Wide-field LOFAR-LBA power-spectra analyses

Gehlot, Bharat; Koopmans, L.V.E.; de Bruyn, A. G.; Zaroubi, Saleem; Brentjens, M. A.; Asad, K. M. B.; Hatef, Mehdi; Jelic, Ratomir; Mevius, M.; Offringa, A. R.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Wide-field LOFAR-LBA power-spectra analyses: impact of calibration, polarization leakage, and ionosphere

B. K. Gehlot,[1]⋆ L. V. E. Koopmans,[1]⋆ A. G. de Bruyn,[1,2]† S. Zaroubi,[1,3] M.
A. Brentjens,[2] K. M. B. Asad,[1,4,5,6] M. Hatef,[1,2] V. Jelić,[2,7] M. Mevius,[1,2] A.
R. Offringa,[2] V. N. Pandey[1,2] and S. Yatawatta[1,2]

[1]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700AV Groningen, the Netherlands*
[2]*ASTRON, PO Box 2, NL-7990AA Dwingeloo, the Netherlands*
[3]*Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel*
[4]*Department of Physics, University of the Western Cape, Cape Town 7535, South Africa*
[5]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*
[6]*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands 7405, South Africa*
[7]*Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia*

ABSTRACT

Contamination due to foregrounds (Galactic and extragalactic), calibration errors, and ionospheric effects poses major challenges in the detection of the cosmic 21 cm signal in various epoch of reionization (EoR) experiments. We present the results of a pilot study of a field centred on 3C196 using LOFAR (LOw Frequency ARray) low-band (56–70MHz) observations, where we quantify various wide field and calibration effects such as gain errors, polarized foregrounds, and ionospheric effects. We observe a 'pitchfork' structure in the 2D power spectrum of the polarized intensity in delay-baseline space, which leaks into the modes beyond the instrumental horizon [EoR/CD (cosmic dawn) window]. We show that this structure largely arises due to strong instrumental polarization leakage (∼30 per cent) towards Cas A (∼21 kJy at 81 MHz, brightest source in northern sky), which is far away from primary field of view. We measure an extremely small ionospheric diffractive scale ($r_{\rm diff} \approx 430$ m at 60 MHz) towards Cas A resembling pure Kolmogorov turbulence compared to $r_{\rm diff} \sim 3{-}20$ km towards zenith at 150 MHz for typical ionospheric conditions. This is one of the smallest diffractive scales ever measured at these frequencies. Our work provides insights in understanding the nature of aforementioned effects and mitigating them in future CD observations (e.g. with Square Kilometre Array-low and Hydrogen Epoch of Reionization Array) in the same frequency window.

Key words: polarization – atmospheric effects – methods: statistical – techniques: interferometric – techniques: polarimetric – dark ages, reionization, first stars.

1 INTRODUCTION

The first stars and galaxies formed during the so-called cosmic dawn (CD) spanning redshifts $30 \gtrsim z \gtrsim 15$ (Pritchard & Furlanetto 2007). The ultraviolet and X-ray radiation from these first stars started to heat and ionize the neutral Hydrogen (HI hereafter) in the surrounding inter-galactic medium (IGM), continuing until hydrogen gas in the Universe transitioned from being fully neutral to become fully ionized (Madau, Meiksin & Rees 1997). Substantial ionization of

the IGM only occurred at $z \lesssim 15$ and this process completed around $z \sim 6$. This era in the history of the Universe is known as the epoch of reionization (EoR).

Current constraints on the redshift range of the reionization are inferred from indirect probes such as high-redshift quasar spectra (Becker et al. 2001; Fan et al. 2003, 2006), the optical depth for Thomson scattering from Cosmic Microwave Background (CMB) polarization anisotropy (Page et al. 2007; Komatsu et al. 2011; Hinshaw et al. 2013; Planck Collaboration XLVII 2016), IGM temperature measurements (Theuns et al. 2002; Bolton et al. 2010), Lyman break galaxies (Pentericci et al. 2011; Ono et al. 2012; Schenker et al. 2012), the kinetic Sunyaev–Zel'dovich effect (Zahn et al. 2012), and high-redshift gamma-ray bursts (Wang 2013). The

⋆ E-mail: gehlot@astro.rug.nl (BG) koopmans@astro.rug.nl (LK)
† Deceased (July 9, 2017)

most recent constraint on the upper limit of reionization redshifts comes from Planck Collaboration XLVII ([2016](#)), suggesting that the Universe is ionized at $\lesssim 10$ per cent level for $z \gtrsim 10$ and substantial reionization happened during redshifts between $z = 7.8$ and 8.8. Although these probes shed some light on the timing and duration of the reionization, there is very little known about the evolution of IGM during reionization, nature of sources of the ionizing radiation, and their evolution.

Observations of 21 cm hyperfine transition of HI at high redshifts promises to be an excellent probe of the HI distribution in IGM during EoR (Madau, Meiksin & Rees [1997](#); Shaver et al. [1999](#); Furlanetto, Oh & Briggs [2006](#); Pritchard & Loeb [2012](#); Zaroubi [2013](#)). Several ongoing and upcoming experiments such as the LOw Frequency ARray[1] (LOFAR; van Haarlem et al. [2013](#)), the Giant Meterwave Radio Telescope[2] (GMRT; Paciga et al. [2011](#)), the Murchison Widefield Array[3] (MWA; Bowman et al. [2013](#); Tingay et al. [2013](#)), the Precision Array for Probing the Epoch of Reionization[4] (PAPER; Parsons et al. [2010](#)), the 21 Centimetre Array (21CMA; Zheng et al. [2016](#)), the Hydrogen Epoch of Reionization Array[5] (HERA; DeBoer et al. [2017](#)), and the Square Kilometre Array[6] (SKA; Mellema et al. [2013](#); Koopmans et al. [2015](#)) aim to detect the redshifted 21 cm emission from the EoR. Although the above instruments focus largely on detecting the EoR, LOFAR low-band antenna (LOFAR-LBA), and the upcoming HERA, SKA-low, LEDA[7] (Large-aperture Experiment to detect Dark Ages; Price et al. [2017](#)), and NENUFAR[8] (New Extension in Nançay Upgrading loFAR; Zarka et al. [2012](#)) also observe at frequency range of 50–80 MHz that corresponds to a part of the redshift range of the CD ($30 \gtrsim z \gtrsim 15$). In this paper, we focus on challenges for observing the CD with LOFAR, and the future SKA-low that will largely have a similar layout. Since these telescopes operate at a lower frequency range (50–80 MHz), they will face challenges (foregrounds and ionosphere) similar to EoR experiments but more severe in strength.

