paciga et al. 2013), 21CMA (Zheng et al. 2016) and the SKA-low (Koopmans et al. 2015). At higher frequencies ($\sim 200-1000$ MHz), CHIME (Bandura et al. 2014), HIRAX (Crichton et al. 2022), TIAN-LAI (Chen 2011), PUMA (Slosar et al. 2019), and SKA-mid (Santos et al. 2015) aim for BOA detections at $z \lesssim 3$.

The common challenge shared by these experiments are the very bright foreground emissions which overwhelm the faint cosmic HI 21 cm signal by many orders of magnitude. The line-of-sight fluctuations of the cosmic line signal offer a means by which it can, in principle, be separated from the spectrally smooth continuum foregrounds (Datta et al. 2010; Parsons et al. 2012b; Thyagarajan et al. 2013; Liu et al. 2014). This separation requires spectral calibration inaccuracies to be smaller than

* Corresponding authors: pkeller@strw.leidenuniv.nl, bn204@cam.ac.uk, Nithyanandan.Thyagarajan@csiro.au

the spectral fluctuations of the cosmic line signal. Instrumental systematics and modelling uncertainties can prevent conventional calibration methods from achieving the necessary spectral fidelity, thereby ruling out a detection of the cosmic line signal (Barry et al. 2016; Ewall-Wice et al. 2017; Byrne et al. 2019, 2021).

As an alternative to standard analysis approaches, Thyagarajan et al. (2018) propose using interferometric closure phases to detect the spectral fluctuations of the cosmic HI 21 cm signal. Closure phases are independent of antenna-based, direction-independent calibration, making them insensitive to a wide range of instrumental systematics (Thompson et al. 2017; Thyagarajan et al. 2022; Samuel et al. 2022). By evading the conventional calibration step, the closure phase approach simplifies the data processing and reduces the chance of introducing processing-based systematics.

Thyagarajan & Carilli (2020) motivate a closure phase based estimator of the power spectrum of the cosmic line signal. Variations of this estimator were used to set approximate upper limits on the EoR HI 21 cm power spectrum with data from HERA (Thyagarajan et al. 2020; Keller et al. 2023), and the MWA (Tiwari et al. 2024). The estimators used in these papers differ mainly by the heuristic scaling used to map phase fluctuations to fluctuations in brightness temperature. The lack of consensus on this scaling reveals a hitherto limited understanding of the relationship between closure phases and the cosmic HI 21 cm power spectrum. In this paper, we aim to derive a more accurate description of this relationship from which the correct scaling of the power spectrum estimator arises naturally.

The paper is structured as follows. Section 2 derives a mapping between closure phases and the cosmic power spectrum in the limit of small fluctuations. Section 3 describes the methods and simulations used to validate the power spectrum estimator. The results of this validation work are presented and discussed in Section 4. Section 5 concludes the paper.

2. Mathematical foundations

The fundamental quantity of radio interferometric measurements is the complex visibility, formed by correlating the electric signals between pairs of antennas. Under certain commonly assumed conditions, the visibility can be shown to represent a Fourier component of the orthographically projected apparent sky intensity (e.g. Thompson et al. 2017):

$$V(\mathbf{b}_\nu, \nu) = \iint_{-\infty}^{+\infty} dldm A(\hat{\mathbf{s}}, \nu) I(\hat{\mathbf{s}}, \nu) e^{-2\pi i \hat{\mathbf{s}} \cdot \mathbf{b}_\nu} \quad (1)$$

$$= \tilde{A}(\mathbf{b}_\nu, \nu) * \tilde{I}(\mathbf{b}_\nu, \nu) \quad (2)$$

where ν is the frequency, \mathbf{b}_ν is the antenna baseline vector in units of wavelengths, $\hat{\mathbf{s}} = (l, m)$ are the orthographic sky coordinates, $I(\hat{\mathbf{s}}, \nu)$ is the specific intensity, and $A(\hat{\mathbf{s}}, \nu)$ describes the antenna primary beam and other direction-dependent effects, if applicable. The second line uses the convolution theorem, where the convolution is denoted by an asterisk and the tildes indicate Fourier-transformed quantities. At the frequencies of interest to 21 cm cosmology, the ionosphere leaves a strong directional and time-dependent imprint on the apparent sky. However, at baselines shorter than a few kilometres all antennas essentially see the same part of the ionosphere, so that the ionospheric phase shifts cancel in the signal correlation process. In this work, we consider only baselines shorter than 200 m for which these

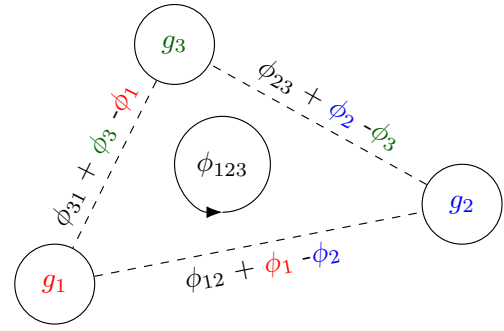


Fig. 1. Illustration of the closure phase concept. The true visibility phases, ϕ_{pq} , are corrupted by direction-independent antenna-based gains, ϕ_p . Summing the corrupted phases in a loop eliminates the gain phases. The resulting quantity, ϕ_{123} , is the closure phase.

assumptions are reasonable (de Gasperin et al. 2018). For notational convenience, we omit measurement noise throughout this work.

2.1. Closure phase

Actual interferometric measurements are corrupted by various effects along the signal propagation path. The dominant corruptions can usually be represented as antenna-based, direction-independent complex gain factors. The visibility, V_{pq} , measured by antennas p and q , can thus be written as

$$V_{pq} = g_p g_q^* \widehat{V}_{pq} = |g_p| |g_q| \left| \widehat{V}_{pq} \right| e^{i(\phi_{pq} + \phi_p - \phi_q)}, \quad (3)$$

where \widehat{V}_{pq} is the uncorrupted visibility, $\phi_{pq} = \arg(\widehat{V}_{pq})$ is the sky-based phase term, g_p are the gains and $\phi_p = \arg(g_p)$ their respective phases, and the asterisk denotes complex conjugation. The frequency and time dependence of the gains and the visibilities are left implicit in this equation.

Consider a triad formed by antennas m , p and q . The closure phase is obtained by summing the visibility phases measured on this triad in a closed loop:

$$\begin{aligned} \phi_{mpq} &= \arg(V_{mp}) + \arg(V_{pq}) + \arg(V_{qm}) \\ &= \arg(\widehat{V}_{mp}) + \arg(\widehat{V}_{pq}) + \arg(\widehat{V}_{qm}) \\ &= \phi_{mp} + \phi_{pq} + \phi_{qm}, \end{aligned} \quad (4)$$

where the gain phases cancel out in the second line. The closure phase is therefore free of antenna-based, direction-independent corruptions and depends only on sky-based terms (Jennison 1958; Thyagarajan & Carilli 2022).

2.2. Perturbation theory

The effect of the cosmic line signal on the closure phase is best described using perturbation theory. Thyagarajan & Carilli (2020) provide an extensive discussion on closure phase perturbations. For completeness, we repeat the main points relevant to this paper.

The uncorrupted visibilities are typically dominated by bright continuum foreground sources. Hence, we can treat the weak cosmic signal as a perturbation to the uncorrupted foreground signal. The measured visibility can then be written as

$$V_j = V_j^F + \delta V_j, \quad (5)$$

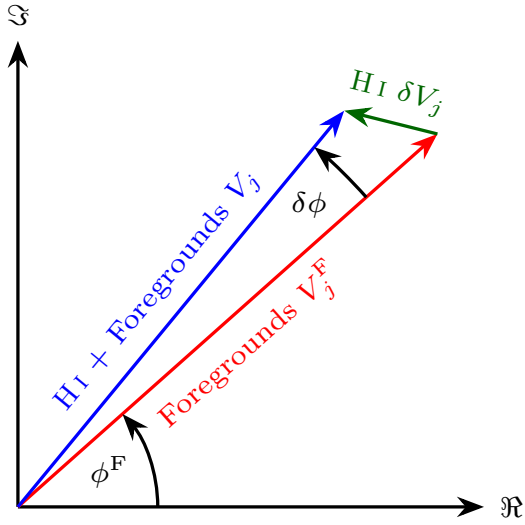


Fig. 2. Exaggerated illustration of the visibility phase perturbations due to the weak cosmic signal δV_j (green arrow). The horizontal and vertical axes represent the real and complex axes of the complex plane, respectively. The red arrow represents the foreground visibility, V_j^F , and the blue arrow represents the sum of the foreground and cosmic H I 21 cm visibilities, $V_j = V_j^F + \delta V_j$. The resulting phase perturbation is $\delta\phi$.

where the subscript j indexes the baseline, the superscript F denotes the foregrounds and δV_j is the perturbation such that $|\delta V_j/V_j^F| \ll 1$ (see Figure 2). The change from indexing quantities by antenna pairs to indexing them by their baseline is notationally more convenient for this section.

Remembering that $\phi_j = \tan^{-1}(\Im\{V_j\}/\Re\{V_j\})$ and expanding to first order in δV_j , one finds

$$\phi_j \approx \phi_j^F + \Im \left\{ \frac{\delta V_j}{V_j^F} \right\}, \quad (6)$$

where ϕ_j^F is the unperturbed visibility phase. The perturbed closure phase ϕ_ν is then simply the sum of three perturbed visibility phases:

$$\phi_\nu \approx \phi_\nu^F + \sum_{j=1}^3 \Im \left\{ \frac{\delta V_j}{V_j^F} \right\} = \phi_\nu^F + \frac{1}{2i} \sum_{j=1}^3 \left[\frac{\delta V_j}{V_j^F} - \frac{\delta V_j^*}{V_j^{F*}} \right], \quad (7)$$

where ϕ_ν^F is the unperturbed closure phase and the summed-over baselines form a closed loop in real space. The right side follows from the complex identity $\Im\{z\} = (z - z^*)/2i$. Thyagarajan & Carilli (2020) find that the fractional inaccuracy of this first-order approximation is smaller than $\sim 10^{-4}$ for the expected strength of the EoR H I 21 cm signal.

