CAMBRIDGE
UNIVERSITY PRESS

Research Article

21-cm Epoch of reionisation power spectrum with closure phase using the Murchison Widefield Array

Himanshu Tiwari[1,2] , Nithyanandan Thyagarajan[2] , Cathryn Trott[1,3] , and Ben McKinley[1,3]

[1]International Centre for Radio Astronomy Research (ICRAR), Curtin University, Bentley, Perth, WA, Australia, [2]Commonwealth Scientific and Industrial Research Organisation (CSIRO), Space & Astronomy, Bentley, WA, Australia and [3]ARC Centre of Excellence for All Sky Astrophysics in Three Dimensions (ASTRO-3D), Australia

Abstract

The radio interferometric closure phases can be a valuable tool for studying cosmological HI from the early Universe. Closure phases have the advantage of being immune to element-based gains and associated calibration errors. Thus, calibration and errors therein, which are often sources of systematics limiting standard visibility-based approaches, can be avoided altogether in closure phase analysis. In this work, we present the first results of the closure phase power spectrum of HI 21-cm fluctuations using the Murchison Widefield Array (MWA), with ~ 12 h of MWA phase II observations centred around redshift, $z \approx 6.79$, during the Epoch of Reionisation. On analysing three redundant classes of baselines – 14, 24, and 28 m equilateral triads, our estimates of the 2σ (95% confidence interval) 21-cm power spectra are $\lesssim (184)^2 pseudo$ mK^2 at $k_{||} = 0.36 pseudo$ hMpc^{-1} in the EoR1 field for the 14 m baseline triads, and $\lesssim (188)^2 pseudo$ mK^2 at $k_{||} = 0.18 \, pseudo$ hMpc^{-1} in the EoR0 field for the 24 m baseline triads. The '$pseudo$' units denote that the length scale and brightness temperature should be interpreted as close approximations. Our best estimates are still 3-4 orders high compared to the fiducial 21-cm power spectrum; however, our approach provides promising estimates of the power spectra even with a small amount of data. These data-limited estimates can be further improved if more datasets are included into the analysis. The evidence for excess noise has a possible origin in baseline-dependent systematics in the MWA data that will require careful baseline-based strategies to mitigate, even in standard visibility-based approaches.

Keywords: Radio interferometry; MWA; cosmic reionisation

1. Introduction

Epoch of Reionisation (EoR) is the period when the first stars and galaxies were formed in the early Universe between $15 < z < 5.3$ and contributed to reionising the predominantly neutral intergalactic medium on cosmic scales (Bosman et al. 2022; Zhu et al. 2022, 2024). It is also one of the least understood periods in the history of the Universe, mainly due to the lack of radiation influx from the first stars and galaxies, which are locally absorbed by the intervening medium. The hyperfine ground state of the atomic Hydrogen (HI) produces a weak transition of $\sim 1\,420$ MHz, popularly known as the 21-cm line. It is considered a very promising probe of the EoR due to the abundance of Hydrogen in the early Universe. The intervening medium is largely transparent to the redshifted 21-cm line; therefore, it provides one of the best avenues to infer the astrophysical properties of the IGM and the cosmology of the early Universe. As the neutral IGM gets ionised, it weakens the strength of the 21-cm signal. One can interpret the stages of cosmic reionisation by estimating the depletion in the redshifted 21-cm signal through cosmic time (see Furlanetto,

Sokasian, & Hernquist 2004; Pritchard & Loeb 2012; Mesinger 2016, for review).

To detect this forbidden transition from the early Universe, several radio instruments such as Murchison Widefield Array (MWA) (Tingay et al. 2013), Hydrogen Epoch of Reionization Array (HERA) (DeBoer et al. 2017), LOw-Frequency ARray (LoFAR) (van Haarlem et al. 2013), Giant metrewave Radio Telescope (GMRT) (Paciga et al. 2013), Precision Array for Probing the Epoch of Reionization (PAPER) (Pober et al. 2011), Long Wavelength Array (LWA) (Eastwood et al. 2019), Experiment to Detect the Global EoR Signature (EDGES) (Bowman, Rogers, & Hewitt 2008; Bowman & Rogers 2010; Bowman et al. 2018), Shaped Antenna measurement of the background RAdio Spectrum (SARAS) (Patra et al. 2012; Singh et al. 2018, 2022), Broad-band Instrument for Global HydrOgen ReioNization Signal (BIGHORNS) (Sokolowski et al. 2015), Large Aperture Experiment to Detect the Dark Age (LEDA) (Bernardi et al. 2016), Dark Ages Radio Explorer (DARE) (Sigel et al. 2013), Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI-HI) (Voytek et al. 2014), Probing Radio Intensity at High-Z from Marion (PRIZM) (Philip et al. 2019) were built or are under construction. These instruments can either aim to detect the sky-averaged 21-cm signal spectrum (Global signal) or measure its spatial fluctuations. The former category of instruments can detect the overall IGM properties, whereas the latter can provide a detailed study of the three-dimensional topology of

Corresponding author: Himanshu Tiwari; Email: himanshu.tiwari@postgrad.curtin.edu.au

the EoR regime. The spatial signatures are quantified through statistical measures such as the power spectrum, which can probe the 21-cm signal strength as a function of cosmological length scales (k-modes). Alongside, the three-dimensional topology of the EoR can be studied via a two-dimensional 21-cm power spectrum (Barry et al. 2019; Trott et al. 2020; Mertens et al. 2020; The HERA Collaboration et al. 2021; Munshi et al. 2023), which shows the variation of the 21-cm power spectrum along the line of sight and transverse axis.

The significant challenges of detecting the 21-cm signal come from the foregrounds, ionospheric abnormalities, instrumental systematics, and radio frequency interference (RFI), emitting in the same frequency range as the redshifted 21-cm signal from the EoR. In an ideal situation avoiding or minimising all the above factors, we still require to calibrate the instrument against the bright foregrounds that require calibration accuracy of $\gtrsim 10^5 : 1$ by the radio instruments to reach the required HI levels. Calibration accuracy is especially important for EoR observations using the MWA because of the presence of sharp periodic features in the bandpass produced by the polyphase filter bank used. The inability to accurately correct for these element-based bandpass structures significantly affects the power spectrum estimates (Beardsley et al. 2016; Barry et al. 2019; Trott et al. 2020; Patwa, Sethi, & Dwarakanath 2021; Yoshiura et al. 2021).

The radio interferometric closure phase has emerged as an alternative and independent approach to studying the EoR while addressing the calibration challenges. The main advantage of this approach is the immunity of closure phases to the errors associated with the direction-independent, antenna-based gains. Thus, calibration is not essential in this approach (Carilli et al. 2018; Thyagarajan & Carilli 2020). The closure phase in the context of the EoR was first investigated by Thyagarajan, Carilli, & Nikolic (2018), Carilli et al. (2018) and further employed on the HERA data by Thyagarajan et al. (2020), Keller et al. (2023). The method has shown significant promise in avoiding serious calibration challenges, and with detailed forward modelling, one can ideally quantify the 21-cm power spectrum. This paper is the first attempt to utilise a closure phase approach on MWA phase II observations. We followed the methods investigated by Thyagarajan et al. (2020), Keller et al. (2023) and applied them to our datasets. This paper is organised as follows. In Section 2, we discuss the background of the closure phase. Sections 3 and 4 of this paper explain the observations and forward modelling with simulations of the foregrounds, HI, and noise. In Section 5, we discuss the data processing and rectification, and finally, we present our results in Section 6 and discuss them in Section 7.

2. Background

In this section, we review the background of the interferometric closure phase in brief (refer to Thyagarajan, Carilli, & Nikolic 2018; Thyagarajan & Carilli 2020, for a complete mathematical understanding of this approach). The measured visibility between two antenna factors at a given baseline (V_{ij}^{m}) can be defined as the sum of true sky visibility and noise:

$$V_{ij}^{\mathrm{m}}(\nu) = \mathbf{g}_i(\nu) V_{ij}^{\mathrm{T}}(\nu) \mathbf{g}_j^*(\nu) + V_{ij}^{\mathrm{N}}(\nu); \tag{1}$$

where $\mathbf{g}_i(\nu)$, $\mathbf{g}_j(\nu)$ denote the element-based gain terms, $\{*\}$ represents the complex conjugate, $V_{ij}^{\mathrm{T}}(\nu)$ is the true sky visibility, and $V_{ij}^{\mathrm{N}}(\nu)$ is the noise in the measurement. The indices $\{ij\}$ correspond to the antennae $\{a, b\}$ forming a baseline. The true sky

visibility can be further decomposed into the foregrounds and faint cosmological HI visibilities:

$$V_{ij}^{\mathrm{T}}(\nu) = V_{ij}^{\mathrm{FG}}(\nu) + V_{ij}^{\mathrm{HI}}(\nu) \tag{2}$$

In general, the foreground $\geq 10^4$K orders of higher magnitude than the HI, and to reach the sensitivity limit of HI, the gains ($\mathbf{g}_i's$) are required to be precisely calibrated up to the HI levels. It presents challenges to the direct visibility-based HI power spectrum analysis as it requires accurate modelling of the foreground and mastering the calibration techniques. In radio interferometry, the term 'closure phase' is assigned to the phase derived from the product of $N \geq 3$ closed loops of antenna visibilities [Jennison (1958)]. When $N = 3$, it is also sometimes referred to as the bispectrum phase in the literature, which can be defined as,

$$\begin{aligned}
\phi_\nabla^{\mathrm{m}}(\nu) &= \arg \prod_{ij=1}^{3} V_{ij}^{\mathrm{m}}(\nu) \\
&= \arg \prod_{ij=1}^{3} \left[\mathbf{g}_i(\nu) V_{ij}^{\mathrm{T}}(\nu) \mathbf{g}_j^*(\nu) + V_{ij}^{\mathrm{N}}(\nu) \right] \\
&= \arg \prod_{ij=1}^{3} V_{ij}^{\mathrm{T}}(\nu) + \text{noise-like terms} \tag{3}
\end{aligned}$$

where $\{ij\}$ runs through antenna-pairs $\{ab, bc, ca\}$, and the gains of individual antenna elements ($\mathbf{g}_i's$) gets eliminated in the closure phase, leaving only the true sky phase as the sole contributing factor in the closure phase.

The closure phase delay spectrum technique also exploits the fact that the foregrounds predominantly obey a smooth spectral behaviour, whereas the cosmological HI creates spectral fluctuations. Thus, in the Fourier delay domain, the foreground signal strength (power) gets restricted within the lower delay modes. In contrast, the HI power can be observed at the higher delay modes, creating the distinction between these two components in the Fourier domain, from which the faint HI can be detected. Because of its advantages in avoiding element-based calibration errors and processing simplicity, it promises to be an independent alternative to estimating the 21-cm power spectrum (Thyagarajan, Carilli, & Nikolic 2018; Carilli et al. 2018; Thyagarajan & Carilli 2020; Thyagarajan et al. 2020; Keller et al. 2023).

The delay spectrum of the closure phase can be estimated by taking the Fourier transform of the complex exponent of the closure phase with a window function along frequency,

$$\tilde{\Psi}_\nabla(\tau) = V_{\mathrm{eff}} \int e^{i\phi_\nabla(\nu)} W(\nu) e^{2\pi i \nu \tau} d\nu, \tag{4}$$

where τ represents delay, the Fourier dual of the sampling frequencies, and V_{eff} is the effective visibility, which can be obtained through the model foreground visibilities or a calibrated visibility. In our work, we used the former estimated through foreground simulations, which are discussed in the next section. Note that we only take the amplitude of the V_{eff}, which acts as a scaling factor in the delay spectrum. $W(\nu)$ is the spectral window function; we used a Blackman–Harris window (Harris 1978) modified to obtain a dynamic range required to sufficiently suppress foreground contamination in the EoR window (Thyagarajan et al. 2016):

$$W(\nu) = W_{\mathrm{BH}}(\nu) * W_{\mathrm{BH}}(\nu),$$

where $W_{\mathrm{BH}}(\nu)$ is the Blackman–Harris window function and $\{*\}$ represents the convolution operation. For a given triad, V_{eff} is

estimated from the sum of inverse variance visibilities weighted over the window,

$$V_{\text{eff}}^{-2} = \sum_{ij=1}^{3} \frac{\int V_{ij}(\nu)^{-2} W(\nu) d\nu}{\int W(\nu) d\nu} . \tag{5}$$

The inverse squaring ensures the appropriate normalisation of visibilities, V_{ij}, taking noise into account. From the delay spectrum, we estimate the delay cross power spectrum by taking the product of the two delay spectra and converting it into [*pseudo* mK$^2 h^{-3}$Mpc^3] units by assimilating the cosmological and antenna-related factors as:

$$P_\nabla(k_{||}) = 2\mathcal{R}\{\tilde{\Psi}_\nabla(\tau) * \overline{\tilde{\Psi}'_\nabla(\tau)}\} \times$$
$$\left(\frac{1}{\Omega B_{\text{eff}}}\right)\left(\frac{D^2 \Delta D}{B_{\text{eff}}}\right)\left(\frac{\lambda^2}{2k_B}\right)^2 [\textit{pseudo} \text{ mK}^2 h^{-3}\text{Mpc}^3] , \tag{6}$$

where Ω is the antenna beam squared solid angle (Parsons et al. 2014), B_{eff} is the effective bandwidth of the observation, and D and ΔD are the cosmological comoving distance and comoving depth corresponding to the central frequency and the bandwidth, respectively. $k_{||}$ is the wavenumber along the line of sight (Morales & Hewitt 2004):

$$k_{||} = \frac{2\pi \tau B_{\text{eff}}}{\Delta D} \approx \frac{2\pi \tau \nu_r H_0 E(z)}{c(1+z)^2} , \tag{7}$$

where ν_r is the redshifted 21-cm frequency and H_0 and $E(z)$ are standard terms in cosmology.

ΩB_{eff} is related to the cosmological volume probed by the instrument and is defined as (Thyagarajan et al. 2016):

$$\Omega B_{\text{eff}} = \iint |A(\hat{s}, \nu)|^2 |W(\nu)|^2 d^2\hat{s} d\nu , \tag{8}$$

where $A(\hat{s}, \nu)$ is the frequency-dependent, directional power pattern of the antenna pair towards the direction, $\hat{s}$, and $W(\nu)$ is the spectral window function. However, we approximated by separating the integrals to obtain Ω as (Parsons et al. 2014)

$$\Omega = \int |A(\hat{s}, \nu_r)|^2 d^2\hat{s} \tag{9}$$

and effective bandwidth, B_{eff}, as (Thyagarajan et al. 2016)

$$B_{\text{eff}} = \epsilon B = \int_{-B/2}^{+B/2} |W(\nu)|^2 d\nu , \tag{10}$$

where ϵ is the spectral window function's efficiency and $B = 30.72$ MHz is MWA's instantaneous bandwidth. For the MWA observations at the chosen band in this study, $\Omega \approx 0.076$ Sr. For the modified Blackman–Harris window function adopted here, $\epsilon \approx 0.42$ and hence, $B_{\text{eff}} \approx 12.90$ MHz.

The '*pseudo*' in Equation (6) is used to note that the power spectrum estimated via the closure phase method is an approximate representation of the visibility-based power spectrum (Thyagarajan & Carilli 2020). Further, we used the power spectrum as defined in Thyagarajan et al. (2020) with the scaling factor 2 instead of 2/3[a] to correct for the effective visibility estimates.