The expected 21 cm signal from $z = 30$ to 15 is extremely faint with $\Delta_{21}^2 \sim (5-6\,\mathrm{mK})^2$ (Furlanetto et al. [2006](#); Pritchard & Loeb [2012](#)). This signal is buried deep below Galactic and extragalactic foreground emission which dominate the sky at these low frequencies (50–80 MHz). The foreground emission is ~ 4 orders of magnitude larger in strength than the 21 cm signal and has a brightness temperature of several Kelvins (Bernardi et al. [2010](#)) (on relevant angular scales) at high Galactic latitudes where LOFAR-EoR observing fields are located. Even if the foregrounds are removed with great accuracy, the noise per voxel in the images cubes after hundreds of hours of integration will still be orders of magnitude higher than the expected signal. Therefore, the current experiments (both EoR and CD) are aiming for a statistical detection of the signal instead of directly mapping out HI in IGM at high redshifts. The LOFAR-EoR Key Science Project (KSP) currently predominantly focuses on a statistical detection of the redshifted 21 cm signal from $z = 7$ to 12 (110 to 180 MHz) using LOFAR high-band antenna (LOFAR-HBA) observations and measure its power spectrum as a function of redshift (Patil et al. [2017](#)). Contamination due

to the (polarized) foregrounds, ionospheric propagation effects, and systematic biases (e.g. station-beam errors) pose considerable challenges in the detection of this signal. It is crucial to remove these bright foregrounds and mitigate other effects accurately in order to obtain a reliable (accurate and precise) estimate of the 21 cm power spectrum. This requires a detailed understanding of the nature of these effects and the errors associated with these effects. Several contamination effects in LOFAR EoR observations (HBA, 110–180 MHz) have been studied in great detail, such as polarization leakage (see Asad et al. [2015](#), [2016](#), [2018](#)), systematic biases (see Patil et al. [2016](#)), ionospheric effects (see Vedantham & Koopmans [2015](#), [2016](#); Mevius et al. [2016](#)), LOFAR radio frequency interference (RFI) environment (Offringa et al. [2010](#); Offringa, van de Gronde & Roerdink [2012](#); Offringa et al. [2013a](#),[b](#)), calibration and effects of beam errors (Kazemi et al. [2011](#); Kazemi, Yatawatta & Zaroubi [2013](#); Kazemi & Yatawatta [2013](#); Yatawatta [2013](#), [2015](#); Yatawatta [2016](#)).

In this work, we study some of the aforementioned effects at low frequencies using LOFAR-LBA observations of a field centred on 3C196 (3C196 field hereafter) at lower frequency (56–70 MHz), covering part of the CD, where both the foregrounds and ionospheric effects are known to be even stronger. LOFAR-HBA observations of the 3C196 field show bright polarized emission of $\sim$ few Kelvins with complicated and rich morphology (Jelić et al. [2015](#)). We address the broad-band nature of the excess noise due to systematic biases, polarized foregrounds, and ionospheric effects. A similar analysis has been done by Ewall-Wice et al. ([2016](#)) using low-frequency MWA observations (75–112 MHz), which addresses the MWA RFI environment, instrumental, and ionospheric effects at these frequencies. Our analysis provides improved insight on the spectral behaviour of the associated errors as well as the level of these contamination effects in ongoing and upcoming experiments to detect the HI signal from the CD era at low frequencies (50–80 MHz).

The paper is organized as follows . In Section 2, we briefly describe the data processing steps. In Section 3, we discuss the differential Stokes power spectrum method to study excess noise and its behaviour for different calibration strategies. In Section 4, we discuss the delay power spectrum method to study the polarized foregrounds and polarization leakage. We also discuss the effect of different calibration strategies and source subtraction on polarization leakage. In Section 5, we discuss the ionospheric effects at low frequencies using cross coherence method. In Section 6, we provide conclusions and summary of the analysis in this work.

2 OBSERVATIONS AND DATA PROCESSING

We have used LOFAR-LBA observations of the 3C196 field for our analysis, it being one of the two primary observation windows of the LOFAR EoR KSP. 3C196 is a relatively compact (4 arcsec) bright radio source placed at the centre of the field and serves as a bandpass calibrator. Observed data were processed using the standard LOFAR software pipeline (see e.g. LOFAR imaging cookbook[9]). The observational set-up and the steps for data processing are briefly described in the following subsections. Fig. [1](#) shows a flow chart of the data processing steps.