The equations above establish that the strength of phase perturbations is proportional to the complex ratio of the perturbing visibilities to the foreground visibilities. Since it is the imaginary parts of these ratios that are relevant, only the perturbing components that are perpendicular to the foreground visibilities in the complex plane induce phase perturbations. Thyagarajan & Carilli (2020) therefore interpret closure phase perturbations as a measure of the relative dissimilarity between the cosmic signal and the foregrounds. Furthermore, they point out that, statistically, only half of the fluctuations of the cosmic signal are recovered in phase, with the other half contained in the

amplitudes. As we work towards an estimator of the cosmological power spectrum, this factor of two appears naturally in the power spectrum scaling.

2.3. Closure phase delay spectrum

The delay spectrum approach is a commonly used method of 21 cm cosmology data analysis, which allows one to access the line-of-sight fluctuations of the cosmic line signal while isolating the smooth continuum emissions of bright foregrounds (Datta et al. 2010; Parsons et al. 2012b). As demonstrated by Thyagarajan et al. (2018), a similar approach can be used to recover the cosmic signal from the closure phase. In analogy to the visibility delay spectrum, Thyagarajan et al. (2018) define the closure phase delay spectrum at a given redshift, z , as a windowed Fourier transform across frequency:

$$\tilde{\Psi}_\nu(\tau; \nu_0) = \int_{-\infty}^{+\infty} d\nu W(\nu - \nu_0) e^{i\phi_\nu(\nu)} e^{2\pi i \tau \nu}, \quad (8)$$

where $W(\nu)$ is a spectral tapering function and $\nu_0 = \nu_{\text{H I}}/(1+z)$ is the centre frequency of the power spectrum estimation. The delay transform in this equation is performed on the complex exponential of the closure phase rather than on the closure phase directly to avoid discontinuities arising from the phase branch cuts, i.e. at $\pm\pi$ if the phase is defined on $(-\pi, \pi]$. In the presence of a small closure phase perturbation, $\delta\phi_\nu$, we can approximate the complex exponential of the closure phase to first order as

$$e^{i(\phi_\nu^F(\nu) + \delta\phi_\nu(\nu))} \approx e^{i\phi_\nu^F(\nu)} (1 + i\delta\phi_\nu(\nu)). \quad (9)$$

Inserting this along with Equation (7) into the expression for the delay spectrum (Equation (8)) and using the convolution theorem, we find (cf. Thyagarajan & Carilli 2020)

$$\tilde{\Psi}_\nu(\tau) = \tilde{\Psi}_\nu^F(\tau) + \frac{1}{2} \sum_{j=1}^3 (\tilde{\Xi}_j(\tau) * \delta\tilde{V}_j(\tau) - \tilde{\Xi}_j^*(-\tau) * \delta\tilde{V}_j^*(-\tau)), \quad (10)$$

where

$$\tilde{\Psi}_\nu^F(\tau) = \int_{-\infty}^{+\infty} d\nu W(\nu - \nu_0) e^{i\phi_\nu^F(\nu)} e^{2\pi i \tau \nu}, \quad (11)$$

$$\tilde{\Xi}_j(\tau) = \int_{-\infty}^{+\infty} d\nu \frac{W(\nu - \nu_0)}{V_j^F(\nu)} e^{i\phi_\nu^F(\nu)} e^{2\pi i \tau \nu}, \quad (12)$$

$$\delta\tilde{V}_j(\tau) = \int_{\nu_{\min}}^{\nu_{\max}} d\nu \delta V_j(\nu) e^{2\pi i \tau \nu} \quad (13)$$

describe a set of Fourier transforms of the various frequency dependent terms in Equations (7)–(9). In Equation (10), the superscript asterisk denotes complex conjugation and the regular asterisk denotes a convolution. In the last equation, the integration limits define the frequency range, (ν_{\min}, ν_{\max}) , in which $W(\nu - \nu_0)$ is non-zero.

2.4. Estimating the cosmological power spectrum

The simplest statistic to describe the fluctuations of the cosmological background signal is the power spectrum, $P(\mathbf{k})$, defined

through the relation

$$\langle \widetilde{\delta I}(\mathbf{k}) \widetilde{\delta I}^*(\mathbf{k}') \rangle = \left(\frac{2k_B}{\lambda^2} \right)^2 (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') P(\mathbf{k}), \quad (14)$$

where the angular brackets denote an ensemble average, \mathbf{k} and \mathbf{k}' are comoving wave numbers, $\widetilde{\delta I}(\mathbf{k})$ is the Fourier transform of the specific intensity of the cosmic background signal, k_B is Boltzmann's Constant, $\lambda = c/\nu$ is the wavelength, and $\delta(\mathbf{k})$ is a Dirac delta function. The first bracket on the right side of Equation (14) is the Rayleigh-Jeans factor, which converts specific intensities to brightness temperatures.

The following paragraphs derive a relationship between the closure phase delay spectrum and the cosmological power spectrum for arbitrary triad shapes. This can be achieved by computing the ensemble average¹ of the squared closure phase delay spectrum. In doing so, there are various cross-correlations among the terms in Equation (10). Assuming that the cosmic line signal is statistically isotropic and uncorrelated with the foregrounds, the only non-zero correlations are of the form

$$\begin{aligned} C(\mathbf{b}, \mathbf{b}', \tau, \tau') &= \langle (\widetilde{\Xi}_j(\tau) * \widetilde{\delta V}_j(\tau)) (\widetilde{\Xi}_j^*(\tau') * \widetilde{\delta V}_j^*(\tau')) \rangle \\ &= \iint_{-\infty}^{+\infty} d\eta d\eta' \langle \widetilde{\Xi}_j(\tau - \eta) \widetilde{\Xi}_j^*(\tau' - \eta') \rangle \langle \widetilde{\delta V}_j(\eta) \widetilde{\delta V}_j^*(\eta') \rangle. \end{aligned} \quad (15)$$

The second ensemble average in the integrand, i.e. the correlation of the visibility perturbations, can be approximated as

$$\begin{aligned} \langle \widetilde{\delta V}_j(\tau) \widetilde{\delta V}_j^*(\tau') \rangle &\approx \iint_{-\infty}^{+\infty} d^2u d^2u' \widetilde{A}(\mathbf{b} - \mathbf{u}) \widetilde{A}^*(\mathbf{b}' - \mathbf{u}') \\ &\times \langle \widetilde{\delta I}(\mathbf{u}, \tau) \widetilde{\delta I}^*(\mathbf{u}', \tau') \rangle, \end{aligned} \quad (16)$$

where we used Equation (2), and implicitly evaluated the effective baseline vector and the antenna beams at ν_0 by making the assumption that they are constant across the bandwidth probed by $W(\nu - \nu_0)$. This is usually a good assumption, as the bandwidth is typically chosen to be a few per cent of the centre frequency to limit the cosmic evolution within the band (see [Datta et al. 2014](#); [Liu et al. 2014](#)).

The observational quantities τ and \mathbf{b} can be mapped to the cosmological line-of-sight and transverse wave numbers at redshift z via the relations $k_{\parallel} = 2\pi\tau/X$ and $\mathbf{k}_{\perp} = 2\pi\mathbf{b}_{\perp}/Y$, where (e.g. [Morales & Hewitt 2004](#); [Parsons et al. 2012b](#); [Liu et al. 2014](#))

$$X = \frac{c(1+z)^2}{\nu_{\text{HI}} H(z)} \quad \text{and} \quad Y = D_c(z), \quad (17)$$

respectively, $H(z) = H_0 \sqrt{\Omega_M(1+z)^3 + \Omega_K(1+z)^2 + \Omega_\Lambda}$ is the Hubble parameter expressed in standard terms of cosmology, ν_{HI} is the rest frequency of the HI 21 cm line, c is the speed of light, and

$$D_c(z) = \int_0^z dz \frac{c}{H(z)} \quad (18)$$

is the comoving line-of-sight distance to the point of the HI 21 cm emission at the frequency ν_0 ([Hogg 1999](#)).

¹ The ensemble average is in practice replaced with an average over different realisations of the sky signal, i.e. over observations of independent portions of the sky, and different triad shapes. The resulting estimate has an associated uncertainty due to cosmic variance.

Using the above relations together with the definition of the power spectrum (Equation (14)), we can express Equation (16) as²

$$\begin{aligned} \langle \widetilde{\delta V}_j(\tau) \widetilde{\delta V}_j^*(\tau') \rangle &\approx \left(\frac{2k_B}{\lambda^2} \right)^2 \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} d^2k_{\perp} \delta(k_{\parallel} - k'_{\parallel}) P(\mathbf{k}_{\perp}, k_{\parallel}) \\ &\times \widetilde{A}\left(\mathbf{b} - \frac{Y\mathbf{k}_{\perp}}{2\pi}\right) \widetilde{A}^*\left(\mathbf{b}' - \frac{Y\mathbf{k}_{\perp}}{2\pi}\right). \end{aligned} \quad (19)$$

Inserting this into Equation (15) gives

$$\begin{aligned} C(\mathbf{b}, \mathbf{b}', \tau, \tau') &= \left(\frac{2k_B}{\lambda^2} \right)^2 \frac{1}{(2\pi)^3} \iint_{-\infty}^{+\infty} d^2k_{\perp} dk_{\parallel} P(\mathbf{k}_{\perp}, k_{\parallel}) \\ &\times \left\langle \widetilde{\Xi}_j\left(\tau - \frac{Xk_{\parallel}}{2\pi}\right) \widetilde{\Xi}_j^*\left(\tau' - \frac{Xk_{\parallel}}{2\pi}\right) \right\rangle \\ &\times \widetilde{A}\left(\mathbf{b} - \frac{Y\mathbf{k}_{\perp}}{2\pi}\right) \widetilde{A}^*\left(\mathbf{b}' - \frac{Y\mathbf{k}_{\perp}}{2\pi}\right). \end{aligned} \quad (20)$$

The second line of this equation is close to zero if $|\tau - \tau'| \Delta\nu > 1$, where $\Delta\nu$ is the width of the effective spectral window function, $W(\nu - \nu_0)/V_j^F(\nu)$. Here, we are primarily interested in the case where $\tau' = -\tau$ (cf. Equation (10)) for which $|C(\mathbf{b}, \mathbf{b}', \tau, -\tau)| \approx 0$ at τ -modes that are not contaminated by foreground power. Similarly, the correlation scale of the terms in the third line of Equation (20) depends mainly on the width of the primary beam and the difference between the baseline vectors, $\Delta b = |\mathbf{b} - \mathbf{b}'|c/\nu$. The primary beam width is $\Delta\theta \approx c/(D\nu)$, where D is the diameter of an array element, and therefore the width of $\widetilde{A}(\mathbf{b})$ is approximately $\Delta\theta^{-1} \approx D\nu/c$. The correlation in Equation (15) is therefore small if $\Delta b/D > 1$ ([Bharadwaj & Sethi 2001](#)) and in the case of nonidentical baselines it is only relevant for the shortest baselines of very compact arrays such as HERA (to be discussed in Section 4.4).