3. Observations

In this work, we used 493 zenith-pointed MWA phase II compact observations from September 2016 under the MWA project
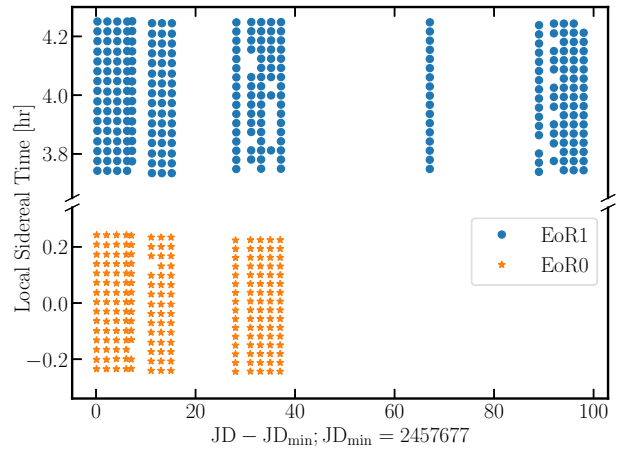
Figure 1. observations used in this analysis: *blue* circles and *orange* stars represent the individual observations made in the respective EoR fields.

`EoR-HighSeason`. Each observation lasts 112 s in the MWA high-band frequency range of 167–197 MHz. Amongst these observations, 198 and 295 target the EoR0 (RA, Dec = 0h, -27deg) and EoR1 (RA, Dec = 4h, -30deg) fields, respectively. Fig. 1 shows the Local Sidereal Time (LST) and Julian Date (JD) of the observations. It amount to ≈6 and 9 h of total observing time on the EoR0 and EoR1 fields, respectively. The MWA raw visibilities are stored in the `gpu-fits` format. To convert them into measurement sets (MS) or uvfits, we used `Birli`[b] (an MWA-specific software that can perform data conversion, averaging in frequency and time, flagging, and other preprocessing steps). Using `Birli`, we averaged the raw visibilities for 8 s at a frequency resolution of 40 kHz. Finally, we output the raw (uncalibrated) visibilities as standard `uvfits`.

Note that we are required to keep all frequency channels for the closure phase analysis; thus, we avoided flagging channel-based RFI (e.g. DTV) and coarse band edge channels (around every 1.28 MHz), which are usually affected by the bandpass[c].

4. Simulations

The closure phases are not linear in the visibilities; thus, forward modelling is key to understanding the data. We incorporated simulations of the foregrounds (FG), HI, and antenna noise to provide cross-validation and comparison with the data. Forward modelling can help identify the excess noise and systematic biases in the data and provide an idealistic estimate for comparison.
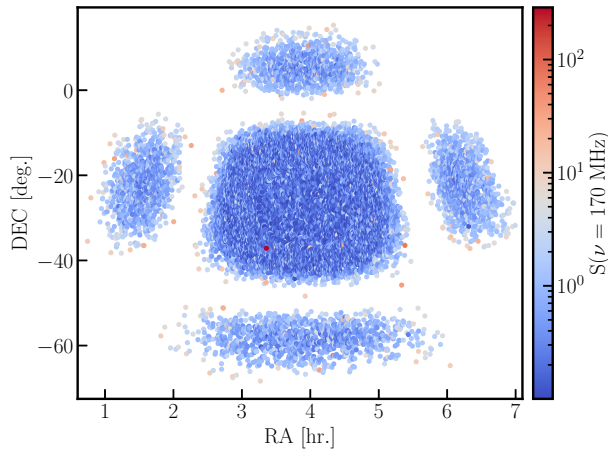
4.1 foregrounds

The simulated FG are generated with the same parameters (i.e. matching LST, frequency, and time resolution) as the observing data. In the first step, we generated the sky maps corresponding to each individual observation. We used the top 20 000 brightest radio sources from the PUMA catalogue (Line et al. 2017) in the observing fields (EoR0, EoR1). The catalogue includes point sources, Gaussians and shapelets. Note that our sky model does

---

[a]The definition of V_{eff}^2 already accounts for the factor of 3.

[b]https://github.com/MWATelescope/Birli.
[c]The MWA's signal processing chain contains filterbanks that yield 24 coarse channels of 1.28 MHz over the full 30.72 MHz band. The fine polyphase filterbank shape results in poor bandpass characteristics at the coarse channel edges (Trott et al. 2020).

Figure 2. Beam attenuated sky-map of 20 000 sources in the EoR1 field used in the foreground simulation. The corresponding Stokes-I flux density at 170 MHz is shown with the colour scale. The sources are only shown as point sources (single component) in the above figure.
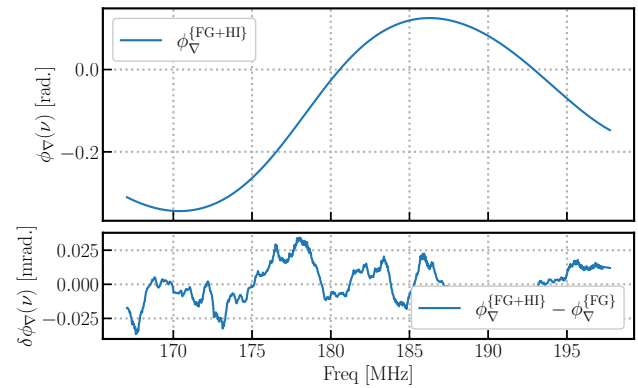

Figure 3. Top: closure phase of the sum of the visibilities of the foreground and HI simulation {FG+HI} for a single triad of 14 m baseline length. Bottom: the difference between the closure phases of {FG+HI} and FG-only simulation, showing the sub-milliradian-level fluctuation of the embedded HI-signal in the closure phase.

not account for the diffuse sky emission. Then, we generated the foreground sky visibilities and converted them to MWA-style `uvfits`. Initially, we experimented with various source counts (e.g. 15 000, 25 000, and 45 000) and their effect on the closure phase. We found the variation in closure phases saturates beyond 20 000. Therefore, we settled for 20 000 source counts in favour of faster computation. However, it should be noted that pinpointing the exact number of source counts where the closure phase saturates is challenging to find. The entire task of sky-map generation and foreground visibility estimation was accomplished using `Hyperdrive`.[d] The sky visibilities are generated using fully embedded antenna element (FEE) beam with *real*-MWA observing scenario where the information of dead dipoles (if present during the observation) and antenna gains are incorporated in the simulation. Fig. 2 shows 20 000 sources around the EoR1 field for a given observation. For simplicity, only the single component of the sources (or point sources flux density) is shown in the figure.

4.2 Neutral hydrogen

Next, we estimate the HI visibilities as observed by MWA. In the limits of the cosmic and sample variance, the characteristic fluctuations in the HI-signal can be assumed to be the same across the sky; therefore, we can avoid simulating HI box multiple times; instead, we can use a single HI simulation box. The HI simulation was generated using 21cmFAST (Mesinger, Furlanetto, & Cen 2010) with a simulation box size of 1.5 cGpc corresponding to $50° \times 50°$ in the sky at a redshift of 6. Then, we passed the simulated voxel data cube to `WODEN`[e] (Line 2022) to generate the MWA-style visibilities of the HI and output it as `uvfits`. The HI visibilities were first generated for each 1.28 MHz coarse band separately with the matching frequency and time resolution of 40 kHz and 8 s of the foreground simulations and then manually stitched together to get the total 30.72 MHz bandwidth. The final HI visibilities were passed to the processing pipeline

for further analysis. The foreground and HI visibilities are added together. We computed the closure phase spectra of the foregrounds as well as of the HI imprinted on the foregrounds. Fig. 3 shows the smooth foreground spectra in the closure phase and the fluctuations (~ 0.01 milliradian) introduced by the presence of HI.

4.3 Noise

The total noise consists of sky noise and receiver noise components. The receiver temperature for the MWA was assumed to be $T_{rx} = 50$ K (Ung et al. 2020), while sky temperature follows a power law in our observing frequency range (Contents of Volume 2006),

$$T_{sky} = T_0 \left(\frac{\nu}{\nu_0} \right)^{\alpha} ; T_0 = 180 \text{ K}, \nu_0 = 180 \text{ MHz}, \alpha = -2.5 ,$$

$$T_{sys} = T_{sky} + T_{rx} .$$

From the system temperature, we estimated the system-equivalent flux density (SEFD) using Thompson, Moran, & Swenson (2017),

$$\text{SEFD} = \frac{2k_B T_{sys}}{A_{eff}}$$

and the RMS,

$$\sigma(\nu) = \frac{\text{SEFD}}{\sqrt{\Delta\nu \Delta t}} , \qquad (11)$$

where k_B is Boltzmann's constant, A_{eff} is the effective collecting area of the telescope, and $\Delta\nu, \Delta t$ are the frequency resolution and integration time, respectively. The $\sigma(\nu)$ is used to generate the Gaussian random noise and converted into the complex noise visibilities with a normalisation factor of $1/\sqrt{2}$ in the real and imaginary parts. Finally, the noise visibilities were added to the corresponding foreground and HI visibilities to get the Model of the sky signal.

4.4 Baseline-dependent gains

The Equation (1) modifies to $V_{ij}^{mod}(\nu) = V_{ij}^m(\nu)\mathbf{g}_{ij}(\nu)$; in scenarios incorporating baseline-dependent gains, where $\mathbf{g}_{ij}(\nu)$ represents the baseline-dependent gain factor. We introduced the

[d]https://github.com/MWATelescope/mwa_hyperdrive.
[e]https://github.com/JLBLine/WODEN.
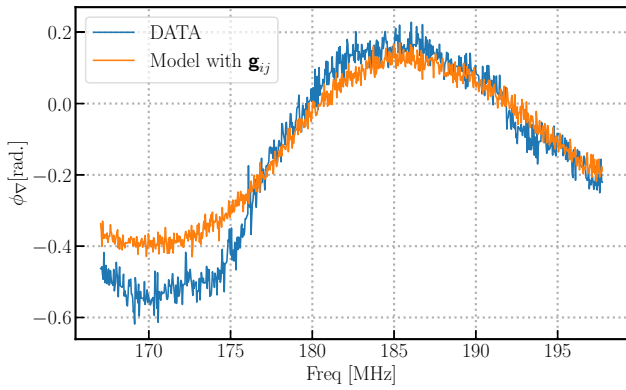
Figure 4. Comparing closure phase spectrum Data (*blue*) and Model with baseline-dependent gains (*orange* line) for EoR1, 28 m baseline length.

baseline-dependent gains using a simple uniform distribution in the $\mathbf{g}_{ij}(\nu)$ phase with unity amplitude. The scaling factor introduced in the $\mathbf{g}_{ij}(\nu)$ sampling is set to approximately match the RMS phase of the binned averaged closure phase of the DATA. We chose brute force method to find the optimal scaling factor for a given EoR field, with a the single scaling factor for given EoR field. Fig. 4 shows the binned averaged closure phase of DATA and Model with $\mathbf{g}_{ij}$. The contribution due to the baseline dependent gains on the binned averaged closure phase about 0.05 rad in both EoR0 and EoR1 fields, which is the RMS of the ratio between the Model with $\mathbf{g}_{ij}$ and Model closure phases. From here onwards, we use two variants of models in the analysis, the first is a forward Model without baseline dependent gains and the second is a Model with $\mathbf{g}_{ij}$.

5. Data processing

The following section provides the basic data processing steps we incorporated into this analysis. The complete schematic flowchart of the data structure is shown in Fig. A6.

5.1 DATA binning

In the first step, the repeated night-to-night observations are sorted based on the Local Sidereal Time (LST) and Julian Date (JD); see Fig. 1. We determined the 14, 24, and 28 m redundant baseline triads from both Hexagonal configurations of MWA (see an example in Fig. 5 (right panel) for which the visibilities are estimated). A given triad {a, b, c} includes $N_{\mathrm{vis}} = 3$ visibilities which correspond to $\{V_{\mathrm{ab}}, V_{\mathrm{bc}}, V_{\mathrm{ca}}\}$. The number of triads ($N_{\mathrm{triads}}$) varies depending on the baseline length. In our case, the N_{triads} are 47, 32, 29 for 14, 24, and 28 m baselines, respectively. Please note that, when accurately measured, the 24 m baseline is $14\sqrt{(3)} \approx 24.25$ m; however, we chose the former for the simple denomination. On the other hand, the 14 and 28 m baselines are nearly accurate for the antenna positional tolerance of MWA tiles. Each dual-polarisation observation was made for 112 s, which included $N_{\mathrm{timestamps}} = 14$ each with 8 s of averaged data and a frequency resolution of 40 kHz, which provides a total of $N_{\mathrm{channels}} = 768$ frequency channels with a bandwidth of 30.72 MHz. The entire observations can be restructured into;

$$N_{\mathrm{obs}} \equiv \{N_{\mathrm{LST}}, N_{\mathrm{JD}}, N_{\mathrm{timestamps}},$$
$$N_{\mathrm{pol}}, N_{\mathrm{triads}}, N_{\mathrm{vis}}, N_{\mathrm{channels}}\} \quad (12)$$

5.2 RFI flagging

The MWA high band (167–197 MHz) lies in the digital television (DTV) broadcasting band; thus, we expect RFI to be present in our dataset (Offringa et al. 2015), which in some cases can completely dominate the useful data from the observations. As mentioned in the previous sections, since our analysis required keeping all the frequency channels from our datasets, we did not perform any frequency channel-based RFI flagging in the data preprocessing step. Instead, we incorporated SSINS (Wilensky et al. 2019), which is designed to detect faint RFI in the MWA data, to either discard the entire frequency band or keep it based on the RFI occupancy of the dataset. Instead of assuming a persistent RFI along a frequency channel, we check for the RFI along the observation time (i.e. along the $N_{\mathrm{timestamps}}$ axis).

The flagging was performed based on the RFI z-score of an observation. Note that the z-score was estimated at successive adjacent timestamps to measure if any faint or persistent RFI was present in the data across all timestamps (see Fig. 6). We took a z-score threshold of 2.5, below which the data was considered good, and the channels where the z-score exceeded the threshold were considered RFI-affected. Then, we independently estimated the level of such RFI-affected channels along the frequencies at each timestamp and checked if the RFI occupancy at a given timestamp was more than 5%. As a first step in selecting good timestamps, we chose an RFI occupancy level of 5% as a threshold. We discarded the entire timestamp if the RFI occupancy exceeded this threshold. Fig. 7 presents RFI occupancy for the entire dataset. Since SSINS z-scores are estimated relative to the successive adjacent timestamps, it might be difficult to quantify whether the RFI leakage between the adjacent timestamps (where one is good and another is bad) is there or not. Therefore, in the second step, we again flagged all the timestamps based on whether they had a bad neighbouring timestamp. The flagging provides a masked array, further propagated through other data processing steps.