---

[1] http://www.lofar.org/

[2] http://gmrt.ncra.tifr.res.in/

[3] http://www.mwatelescope.org/

[4] http://eor.berkeley.edu/

[5] http://reionization.org/

[6] http://skatelescope.org/

[7] http://www.tauceti.caltech.edu/leda/

[8] https://nenufar.obs-nancay.fr/

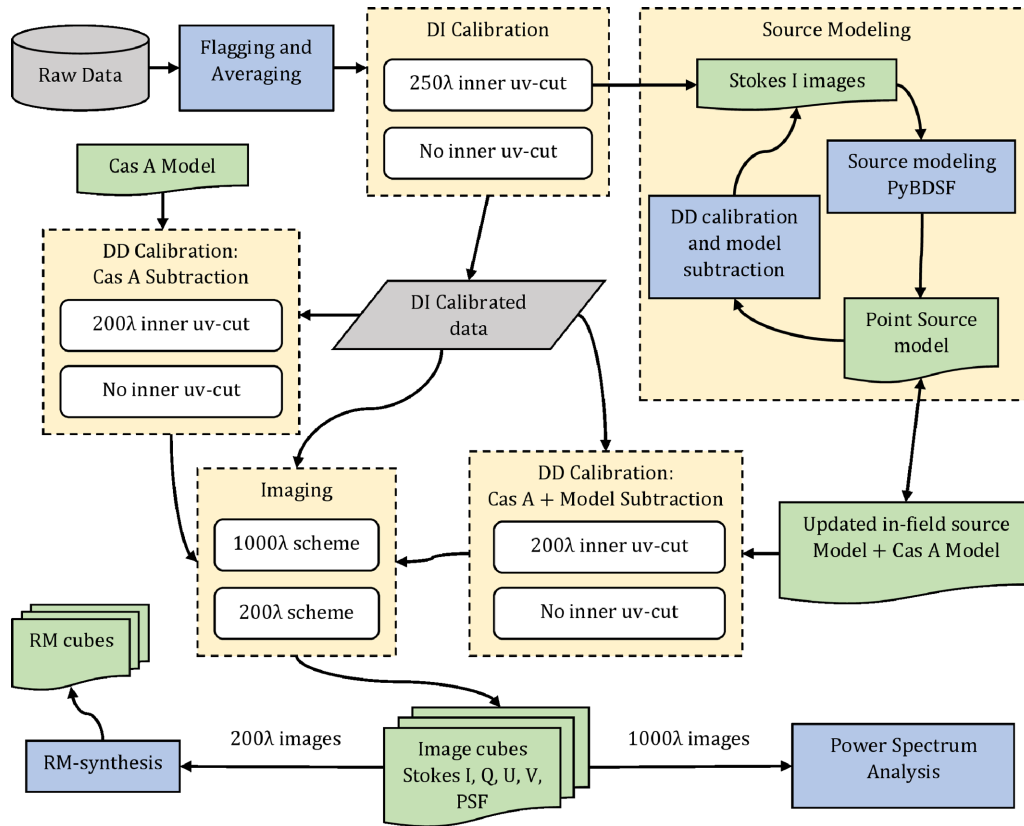[9] https://www.astron.nl/radio-observatory/lofar/lofar-imaging-cookbook

Figure 1. Flow chart illustrating the steps involved in data processing. Solid rectangles represent the processes. Rounded rectangles represent various schemes within a process. Dotted rectangles represent processes with subprocesses and/or multiple processing schemes. Parallelograms represent the stored visibility data. Boxes with curved bottom represent data products of the pipeline, such as sky-model and image cubes. Arrows represent the data flow.

2.1 LOFAR-LBA system

The LOFAR array has 38 stations in the Netherlands, out of which 24 stations (also known as core stations) are spread within a core of 2 km radius, and 14 stations (known as remote stations) are spread across 40 km east–west and 80 km north–south area in northeastern part of the Netherlands. Each LOFAR station has 96 low-band dual-polarization dipole antennas spread within an area of 87 m diameter. LBA dipoles have an arm length of 1.38 m, which corresponds to a resonance frequency of 52 MHz. LBAs are designed to operate in the frequency range of 10–90 MHz, but the operational bandwidth of LBA is limited to 30–80 MHz to avoid strong RFI below 30 MHz and RFI due to proximity to the FM band above 80 MHz. At a given time, signals from only 48 out of 96 LBA dipoles can be processed. The signals from these 48 dipoles are digitized and beam formed to produce a station beam which is steered digitally to track a fixed phase centre in the sky. The LOFAR–LBA system offers three different LBA dipole configurations, viz. LBA_INNER where 48 innermost dipoles (array width ∼30 m) are beam-formed, LBA_OUTER where 48 outermost dipoles (array width ∼87 m) are beam-formed, and LBA_SPARSE where half of the innermost 48 dipoles, plus half of the outermost 48 dipoles (array width ∼87 m) are beam-formed. These different configurations provide different field of view (FoV) areas as well as different sensitivities due to mutual coupling between the dipoles. The data are digitized by the receivers with 200 MHz sampling clock, providing an RF bandwidth of 96 MHz. The digitized data are transported to the GPU correlator via a fibre optics network. The correlator generates visibilities with

3 kHz frequency resolution (64 channels per sub-band) and 1 s integration and stores them in a measurement set (MS) format. Readers may refer to van Haarlem et al. (2013) for more information about LOFAR capabilities.

2.2 Observations

We use 8 h of synthesis observation data [L99269 (LOFAR Cycle 0): 2013 March 2–3] of 3C196 field (pointing/phase centre: RA = $08^h13^m36^s$, Dec. = $+48°13'03''$, epoch = J2000) using the LOFAR-LBA system. The field was observed with 37 LOFAR-LBA stations in the Netherlands (70 m to 80 km baseline) operating in the frequency range of 30–78 MHz. The correlations of voltages from antenna pairs were recorded with 1 s time resolution and 3 kHz frequency resolution. The recorded data consists of 248 sub-bands, and each sub-band has 195.3 kHz width and consists of 64 channels. We used only 56–70 MHz band in our analysis, which is the most sensitive region of the LBA band and is relatively free from RFI. Four out of eight observation hours are used in our analysis and we discarded the visibilities for the first 2 h and last 2 h of observation. The choice of this 'hard cut' is based on the quality of station based gain solutions after direction independent calibration step. We observed that the phases of the gain solutions were varying rapidly as a function of time in the beginning and at the end of the observations. The rapid variation of phases of gain solutions represents strong ionospheric activity that leads to strong amplitude scintillation. The observational details of the data are summarized in Table 1.

Table 1. Observational details of the data.