The following discussion uses the approximation $|C(\mathbf{b}, \mathbf{b}', \tau, \tau)| \approx 0$ for $\mathbf{b} \neq \mathbf{b}'$, allowing us to write the closure phase delay power spectrum of the perturbing signal as

$$\begin{aligned} \langle |\widetilde{\Psi}_{\nabla}^H(\tau)|^2 \rangle &= \frac{1}{4} \sum_{j=1}^3 q_j [C(\mathbf{b}_j, \mathbf{b}_j, \tau, \tau) + C(\mathbf{b}_j, \mathbf{b}_j, -\tau, -\tau)] \\ &= \frac{1}{2} \left(\frac{2k_B}{\lambda^2} \right)^2 \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} d^2k_{\perp} dk_{\parallel} W_{\nabla}(\mathbf{k}_{\perp}, k_{\parallel}; \tau) P(k), \end{aligned}$$

where

$$W_{\nabla}(\mathbf{k}_{\perp}, k_{\parallel}; \tau) = \sum_{j=1}^3 q_j \left\langle \left| \widetilde{\Xi}_j\left(\tau - \frac{Xk_{\parallel}}{2\pi}\right) \widetilde{A}\left(\mathbf{b}_j - \frac{Y\mathbf{k}_{\perp}}{2\pi}\right) \right|^2 \right\rangle \quad (21)$$

is the window function, $\widetilde{\Psi}_{\nabla}^H(\tau) = \widetilde{\Psi}_{\nabla}(\tau) - \widetilde{\Psi}_{\nabla}^F(\tau)$ is the closure phase delay spectrum of the perturbing signal, $q_j = 2$ if another baseline vector within the triad is identical to the baseline vector j (as is possible for linear triads) and $q_j = 1$ otherwise, $k = (k_{\parallel}^2 + k_{\perp}^2)^{1/2}$ is the magnitude of the comoving wavenumber and the second equality uses the fact that $P(k)$ is invariant under the orientation of $(\mathbf{k}_{\perp}, k_{\parallel})$ due to the isotropy of the cosmic signal. This expression shows that the closure phase delay

² The $(2\pi)^3$ arises from the fact that $\widetilde{\delta I}(\mathbf{k})$ and $\widetilde{\delta I}(\mathbf{b}, \tau)$ use different Fourier conventions (see [Liu & Shaw 2020](#)).

power spectrum is proportional to an average of $P(k)$ weighted by the window function. In particular, each baseline of a closure triad maps to its own \mathbf{k}_\perp -mode, as is represented by the sum over baselines in the expression of the window function. This means that there is no exact one-to-one mapping between non-equilateral closure triads and k_\perp -space. The topography of the window function is determined further by the bandwidth, the spectral tapering, and the antenna primary beam. As a result, power spectrum estimates are always affected by some amount of ‘mode-mixing’, even if they are derived from equilateral triad measurements or from visibilities. If we disregard the more subtle effects of mode-mixing, we can indeed associate an equilateral triad with a unique $|k_\perp|$ -mode since there is only one baseline length involved. For compact arrays such as HERA, the finite bandwidth often dominates the mode-mixing (to be discussed in Section 2.6), even when a triad comprises baselines of different lengths. Moreover, the window function is often sufficiently localised around some value of k for $P(k)$ to be approximated as constant across the window. In such cases, we can take $P(k)$ out of the integral and estimate the cosmological power spectrum as

$$\widehat{P}_\nu(\tau) = 2 \left(\frac{\lambda^2}{2k_B} \right)^2 \left(\frac{1}{\Omega_\nu} \right) |\widehat{\Psi}_\nu^H(\tau)|^2. \quad (22)$$

The quantity Ω_ν is the volume of the window function and can be written as

$$\begin{aligned} \Omega_\nu &= \int_{-\infty}^{+\infty} \frac{d^2 k_\perp dk_\parallel}{(2\pi)^3} \sum_{j=1}^3 q_j \left| \widetilde{\Xi}_j \left(\tau - \frac{Xk_\parallel}{2\pi} \right) \widetilde{A} \left(\mathbf{b}_j - \frac{Y\mathbf{k}_\perp}{2\pi} \right) \right|^2 \\ &= \frac{1}{XY^2} \int_{-\infty}^{+\infty} dldm |A(\hat{\mathbf{s}})|^2 \sum_{j=1}^3 q_j \int_{-\infty}^{+\infty} dv' \left| \frac{W(v')}{V_j^F(v')} \right|^2, \end{aligned}$$

where on the second line we split the integral into its transverse and line-of-sight components, changed the integration variables, and used Plancherel’s theorem. The expression for the window function motivates the definitions of the beam solid angle, Ω_{pp} , the effective bandwidth, B_{eff} , and the effective visibility, V_{eff} , as

$$\Omega_{pp} = \int_{-\infty}^{+\infty} dldm |A(\hat{\mathbf{s}}, \nu)|^2, \quad (23)$$

$$B_{\text{eff}} = \int_{-\infty}^{+\infty} dv |W(\nu)|^2, \quad \text{and} \quad (24)$$

$$V_{\text{eff}}^{-2} = \sum_{j=1}^3 q_j (\widehat{V}_j^F)^{-2}, \quad (25)$$

respectively, where

$$(\widehat{V}_j^F)^{-2} = \frac{1}{B_{\text{eff}}} \int_{-\infty}^{+\infty} dv \left| \frac{W(\nu)}{V_j^F(\nu)} \right|^2. \quad (26)$$

Using these expressions and writing out the scaling parameters X and Y , we can write Equation (22) as

$$\widehat{P}_\nu(\tau) = 2 \left(\frac{\lambda^2}{2k_B} \right)^2 \left(\frac{c(1+z)^2 D_c^2}{H(z)\nu_{\text{HI}}} \right) \left(\frac{V_{\text{eff}}^2}{\Omega_{pp} B_{\text{eff}}} \right) |\widehat{\Psi}_\nu^H(\tau)|^2. \quad (27)$$

The normalisation in the above expression differs from that of the standard visibility delay power spectrum in that it contains an additional factor of two due to the fact that only half

of the HI fluctuations are recovered in the visibility phases (see Table 1)³. The effective visibility can be thought of as that part of Ω_ν which differs from the conventional scaling of the visibility-based power spectrum estimate. While providing the estimated power spectrum with the correct units, V_{eff} also counteracts the inverse proportionality of the closure phase perturbations to the foreground visibility (see Equation (7)). Since V_{eff} is obtained by averaging $V_j^F(\nu)$ in inverse quadrature, it tends towards the lowest visibility amplitude within the closure triad, where perturbations to the visibility phase are the strongest (see Equation (6) and Figure 3). In practice, $V_j^F(\nu)$, and hence V_{eff} , can be obtained from calibrated or modelled data. The requirements on the fractional accuracy of the calibration solutions or the model are not as high as in standard approaches because the sole purpose of estimating $V_j^F(\nu)$ is to obtain the power spectrum scaling and possibly to compute window functions, but not to correct the data in any other way that could impart spectral structure on the data (cf. Thyagarajan & Carilli 2020).

2.5. Comparison with previous work

The effective visibility was first introduced by Thyagarajan & Carilli (2020) as a heuristic amplitude scaling to gauge the cosmic HI signal fluctuations recovered in the closure phase to those in visibilities. However, they estimated \widehat{V}_j^F as a windowed average of $V_j^F(\nu)$ over frequency (see Table 1). In contrast, our derivation shows that \widehat{V}_j^F should be obtained by averaging $|V_j^F(\nu)|$ in inverse quadrature across the frequency band (Equation (26)). This form of V_{eff} arises naturally from the normalisation of the effective delay window (Equation (12)) of the closure phase delay spectrum. The difference between the two definitions of V_{eff} should be subtle if $V_j^F(\nu)$ varies slowly and smoothly with frequency, but could otherwise lead to substantial biases in the estimated power spectrum. The estimator presented here is free of such biases, provided that the assumptions of the derivation are met, i.e. $V_j^F(\nu) \gg V_j^H(\nu)$, and the correlations $C(\mathbf{b}_j, \mathbf{b}_{j'}, \tau, \tau')$ are small between non-identical baselines in a closure triad.

2.6. Mode-mixing

In deriving $\widehat{P}_\nu(\tau)$, we assumed that the true cosmological power spectrum is approximately constant within the region probed by the window function. Generally, this is not necessarily the case, but the scaling of $\widehat{P}_\nu(\tau)$ nevertheless accounts for the correct normalisation of the window function. The window function provides information on how different power spectrum modes contribute to $\widehat{P}_\nu(\tau)$ and is therefore essential for the correct interpretation of the estimated power spectrum.