5.3 Triad filtering

The presence of faulty tiles or dipoles/antennae corrupts the voltages recorded by the correlator; therefore, we are required to cross-check the visibilities at each antenna triad. The easiest way was to perform a geometric median-based rectification on the closure phase. We performed a two-step median rectification on the closure phase. For a given observation, the data structure of the triad filtering can be shown as follows:

$$\phi_\nabla \equiv \{N_{\mathrm{pol}}, N_{\mathrm{triads}}, N_{\mathrm{timestamps}}, N_{\mathrm{channels}}\}$$

First, we estimated the median absolute deviation (MAD = Median($|X_i - \bar{X}|$)) of the closure phases against the mean along the N_{triads},
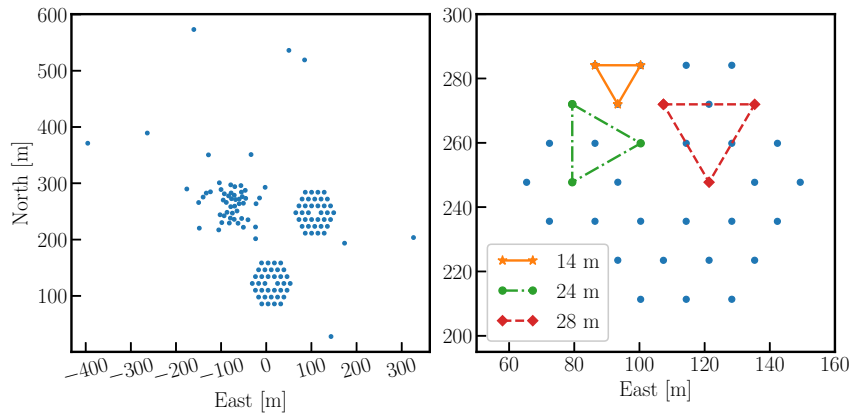
$$\phi_\nabla^{\mathrm{MAD}} \equiv \{N_{\mathrm{pol}}, N_{\mathrm{triads}}, N_{\mathrm{timestamps}}, N_{\mathrm{channels}}\}$$

and then we estimated the mean of the MAD (i.e. $\mathrm{MAD}_{\mathrm{mean}}$) along the N_{channels} axis. Finally, we estimated the MAD of the $\mathrm{MAD}_{\mathrm{mean}}$.
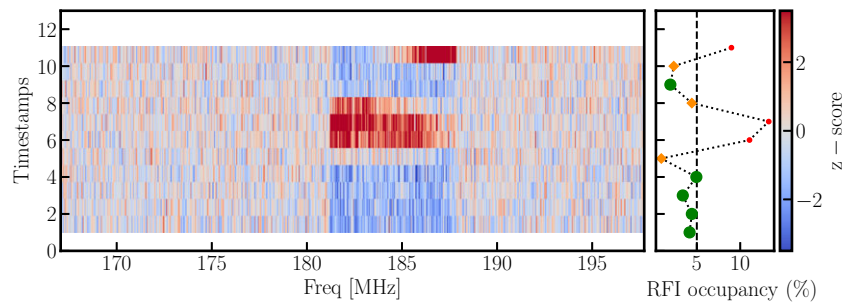
$$\mu\{\phi_\nabla^{\mathrm{MAD}}\} \equiv \{N_{\mathrm{pol}}, N_{\mathrm{triads}}, N_{\mathrm{timestamps}}, 1\}$$

This step provides a single value of the $\mathrm{MAD}_{\mathrm{mean}}$ for a given triad at every timestamp.

$$\mathrm{MAD}\{\mu\{\phi_\nabla^{\mathrm{MAD}}\}\} \equiv \{N_{\mathrm{pol}}, N_{\mathrm{triads}}, N_{\mathrm{timestamps}}, 1\}$$

Figure 5. Left: MWA phase II compact configuration. Right: Northern hexagon showing four equilateral triad configurations, 14 metres *(orange stars)*, 24 metres *(green circles)*, and 28 metres *(red diamonds)*.

Figure 6. *Left:* SSINS *z*-score for a visibly RFI-affected obsID for XX-polarisation and cross baselines. The first and last timestamps are avoided in the analysis. *Right:* A $z-$score threshold of 2.5 was chosen to identify RFI-affected channels and timestamps, then at the first iteration of RFI flagging (whole timestamp) was chosen based on the RFI occupancy of 5%. The timestamps corresponding to the *red* points were completely discarded in the first step, and the rest of the *orange* diamonds were considered good. *Right:* Since the two adjacent timestamps ($T_{[i+1]}, T_i$) is used to estimate the *z*-score, additional timestamps were flagged in the second iteration to finally get the good timestamps shown with *green* dots.
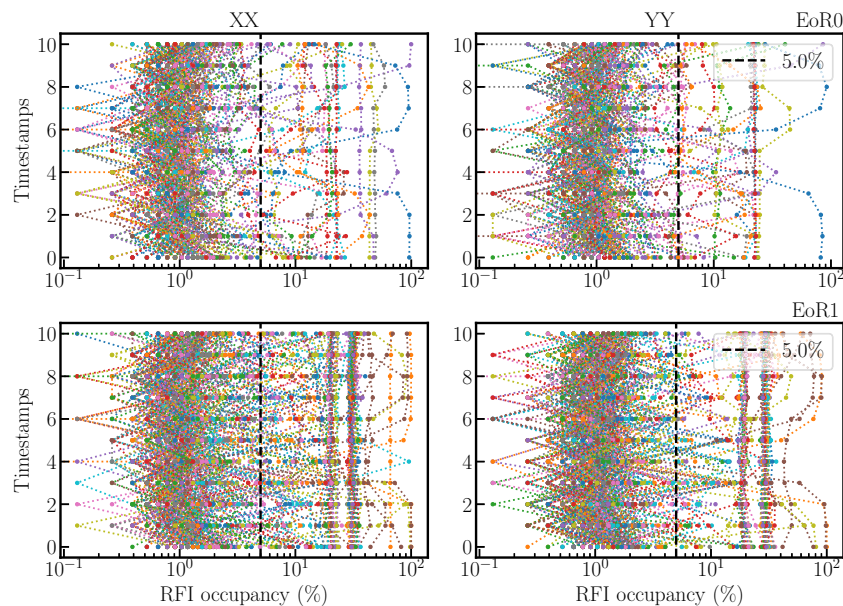
Figure 7. RFI occupancy levels of the full dataset used in this work were obtained using SSINS. For a *z*-score threshold of 2.5, the majority of the dataset (EoR0: XX-82.5%, YY-82.8%; EoR1: XX-78.8%, YY-80.9%) shows RFI occupancy < 5%.
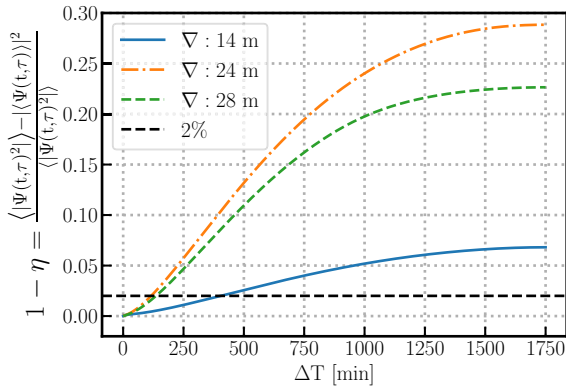
Figure 8. Fractional power loss for different triad configurations (shown by *coloured* lines). The 2% threshold level is shown with the *black* dashed horizontal line. The intersection between the threshold line and different triad loss provides an estimate of the time the data can be averaged coherently along LST. The coherent averaging time for the 14, 24, and 28 m baseline triads is roughly 408, 130, and 120 s, respectively.

Finally, we masked the triads if the mean of the MAD is greater than $1\sigma \approx 1.4826$ of MAD and considered the triad performing poorly at that given timestamp.

5.4 Coherent time averaging

The coherent averaging gives us an estimate of a timescale up to which the sky signal can be assumed identical and averaged coherently to improve sensitivity. It can be estimated by measuring the variation in the sky signal with time for a fixed pointing. Indeed, these vary with instrument and frequency of observation since the beam sizes are different. To check this with MWA, we used a continuous drifted sky simulation (FG and HI) under ideal observing settings (i.e. unity antenna gains, equal antenna element elevation from the ground) for ≈ 0.5 h while keeping a fixed zenith pointing. The sky moves about $\approx 7.5°$ in 0.5 h, which is less than the MWA beam size of $\approx 9° - 7.5°$ at the shortest (14 m) triad, thus justifying the simulation time range of 0.5 h.

We simply added the ideally simulated visibilities (FG and HI) and estimated the closure phase power spectrum as a function of time for higher delay ($|\tau| > 2\,\mu$s). We used a fractional signal loss of 2% to measure the coherence threshold, a similar approach used by [Keller et al.(2023)]. The fractional loss in power is defined as,

$$1 - \eta = \frac{\langle |\psi_\nabla(t, \tau)^2| \rangle - |\langle \psi_\nabla(t, \tau) \rangle|^2}{\langle |\psi_\nabla(t, \tau)^2| \rangle} \quad (13)$$

The choice of $|\tau| > 2\,\mu$s is to choose the timescale based on the loss of HI signal, which is where it would be significant, namely, the higher delay modes. Fig. 8 shows the fractional loss of coherent HI signal power as a function of averaging timescale for three triad configurations. We found the coherence averaging time for $\{\nabla:14, 24, 28\}$ triads to be approximately 408, 130, and 120 s, respectively.

6 Results

The bandpass structure of the MWA leaves significant systematics in the edge channels. Usually, these edge channels get flagged in the early data preprocessing step. However, we did not apply the
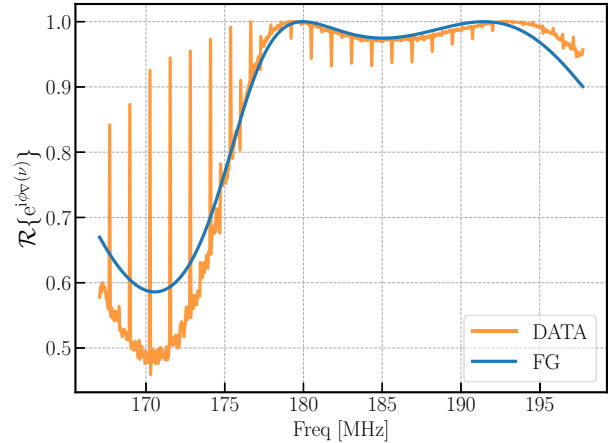
Figure 9. The real part of the complex exponent of the closure phase for XX polarisation 28 m baseline. The data matches with the foreground simulations. For the sanity check, the foreground simulation of the same is plotted over the data. Despite eliminating element-based bandpass gains, the data contains periodic spikes corresponding to the 1.28 MHz coarse channel edges of the MWA bandpass, indicating possible systematics of baseline-dependent origin.

flags since we require the full observing band for the delay spectrum analysis. The closure phase must be free of phase errors associated with the individual antenna elements. Thus, we expected the element-based bandpass structure to be removed in the closure phase. However, we still observed some residual bandpass structures in the closure phases; see Fig. 9. It can happen if there are some baseline-dependent gains present in the data in addition to antenna-dependent gains. We validated our hypothesis by developing a simple bandpass simulator, where we introduced an additional bandpass structure to MWA data that consisted of both element-based and baseline-dependent terms. The bandpass persists if baseline-dependent gains are present in the data, which otherwise gets completely removed if only antenna-based gains are present in the data (see Appendix Fig. A4). In total, about 128 frequency channels are affected by the bandpass, which accounts for about 16% of the total bandwidth.

6.1 Mitigation of baseline-dependent bandpass effects

We investigated two approaches to overcome the presence of baseline-dependent edge channel effects, the first being the Non-uniform Fast Fourier transform (NFFT), where we tried to avoid the bandpass-affected channels in the Fourier transform. The second Gaussian Process Regression (GPR) based data inpainting to estimate the missing channel information at the location of the spikes in the bandpass spectrum.

6.1.1 Gaussian Process Regression (GPR)

GPR (Wiener 1949; Rybicki & Press 1992; Robertson et al. 2015) is a popular supervised machine learning method. GPR has previously been used in foreground subtraction (Mertens, Ghosh, & Koopmans 2018; Ghosh et al. 2020), and data inpainting (Trott et al. 2020; Kern & Liu 2021) and characterisation (Pagano et al. 2023) for 21-cm cosmology studies. We follow a similar formalism to Kern & Liu (2021). In the first step, we identified all the bad edge channels in the bandpass and flagged them in the
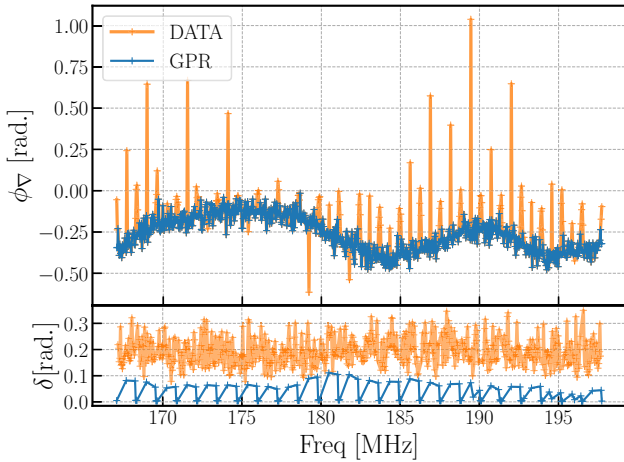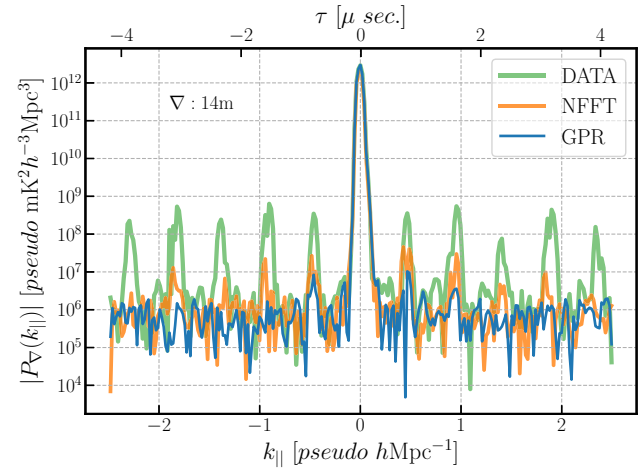
Figure 10. Top: bin-averaged closure phase of the data shown by *blue* and GPR reconstruction *orange* line. Bottom: RMS of the difference of closure phase. The edge channels were removed while calculating the RMS in the data, whereas only the edge channels were retained for the GPR.

Figure 11. Absolute cross-power spectra of the DATA, NFFT, and GPR are shown with *coloured* lines. MWA DATA suffers from baseline-dependent bandpass structure (see the regularly spaced spikes, which correspond to ≈ 1.28 MHz). NFFT significantly dampens the bandpass, but the spikes persist in the spectrum. GPR reconstruction shown by the *blue* line provides the cleanest spectra.

closure phase. As mentioned before, edge channel contamination is present in the MWA data. The data at 40 kHz resolution has 768 frequency channels and 128 contaminated edge channels. We implemented GPR on the complex exponent of the closure phase $\exp(i\phi_\nabla(\nu))$, with the real and imaginary components separately. The GPR implementation was rather simplistic since the complex phase varies between [0, 1]. We used the Matérn kernel as model covariance in our analysis. To optimise the kernel hyperparameters, we used the `scipy`-based `L−BFGS` (Liu & Nocedal 1989) to find the minima of the objective function. The optimisation was reiterated over ten times to ensure the kernel hyperparameters' convergence. Note that for a given frequency range, we applied GPR to the entire N_{obs} (see Equation 12) separately; therefore, the kernel Hyperparameters are also different for each closure phase frequency spectrum.

Fig. 10 shows the closure phase of the data and GPR reconstruction. In the top panel, the data can be seen with spikes at regularly spaced edge channels of ≈ 1.28 MHz intervals. The interpolated values of the closure phase are plotted over the data. The bottom panel shows the difference in the RMS in the closure phase along the frequency axis. Note that, while estimating the relative difference in the data closure phase, we avoided the noisy edge channels, whereas, for the GPR case, we only included the relative difference near the edge channels. This will enable us to query whether the GPR closure phase has a similar variation across the frequency compared to the data. It can be seen that the RMS of the data is higher than the GPR values, which means that the GPR has performed quite well. We used the Python-based module GPy[f] for the GPR implementation.

6.1.2 Non-uniform Fast Fourier transform (NFFT)

NFFT is a well-known method to get the Fourier transform of the data with missing samples (Dutt & Rokhlin 1993; Beylkin 1995). To estimate the FFT of bandpass-affected data, first, we removed the 128 edge channels from the data and estimated the $\Psi_\nabla(\nu)$

(a Fourier conjugate of Equation 4), which is then supplied to the NFFT function to get $\Psi_\nabla(\tau)$. We used a Python-based NFFT[g] package to develop the NFFT functions.