Parameter	Value
Telescope	LOFAR LBA
Observation cycle and ID	Cycle 0, L99269
Antenna configuration	`LBA_INNER`
Number of stations	37 (NL stations)
Observation start time (UTC)	March 2, 2013; 17:02:52
Phase centre (α, δ; J2000)	$08^h13^m36^s$, $+48°13'03''$
Duration of observation	8 h
Frequency range	30–78 MHz
Primary beam full width at half-maximum (at 60 MHz)	$9.77°$
Field of View (at 60 MHz)	75 deg^2
SEFD (at 60 MHz)	$\sim$26 kJy
Polarization	Linear X–Y
Time, frequency resolution:	
Raw Data	1 s, 3 kHz
After flagging and averaging	5 s, 183.1 kHz

2.3 Flagging and averaging

The first step of the processing is flagging of RFI-corrupted data. RFI mitigation usually works best on the highest resolution data in order to minimize any information loss. Thus, this step is performed on the raw data with the time and frequency resolution of 1 s and 3 kHz. RFI mitigation is performed using the `AOFlagger` software (Offringa et al. 2010, 2012). Two channels on either edge of every sub-band are discarded in order to avoid edge effects due to the polyphase filter, resulting in a final width of 183.1 kHz per sub-band. Note that the separation between two consecutive sub-bands is still 195.3 kHz. After flagging, the remaining data are averaged to 5 s and to 183.1 kHz sub-band resolution. These resolutions are chosen such that the time and frequency smearing is limited to the longer baselines and does not affect the baselines of interest ($\lesssim$1000λ). In addition, we flagged 3 stations CS013LBA, CS030LBA, and RS409LBA, which have 4, 6, and 10 non-working dipoles, respectively, on the basis of their poor quality of the direction-independent gain solutions.

2.4 Calibration

The sky observed by LOFAR is distorted by the characteristics of the instrument (station beam, global band pass, clock drift, etc.) and the environment (ionosphere). Calibration of a radio telescope refers to the estimation of the errors that corrupt the visibilities measured by the telescope and to obtain an accurate estimate of the visibilities from the observed data. The influence of the instrument and the environment on the measured visibilities can be described by the radio interferometer *measurement equation* (Hamaker, Bregman & Sault 1996; Smirnov 2011a,b). The effects that corrupt the observed visibilities can be divided into two categories: (a) *direction-independent effects* (DIEs) and (b) *direction-dependent effects* (DDEs).

DIEs are instrument-related effects which are independent of the sky direction. These include complex antenna gains and frequency band-pass, as well as, a single phase and amplitude correction for the average ionosphere above each station. The DDEs vary as a function of the sky direction. These are, for example, caused by antenna voltage patterns, ionospheric phase fluctuations, and Faraday rotation.

2.4.1 Direction-independent calibration

Direction-independent calibration refers to the estimation of a single instrumental gain for each beam-formed interferometric element (a station in the context of LOFAR). LOFAR station gain is described by a complex 2×2 Jones matrix and represents two linear polarizations. The QSO 3C196 is a very bright radio source with known flux (130 Jy at 74 MHz; Kassim et al. 2007) and is located at the phase centre of the field. It can be used as a flux calibrator to determine the station band-pass gains. We use a model of 3C196[10] which has 4 Gaussian components to describe the source, and the source spectrum is described by a second-order log-polynomial. This model was iteratively derived using LOFAR-HBA (full Dutch array with baseline range of 100 m to 120 km) observation data of 3C196 over the frequency range of 115–185 MHz. The parameters of the model components including the spectral indices were obtained by fitting in the visibility domain. The source model includes the flux at large angular scales and also represents the high-resolution structures (with arcsec accuracy). We compared the 3C196 model flux extrapolated at lower frequencies with other 3C196 observations at 74 MHz using Very Large Array (VLA; Kassim et al. 2007) and at 60 MHz using Serpukov radio telescope (Aslanian et al. 1968). The model flux matches the VLA observation within 4 per cent error and the Serpukov radio-telescope observation within 2 per cent error. Hence, the model performs well at the frequencies of interest. We use the Black Board Selfcal (`BBS`) package (Pandey et al. 2009) to obtain and subsequently apply the gain solutions for 30 s and 183.1 kHz intervals. 3C196 is subtracted in this step, and the residual visibilities are used for further processing. We use two different strategies for DI calibration: (1) using the baselines which are $\geq$250λ for calibration ('250λ cut', hereafter), to avoid model incompleteness due to diffuse emission (see Patil et al. 2016, 2017), and (2) using all baselines ('no cut', hereafter) for calibration. The reasoning behind this is to reflect the effect of including/excluding small baselines and inclusion/exclusion of unmodelled diffuse flux on the calibration products. This is further explained in later sections. Parameters for the DI calibration steps are listed in Table 2.

2.4.2 Direction-dependent calibration and source subtraction

The low-frequency radio sky is dominated by Galactic diffuse foregrounds (synchrotron, free–free emission) and extragalactic compact sources (radio galaxies, supernova remnants). The Galactic diffuse emission at high Galactic latitudes dominates only on small baselines ($\leq$50λ) and LOFAR-LBA has very few baselines $\leq$50λ at lower frequencies causing lesser sensitivity. Hence, the diffuse emission is mostly undetectable in the images. LOFAR-LBA images are dominated by the extragalactic compact sources which need to be removed in order to obtain a clean power spectrum relatively free from the foregrounds. The signal arriving from different directions, however, is corrupted by direction-dependent errors, which arise from wave propagation effects through the ionosphere and the primary beam (i.e. gain errors per station receiver element). These effects can produce artefacts around bright sources making it difficult to subtract them without leaving strong artefacts in the images. These effects can be accounted for during source subtraction by using direction-dependent (DD) calibration. This requires obtaining the gain solutions in multiple directions. We use `SAGECal`

---

[10]V. N. Pandey via private communication

Table 2. Calibration parameters.