The mode-mixing in a closure phase based analysis differs from that in a visibility-based analysis in two particular ways (see Table 1). Firstly, the line-of-sight window function depends on the inverse foreground visibilities and the foreground closure phase through the term $\Xi_j(\nu) = \exp\{i\phi_\nu^F(\nu)\}/V_j^F(\nu)$ (cf.

³ Thyagarajan et al. (2020) and Keller et al. (2023) include an additional factor of 1/3 in the power spectrum scaling to account for the contribution of three visibility phases to the closure phase. This would be justified if V_{eff} was defined by an average rather than a sum of \widehat{V}_j^F over baselines in inverse quadrature. Since this is not how Thyagarajan et al. (2020) define V_{eff} , it appears that the additional factor of 1/3 stems from a bookkeeping mistake.

Table 1. Comparison of terms used in different delay spectrum approaches.

Term	Visibility	Closure phase prev.	This work
$\widehat{P}(\tau)$	$\propto \widetilde{V}^H(\tau) ^2$	$\propto \frac{2}{3} V_{\text{eff}} \widetilde{\Psi}_\nabla^H(\tau) ^2$	$\propto 2 V_{\text{eff}} \widetilde{\Psi}_\nabla^H(\tau) ^2$
$(\widehat{V}_j^F)^{-2}$	–	$\left(\frac{\int d\nu W(\nu) V_j^F(\nu)}{\int d\nu W(\nu)} \right)^{-2}$	$\frac{1}{B_{\text{eff}}} \int d\nu \left \frac{W(\nu)}{V_j^F(\nu)} \right ^2$
$W(\mathbf{k}_\perp, k_\parallel; \tau)$	$\left\langle \left \widetilde{W} \left(\tau - \frac{Xk_\parallel}{2\pi} \right) \widetilde{A} \left(\mathbf{b} - \frac{Y\mathbf{k}_\perp}{2\pi} \right) \right ^2 \right\rangle$	$\sum_{j=1}^3 q_j \left\langle \left \widetilde{\Xi}_j \left(\tau - \frac{Xk_\parallel}{2\pi} \right) \widetilde{A} \left(\mathbf{b}_j - \frac{Y\mathbf{k}_\perp}{2\pi} \right) \right ^2 \right\rangle$	$\sum_{j=1}^3 q_j \left\langle \left \widetilde{\Xi}_j \left(\tau - \frac{Xk_\parallel}{2\pi} \right) \widetilde{A} \left(\mathbf{b}_j - \frac{Y\mathbf{k}_\perp}{2\pi} \right) \right ^2 \right\rangle$

Notes. This table compares terms used in the standard delay spectrum approach (first column), the closure phase approach according to Thyagarajan & Carilli (2020) (second column) and that according to this work (third column). From top to bottom, the rows show the terms used to estimate the power spectrum, the visibility-like term used to compute the effective visibility (cf. Equation (25); relevant only to the closure phase approaches), and the cylindrical window function. For the power spectrum terms, we have omitted parts that are common to all three approaches and use $\widetilde{V}^H(\tau)$ to denote the standard visibility delay spectrum of the cosmic HI signal. In the zeroth (‘Term’) column, we have omitted the ∇ subscript to make the terms representative of both the visibility and the closure phase approaches. See the text for the definitions of the other symbols used in this table.

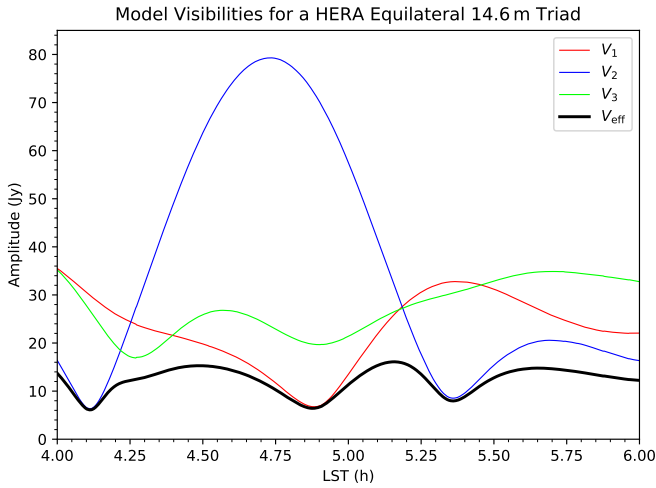


Fig. 3. Model visibility amplitudes for a HERA equilateral 14.6 m triad (red, blue and green) as a function of local sidereal time (LST). The thick black line represents the effective visibility, V_{eff} , which provides the scaling of the closure phase delay power spectrum. This example neglects the frequency dependence of the visibility amplitudes.

Equation (12)). This additional spectral component broadens the line-of-sight window function, resulting in a stronger mixing of neighbouring k_\parallel -modes. Secondly, $\widehat{P}_\nabla(\tau)$ receives contributions from up to three distinct baselines, which results in additional k_\perp mode-mixing. In order to associate a given triad and delay mode with a unique k -value, we propose computing a representative k_\perp by defining an effective baseline length in place of \mathbf{b}_ν (cf. Equation (25)):

$$b_{\text{eff}} = V_{\text{eff}}^2 \sum_{j=1}^3 b_j q_j (\widehat{V}_j^F)^{-2}. \quad (28)$$

This definition is motivated by the fact that the contribution of each baseline to the closure phase delay spectrum is inversely proportional to the foreground visibility amplitude (see Equation (6)). For HERA baselines, this approach is sufficient, since we are generally in the regime where $k_\perp \ll k_\parallel$ and hence $k \sim k_\parallel$. Henceforth, we use b_{eff} to compute unique k -values for any given triad and delay mode, as this allows us to plot closure phase power spectra as a function of k and make direct comparisons with standard power spectra. Note that $\widehat{P}_\nabla(k)$ does not corre-

spond exactly one-to-one to the cosmological power spectrum, but may approximate it to high accuracy if the window function is sufficiently localised. We investigate the accuracy of using $\widehat{P}_\nabla(k)$ as an estimator of the cosmological power spectrum in Section 4.

Finally, it is instructive to determine observational parameters at which the mixing of k_\parallel -modes and that of k_\perp -modes take on an equal extent. Consider a typical bandwidth of 10 MHz and a central frequency of 165 MHz ($z = 7.9$) as used in previous HERA studies (HERA Collaboration 2022, 2023; Thyagarajan et al. 2020; Keller et al. 2023). The resolution of k_\parallel is then at most $0.056 h \text{ Mpc}^{-1}$ (see Equation (17)), but may be coarser depending on the spectral tapering used and the spectral properties of the foregrounds. Taking the same number for the resolution of k_\perp and translating it to a difference in baseline lengths, we find a value of ~ 100 m. In this example, bandwidth is the dominant source of mode-mixing for baseline length differences smaller than 100 m. For longer baseline length differences k_\perp -mixing starts to dominate.

2.7. Foregrounds

Astrophysical foregrounds have smooth spectra at frequencies of interest to cosmological HI 21 cm studies. Their imprint on the delay spectrum is therefore expected to be confined to low delay modes. However, the chromatic response of an interferometer makes the foreground continuum appear unsmooth, which broadens the foreground dominated region in delay space. For flat-spectrum foregrounds and in the absence of instrumental systematic effects, foreground power in the standard visibility delay spectrum is bounded by the horizon limit, defined as $\tau_{h,j} = |\mathbf{b}_j|/c$ (Parsons et al. 2012b; Liu et al. 2014). The foreground contaminated region of a delay spectrum is termed the ‘wedge’ due to its wedge-shaped appearance in cylindrical (k_\perp, k_\parallel)-space (Datta et al. 2010).

Thyagarajan & Carilli (2020) find two additional effects that broaden the foreground component in the closure phase delay spectrum. Firstly, the foreground phase term, $\exp(i\phi_\nabla^F)$, is a triple product of three individual visibility phase terms, resulting in a triple convolution in delay space. This effect proliferates foreground power up to $\tau_{\nabla,h} = \sum_j |\mathbf{b}_j|/c$. Secondly, higher-order terms in the closure phase approximation cause foreground power to leak up to arbitrarily high delay modes. These higher order terms are typically small enough to be neglected (Thyagarajan & Carilli 2020). In this work, we assume delay modes $|\tau| > \tau_{\nabla,h} + \tau_b$ to be uncontaminated by foregrounds,

where $\tau_b = 0.39 \mu\text{s}$ is a buffer to account for other spectral effects such as foreground and beam chromaticity (see Section 2.8), and the effects of computing the delay spectrum over a finite bandwidth. This value ensures a dynamic range of 10^{10} outside the buffer region and was empirically determined by investigating the delay window functions obtained from simulated data (to be discussed in Section 3). The value of τ_b is ultimately dependent on the effective bandwidth used to compute the delay spectrum (cf. Thyagarajan et al. 2016).

2.8. Beam chromaticity

So far, we have assumed the frequency dependence of the primary beam to be negligible (cf. Equation (16)). This is a commonly made assumption and usually approximates the primary beam well enough for a simple power spectrum estimation. However, in principle, there is no need to make this approximation (see Thyagarajan et al. 2016). This leads to a modification of $W_{\nabla}(\mathbf{k}_{\perp}, k_{\parallel}; \tau)$, where the delay transformed primary beam is convolved with $\tilde{\Xi}_j(\tau)$. The volume of the new window function is then

$$\Omega'_{\nabla} = \frac{1}{XY^2} \sum_{j=1}^3 q_j \int_{-\infty}^{+\infty} dl dm dv' \left| A(\hat{\mathbf{s}}, \nu) \frac{W(\nu')}{V_j^F(\nu')} \right|^2. \quad (29)$$

In this expression, the terms due to the beam, the bandwidth, and the foregrounds can no longer be separated into a product of Ω_{pp}^{-1} , B_{eff}^{-1} , and V_{eff}^2 . In this work, we use the achromatic approximation, as we find that the relative difference between Ω'_{∇} and Ω_{∇} is only $\sim 10^{-3}$, which is at least an order of magnitude below our targeted accuracy. Moreover, the approximation is computationally more efficient and better aligned with the previous closure phase studies.