The absolute values of the closure phase delay cross-power spectrum are shown in Fig. 11. It can be seen that the data is highly affected by the excess systematics in the power, evident by the periodic spikes. NFFT significantly reduces the bandpass systematics but does not eliminate it entirely. Since the GPR performed best between the two methods, we adopted only the GPR for the later analysis. From now on, we will be using 'GPR-reconstructed DATA' as 'DATA' for the entirety of the paper.

6.2 Cross-power spectrum estimation

We proceeded to the power spectrum estimation after eliminating a significant part of the baseline-dependent bandpass contamination using GPR inpainting. Based on the LST-JD of the observations (see Fig. 1, we split the dataset equally along the JD axis (i.e. $N_{JD} = 2$). We binned the data along LST according to the coherent averaging time of MWA, which we already estimated for the redundant 14 m, 24 m, and 28 m baselines, resulting in 5, 14, and 17 LST bins, respectively. The first level of weighted averaging (i.e. coherently averaging) was done to the bispectrum (visibility triple product) and effective foreground visibilities, (V_{eff}), that lie within the same LST bin. The weights are the number of good observations left after being rectified by the RFI flagging and triad filtering within the given LST bin for a given polarisation and triad.

In the next step, we estimated the delay spectra, $\Psi_\nabla(\tau)$, from the phase of the LST binned averaged closure spectra at each LST bin. It resulted in delayed spectra data structure,
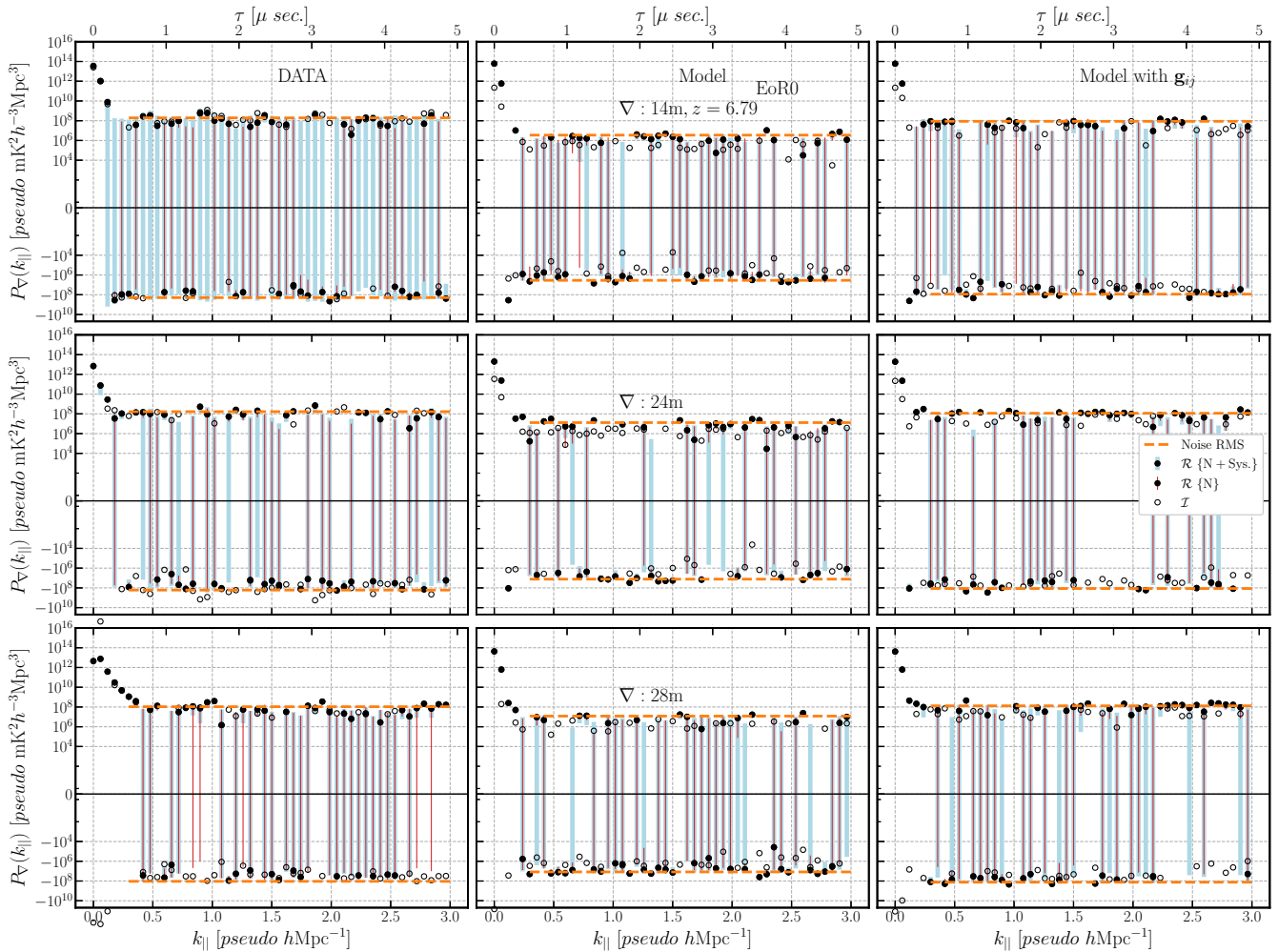
$$\{N_{LST}, N_{JD}, N_{pol}, N_{triads}, N_{delays}\}; N_{delays} = N_{channels}$$

Finally, we estimated the cross-power between the unique triad pairs of the delay spectra across the two JD bins according to

---

[f]https://github.com/SheffieldML/GPy.

[g]https://github.com/jakevdp/nfft.

Figure 12. Cross power spectrum of the closure phase delay spectrum for EoR0 observing field. The left panel represents the DATA, middle panel Model {FG + HI + noise} and the right Model with $\mathbf{g}_{ij}$. The top, middle, and bottom panels show the power spectra for 14, 24, and 28 m baseline lengths, respectively. The real part (*filled circles*) denotes the power, while the imaginary (*hollow circles*) represents the systematic level in the data. 2σ uncertainties are shown for two scenarios; the noise+Systematics are shown with *skyblue* error bars while noise-only is shown with *red* error bars. The RMS level for $k_{||} \geq 0.15$ *pseudo* $h\mathrm{Mpc}^{-1}$ is shown using *orange* dashed line. The noise+Systematics uncertainties are >10 times the noise-only uncertainties at low delays ($k_{||} < 0.5$ *pseudo* $h\mathrm{Mpc}^{-1}$), while fluctuating between > 4–8 times at higher delays.

Equation (6) where

$$\tilde{\Psi}_\nabla(\tau).\overline{\tilde{\Psi}'_\nabla(\tau)} = \frac{1}{^{N_{\mathrm{triads}}}C_2} \sum_{i,j}^{N_{\mathrm{triads}}} \tilde{\Psi}_\nabla(i,\tau).\overline{\tilde{\Psi}_\nabla(j,\tau)}, \quad i > j, \quad (14)$$

where i, j runs over N_{triads} (upper triangle of the $\{i, j\}$ pairs) from the first and second JD bins. The normalising factor of 2 arises since phases only capture half the power in the fluctuations. After this operation, the data structure of the binned averaged cross $P_\nabla(k_{||})$ becomes

$$\{N_{\mathrm{LST}}, {}^{N_{\mathrm{JD}}}C_2, N_{\mathrm{pol}}, {}^{N_{\mathrm{triads}}}C_2, N_{\mathrm{delays}}\}; {}^{N_{\mathrm{JD}}}C_2 = 1.$$

We took the weighted mean (i.e. incoherently averaged) along the triad pairs, where the weights were propagated from the previous step (refer to data flowchart Fig. A6 for details). As the sky varies with LST, we applied the inverse variance weights along the LST axis and averaged them to get the final estimates of the power spectrum. The same operation was done for the imaginary part of

the data to get an estimate of the level of systematics in the power spectrum.

Ultimately, we incoherently averaged the two polarisations and downsampled the original delays to the effective bandwidth of the applied window function (i.e. ≈ 12.9 MHz). We used Scipy-based BSpline to interpolate at the new downsampled delays. The final estimates of the closure phase power spectra are shown in Figs. 12 and 13 for the EoR0 and EoR1 fields, respectively.

6.3 Error estimation

The uncertainties on the power spectrum can be estimated in multiple ways. We primarily focused on estimating the uncertainties in two ways, the first being 'noise+systematics' and the second only noise, where we tried to mitigate the systematics.

6.3.1 Noise+systematics

To estimate uncertainties in power, we increased the number of samples that go into the uncertainty estimation by splitting the
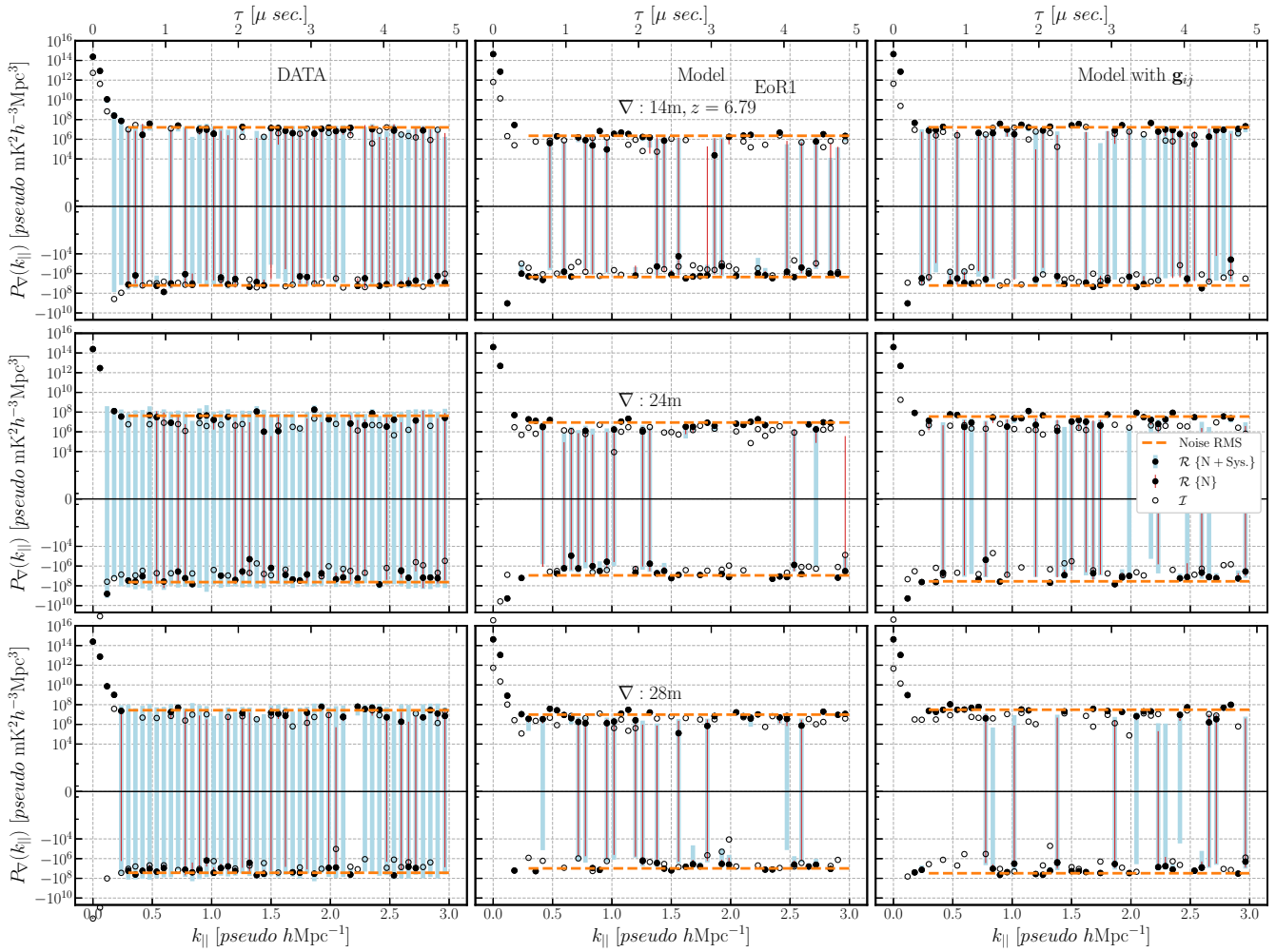
Figure 13. Same as Fig. 12 but for EoR1 observing field.

JD axis into four parts (i.e. $N_{JD} = 4$). This led to the data ($\Psi_\nabla(\tau)$) structure being $\{N_{LST}, N_{JD} = 4, N_{pol}, N_{triads}, N_{delays}\}$. We similarly estimated the cross-power of the N_{triads} along the unique pairs of N_{JD} axis. This operation provided us with $\{N_{pol}, N_{LST}, {}^{N_{JD}}C_2 = 6, {}^{N_{triads}}C_2, N_{delays}\}$ unique power spectra. The weighted mean power was estimated by along ${}^{N_{triads}}C_2$ where the weights are coming from the number of good observations that went into the Ψ_∇ for a given triad.

Next, We estimated the weighted variance on the power using the standard error of the weighted mean provided in Cochran (1977),

$$
\begin{aligned}
(\text{SEM}_{\text{wtd}})^2 = \frac{n}{(n-1)(\sum w_i)^2} & \Big[\sum (w_i X_i - \bar{w}\bar{X}_{\text{wtd}})^2 \\
- 2\bar{X}_{\text{wtd}} \sum (w_i - \bar{w})(w_i X_i - \bar{w}\bar{X}_{\text{wtd}}) & + \bar{X}_{\text{wtd}}^2 \sum (w_i - \bar{w})^2 \Big],
\end{aligned}
\tag{15}
$$

where w_i are the weights, and n represents the weight count. Note that since the $N_{JD} = 4$, we are required to normalise the variance.

Fig. A6 illustrates the detailed data structure flow for the noise estimation.

6.3.2 Noise

The uncertainties for the only noise case follow a similar procedure as the previous one, with the same JD split (i.e. $N_{JD} = 4$), the cross-power is estimated, which led to the data structure $\{N_{pol}, N_{LST}, {}^{N_{JD}}C_2 = 6, {}^{N_{triads}}C_2, N_{delays}\}$. After this operation, we took the difference in the power spectra between the unique pairs of JD (i.e. the same sky signal). The only noise scenario can be understood assuming the cross-power between the two independent JD bins correlates with the common signal and systematics across the triads. Assuming the power is coherent within the LST bin, the successive unique difference in the power within the LST bin eliminates the correlated power and systematics, leaving only the noise-like residuals. Also, since the differences eliminate the sky signal, the LST variation in the difference power is minimal; thus, we collapsed our datasets into a single axis and estimated the weighted standard deviation of the differenced power to get the final noise-like uncertainties. Finally, again, we took the weighted variance using Equation (15).