Direction-independent calibration		
Parameter	Value	Comments
Flux calibrator	3C196	J2000: $08^h13^m36^s$, $+48°13'03'$
Sky-model components	4	Gaussian
Source spectral order (n)	2	Log-polynomial spectra;
		$\log P_n = \log S_\circ + \sum_{i=1}^{n} \alpha_i \left[\log \left(\frac{\nu}{\nu_\circ} \right) \right]^n$
Calibration baselines	1. $\geq 250\lambda$	Two strategies
	2. All ($\geq 0\lambda$)	
Solution type	Full Jones	Solves for all polarizations
Solution interval:		
Time	30 s	
Frequency	183.1 kHz	
Direction-dependent calibration		
Sky-model components	188	Compact; with apparent fluxes
Cas A model components	25	11 Gaussian + 14 compact; with apparent fluxes
Source spectral order (n)	1	log-polynomial spectra;
		$\log P_n = \log S_\circ + \sum_{i=1}^{n} \alpha_i \left[\log \left(\frac{\nu}{\nu_\circ} \right) \right]^n$
Calibration directions	5	4 within FoV and 1 on Cas A
Calibration baselines	1. $\geq 200\lambda$	Two strategies
	2. All ($\geq 0\lambda$)	
Solution type	Full Jones	Solves for all polarizations
Solution interval:		
Time	5 min	
Frequency	183.1 kHz	

(Kazemi et al. 2011, 2013; Kazemi & Yatawatta 2013; Yatawatta 2015; Yatawatta 2016) for DD calibration and source subtraction. Note that we do not perform consensus optimization (SAGECal-CO which is a more recent addition to SAGECal) while solving for the gains, but solve for each sub-band independently from the other sub-bands. The sources in the calibration model are removed by multiplying the obtained gain solutions with the predicted visibilities and subtracting the product from the observed visibilities. In the DD calibration step, we provide a sky-model consisting of 188 compact sources within the primary beam FoV (in-field model, hereafter) with a flux density range between 300 mJy and 11 Jy (described in later section) and Cas A (25 components[11]) containing positions, apparent fluxes, and spectral indices of the sources as an input for SAGECal. We solve for five directions: four directions are within the primary beam (each quadrant) and one is towards Cas A. Choosing only four directions within the primary beam optimizes signal-to-noise ratio (SNR) in each direction as well as minimizes the image noise which allows us to subtract more fainter sources compared to more number of directions. We choose the solution interval of 5 min and 183.1 kHz for the gain solutions. Subtracting Cas A is important, because the bright sources such as Cas A and Cyg A can cause significant sidelobe noise in the images even if these are far outside ($\gtrsim 40°$) the primary beam. Cyg A ($\sim 90°$ away from the phase centre) does not affect our observations because it is close to the horizon during the entire observation (discussed later). The residual visibilities after the DD calibration step and source subtraction are stored and imaged for further analysis. We also perform an alternative DD calibration step where we only sub-

tract Cas A and image the residuals. In both cases, we choose two calibration strategies: (1) with $\geq 200\lambda$ baselines ('200λ cut', hereafter) to avoid diffuse emission (absent in the sky-model) biasing the gain solutions and (2) using all baselines. Note that we use $\geq 200\lambda$ baselines in DD calibration instead using $\geq 250\lambda$ baselines as in DI calibration. We noticed that choosing $\geq 250\lambda$ cut in DD calibration produces noisy gain solutions across several sub-bands causing comparatively higher image rms values in these sub-bands. We think that the main reason behind these noisy gain solutions is the low SNR in each direction which is solved for in DD calibration step. The SNR increases when we add more baselines to the calibration step by lowering the calibration cut to $\geq 200\lambda$ still without adding significant unmodelled diffuse emission in the calibration. The $\geq 200\lambda$ cut results in images with lower image rms values compared to the former. Parameters for the DD calibration steps are listed in Table 2.

2.5 Imaging

We use the WSClean (Offringa et al. 2014) package to image the visibilities. WSClean is a CPU-based imager and produces Stokes I, Q, U, V, and point spread function (PSF) images as output. We image the visibilities after DI-calibration step and DD calibration step for both calibration strategies. We use two different imaging schemes, viz. 1000λ imaging and 200λ imaging. The 1000λ imaging scheme employs $0-1000\lambda$ baselines for imaging, and the output images are used for the power spectrum analysis. The 200λ imaging scheme uses $0-200\lambda$ baselines for imaging, and the output images are used to perform the RM synthesis (see the Appendix). Both schemes use 'uniform' weighting to achieve a cleaner side-lobe response. Although 'natural' weighting scheme produces images with higher SNR values compared to 'uniform' weighting scheme, it produces a biased result in uv-space which has to be re-normalized to remove the effect of the gridding weights (i.e. tapering). Making

---

[11] Cas A model is derived from a single sub-band Cas A image (with 40 arcsec restored beam size) produced using LOFAR-LBA at 52 MHz (Asgekar et al. 2013). We used the source spectrum with spectral index of -0.77 (Baars et al. 1977).

Table 3. Imaging parameters.

Parameter	value	
Imaging scheme	1000λ	200λ
Imaging baselines	0–1000λ	0–200λ
Frequency range	56–70MHz	56–70MHz
Weighting scheme	Uniform	Uniform
Spatial resolution	2.75 arcmin	13.75 arcmin
Pixel size	45 arcsec	3 arcmin
Number of pixels	1200 × 1200	192 × 192

a power spectrum from natural weighted images requires dividing the flux density in each *uv*-cell by its sampling density to get an unbiased power spectrum. It is mathematically almost equivalent to uniform weighting, which of course also performs this division. The only difference is when the 'kernels' (antialising, beam, etc.) are applied. Since the *uv*-coverage of LOFAR over the measured *uv*-cells is almost uniform, making power spectra from uniform and natural images results in similar power spectra. The reason we have used uniform weighting here is that uniform weighted images are easier to interpret and produce unbiased power spectra. The imaging parameters for both schemes are listed in Table 3. Fig. 2 shows the DI calibrated (using 250λ) dirty and cleaned Stokes *I* continuum image (56–70MHz) of 3C196 field where 3C196 has been subtracted off. Cas A has also been subtracted off using DD calibration (using 200λ scheme). We performed the multifrequency deconvolution using WSClean with a cleaning threshold of 50 mJy.