It is also worth noting that the frequency-dependence of the antenna beams can add spectral structure to the otherwise smooth foregrounds, especially when they lie in the sidelobes of the primary beam. This effect has been investigated by Thyagarajan et al. (2016) for the standard visibility approach and by Charles et al. (2022) for the closure phase approach. Both studies find that beam chromaticity induces foreground leakage beyond the wedge and report values of $0.2\text{--}0.3 h \text{Mpc}^{-1}$, below which the leakage may preclude the detection of the cosmological signal.

3. Validation

To validate our estimator, we simulated visibilities for 7 HERA antennas using the `matvis`⁴ simulator module through `hera_sim`⁵ (Kittiwisit et al. 2025). The antennas were chosen to include a variety of triad shapes, such as equilateral, linear, and scalene triads. The baseline lengths range from 14.6 m to 116.8 m, allowing us to investigate the mixing of different transverse scales in triads with extreme baseline length ratios. In total, there are 15 triads, which form a complete and independent set of closure phases (cf. Thompson et al. 2017). Here, we used only triads that include antenna 0 (see Figure 4).

3.1. Sky models

Our simulations used sky models consisting of foreground sources and a cosmic line signal, weighted by a frequency-

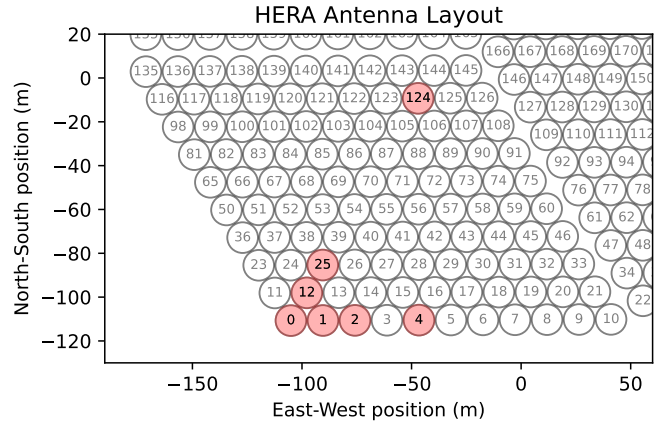


Fig. 4. HERA antennas used for validation (red). We only used triads that include antenna 0 to compute a complete and independent set of closure phases.

dependent model of the HERA primary beam (Fagnoni et al. 2021). For the foreground model, we selected point sources from the GLEAM catalogue (Hurley-Walker et al. 2017) with flux densities greater than 50 mJy at 200 MHz. For the cosmic line signal, we used two different models:

- Model 1: Gaussian random field with ionised regions. This is a toy model of the EoR that starts with a Gaussian random temperature field, $T(l, m)$, with a power-law spectrum $\sim k^{-\alpha/2}$, mean δT_0 , and standard deviation σ_T . We then add 'ionised regions' by setting the temperature field to zero where $T(l, m) > T_c$. Here, we chose $\alpha = 2$, $\sigma_T = 3 \text{ mK}$ and $T_c = \delta T_0 = 10 \text{ mK}$, which results in a temperature field that is approximately half ionised. The advantage of this simple toy model is that it allows us to efficiently generate a large number of independent samples. We computed 300 independent realisations of this model spanning 4.8 Gpc and 3072 cells along each transverse axis, thus covering around 30 square degrees of the sky at the redshifts of the EoR.
- Model 2: 21cmFast light cones. This is a more realistic representation of the HI 21cm reionisation signal generated with the 21cmFast⁶ simulator (Mesinger et al. 2011; Murray et al. 2020). We ran 21cmFast with an ionising efficiency $\zeta = 20$, a minimum virial temperature of star forming haloes $T_{\text{vir}} = 10^4 \text{ K}$, and left the other astrophysical parameters at their default setting. In our simulations, we did not account for spin temperature fluctuations. Due to the computational cost of these simulations, we generated only 12 light cones with volumes of 1.6^3 Gpc^3 and 1024^3 cells. At EoR redshifts these light cones cover only around 10 square degrees of the sky. This is not enough to fill the primary beam of a HERA antenna (FWHM $\sim 10^\circ$; Fagnoni et al. 2021). For this reason, we followed the same procedure as in Thyagarajan et al. (2020) and Keller et al. (2023) and tiled the light cones three times along each transverse axis. Each tile is a reflected version of its neighbouring tiles to avoid discontinuities at the tiling boundaries.

Following the same procedure as in Thyagarajan et al. (2020) and Keller et al. (2023), we smoothed and downsampled the above models to an angular resolution of $\sim 7'$, resulting in 256

⁴ <https://github.com/HERA-Team/matvis>

⁵ <https://hera-sim.readthedocs.io/en/latest/>

⁶ <https://21cmfast.readthedocs.io/en/latest/>

pixels per transverse axis. This downsampling is appropriate for the $\sim 0.5^\circ$ FWHM of the HERA synthesised beam and reduces the computational cost of the subsequent visibility computation.

3.2. Power spectrum computation

We simulated visibilities for the foregrounds visible to HERA between $0.0 \text{ h} < \text{RA} < 6.25 \text{ h}$ and $21.5 \text{ h} < \text{RA} < 24.0 \text{ h}$, corresponding to sky regions that are devoid of bright galactic diffuse emissions. We performed a direct Fourier transform on the beam weighted EoR cubes to obtain visibilities which we then added to a uniformly selected sample of foreground visibilities.

We computed delay spectra using a standard fast Fourier transform (FFT) with a modified Blackman-Harris window as our spectral tapering function (Blackman & Tukey 1958; Thyagarajan et al. 2016). The modified Blackman-Harris window is a regular Blackman-Harris function convolved with itself for increased spectral dynamic range. The drawback of the modified spectral tapering function is that it has a broader response in delay space (noise equivalent width: $w_{\text{ne}} \approx 2.88$; Thyagarajan et al. 2016). For this reason, we downsampled the delay spectra by a factor of 3 along the delay axis to contain only samples that are approximately independent. The band over which we computed delay spectra is 152–167 MHz, corresponding roughly to Band 2 used in HERA Collaboration (2023) to set upper limits on the EoR HI 21 cm signal at redshift $z = 7.9$. We then computed power spectra according to Equation (27), where we used the cosmological parameters from Planck Collaboration VI (2020) to compute the scaling and obtained V_{eff} directly from the simulated foreground visibilities, and $\Omega_{\text{pp}} \approx 0.016$ from the HERA primary beam model (Fagnoni et al. 2021).

When averaging power spectra, we flagged instances where $V_{\text{eff}} < 1 \text{ Jy}$ to ensure that the assumptions of the closure phase approximation are met (cf. Section 2.2). In practice, when estimating V_{eff} from calibrated data, it is desirable to have enough signal-to-noise ratio in the visibility amplitudes to avoid biases in the power spectrum scaling. Moreover, as discussed in the next section, low visibility amplitudes can lead to a wider delay response and increased foreground leakage. The flagging on V_{eff} helps avoid such instances.

Alongside power spectra, we also computed the ‘C-terms’ (Equation (15)) and window functions (i.e. Equation (21) and the components thereof) using FFTs.

4. Results

This section presents the results of our validation efforts. We discuss the window functions first, as these offer the necessary information about the amount of mode-mixing in a given measurement to decide whether to include it in further analysis. We then test the accuracy of our power spectrum estimator against the standard visibility estimator and compare them with the intrinsic power spectra of our models. Lastly, we discuss the effect of inter-mode C-terms in our estimator.

4.1. Delay window functions

We investigate the effective delay window function, $\widetilde{\Xi}(\tau) = \sum_j q_j \widetilde{\Xi}_j(\tau)$, which determines the amount of line-of-sight mode-mixing in the closure phase delay spectrum. Unlike in the standard visibility delay spectrum, this delay window function is inherently foreground dependent. Spectral structure in the foreground visibility phases can lead to intolerable amounts of line-

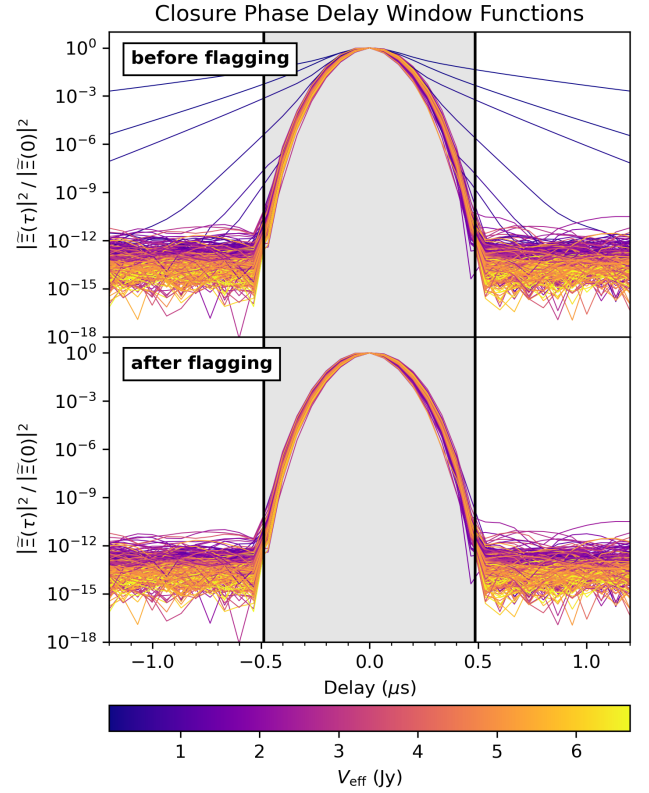


Fig. 5. Normalised closure phase delay window functions for an equilateral 14.6 m triad before (top) and after (bottom) flagging instances with low dynamic range ($< 10^{10}$). The colour indicates the corresponding value of the effective visibility, showing that low dynamic range windows tend to be associated with low visibility amplitudes and hence low values of V_{eff} . The shaded region indicates delays which lay within the horizon limit of the closure phase delay window function plus buffer.

of-sight mode-mixing and foreground leakage beyond the horizon limit. This is especially true for low visibility amplitudes, where small spectral variations in the visibilities can have a strong effect on $\Xi_j(\nu)$. The amount of spectral structure in $\Xi_j(\nu)$ therefore depends on the baseline length and orientation, as well as the apparent sky.