Table 1. 2-sided KS test comparison between the Data and Model, Data and Model with g_{ij} at $k_{||} > 0.15$ (*pseudo* hMpc^{-1}).

field	∇	Model		Model with g_{ij}	
		p-value	statistic	*p*-value	statistic
	14 m	6×10^{-6}	0.51	0.24	0.21
EoR0	24 m	8×10^{-4}	0.40	0.50	0.17
	28 m	2×10^{-6}	0.53	0.09	0.25
	14 m	2×10^{-3}	0.38	0.84	0.13
EoR1	24 m	3×10^{-2}	0.29	0.68	0.15
	28 m	8×10^{-3}	0.34	0.24	0.21

6.4 Validation

We performed a two-sided KS test on the closure phase power spectra of the data and two model variants for the statistical comparison. The null hypothesis was rejected in all scenarios with the Model (without baseline dependent gains); however, it failed to reject the null hypothesis in all scenarios when using the Model with g_{ij}. The former implies that the data and the model uncertainties were unlikely to be drawn from the same distribution, whereas the latter concludes the contrary. The test results are provided in Table 1. The test results for the Model are not unexpected since we can see that the RMS floor levels of the data and Model are sufficiently different, which match the data and Model with g_{ij}. We modelled the excess RMS in the data as arising from baseline-dependent gain factors, although it is believed to have originated from the systematics and residual RFI. Note that we do not claim that the excess noise in the data is solely due to the baseline dependent systematics; however, if the argument is true, the baseline dependent gains introduced in our analysis suffice for the excess power in the data.

6.5 21-cm power spectrum

We estimated the final dimensionless 21-cm power spectrum from the closure phase power spectrum. The closure phase delay power spectrum can be written into a 21-cm power spectrum ('*pseudo*') as follows:

$$\Delta_\nabla^2(k) = \frac{k^3 P_\nabla(k_{||})}{2\pi^2} [pseudo \text{ mK}^2] \tag{16}$$

where $k^2 = k_\perp^2 + k_{||}^2$, with $k_\perp = \frac{2\pi |b_\nabla|}{\lambda D}$, where b_∇ is the baseline length of the triad, and D is the cosmological comoving distance. Note that the 21-cm power spectrum estimates from the closure phase power spectrum should not be interpreted as true but rather the approximate representation of the actual 21-cm power spectrum (Thyagarajan et al. 2020; Keller et al. 2023). The power spectra converted to cosmological units for EoR0 and EoR1 fields are shown in Figs. 14 and 15, respectively.

Assuming the convergence to normal distribution due to the central limit theorem, we estimated 2σ (95% confidence intervals (CI)) using the uncertainties since our sample size was sufficient (>30). The upper limits on the 21-cm power spectrum (*pseudo* mK^2) were then estimated $\{\Delta_{\nabla\text{UL}}^2 = (\mu_{\Delta_\nabla^2} \pm$ CI) [*pseudo* mK^2]$\}$ for both the EoR0 and EoR1 fields are provided in the Table 2, and the full table in the (Appendix 1.5).

7. Discussion

We used the closure phase delay spectrum technique to obtain an independent estimate of the 21-cm power spectrum for the MWA phase II observations. These observations were centred on the EoR0 and EoR1 fields and were zenith-pointed, similar to the observing strategy of HERA. Our analysis revealed that MWA observations are possibly suffer from a baseline-dependent band-pass structure, which is especially noticeable in the edge channels. This bandpass structure results in structured bumps in the delay power spectrum (see Fig. 11), significantly contaminating the power spectrum. To address this issue, we explored two methods: Gaussian Process Regression (GPR) and Non-uniform Fast Fourier Transform (NFFT), to inpaint and mitigate the impact of the bandpass-affected edge channels on our power spectrum estimation. However, we decided against adopting the NFFT method because, although it reduced the magnitude of the bandpass, the bandpass remained evident in the NFFT delay spectra (see Fig. 11). Finally, we estimated the 21-cm power spectra using closure phase delay spectra. Additionally, we performed forward modelling in parallel with the observations to gain insights into the nature of the power spectrum under ideal observing conditions. The main findings of our analysis are summarised below.

When we averaged closure phases across multiple timestamps within the same Local Sidereal Time (LST) bin, we noticed a significant residual bandpass structure, particularly noticeable in the edge channels (see Fig. 9). Since closure phases are unaffected by element-based bandpass variations, we concluded that these bandpass issues cannot be simplified into element-based terms and could instead be dependent on the baseline. To test this hypothesis, we simulated the same effect on foreground visibilities (see Fig. A4). We explored data inpainting techniques to address these systematic errors to estimate closure phases in channels contaminated by baseline-dependent bandpass systematics (see Fig. 11). It is important to note that while these baseline-dependent issues are most noticeable in the edge channels, they could potentially affect all frequency channels since closure phases do not eliminate them. These issues also impact standard visibility-based power spectrum analysis methods. Understanding how the antenna layout contributes to such systematic errors is crucial for executing baseline-based mitigation strategies. Further investigation is needed to fully understand the extent to which these systematic errors affect MWA EoR power spectrum estimates. With a simple baseline dependent gains in the forward modelling, we aim to address the anomalies present in the DATA.

On comparing the final closure phase power spectrum of the DATA and Model for the EoR0 field (see Fig. 12), we found that the peak power (at $\tau = 0\mu s$) (i.e. $\approx 10^{14} pseudo$ mK$^2 h^{-3}$Mpc^3) of the DATA and Model only roughly matches for the 14 m triads, however the same for 24 and 28 m triads differ significantly. During the initial closure phase estimation stage, we found EoR0 data having multiple phase wraps, which could be due to the presence of some systematics or residual RFI. It caused an overall shift in the peak power away from zero delays in the coherent averaging. This effect can be seen in the 28 m triad, which shows greater power next to the zeroth delay (see Fig. 12). The RMS floor level is between 1-2 orders of magnitude higher in the DATA ($\approx 10^8$ *pseudo* mK$^2 h^{-3}$Mpc^3) compared to the Model ($\approx 10^7 pseudo$ mK$^2 h^{-3}$Mpc^3) and Model with g_{ij} ($\approx 10^8$ *pseudo* mK$^2 h^{-3}$Mpc^3). This excess power in the data compared to the Model may arise from a smaller DATA sample size in

Table 2. 2σ upper limits on 21-cm power spectrum (*pseudo* mK^2). The two estimates correspond to only-noise and noise+Systematics case.

	$\Delta^2_{\nabla \mathrm{UL}}$ (*pseudo* mK^2)											
Field	EoR0						EoR1					
Baseline	∇:14 m		∇:24 m		∇:28 m		∇:14 m		∇:24 m		∇:28 m	
k [*pseudo* hMpc^{-1}]	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.
0.18	–	$(392^*)^2$	$(188^*)^2$	$(207^*)^2$	–	–	–	$(526)^2$	–	$(361)^2$	–	–
0.24	$(347^*)^2$	$(420)^2$	–	–	–	–	–	$(427)^2$	–	$(458)^2$	$(236)^2$	$(314)^2$
0.30	–	$(534)^2$	–	–	–	–	$(218)^2$	$(263)^2$	–	$(503)^2$	–	$(512)^2$
0.36	$(490)^2$	$(608)^2$	–	–	–	–	$(184)^2$	$(330)^2$	–	$(849)^2$	–	$(572)^2$
0.42	–	$(1\,562)^2$	–	$(1\,065)^2$	$(732^*)^2$	$(708^*)^2$	$(474)^2$	$(434)^2$	–	$(1\,037)^2$	–	$(762)^2$
1.50	$(5\,275)^2$	$(7\,175)^2$	$(2\,466)^2$	$(3\,263)^2$	$(2\,525)^2$	$(3\,751)^2$	$(2\,926)^2$	$(2\,748)^2$	$(739)^2$	$(3\,695)^2$	$(3\,758)^2$	$(4\,835)^2$

^{a}k-modes where the uncertainty brackets do not include zero power are masked and shown with dashes (−).

^{b}best limits for a given baseline triad are shown with filled Gray box and unfilled boxes.

^{c}limits quoted with an asterisk (∗) might be affected by systematics or persistent residual RFI.
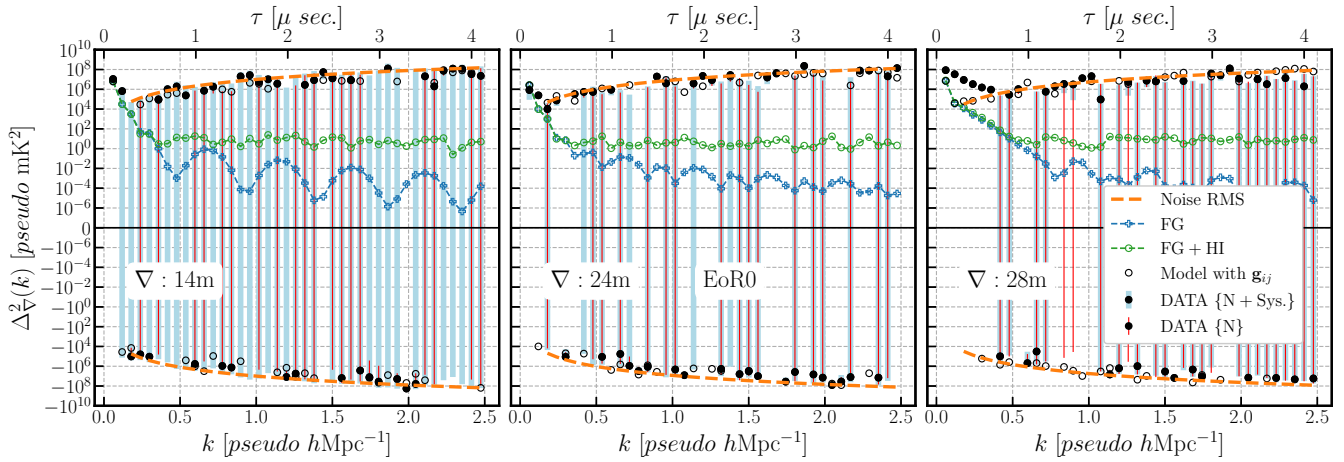
Figure 14. 21-cm power spectrum from 14 m (left), 24 m (middle), and 28 m equilateral triads for the EoR0 field.
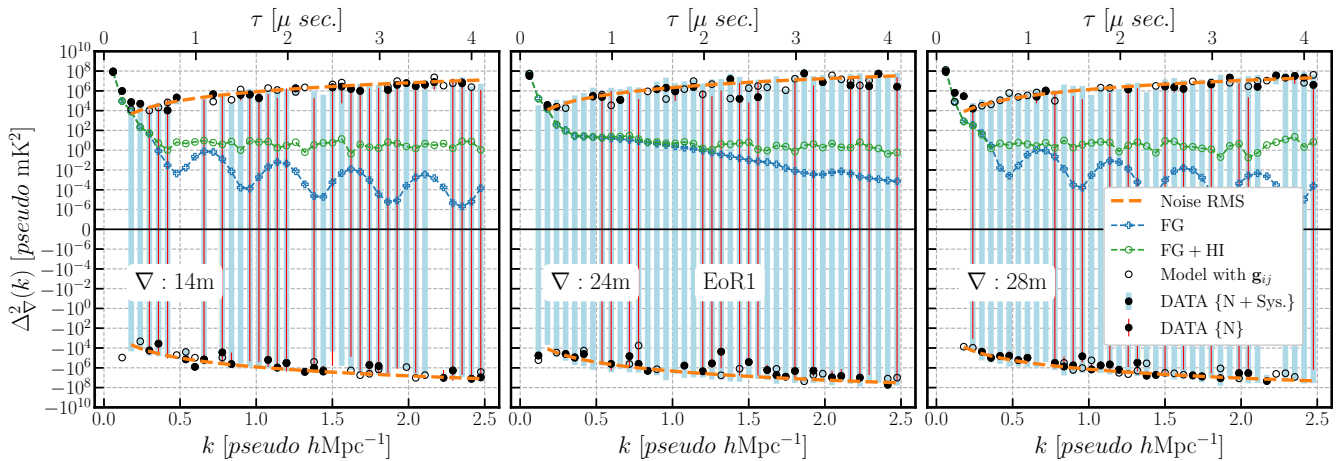
Figure 15. Same as Fig. 14 but for the EoR1 field.

the EoR0 field or systematics and residual RFI. Using a baseline-dependent gain factor in the simulation, we aimed to incorporate such systematics. We performed a 2-sided KS test on the DATA and Model at $k_{||} > 0.15 \, pseudo \, h\mathrm{Mpc}^{-1}$, which shows rejection of the null hypothesis for the likelihood of DATA and Model drawn from the same distribution at all baseline cases, which is expected since both differ significantly. In contrast, the KS-test setifies the null hypothesis when comparing DATA and Model with $\mathbf{g}_{ij}$.

In the closure phase power spectrum of the EoR1 field, the peak power of the DATA and Model ($\approx 10^{15} \, pseudo \, \mathrm{mK}^2 h^{-3}\mathrm{Mpc}^3$) match for all triads. The RMS floor level between the Model and DATA gets significantly better compared to the EoR0 field. They nearly match in all cases, except for the 14 m triads where the difference is approximately an order of magnitude higher in the DATA compared to the Model (see Fig. 13). It shows that we can improve our estimates of the power spectrum with an increased number of datasets. Thus, the analysis is data-limited for the amount of the data used. However, similar to the EoR0 case, the 2-sided KS test rejected the null hypothesis for all cases on Model, whereas fail to reject the null hypothesis in favour of the two distributions Model with $\mathbf{g}_{ij}$ might be drawn from the same distribution.

Since our observations lie in the middle of the DTV broadcasting band, we further investigated for residual RFI in our data. We shifted our entire analysis to the lower frequency band (167–177 MHz), avoiding the central DTV-affected band (although, as shown in figure no. 2 in Offringa et al. 2015, the DTV RFI nearly covers the entire EoR high-band observations). This analysis can help understand whether the residual RFI or other systematics are present in the data. Note that since we reduced our bandwidth by nearly three, we reduced our sensitivity by the same factor in the delay power spectrum. Therefore, the direct comparison of the mean power at $k_{||} > 0.15 \, pseudo \, h\mathrm{Mpc}^{-1}$ might not be valid with the previous result. The closure phase power spectrum for the shifted spectrum is shown in Figs. A2 and A3. We can see significant improvements in the peak power of the DATA and Model compared to previous results. However, the overall RMS level was increased by an order of magnitude in all cases, possibly due to the lesser sensitivity (lower sample size). The DATA peak power at $\tau = 0 \, \mu s$, especially in the EoR0 field, now matches the Model. Thus, we can justifiably argue that DTV RFI, which is expected to be prominent near 180 MHz, significantly contributed to the systematics present in the EoR0 DATA. On the other hand, EoR1 DATA seem relatively similar in both analyses, thus indicating that the systematics (such as persistent RFI) other than DTV RFI might be present in the data. Our findings are also confirmed when performing the KS-test on the DATA and Model, which shows non-rejection of the null hypothesis in all EoR1 field scenarios. We also compared the results with Model with $\mathbf{g}_{ij}$ which are shown in Table 3 shows KS-test outcomes of the Data and Model and Data and Model with $\mathbf{g}_{ij}$ on the shifted spectrum.