2.6 Source modelling

Radio galaxies, galaxy clusters, and supernova remnants are the discrete foreground sources observed at low radio frequencies. We used the Python Blob Detection and Source Finder (PyBDSF) software (Mohan & Rafferty 2015) to model the bright compact sources in the 3C196 field. Source modelling is an iterative process where DI-calibrated images are used to model the sources above a particular SNR threshold and determine their frequency spectra. The resulting sky-model from PyBDSF is used to perform DD calibration (see Section 2.4.2) and source subtraction (see Section 2.4.2) on the DI-calibrated visibilities. The Stokes*I*images of the residual visibilities are again modelled with PyBDSF to include fainter sources. This process is repeated until the confusion limit (∼90 mJy at 60 MHz) is reached in a single sub-band. We have not applied any beam model[12] to the data prior to modelling, which means that the modelled fluxes are apparent and averaged over the on-sky time. We create 15° × 15° images (centred around the primary beam) with pixel size of 45 arcsec using the 1000λ imaging scheme (see Section 2.5 and Table 3) for source modelling. We use a comb configuration with 12 sub-bands evenly spread across 56 to 70 MHz for spectral index estimation. The final sky-model contains source positions, apparent fluxes and source spectra for 188 compact sources which have flux densities ≳2.5 times the rms noise in a single sub-band image. We use 1000λ

---

[12]Beam model for LOFAR-LBA is a very recent addition to the LOFAR data processing pipeline and is still being improved. There were no beam models available for LBA when we performed most of the analysis. We only used the beam model in simulations (discussed later in Section 4.1.3). The current version of the LOFAR-LBA beam model is derived from the electromagnetic simulations of LOFAR-LBA dipoles. Beam model for LBA will be taken into account in future analyses.

imaging scheme for the modelling purpose because the images produced with baselines greater than 1000λ produces artefacts in the residual images after source subtraction as well as the PSF is more symmetric in 1000λ images compared to the other cases.

3 DIFFERENTIAL POWER SPECTRUM

Azimuthally averaged power spectrum of the difference between the Stokes images of adjacent sub-bands (differential Stokes images, hereafter) may be used to quantify the effects which are non-smooth in frequency (on sub-band level) such as instrumental and calibration effects. In an ideal scenario, total signal in a Stokes *I* image at a given frequency can, to the first order, be expressed as a sum of the total sky signal convolved with the PSF and additive noise (see e.g. Patil et al. 2016). Let I_1 and I_2 be the Stokes*I*images at two consecutive frequency sub-bands, and V_1 and V_2 be the Stokes *V* images, respectively. We can write

$$I_1 = S_1 * P_1 + N_1^I, \text{ and } V_1 = N_1^V, \tag{1}$$

$$I_2 = S_2 * P_2 + N_2^I, \text{ and } V_2 = N_2^V, \tag{2}$$

where S is the sky signal, P is the PSF, N^I and N^V represent the noise in Stokes *I* and *V* images. We assume that the signal from the sky does not change within the 195 kHz frequency separation, which is the separation between two consecutive frequency sub-bands, i.e. $S = S_1 \approx S_2$. By making such assumption, we expect all the effects contributed by the foregrounds and ionosphere (assuming smoothness in frequency) to drop out, but the effects which are non-smooth in frequency on sub-band level are expected to remain. Therefore,

$$\Delta I = I_1 - I_2 = S * (P_1 - P_2) + (N_1^I - N_2^I), \tag{3}$$

$$\Delta V = V_1 - V_2 = (N_1^V - N_2^V). \tag{4}$$

In Fourier space, equations (3) and (4) can be written as

$$\tilde{\Delta I} = \tilde{S} \times d\tilde{P} + \tilde{N_1^I} - \tilde{N_2^I}, \tag{5}$$

$$\tilde{\Delta V} = \tilde{N_1^V} - \tilde{N_2^V}, \tag{6}$$

where the tilde represents Fourier transform (FT) and $dP = P_1 - P_2$ is the differential PSF due to slightly different *uv*-coverage. The spatial power-spectrum of the difference, $|\tilde{\Delta I}|^2$, is divided into M annuli of width $\delta b = 19.1$ m in the *uv*-plane, and all the points within an annulus are averaged to obtain an estimate of the power. The differential power spectrum can finally be written as

$$P_{\Delta I} = \langle |\tilde{\Delta I}|^2 \rangle = |\tilde{S}|^2 |d\tilde{P}|^2 + \langle |\tilde{N_1^I}|^2 \rangle + \langle |\tilde{N_2^I}|^2 \rangle, \tag{7}$$

$$P_{\Delta V} = \langle |\tilde{\Delta V}|^2 \rangle = \langle |\tilde{N_1^V}|^2 \rangle + \langle |\tilde{N_2^V}|^2 \rangle \tag{8}$$

where $P_{\Delta I}$ and $P_{\Delta V}$ represent azimuthally averaged Stokes *I* and *V* power spectra, respectively. We use the DD calibrated residual images produced using the 1000λ imaging scheme with sub-band frequencies $\nu_1 = 59.7641$ MHz and $\nu_2 = 59.9594$ MHz to calculate $P_{\Delta I}$ and $P_{\Delta V}$. The selected sub-bands lie in the most sensitive region of the frequency band and are free from RFI. We estimate the power spectra for both calibration strategies.