Figure 5 shows $|\widetilde{\Xi}(\tau)|^2$ for an equilateral 14.6 m triad and 300 different realisations of the foreground sky, coloured by the corresponding values of V_{eff} . As expected, some of the window functions are severely limited in dynamic range. Furthermore, the plot confirms our intuition about the scaling of the dynamic range with V_{eff} . We propose flagging instances where a certain dynamic range is not achieved within a given delay range. We determine a suitable delay range by considering the expected width of the window function. Since $\Xi_j(\nu)$ consists of two visibility phase terms (ϕ_j cancels out in the division by V_j^F), its maximum delay is approximately $\tau_{\text{wh},j} = \tau_{\nu,h} - \tau_{h,j}$ in the limit of spectrally flat visibility amplitudes. Hence, we define the horizon limit of the closure phase delay window function simply as $\tau_{\text{wh}} = \max(\{\tau_{\text{wh},1}, \tau_{\text{wh},2}, \tau_{\text{wh},3}\})$. Motivated by the fact that foregrounds overpower the EoR HI 21 cm line by a factor of $\sim 10^5$, we flag instances where the dynamic range of the squared window function is smaller than 10^{10} at delays $|\tau| < \tau_{\text{wh}} + \tau_b$ (see lower graph in Figure 5). For our foreground model, the flagging fractions range from 4% to 30%. The flagging fraction tends to increase with the lengths of the triad baselines

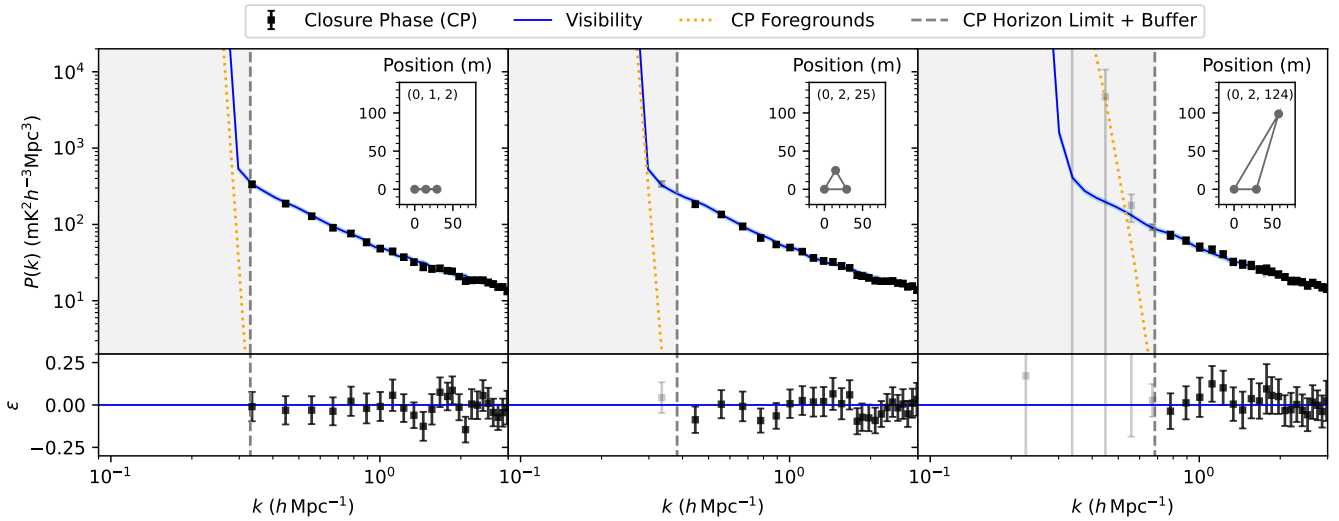


Fig. 6. Model 1 power spectrum estimates for three triads (see inserts). The top panel shows power spectra estimated from visibilities (solid blue line) and closure phases (black scatter points). The dotted orange line shows the closure phase power spectrum for a sky consisting of foregrounds only. The lower panel shows the fractional difference, ϵ , between the visibility and closure phase power spectra. The shaded regions mark the foreground wedge. All uncertainties due to cosmic variance are 2σ .

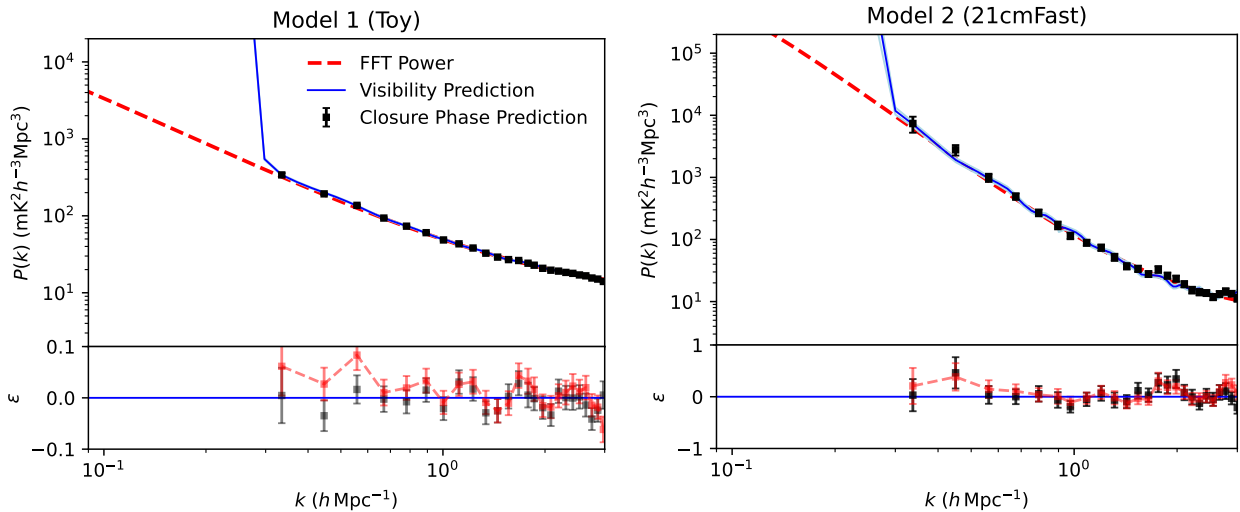


Fig. 7. Same as Figure 6 but averaged across all 15 triads in bins of k . The left and right plots show the estimated power spectra for Model 1 and Model 2, respectively. The red dashed line shows the power spectrum obtained from doing a 3D FFT on the full resolution cubes.

due to their increased chromaticity, which tends to broaden the delay window function and limit the spectral dynamic range (see Section 2.6).

4.2. Power spectra

Figure 6 shows the averaged visibility and closure phase power spectra of Model 1 for three different triads: a linear east-west triad with two 14.6 m baselines, an equilateral 29.2 m triad, and a scalene triad with one short baseline (29.2 m) and two long baselines (105.3 m and 116.8 m). The flagging fraction was 6.6%, 7%, and 30%, respectively (see previous section). The three visibility power spectra of a triad were averaged in bins of k to obtain the plotted visibility power spectrum (solid blue line). The closure phase delay power spectrum of the foregrounds is also shown (dotted orange line). All uncertainties in the plots are 2σ .

The broadening of the foreground wedge with increasing triad size is clearly discernible; while the wedge extends to $k \sim 0.3 - 0.4 h\text{Mpc}^{-1}$ for the linear and equilateral triads, it extends to $k \sim 0.7 h\text{Mpc}^{-1}$ for the long-baseline triad. For the latter, the increased size of the foreground region of the closure phase power spectrum with respect to that of the visibilities is particularly evident.

As can be seen in the residual plots (lower panel of Figure 6), the closure phase and visibility power spectra agree well within the 2σ uncertainty. The RMS values of the fractional residuals are 3.3%, 4.2% and 5.3%, respectively, and can be expected to decrease further as more samples are averaged.

We averaged the power spectra of all 15 triads to further test the accuracy of our power spectrum estimator. Figure 7 shows averaged power spectrum estimates for Model 1 (left) and Model 2 (right), respectively, alongside a more accurate estimate of the intrinsic power obtained from a 3D FFT of the full-resolution