We estimated the 2σ (95% confidence interval assuming Gaussianity from the convergence to Central Limit Theorem) on the 21-cm power spectrum for both the EoR0 and EoR1 fields. Our best upper limit on the 21-cm power spectrum of $\lesssim (184 \, pseudo \, \mathrm{mK})^2$ came from EoR1 field on 14 m triads at $k = 0.36 \, pseudo \, h\mathrm{Mpc}^{-1}$ with the noise-only uncertainties. In the EoR0 field, our best estimate, $\lesssim (188 \, pseudo \, \mathrm{mK})^2$, came from the 24 m triads at $k = 0.18 \, pseudo \, h\mathrm{Mpc}^{-1}$ again using the noise-only uncertainty. However, as we discussed earlier, the systematics

Table 3. 2-sided KS test outcomes on DTV avoided band. The null hypothesis compares the Data and Model, Data and Model with $\mathbf{g}_{ij}$ at $k_{||} > 0.15 \, [pseudo \, h \, \mathrm{Mpc}^{-1}]$.

field	∇	2-sided KS-test			
		Model		Model with $\mathbf{g}_{ij}$	
		p-value	statistic	p-value	statistic
	14 m	0.04	0.54	0.98	0.15
EoR0	24 m	0.01	0.61	0.30	0.38
	28 m	0.04	0.54	0.13	0.46
	14 m	0.13	0.46	0.90	0.23
EoR1	24 m	0.59	0.30	0.30	0.38
	28 m	0.13	0.46	0.13	0.46

or residual RFI might have still affected these estimates, which we aim to address by introducing baseline-dependent gains in the modelling. It should be noted that, the exact nature of such baseline-dependent gains is not well understood. We have seen that, unlike Foregrounds, which usually gets restricted in lower delay modes, allowing faint HI signal to fluctuate visible at higher delay modes, the systematics equally affect all delay modes. Thus, for the scientific goal of observing milliradian-level sensitivity could be a significant challenge if such baseline dependent gains are present in the DATA. However, with extensive coherent averaging, the effect of such can be reduced. It should also be noted that the exact description of anomalies in the DATA can not be solely due to baseline-dependent systematics. Therefore, we state that if only the baseline-dependent systematics is present in the data, the level of noise introduced by the baseline-dependent gains justifies our DATA. Nevertheless, the level of fiducial HI and FG+HI powers are still lower than the data by 4–5 and 3–4 orders of magnitude, respectively; however, since our analysis is still data-limited, there is significant scope for improving upon the current estimates. Table 2 provides the best 2σ estimates while table Appendix 1.5 provides all estimates for each triad studied here.

8. Summary

We present independent EoR 21-cm power spectrum results from the closure phase analysis of ≈ 12 h of MWA phase II data in the frequency range 167–197 MHz on three redundant baseline groups, namely, 14, 24, and 28 m baselines. Using the closure phase diagnostic, we found evidence for significant baseline-dependent systematics in the MWA data. Our best estimates of the 21-cm power spectrum at $z = 6.79$ are $\lesssim (184 \, pseudo \, \mathrm{mK})^2$ at $k = 0.36 \, pseudo \, h\mathrm{Mpc}^{-1}$ in the EoR1 field using 14 m baseline triads and $\lesssim (188 \, pseudo \, \mathrm{mK})^2$ at $k = 0.18 \, pseudo \, h\mathrm{Mpc}^{-1}$ in the EoR0 field using 24 m baseline triads. Even with the limited amount of data analysed, our closure phase method shows significant promise in independently constraining the 21-cm power spectrum during the EoR. Our results are still data-limited; hence, there is scope for further improvement by including more data in the analysis. Extensive sky modelling, such as accounting for the Galactic diffuse emission, is required before directly comparing the closure phase analysis with the standard visibility-based power spectrum.

Data availability statement. This project was developed using Python and dependent libraries mentioned in the main text. Our data processing pipeline is publically available at `Github`,[h] along with the final processed datasets. The core data used in this work will be made available upon reasonable request.

References

Barry, N., et al. 2019, ApJ, 884, 1. https://doi.org/10.3847/1538-4357/ab40a8.

Beardsley, A. P., et al. 2016, ApJ, 833, 102. https://doi.org/10.3847/1538-4357/1/102. arXiv: 1608.06281 [astro-ph.IM].

Bernardi, G., et al. 2016, MNRAS, 461, 2847. issu: 1365-2966. https://doi.org/10.1093/mnras/stw1499.

Beylkin, G. 1995, ACHA, 2, 363. https://api.semanticscholar.org/CorpusID: 120066320.

Bosman, SARAH E. I., et al. 2022, MNRAS, 514, 55. https://doi.org/10.1093/mnras/stac1046. arXiv: 2108.03699 [astro-ph.CO].

Bowman, J. D., & Rogers, A. E. E. 2010, Natur, 468, 796. https://doi.org/10.1038/nature09601.

Bowman, J. D., Rogers, A. E. E., & Hewitt, J. N. 2008, ApJ 676, 1. https://doi.org/10.1086/528675. arXiv: 0710.2541 [astro-ph].

Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. 2018, Natur, 555, 67. https://doi.org/10.1038/nature25792. arXiv: 1810.05912 [astro-ph.CO].

Carilli, C. L., Nikolic, B., Thyagarajan, N., & Gale-Sides, K. 2018, RS, 53, 845. issu: 0048-6604. https://doi.org/10.1029/2018rs006537.

Cochran, W. G. 1977, Sampling Techniques (John Wiley & Sons).

Contents of Volume. 2006, PhR, 433, 302. issu: 0370-1573. https://doi.org/10.1016/S0370-1573(06)00351-6. https://www.sciencedirect.com/science/article/pii/S0370157306003516.

DeBoer, D. R., et al. 2017, PASP, 129, 045001. https://doi.org/10.1088/1538-3873/129/974/045001. arXiv: 1606.07473 [astro-ph.IM].

Dutt, A., & Rokhlin, V. 1993, SIAM JSC, 14, 1368. https://doi.org/10.1137/0914081.

Eastwood, M. W., et al. 2019, AJ, 158, 84. https://doi.org/10.3847/1538-3881/ab2629.

Furlanetto, S. R., Sokasian, A., & Hernquist, L. 2004, MNRAS, 347, 187. issu: 0035-8711. https://doi.org/10.1111/j.1365-2966.2004.07187.x.

Ghosh, A., et al. 2020, MNRAS, 495, 2813. https://doi.org/10.1093/mnras/staa1331. arXiv: 2004.06041 [astro-ph.CO].

Harris, F. J. 1978, IEEE Proc., 66, 51.

Jennison, R. C. 1958, MNRAS, 118, 276.

Keller, P. M., et al. 2023, arXiv e-prints: arXiv:2302.07969. https://doi.org/10.48550/arXiv.2302.07969. arXiv: 2302.07969 [astro-ph.CO].

Kern, NICHOLAS S., & Liu, A. 2021, MNRAS, 501, 1463. https://doi.org/10.1093/mnras/staa3736. arXiv: 2010.15892 [astro-ph.CO].

Line, J. L. B., Webster, R. L., Pindor, B., Mitchell, D. A., & Trott, C. M. 2017, PASA, 34, e003. https://doi.org/10.1017/pasa.2016.58.

Line, J. L. B. 2022, JOSS, 7, 3676. https://doi.org/10.21105/joss.03676.

Liu, D. C., & Nocedal, J. 1989, MP, 45, 503.

Mertens, F. G., et al. 2020, MNRAS, 493, 1662. issu: 1365-2966. https://doi.org/10.1093/mnras/staa327.

Mertens, F. G., Ghosh, A., & Koopmans, L. V. E. 2018, MNRAS, 478, 3640. https://doi.org/10.1093/mnras/sty1207. arXiv: 1711.10834 [astro-ph.CO].

Mesinger, A., ed. 2016, Understanding the Epoch of Cosmic Reionization (Vol. 423), Astrophysics and Space Science Library. https://doi.org/10.1007/978-3-319-21957-8.

Mesinger, A., Furlanetto, S., & Cen, R. 2010, MNRAS, 411, 955. issu: 0035-8711. https://doi.org/10.1111/j.1365-2966.2010.17731.x.

Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7. https://doi.org/10.1086/424437. arXiv: astro-ph/0312437 [astro-ph].

Munshi, S., et al. 2023, arXiv: 2311.05364 [astro-ph.CO].

Offringa, A. R., et al. 2015, PASA, 32. https://doi.org/10.1017/pasa.2015.7.

Paciga, G., et al. 2013, MNRAS, 433, 639. https://doi.org/10.1093/mnras/stt753. arXiv: 1301.5906 [astro-ph.CO].

Pagano, M., et al. 2023, MNRAS, 520, 5552. https://doi.org/10.1093/mnras/stad441. arXiv: 2210.14927 [astro-ph.IM].

Parsons, A. R., et al. 2014, ApJ, 788, 106. https://doi.org/10.1088/0004-637X/788/2/106. arXiv: 1304. 4991 [astro-ph.CO].

Patra, N., Subrahmanyan, R., Raghunathan, A., & Rao, U. N. 2012, ExA, 36. https://doi.org/10.1007/s10686-013-9336-3.

Patwa, A. K., Sethi, S., & Dwarakanath, K. S. 2021, MNRAS, 504, 2062. https://doi.org/10.1093/mnras/stab989. arXiv: 2104.03321 [astro-ph.CO].

Philip, L., et al. 2019, JoAI, 08, 1950004. https://doi.org/10.1142/S2251171719500041.

Pober, J., et al. 2011, American Astronomical Society Meeting Abstracts, 217, 432.06

Pritchard, J. R., & Loeb, A. 2012, RPP, 75, 086901. https://doi.org/10.1088/0034-4885/75/8/086901. arXiv: 1109.6012 [astro-ph.CO].

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19. https://doi.org/10.1088/2041-8205/802/2/L19. arXiv: 1502.02024 [astro-ph.CO].

Rybicki, G. B., & Press, W. H. 1992, ApJ, 398, 169. https://doi.org/10.1086/171845.

Sigel, D. A., Bach, V. M., Thomson, M. W., Bradley, R. F., Amaro, L. R., Lazio, J., & Burns, J. O. 2013, https://doi.org/10.1109/USNC-URSI-NRSM.2013.6525025.

Singh, S., et al. 2022, Natur, 6, 607. https://doi.org/10.1038/s41550-022-01610-5. arXiv: 2112.06778 [astro-ph.CO].

Singh, S., et al. 2018, ApJ, 858, 54. https://doi.org/10.3847/1538-4357/aabae1. arXiv: 1711.11281 [astro-ph.CO].

Sokolowski, M., et al. 2015, PASA, 32, e004. https://doi.org/10.1017/pasa.2015.3. arXiv: 1501.02922 [astro-ph.IM].

THE HERA Collaboration, et al. 2021, arXiv e-prints: arXiv: 2108.02263. arXiv: 2108.02263 [astro-ph.CO].

Thompson, A. R., Moran, J. M., & Swenson Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy (3rd edn.). https://doi.org/10.1007/978-3-319-44431-4.

Thyagarajan, N., & Carilli, C. 2020, PhRvD, 102, 022001. https://doi.org/10.1103/PhysRevD.102.022001. arXiv: 2005.10274 [astro-ph.CO].

Thyagarajan, N., et al. 2020, PhRvD, 102, 022002. https://doi.org/10.1103/PhysRevD.102.022002. arXiv: 2005.10275 [astro-ph.CO].

Thyagarajan, N., Carilli, C., & Nikolic, B. 2018, PhRvL, 120, 251301. https://doi.org/10.1103/PhysRevLett.120.251301. arXiv: 1805.00954 [astro-ph.CO].

---

[h]https://github.com/himmng/Closure_phase_analysis.

Thyagarajan, N., Parsons, A. R., DeBoer, D. R., Bowman, J. D., Ewall-Wice, A. M., Neben, A. R., & Patra, N. 2016, ApJ, 825, 9. https://doi.org/10.3847/0004-637X/825/1/9.

Tingay, S. J., et al. 2013, PASA, 30. issu: 1448-6083. https://doi.org/10.1017/pasa.2012.007.

Trott, C. M., et al. 2020, MNRAS, 493, 4711. https://doi.org/10.1093/mnras/staa414. arXiv: 2002.02575 [astro-ph.CO].

Ung, D. C. X., Sokolowski, M., Sutinjo, A. T., & Davidson, D. B. 2020, IEEE TAP, 68, 5395. https://doi.org/10.1109/TAP.2020.2980334. arXiv: 2003.05116 [astro-ph.IM].

van Haarlem, et al. 2013, AnA, 556, A2. https://doi.org/10.1051/0004-6361/201220873.

Voytek, T. C., Natarajan, A., Garcia, J. M. J., Peterson, J. B., & Lopez-Cruz, O. 2014, ApJ, 782, L9. https://doi.org/10.1088/2041-8205/782/1/l9.

Wiener, N. 1949, Extrapolation, Interpolation, and Smoothing of Stationary Time Series: with Engineering Applications (The MIT Press). isbu: 9780262257190. https://doi.org/10.7551/mitpress/2946.001.0001. eprint: https://direct.mit.edu/book-pdf /2161108/book_9780262257190.pdf.

Wilensky, M. J., Morales, M. F., Hazelton, B. J., Barry, N., Byrne, R., & Roy, S. 2019, PASP, 131, 114507. issu: 1538-3873. https://doi.org/10.1088/1538-3873/ab3cad.

Yoshiura, S., et al. 2021, MNRAS, 505, 4775. issu: 0035-8711. https://doi.org/10.1093/mnras/stab1560.

Zhu, Y., et al. 2024, MNRAS, 533, L49. https://doi.org/10.1093/mnrasl/slae061. arXiv: 2405.12275 [astro-ph.CO].

Zhu, Y., et al. 2022, ApJ, 932, 76. https://doi.org/10.3847/1538-4357/ac6e60. arXiv: 2205.04569 [astro-ph.CO].

Appendix 1. Appendix

Appendix 1.1. Inner vs Outer triads

The core of MWA Hexagons may be more affected by mutual coupling than the tiles near the edges. Therefore, we tried to perform an independent check on the power differences between the inner and outer tile triads to quantify the cross-talk level in the data. We chose 14m baseline triads to get a higher count and only estimated the closure phase power spectrum from them. We used first and second outer layer of the Hexagon configuration as the outer triads, and third and fourth tile layers as inner triads (see right panel of Fig. 5). In scenarios, where either two tiles from outer while the third from inner, or vice versa, we considered those triads to be outer or inner triads, respectively. Fig. A1 shows the closure phase power spectrum of the inner and outer triads at 14m baseline lengths. We observed the relative percentage difference between the RMS power estimated for $k_{||} > 0.15$ $pseudo\ h$ Mpc^{-1} between inner triads is higher than the outer triads are about 41%, 34% for EoR0 and EoR1 fields, respectively. However, the relative difference with the Model RMS (Figs. A2, A3) $\approx$ 6%, -25% for inner and outer triads in EoR0 field, and $\approx$ 186%, 114% for inner and outer triads in EoR1 field. These when comparing with the mean RMS value of the Model ($\approx 10^8$, $10^7 pseudo$ mK$^2 h^{-3}$Mpc^3 for E0R0, EoR1 fields) consistent with each other.

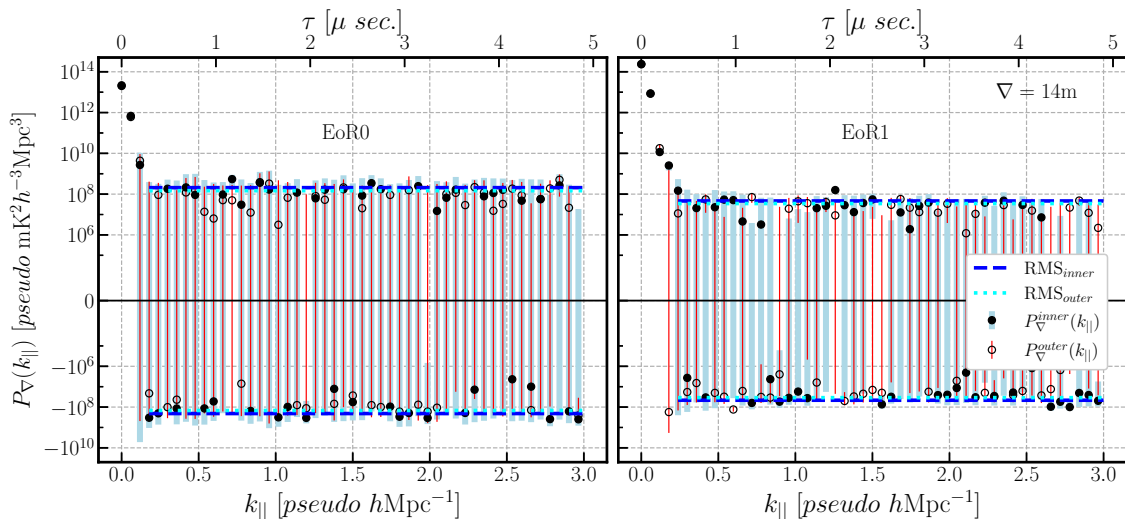**Figure A1.** Power spectrum comparison between the Inner (filled circles) and Outer (empty circles) in 14 m triads for EoR0 (*Left*) and EoR1 (*Right*) fields, respectively.