3.1 Excess noise

The sky signal has negligible circularly polarized component which is assumed to be well below the thermal noise. Because of this,
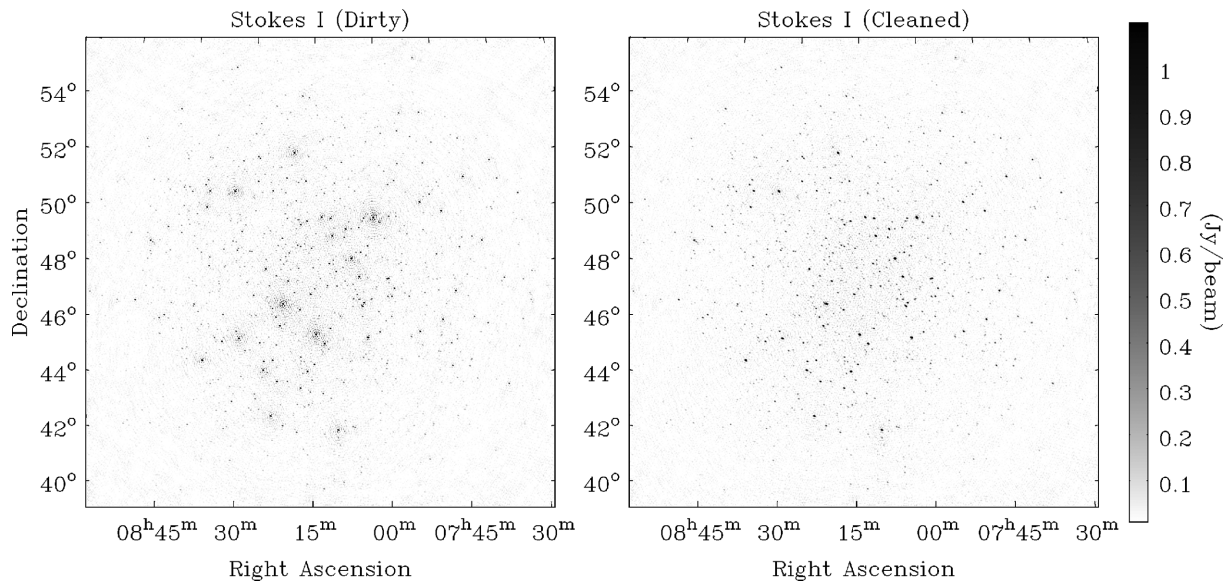
Figure 2. Direction-independent calibrated continuum images (56–70MHz) of the 3C196 field. The left-hand panel shows the dirty image, and right-hand panel shows the cleaned image. Calibration was done using 250λ cut parameters and imaging using $\leq 1000\lambda$ baselines with uniform weights. The image has a point source rms $\sigma_{\text{image}} \sim 27$ mJy, whereas the expected theoretical value of the thermal noise is $\sigma_{\text{th}} \sim 2.1$ mJy for the observation, which is around 13 times smaller than the observed value. σ_{th} can be calculated as $\sigma_{\text{th}} = \text{SEFD}/\sqrt{N(N-1)\Delta\nu\Delta t}$, where SEFD (system equivalent flux density) ~ 26 kJy at 60 MHz, $N = 30$ (corresponding to 1000λ imaging cut), $\Delta\nu = 13.7$ MHz, $\Delta t = 0.9 \times 4$ h (assuming flagged data at 10 per cent level).

Stokes V can be used as a proxy for the thermal noise of the system. However, we observed that the point source rms value in the Stokes V image for a single sub-band at 60 MHz is $\sigma_{\text{image}} \sim 30$ mJy, which is ~ 1.6 times the theoretical value $\sigma_{\text{th}} \sim 18$ mJy (calculated using $\sigma_{\text{th}} = \text{SEFD}/\sqrt{N(N-1)\Delta\nu\Delta t}$, where SEFD ~ 26 kJy at 60 MHz, $N = 30$, $\Delta\nu = 183.1$ kHz, $\Delta t = 0.9 \times 4$ h). We think that this excess Stokes V rms is due to the errors on the gain solutions which are applied to all polarizations during calibration step. Since each sub-band has different realizations of noise, the noise from two different sub-bands does not correlate. Also, the thermal noise in Stokes I and V is expected to be identical (see e.g. van Straten [2009]), which means that they have identical statistical properties (e.g. variance). This leads us to define the excess noise (P_X) in Stokes I as

$$P_X = P_{\Delta I} - P_{\Delta V} = \langle |\Delta\tilde{I}|^2 \rangle - \langle |\Delta\tilde{V}|^2 \rangle . \qquad (9)$$

P_X can be interpreted as excess power in differential Stokes I compared to differential Stokes V. Fig. 3 shows $P_{\Delta I}$ and $P_{\Delta V}$ for the both 200λ cut and all baselines strategies. The right-hand panel of Fig. 3 shows the ratio $P_{\Delta I}/P_{\Delta V} = P_X/P_{\Delta V} + 1$ for the both calibration strategies. We observe that P_X is $\gtrsim 10$ times higher than $P_{\Delta V}$. Ideally, if the noise in Stokes I and V are statistically identical, then $P_X \approx |\tilde{S}|^2 |d\tilde{P}|^2$, which is the contribution due to chromatic PSF. Contribution due to chromatic PSF can be estimated by multiplying $|d\tilde{P}|^2$ with the sky contribution $|\tilde{S}|^2$ in Fourier space (readers may refer to Patil et al. [2016] for detailed calculation of chromatic PSF contribution). Patil et al. ([2016]) showed that the contribution due to chromatic PSF in LOFAR-HBA observations is a small fraction of the excess noise. We also observe a similar behaviour in LBA observations; the chromatic PSF seems to contribute less than 20 per cent to the overall excess noise between sub-bands on the relevant baselines. The sky brightness also varies as a function of frequency (diffuse emission has $\nu^{-2.55}$ dependence), causing a brightness change of ~ 1 per cent for 183.1 kHz difference between sub-bands. This is also a negligible effect compared to the excess noise we have observed, but it might become relevant in deeper experiments. We

see a factor $\gtrsim 10$ larger power (i.e. $\gtrsim 3 \times$ larger rms) in differential Stokes I than Stokes V for both calibration strategies. This ratio is almost constant as a function of baseline length and does not change between the two calibration strategies we employed. Introducing a calibration cut, however, decreases the power on baselines outside the cut and increases it on baselines inside the cut. However, the power in differential Stokes I when calibrated using all baselines seems to decrease on smaller baselines. This decrease in power might occur because a diffuse sky-model is not included in the calibration. There may be several causes of the significant excess power in $P_{\Delta I}$. These factors could include incomplete sky-model, imperfect source subtraction, and ionospheric effects (Barry et al. [2016]; Patil et al. [2016]; Ewall-Wice et al. [2017]).