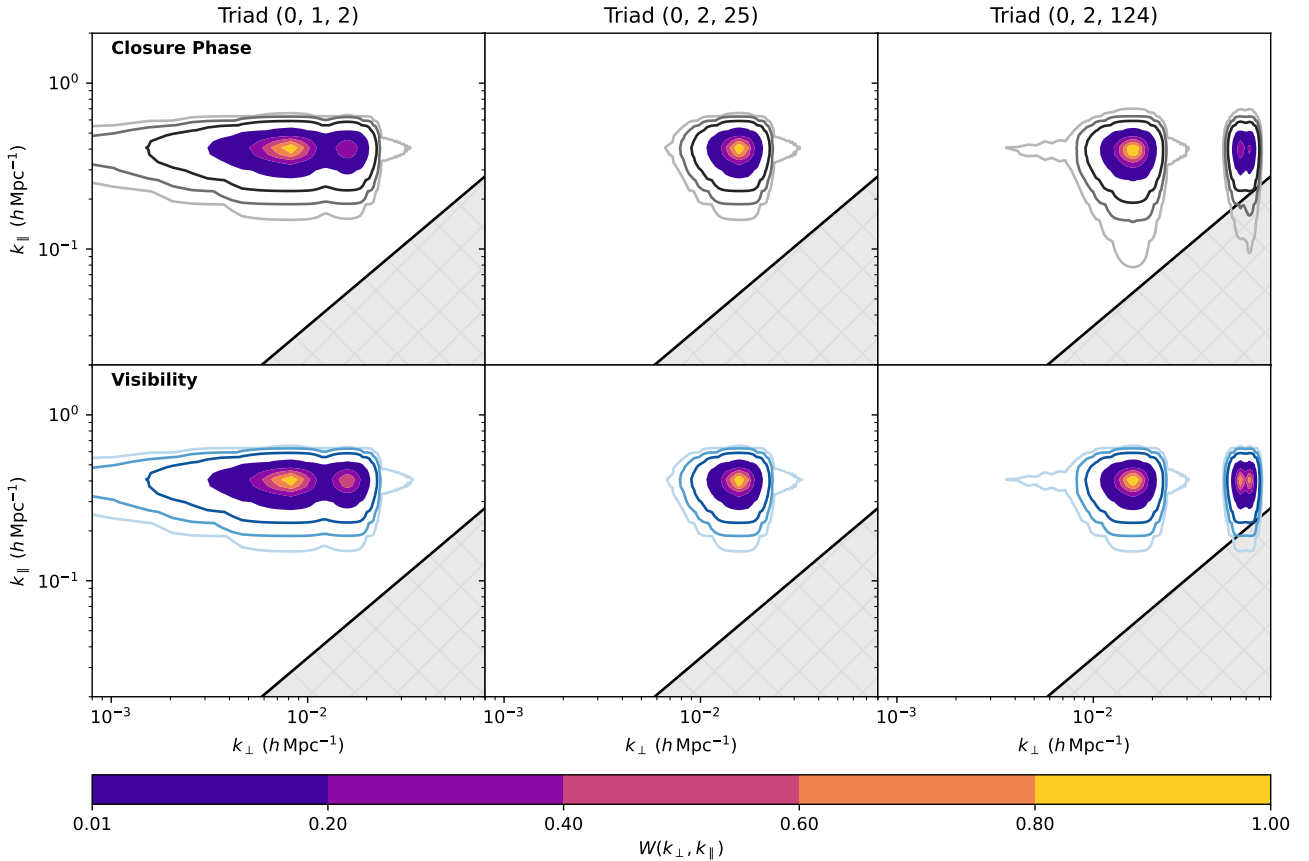


Fig. 8. Normalised cylindrical window functions at $k_{\parallel} = 0.4 \text{ h Mpc}^{-1}$ for closure phases (top row) and visibilities (bottom row). Each column shows the window functions for a different triad: a linear east-west triad with two 14.6 m baselines, an equilateral 29.2 m triad, and a scalene triad with one short baseline (29.2 m) and two long baselines (105.3 m and 116.8 m) (cf. Figure 6). For visualisation purposes, the bottom row shows the sum of the visibility window functions of the three triad baselines. The contour lines are at the levels of 10^{-4} , 10^{-6} , and 10^{-8} . The shaded regions in the bottom right corners indicate the standard wedge region.

model cubes (dotted red lines). The visibility and closure phase results continue to agree with each other at high significance. While there are a few 2σ -outliers in the residual plots of both models, they are still consistent with zero at the 3σ -level. Note that the residuals are strongly correlated, so neighbouring modes of an outlier point are more likely to be outliers. Such accumulation of outliers can be seen, for example, in the right plot of Figure 7 at around $k \sim 2 \text{ h Mpc}^{-1}$. The overall consistency between the visibility and closure phase power spectra suggests that upper limits derived from the two approaches can be directly compared to each other once the relevant scaling factors have been accounted for.

The closure phase delay power spectra accurately estimate the intrinsic power at intermediate and high k -modes, but slightly overestimate it at $k \lesssim 0.6 \text{ h Mpc}^{-1}$ for Model 1. This is probably attributed to the smearing of power by the window function, as has already been noted in similar tests done by Aguirre et al. (2022) and has been discussed in Section 2.6 of this paper.

Although thermal noise was not included in our simulations, we note that the flagging fraction due to low V_{eff} would result in moderate sensitivity detriments of approximately 3%, 4%, and 16%, respectively, assuming uniform sensitivity across the LST-range. In a real data reduction, this would be additional to the commonly applied flags of a visibility-based analysis (e.g. RFI and instrumental flags), although some overlap may be possible. The loss in sensitivity is relatively low, as most of the sensitiv-

ity comes from coherent averaging over consecutive nights and redundant baseline measurements (see Parsons et al. 2012a). We leave an in depth discussion of the effects of thermal noise on the sensitivity of the closure phase power spectrum to future work.

4.3. Cylindrical mode-mixing

We investigate the relative contributions of different $(k_{\perp}, k_{\parallel})$ -modes to the power spectrum estimates. Figure 8 shows cylindrical window functions of the power spectra at $k_{\parallel} = 0.4 \text{ h Mpc}^{-1}$ estimated from closure phases (top row) and from visibilities (bottom row). The columns in the figure correspond to the three triads discussed in the previous section. For better visualisation, the visibility window functions have been averaged across the three baselines.

As expected, the number of peaks in the cylindrical power spectra corresponds to the number of unique baseline lengths in a given triad. Each peak is located at the $(k_{\perp}, k_{\parallel})$ -mode mapped by one of the triad baselines and the delay mode (see Equation (17)). The relative amplitudes of the peaks are not necessarily equal. For example, the peak corresponding to the 29.2 m baseline of the linear triad is only 20–40% of the peaks corresponding to the 14.6 m baselines (top left plot in Figure 8). This is a consequence of the foreground dependence of the window functions: The contribution of each baseline is weighted by the squared inverse of the visibility amplitude (see Equations (12)

and (21)). Apart from this weighting, the closure phase and visibility window functions are almost indistinguishable. The main discrepancies between the closure phase and visibility window functions are seen at modes that are close to the foreground wedge. In our example, this is the case for the scalene triad (right column), where the closure phase window function has a noticeably larger extent in the k_{\parallel} -direction, especially at levels below 10^{-6} . The window function thus extends further into the foreground wedge, thereby increasing the amount of foreground power leakage beyond the horizon limit. This broadening of the window function close to the wedge is caused by the $\exp\{i\phi_{\nu}^{\text{E}}(\nu)\}$ term in the line-of-sight component of the window function (see Equation (12) and discussion in Sections 2.6 and 4.1).

Note that for the baselines used in this work, the scale of mode-mixing along the line-of-sight dimension is nearly an order of magnitude greater than along the transverse dimension. The spherically averaged window functions (that is, averaged in bins of k) is therefore approximately the same for closure phases and visibilities, despite their different imprint along the transverse dimension of the cylindrical power spectrum.

4.4. C-terms

In our derivation of the power spectrum estimator, we assumed that inter-mode correlations are approximately zero (see Equation (20)). Due to the finite sample size, these inter-mode C-terms constitute an additional noise-like component in the estimated power spectrum. We compute the relative contribution of these C-terms to the averaged Model 1 power spectrum (left plot in Figure 7) as

$$\epsilon_c(k) = \frac{\langle \sum_{\text{inter-modes}} C(\mathbf{b}, \mathbf{b}', \tau, \tau') \rangle_k}{\langle \sum_{j=1}^3 C(\mathbf{b}_j, \mathbf{b}_j, \tau, \tau) + C(\mathbf{b}_j, \mathbf{b}_j, -\tau, -\tau) \rangle_k}, \quad (30)$$

where the sum in the numerator is over C-terms for which $\mathbf{b} \neq \mathbf{b}'$ or $\tau = -\tau'$, and the angular brackets denote an average over different sky regions and triads in bins of k weighted by V_{eff}^2 to correctly account for the relative contribution of each term to the final average. As can be seen in Figure 9, $\epsilon_c(k)$ appears noise-like and has a standard deviation of $\sigma_c = 0.014$. We can compare this with the RMS of ~ 0.018 of the fractional residuals in the left plot of Figure 7, indicating that the C-terms account for approximately 60% of the fractional difference between the closure phase and visibility power spectra. The remaining differences are mainly attributed to cosmic variance, which arises from the fact that the two estimates give different weights to different modes of the intrinsic power spectrum via their unique window functions.

For very compact triads, the inter-baseline C-terms should eventually stop averaging down like noise because the modes probed by the triad baselines are not entirely independent. As can be seen in Equation (20), the amount of additional signal recovered by inter-baseline C-terms depends on the primary beam and the difference between baseline vectors. Using the HERA beam model, we estimate the expected signal recovery in inter-baseline C-terms to be around 4% for the shortest HERA baselines. The resulting bias can, in principle, be corrected for if higher accuracy is required. Here, we do not achieve a fractional accuracy high enough on individual triads to make this correction necessary.

As the inter-mode C-terms are at most a few per cent of the cosmological signal strength, they are at least a factor of 10^3 smaller than the thermal-like noise that limits current closure

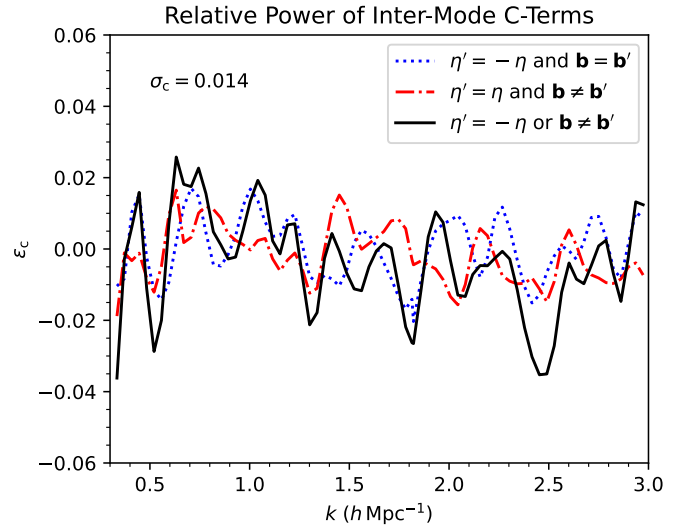


Fig. 9. Relative power of inter-mode C-terms in the averaged Model 1 power spectrum (left plot in Figure 7). The following C-terms are plotted: different delay modes but same baseline vector (blue dotted line), same delay modes but different baseline vectors (red dash-dotted line), and different delay modes and/or different baseline vectors (black line). The variance of the latter is $\sigma_c = 0.014$.

phase upper limits (e.g. Keller et al. 2023). Future detections of the cosmological signal will have to quantify the contribution of inter-mode C-terms to the estimated power spectrum using a forward modelling approach as presented here.