Appendix 1.2. Avoiding DTV

The observations are centred around 180 MHz, which overlap with the Australian Digital Television (DTV) band; therefore, despite the extensive RFI flagging, some RFI could remain in the final processed data. Therefore, a useful test would be to shift the spectral window to lower frequencies, avoiding the DTV band in the analysis.

To do this, we worked with the first ≈ 10 MHz band (167–177 MHz) of the total 30.72 MHz bandwidth and estimated the cross-power spectrum using the same procedure. However, the effective bandwidth for this analysis was reduced to ≈ 3.3 MHz,

thus reducing the sensitivity by the same factor. The idea was to check whether the floor level of the power spectra (refer to Figs. 12, 13), which shows the excess power in the DATA, reduces and matches with the Model in the shifted window. It would suggest that the RFI could be localised in the central portion of the band, which is one of the major contributors to the excess power. However, the power spectra of shifted window (see Figs. A2 and A3) show similar behaviour as that in Figs. 12 and 13, indicating that the level of the RFI might be ubiquitous across the entire observing band along with the systematics in the data which are not being modelled in the simulations.
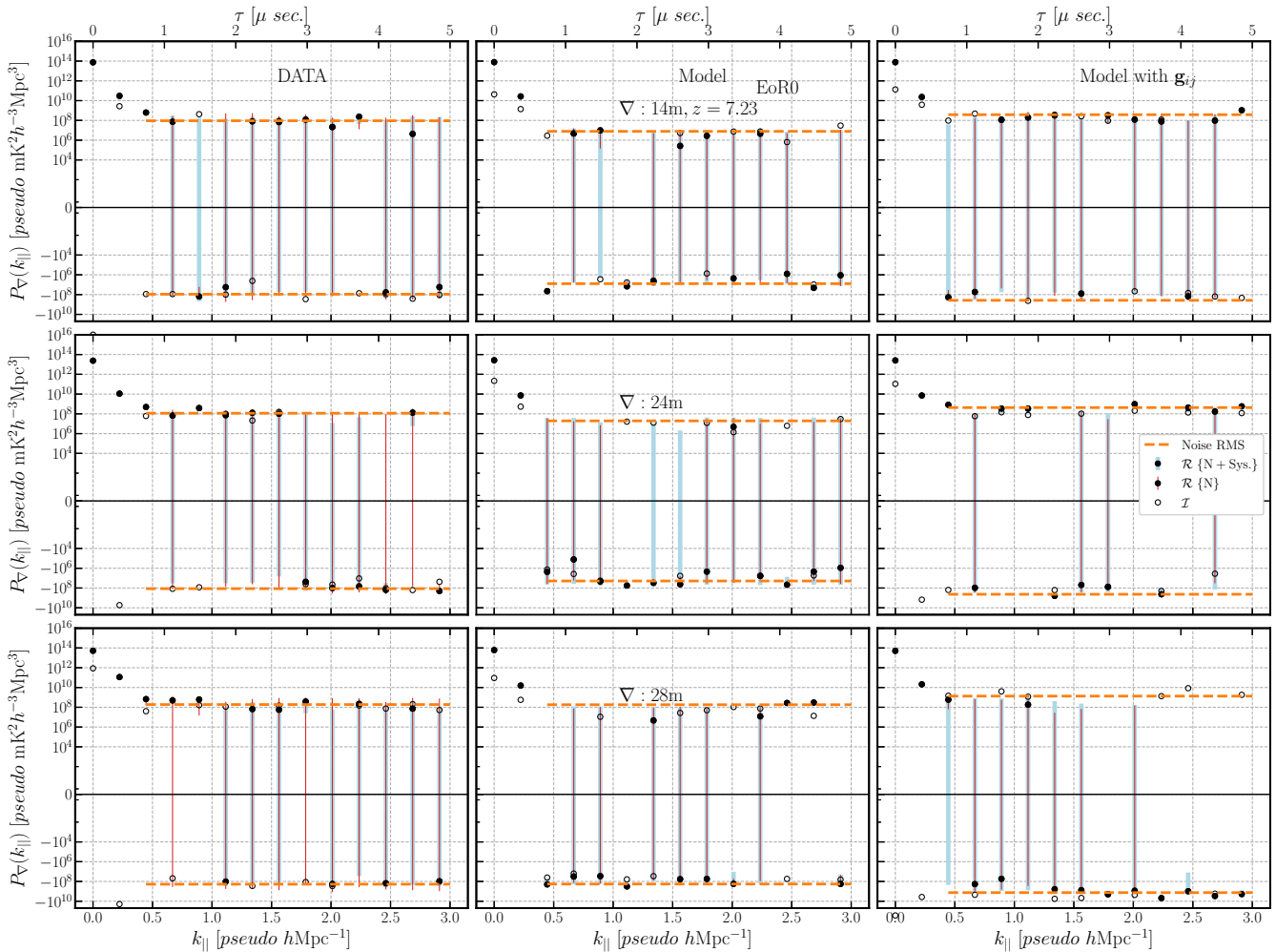
Figure A2. Cross power spectrum of the closure phase delay spectrum for EoR0 observing field when the window function is shifted towards lower frequencies (167–177) MHz to avoid the DTV frequency band around 180 MHz. All symbols, colours, and line styles are the same as in Fig. 12.
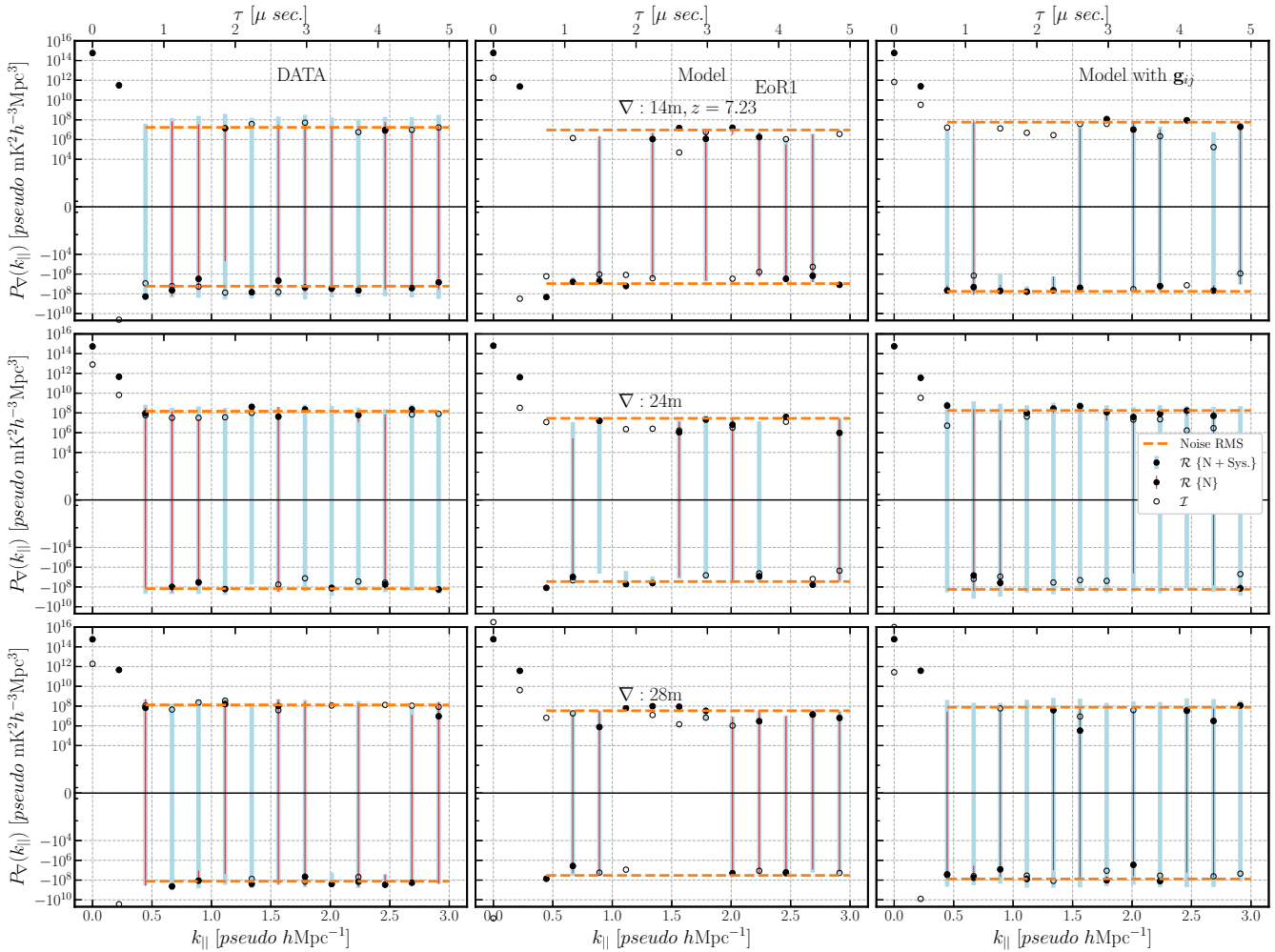
Figure A3. Same as Fig. A2 but for the EoR1 field.

Appendix 1.3. *Diagnosing bandpass systematics*

We checked the closure phases for the baseline-dependent systematics, which persisted in the MWA data by modifying the antenna element-based gains ($\mathbf{g}'_i$s Equation 3). First, we created one set of MWA bandpass using random Gaussian between [0, 1] for each edge-channel frequency. We multiplied these by each visibilities in the correct parity pair order. It provides us with modified visibilities with randomised gains at the edge channels, which mimics the bandpass-affected visibilities. We used simulated flux densities for this procedure since they were produced ideally with unity antenna gains (although it does not affect having unity gains in the first place).

Second, we used two scenarios to modify the bandpass further. In the first, the antenna element-based gains were modified with new gains multiplied by the existing ones. This step verifies how the individual antenna-based gains vanish in the closure phase, illustrated in the top panel of Fig. A4. In the second scenario, instead of multiplicative gains from individual antenna elements (e.g. $\mathbf{g}_i$, $\mathbf{g}_j$), we multiplied an additional baseline-dependent term ($\mathbf{g}_{ij}$) that is not factorisable into element-based terms. It demonstrated that baseline-dependent gains do not cancel in the closure phase; see the same Fig. A4 bottom panel, where the residual difference between the residuals between the closure phase modified with $\mathbf{g}_{ij}$ and the original does not vanish.
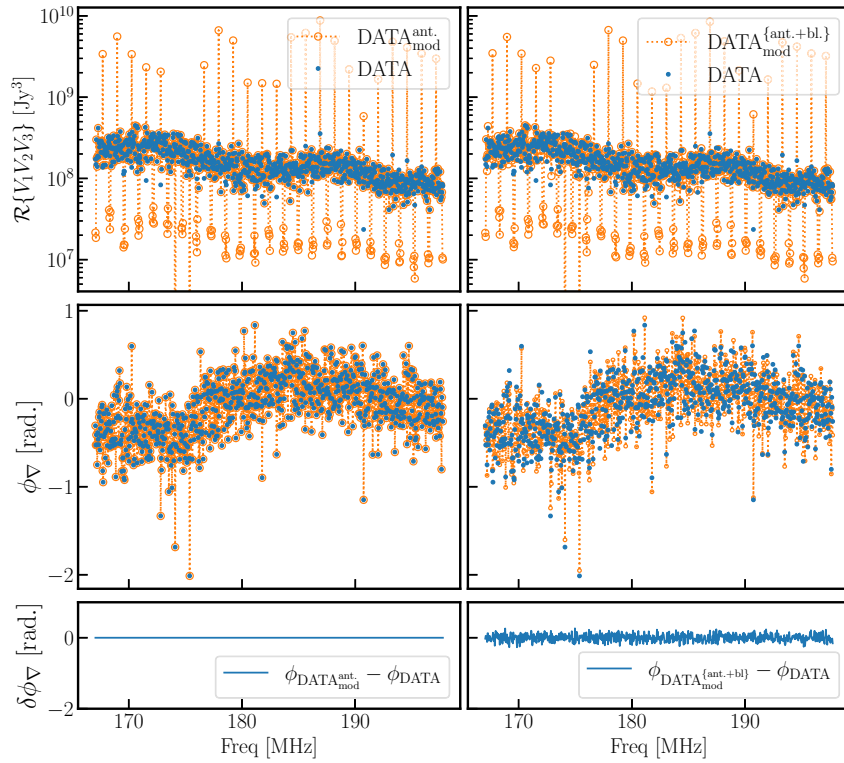
Figure A4. Antenna element gains and baseline-dependent gains inserted in the data at 1.28 MHz regular intervals. The top-bottom panel shows the bispectrum, closure phase, and residual closure phase between the data and modified data. *Left*: Showing antenna element gains getting eliminated in the closure phase. *Right*: baseline-dependent gains persist in the closure phase.

Appendix 1.4. Incomplete modelling

Realistic vs. ideal beams or using fewer foreground sources can affect the final power estimates. Thus, we produced two test foreground simulations, one with the same 20 000 foreground sources but with ideal beam conditions where all dipole gains are active and set to unity (FGI_{20k}) and a second with only 5,000 foreground sources and real beam conditions, dead/missing dipoles incorporated in the beam evaluation (FGR_{5k}), which we compared against 20 000 foreground sources but with real beam conditions (FGR_{20k}). Note that in the main results, we used 20 000 foreground sources with real beams, which are compared against the two test scenarios. The three cases' final closure phase power spectrum,

foreground with real dipole gains with 20 000 sources, foreground with unity dipole gains with 20 000 sources, and foreground with real dipole gains with 5 000 sources, are shown in Fig. A5. We obtained RMS power at ≥ 1.0 *pseudo* $h\mathrm{Mpc}^{-1}$ to differentiate the two scenarios with the foreground with real dipole gains with 20 000 sources. The relative percentage error for unity dipole gains (20 000 sources) at 14, 24 m, and 28 m baselines are 4%, 39%, 0.3%, and for real dipole gains (5 000 sources) at 14, 24, 28 m baselines are 1.7%, 1.9%, 40%, respectively. Comparing with the RMS value of the Model from Figs. A2, A3) ($\approx 10^8$, 10^7 *pseudo* $\mathrm{mK}^2 h^{-3}\mathrm{Mpc}^3$ for E0R0, EoR1 fields), these relative differences in both the scenarios are sufficiently less, thus, consistent with the Model.
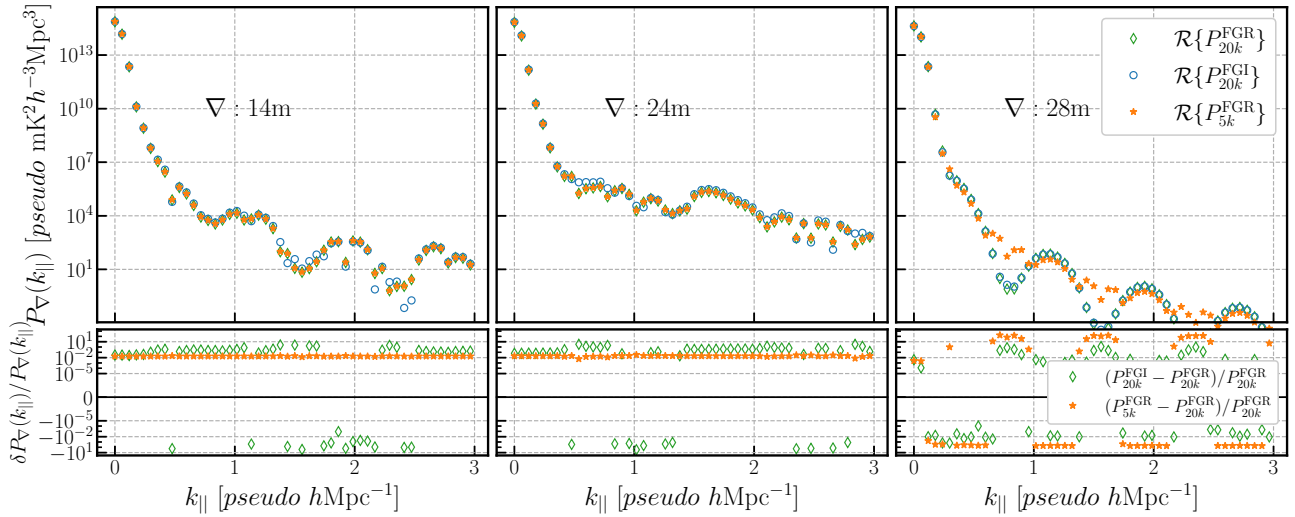
Figure A5. Top: Foreground model power for three scenarios, Model with 20 000 sources with real dipole gains, Model with 20 000 sources with unity (*ideal*) dipole gains, foreground model with 5 000 sources with real dipole gains in EoR1 field. Bottom: The relative difference between the unity dipole gains model (20 000 sources), and real dipole gains (5 000 sources) at 14, 24, 28 m with real dipole gains model (20 000 sources).