5. Summary and conclusion

In this work, we have revised the closure phase delay spectrum technique for estimating the power spectrum of a cosmological line signal in the presence of strong foreground continuum emissions. Our estimator differs from previous works by our definition of the ‘effective visibility’; a scaling factor which gauges spectral phase fluctuations to intensity fluctuations. As part of the derivation, we obtain the window function, which quantifies mode-mixing in our estimator. The effective visibility arises naturally as a part of the normalisation of the window function.

We have validated our estimator against the standard visibility delay spectrum technique using a subset of the HERA telescope and a sky model consisting of realistic point source foregrounds and two different models of the EoR HI 21 cm signal. We find that the power spectrum estimates are consistent with those from visibilities within the uncertainty of cosmic variance, which is smaller than 2% in our simulations. This indicates that the accuracy of the estimator may further increase as more samples are averaged, i.e. by including more triad shapes and orientations or more independent patches of the sky.

The main differences between the closure phase and visibility delay spectrum techniques can be summarised as follows:

1. *Window functions:* The estimated power receives contributions from a range of intrinsic k -modes. This is due to the spatial boundaries of the measurement, set by the finite frequency bandwidth and the primary beam response. The mapping between the estimated and intrinsic power spectra is described by the window function. In contrast to the visibility estimator, the closure phase window function may peak at multiple k_{\perp} if the triad baselines are of different lengths. Additionally, its line-of-sight component depends

on the foreground visibilities, resulting in a broader response in k_{\parallel} -space. As a result, we find that low visibility amplitudes can severely reduce the width and dynamic range of the window function. However, after flagging such instances, we find that the width and shape of the closure phase window function are comparable to that resulting from the visibility-based technique.

2. *C-terms*: Our power spectrum estimates are obtained by squaring and averaging the amplitudes of the closure phase delay spectra. In doing so, there are several inter-baseline and inter-mode cross-terms, which we call ‘C-terms’ (see Equation (20)). We have shown that inter-mode C-terms average down to approximately zero and constitute an additional noise-like component in the closure phase power spectrum.
3. *Higher-order terms*: The closure phase approximation uses linear perturbation theory to map spectral intensity fluctuations to phase fluctuations. Thyagarajan & Carilli (2020) have shown this approximation to reach a fractional accuracy of 10^{-4} for the expected strength of the EoR HI 21 cm signal. This should be subdominant to the above effects and accurate enough for the initial detection and characterisation of the cosmic signal.
4. *Foreground wedge*: As already noted by Thyagarajan & Carilli (2020), the foreground dominated region of the closure phase delay spectrum goes beyond the conventional foreground wedge. The broadening of the foreground wedge results from the multiplication of three visibility phase terms, which translates to a triple convolution in delay-space. The closure phase technique is therefore less sensitive to the lowest k -modes. This effect can potentially be mitigated by applying the delay transform directly to the unwrapped closure phase instead of its complex exponential (See appendix of Thyagarajan & Carilli 2020).

This work confirms the viability of estimating the cosmological HI 21 cm power spectrum with closure phases. However, previous upper limits on the cosmic signal will have to be revised to account for the scaling differences. The improved understanding of mode-mixing will allow future studies to combine different triad shapes to achieve better sensitivity while accounting for the window functions. Our planned future work will focus on finding optimal ways to average the measurements of different closure triads in the presence of noise. Furthermore, we aim to improve the sensitivity of the approach, for example, by using other closure quantities that include visibility amplitude information (e.g. closure invariants; Thyagarajan et al. 2022) and adapting techniques to further mitigate the foreground wedge (cf. Thyagarajan & Carilli 2020).

Acknowledgements. PMK is part of Allegro, the European ALMA Regional Centre node in the Netherlands. Allegro is funded by The Netherlands Organisation for Scientific Research (NWO/EW). The software developed for this analysis uses Python and the publicly-accessible and open-sourced Python packages 21cmFast (Mesinger et al. 2011; Murray et al. 2020), Numpy (van der Walt et al. 2011), SciPy (Virtanen et al. 2020), Astropy (Astropy Collaboration 2022), Matplotlib (Hunter 2007), `matvis` and `hera_sim` (Kittiwisit et al. 2025).

References

Aguirre, J. E., Murray, S. G., Pascua, R., et al. 2022, *ApJ*, 924, 85

- Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, *ApJ*, 935, 167
- Bandura, K., Addison, G. E., Amiri, M., et al. 2014, *SPIE Conf. Ser.*, 9145, 914522
- Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., & Pober, J. C. 2016, *MNRAS*, 461, 3135
- Bharadwaj, S., & Sethi, S. K. 2001, *J. Astrophys. Astron.*, 22, 293
- Blackman, R. B., & Tukey, J. W. 1958, *Bell Syst. Techn. J.*, 37, 185
- Byrne, R., Morales, M. F., Hazelton, B., et al. 2019, *ApJ*, 875, 70
- Byrne, R., Morales, M. F., Hazelton, B. J., & Wilensky, M. 2021, *MNRAS*, 503, 2457
- Charles, N., Bernardi, G., Bester, H. L., et al. 2022, *MNRAS*, 512, 2716
- Chen, X. 2011, *Scientia Sinica Phys. Mech. Astron.*, 41, 1358
- Crichton, D., Aich, M., Amara, A., et al. 2022, *J. Astron. Telescopes Instrum. Syst.*, 8, 011019
- Datta, A., Bowman, J. D., & Carilli, C. L. 2010, *ApJ*, 724, 526
- Datta, K. K., Jensen, H., Majumdar, S., et al. 2014, *MNRAS*, 442, 1491
- de Gasperin, F., Mevius, M., Rafferty, D. A., Intema, H. T., & Fallows, R. A. 2018, *A&A*, 615, A179
- DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, *PASP*, 129, 045001
- Eastwood, M. W., Anderson, M. M., Monroe, R. M., et al. 2019, *AJ*, 158, 84
- Ewall-Wice, A., Dillon, J. S., Liu, A., & Hewitt, J. 2017, *MNRAS*, 470, 1849
- Fagnoni, N., de Lera Acedo, E., DeBoer, D. R., et al. 2021, *MNRAS*, 500, 1232
- Furlanetto, S. R. 2016, *Astrophys. Space Sci. Libr.*, 423, 247
- Hera, C. 2022, *ApJ*, 925, 221
- Hera, C. 2023, *ApJ*, 945, 124
- Hogg, D. W. 1999, arXiv e-prints [arXiv: astro-ph/9905116]
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, 9, 90
- Hurley-Walker, N., Callingham, J. R., Hancock, P. J., et al. 2017, *MNRAS*, 464, 1146
- Jennison, R. C. 1958, *MNRAS*, 118, 276
- Keller, P. M., Nikolic, B., Thyagarajan, N., et al. 2023, *MNRAS*, 524, 583
- Kittiwisit, P., Murray, S. G., Garsden, H., et al. 2025, *RAS Techn. Instrum.*, 4, rzaf001
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1
- Liu, A., & Shaw, J. R. 2020, *PASP*, 132, 062001
- Liu, A., Parsons, A. R., & Trott, C. M. 2014, *Phys. Rev. D*, 90, 023018
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, 411, 955
- Morales, M. F., & Hewitt, J. 2004, *ApJ*, 615, 7
- Murray, S., Greig, B., Mesinger, A., et al. 2020, *J. Open Source Software*, 5, 2582
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, 433, 639
- Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012a, *ApJ*, 753, 81
- Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012b, *ApJ*, 756, 165
- Planck Collaboration VI. 2020, *A&A*, 641, A6
- Pritchard, J. R., & Loeb, A. 2012, *Rep. Prog. Phys.*, 75, 086901
- Samuel, J., Nityananda, R., & Thyagarajan, N. 2022, *Phys. Rev. Lett.*, 128, 091101
- Santos, M., Bull, P., Alonso, D., et al. 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 19
- Slosar, A., Ahmed, Z., Alonso, D., et al. 2019, *Bull. Am. Astron. Soc.*, 51, 53
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 2017, in *Interferometry and Synthesis in Radio Astronomy*, 3rd edn. (Cham: Springer)
- Thyagarajan, N., & Carilli, C. L. 2020, *Phys. Rev. D*, 102, 022001
- Thyagarajan, N., & Carilli, C. L. 2022, *PASA*, 39, e014
- Thyagarajan, N., Udaya Shankar, N., Subrahmanyam, R., et al. 2013, *ApJ*, 776, 6
- Thyagarajan, N., Parsons, A. R., DeBoer, D. R., et al. 2016, *ApJ*, 825, 9
- Thyagarajan, N., Carilli, C. L., & Nikolic, B. 2018, *Phys. Rev. Lett.*, 120, 251301
- Thyagarajan, N., Carilli, C. L., Nikolic, B., et al. 2020, *Phys. Rev. D*, 102, 022002
- Thyagarajan, N., Nityananda, R., & Samuel, J. 2022, *Phys. Rev. D*, 105, 043019
- Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, *PASA*, 30, e007
- Tiwari, H., Thyagarajan, N., Trott, C., & McKinley, B. 2024, *PASA*, 41, e069
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Comput. Sci. Eng.*, 13, 22
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, 556, A2
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, 17, 261
- Zarka, P., Girard, J. N., Tagger, M., & Denis, L. 2012, in *SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, eds. S. Boissier, P. de Laverny, & N. Nardetto, 687
- Zheng, Q., Wu, X.-P., Johnston-Hollitt, M., Gu, J.-H., & Xu, H. 2016, *ApJ*, 832, 190