Appendix 1.5. Data Structure Flowchart

As we proceed through the different averaging steps, the data structure is shown as a flowchart in Fig. A6.

Table A1. Complete table of 2σ upper limit estimates of 21-cm power spectrum (*pseudo* mK^2).

	$\Delta^2_{\nabla\mathrm{UL}}$ (*pseudo* mK^2)											
Field	EoR0						EoR1					
Baseline	∇:14 m		∇:24 m		∇:28 m		∇:14 m		∇:24 m		∇:28 m	
k (*pseudo* h Mpc^{-1})	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.
0.18	–	$(392^*)^2$	$(188^*)^2$	$(207^*)^2$	–	–	–	$(526)^2$	–	$(361)^2$	–	–
0.24	$(347^*)^2$	$(420^*)^2$	–	–	–	–	–	$(427)^2$	–	$(458)^2$	$(236)^2$	$(314)^2$
0.30	–	$(534^*)^2$	–	–	–	–	$(218)^2$	$(263)^2$	–	$(503)^2$	–	$(512)^2$
0.36	$(490^*)^2$	$(608^*)^2$	–	–	–	–	$(184)^2$	$(330)^2$	–	$(849)^2$	–	$(572)^2$
0.42	–	$(1\,562^*)^2$	–	$(1065^*)^2$	$(732^*)^2$	$(708^*)^2$	$(474)^2$	$(434)^2$	–	$(1037)^2$	–	$(762)^2$
0.48	–	$(2391)^2$	$(1\,133)^2$	$(1358)^2$	$(834)^2$	$(943)^2$	–	–	–	$(1\,888)^2$	–	$(1\,082)^2$
0.54	–	$(1\,593)^2$	$(658)^2$	$(772)^2$	–	–	–	–	$(1252)^2$	$(1\,336)^2$	–	$(998)^2$
0.60	$(1\,152)^2$	$(1\,525)^2$	–	–	–	–	–	–	$(1\,531)^2$	$(2\,013)^2$	–	$(1\,129)^2$
0.66	$(1\,329)^2$	$(1\,830)^2$	$(724)^2$	$(932)^2$	$(826)^2$	$(989)^2$	$(687)^2$	$(720)^2$	–	$(1\,610)^2$	–	$(1\,639)^2$
0.72	–	$(2\,298)^2$	–	$(1\,098)^2$	$(1\,436)^2$	$(1\,539)^2$	–	–	$(1\,083)^2$	$(2\,148)^2$	–	$(1\,875)^2$
0.78	$(1\,252)^2$	$(2\,312)^2$	–	–	–	–	$(785)^2$	$(855)^2$	$(958)^2$	$(2\,090)^2$	$(1\,203)^2$	$(1\,509)^2$
0.84	$(1\,407)^2$	$(2\,284)^2$	$(1\,660)^2$	$(1\,771)^2$	–	–	–	$(884)^2$	–	$(2\,008)^2$	–	$(2\,902)^2$
0.90	–	$(7\,324)^2$	–	–	–	–	–	$(1\,391)^2$	–	$(3\,798)^2$	$(1\,193)^2$	$(2\,520)^2$
0.96	–	$(8\,728)^2$	$(3\,935)^2$	$(4\,599)^2$	–	–	$(1\,325)^2$	$(1\,647)^2$	–	$(6\,085)^2$	$(598)^2$	$(2\,326)^2$
1.02	$(3\,608)^2$	$(5\,318)^2$	$(2\,004)^2$	$(2\,214)^2$	–	–	$(1\,045)^2$	$(1\,325)^2$	–	$(3\,104)^2$	–	$(1\,753)^2$
1.08	–	$(5\,048)^2$	–	–	$(1\,961)^2$	$(2\,473)^2$	$(1\,024)^2$	$(1\,212)^2$	–	$(4\,041)^2$	–	$(2\,176)^2$
1.14	$(3\,273)^2$	$(4\,149)^2$	–	$(3\,152)^2$	–	–	$(1\,375)^2$	$(1\,648)^2$	–	$(4\,893)^2$	$(1\,327)^2$	$(2\,708)^2$
1.20	–	$(3\,694)^2$	–	–	$(2\,501)^2$	$(2\,524)^2$	$(1\,828)^2$	$(1\,758)^2$	$(2\,342)^2$	$(3\,550)^2$	–	$(2\,273)^2$
1.26	$(2\,841)^2$	$(5\,036)^2$	–	–	–	–	–	–	$(1\,038)^2$	$(4\,967)^2$	$(2\,468)^2$	$(3\,158)^2$

Table A1. (Continued)

	$\Delta^2_{\nabla \mathrm{UL}}$ (*pseudo* mK^2)											
Field	EoR0						EoR1					
Baseline	∇:14 m		∇:24 m		∇:28 m		∇:14 m		∇:24 m		∇:28 m	
k (*pseudo h* Mpc^{-1})	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.	N.	N.+Sys.
1.32	$(5\,160)^2$	$(6\,225)^2$	$(2\,393)^2$	$(2\,832)^2$	$(2\,412)^2$	$(2\,996)^2$	–	–	$(1\,360)^2$	$(5\,115)^2$	$(1\,885)^2$	$(3\,286)^2$
1.38	$(6\,337)^2$	$(8\,341)^2$	–	$(7\,501)^2$	–	–	$(1\,393)^2$	$(1\,788)^2$	–	$(9\,528)^2$	–	$(3\,420)^2$
1.44	–	–	$(4\,462)^2$	$(4\,316)^2$	$(5\,041)^2$	$(5\,977)^2$	–	$(2\,156)^2$	$(3\,262)^2$	$(6\,583)^2$	–	$(3\,391)^2$
1.50	$(5\,275)^2$	$(7\,175)^2$	$(2\,466)^2$	$(3\,263)^2$	$(2\,525)^2$	$(3\,751)^2$	$(2\,926)^2$	$(2\,748)^2$	$(739)^2$	$(3\,695)^2$	$(3\,758)^2$	$(4\,835)^2$
1.56	$(5\,722)^2$	$(6\,703)^2$	$(3\,258)^2$	$(3\,487)^2$	–	–	–	$(3\,149)^2$	$(3\,273)^2$	$(4\,824)^2$	–	$(5\,720)^2$
1.62	$(6\,298)^2$	$(6\,931)^2$	–	–	–	–	$(2\,313)^2$	$(2\,826)^2$	–	$(5\,113)^2$	$(2\,798)^2$	$(4\,530)^2$
1.68	$(5\,693)^2$	$(7\,257)^2$	–	–	$(3\,419)^2$	$(3\,759)^2$	$(2\,551)^2$	$(2\,965)^2$	–	$(7\,644)^2$	–	$(4\,205)^2$
1.74	–	$(6\,169)^2$	–	–	–	$(4\,612)^2$	$(1\,920)^2$	$(2\,523)^2$	–	$(3\,926)^2$	–	$(5\,890)^2$
1.80	–	$(10\,763)^2$	$(5\,303)^2$	$(5\,298)^2$	–	$(10\,595)^2$	$(2\,145)^2$	$(2\,714)^2$	$(3\,837)^2$	$(7\,290)^2$	$(4\,584)^2$	$(5\,548)^2$
1.86	–	$(19\,820)^2$	–	–	–	–	$(3\,223)^2$	$(3\,471)^2$	–	$(15\,979)^2$	–	$(9\,441)^2$
1.92	–	$(12\,907)^2$	$(6\,168)^2$	$(7\,035)^2$	–	–	$(3\,953)^2$	$(3\,897)^2$	$(3\,023)^2$	$(7\,422)^2$	–	$(9\,452)^2$
1.99	–	–	$(4\,682)^2$	$(5\,501)^2$	$(7\,153)^2$	$(7\,474)^2$	–	$(4\,728)^2$	–	$(8\,259)^2$	$(3\,951)^2$	$(6\,392)^2$
2.05	–	$(12\,125)^2$	–	–	–	–	$(3\,891)^2$	$(4\,085)^2$	–	$(8\,977)^2$	$(2\,479)^2$	$(7\,569)^2$
2.11	$(6\,947)^2$	$(11\,203)^2$	–	–	–	$(7\,096)^2$	–	$(4\,641)^2$	–	$(5\,832)^2$	–	$(8\,127)^2$
2.17	$(8\,365)^2$	$(8\,825)^2$	$(4\,205)^2$	$(5\,379)^2$	–	$(7\,607)^2$	–	–	$(6\,522)^2$	$(9\,015)^2$	–	–
2.23	–	$(13\,263)^2$	–	–	$(6\,100)^2$	$(6\,331)^2$	–	–	$(5\,587)^2$	$(10\,802)^2$	–	–
2.29	–	$(16\,087)^2$	–	–	$(6\,661)^2$	$(8\,300)^2$	$(5\,563)^2$	$(4\,660)^2$	$(4\,315)^2$	$(8\,076)^2$	–	$(9\,638)^2$
2.35	–	$(17\,405)^2$	$(7\,399)^2$	$(8\,814)^2$	$(5\,434)^2$	$(6\,674)^2$	$(5\,612)^2$	$(6\,196)^2$	–	$(14\,833)^2$	–	$(11\,609)^2$
2.41	$(9\,506)^2$	$(13\,119)^2$	$(7\,532)^2$	$(8\,437)^2$	$(8\,020)^2$	$(9\,034)^2$	$(4\,847)^2$	$(5\,121)^2$	–	$(11\,582)^2$	–	$(11\,719)^2$
2.47	$(11\,678)^2$	$(14\,416)^2$	–	–	$(5\,505)^2$	$(7\,577)^2$	$(4\,176)^2$	$(4\,897)^2$	$(6\,363)^2$	$(10\,704)^2$	$(3\,984)^2$	$(8\,565)^2$

[a] k-modes where the uncertainty brackets do not include zero power are masked and shown with dashes (−).
[b] limits quoted with an asterisk (∗) might be affected by systematics or persistent residual RFI.
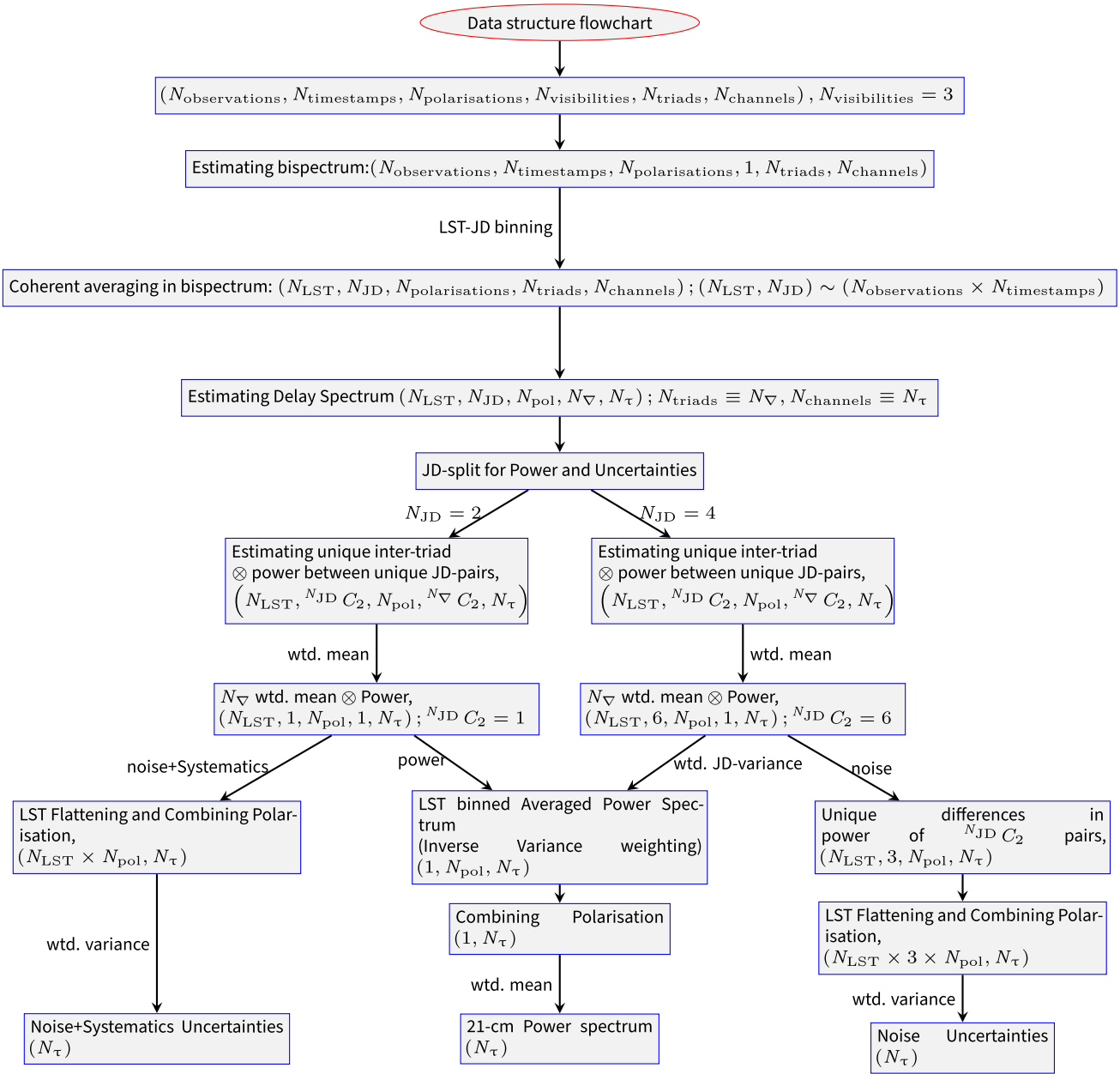
Figure A6. Schematic flow chart of the data structure through processing pipeline.