3.2 Effect of calibration cut

In the calibration scheme employed by Patil et al. ([2016], [2017]), small baselines are excluded in calibration steps. The reasoning behind this is that small baselines ($<100\lambda$) are dominated by diffuse foreground emission, and it is more difficult to model this emission and include it in the calibration model. One way to avoid any unmodelled flux biasing the calibration process is by choosing only those baselines where the diffuse emission is already resolved out. In such calibration schemes, longer baselines are used to obtain the gain solutions which are applied to all the visibilities, including shorter baselines. We compare the excess noise in the differential Stokes I power spectrum in the two calibration strategies we employed. Fig. 3 shows the ratio $P_{\Delta I}(200\lambda \text{ cut})/P_{\Delta I}(\text{nocut})$ and $P_{\Delta V}(200\lambda \text{ cut})/P_{\Delta V}(\text{nocut})$. We observe that both ratios have a discontinuity at the exact location of the calibration cut. The excess noise, suddenly, is $\gtrsim 2$ times higher on baselines $<200\lambda$ than on baselines $>200\lambda$. This has also been observed in LOFAR-HBA observations by Patil et al. ([2016]). This ratio is no longer constant on baselines $\leq 200\lambda$, but shows a slope with increasing excess power at shorter baselines. This effect is not only limited to Stokes I but
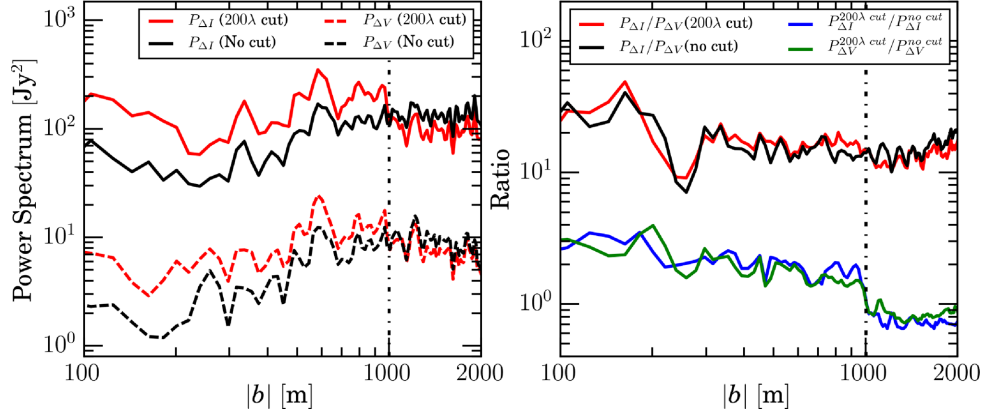
Figure 3. The left-hand panel shows the differential Stokes *I* and *V* power spectra for the two calibration schemes. Solid lines correspond to $P_{\Delta I}$ and dashed lines correspond to $P_{\Delta V}$. The colour codes represent different calibration schemes; red colour corresponds to 200λ cut strategy and black colour corresponds all baselines (no cut strategy). Right-hand panel shows the ratios of different combinations of $P_{\Delta I}$ and $P_{\Delta V}$. The red and black colours correspond to the ratio $P_{\Delta I}/P_{\Delta V}$ for 200λ cut scheme and no cut, respectively. The blue curve corresponds to the ratio $P_{\Delta I}(200\lambda \text{ cut})/P_{\Delta I}(\text{Nocut})$ and the green curve corresponds to the ratio $P_{\Delta V}(200\lambda \text{ cut})/P_{\Delta V}(\text{Nocut})$. The vertical dot–dashed line shows the location of the 200λ baseline at 60 MHz.

also present in Stokes *Q*, *U*, and *V*. We do not show the ratios for *Q*, *U* here. We expect this effect to be purely because of the calibration cut. Because we perform a full Jones gain calibration, we expect this discontinuity to be present in all the Stokes parameters. Given that all Stokes power-spectra increase in the same manner, whereas their ratio with Stokes *V* does not show any sign of change, suggests that this is the result of random errors introduced in the Jones matrices during the calibration process, which are subsequently applied to the sky-model and transferred to the image residuals during model subtraction. The cause of these random gain errors on the longer baselines could be due to sky-model incompleteness or the ionosphere (Barry et al. 2016; Patil et al. 2016; Ewall-Wice et al. 2017). Although differencing between sub-bands is a good first-order sanity check of whether the data reaches the expected noise levels, a more powerful analysis can be carried out by using the combined information in all sub-bands. This is discussed in the next section.

4 DELAY SPECTRUM OF GRIDDED VISIBILITIES

The delay spectrum is a powerful tool to study foregrounds and various contamination effects which can leak foregrounds into the EoR window. A delay spectrum (see e.g. Parsons & Backer 2009; Parsons et al. 2012) is defined as the FT of the visibilities along the frequency axis. Consider the gridded visibilities, $\mathcal{V}(u, v; \nu)$, as a function of baseline coordinates (u, v)[13] and frequency ν. Then

$$\tilde{\mathcal{V}}_S(u, v; \tau) = \int_{-\infty}^{\infty} \mathcal{V}_S(u, v; \nu) e^{-2\pi i \nu \tau} d\nu, \quad (10)$$

$$P_S(u, v; \tau) = |\tilde{\mathcal{V}}_S(u, v; \tau)|^2, \quad (11)$$

where $\tilde{\mathcal{V}}_S(u, v, \tau)$ is the 3D delay spectrum and $P_S(u, v; \tau)$ is the power spectrum in the delay-baseline space. The subscript 'S' refers to one of the Stokes parameters *I*, *Q*, *U*, *V*, or the complex polarized intensity $\mathcal{P} = Q + iU$. The 2D delay power spectrum $P_S(|b|, \tau)$

can be obtained by azimuthally averaging $P_S(u, v, \tau)$ in *uv*-plane, where $|b| = (\sqrt{u^2 + v^2}) \times \lambda$ is the baseline length (in metres) and τ is the delay which corresponds to the geometric time delay between the signal arriving at two different antennas from a given direction. The delay τ can also be written as

$$\tau = \frac{\mathbf{b} \cdot \hat{\mathbf{s}}}{c}, \quad (12)$$

where $\hat{\mathbf{s}}$ is the unit vector towards the direction of the incoming signal, θ is angle between zenith and $\hat{\mathbf{s}}$, and *c* is the speed of light. For $\theta = 90°$, $\tau = |b|/c$; this delay corresponds to the instrumental horizon. A 2D delay spectrum scaled with proper cosmological parameters results in the 2D cosmological power spectrum, which is a widely used statistic in EoR experiments. The 2D cosmological power spectrum can be derived from the delay spectrum as (Parsons et al. 2012; Thyagarajan et al. 2015a)

$$P(k_\perp, k_\parallel) = |\tilde{\mathcal{V}}(|b|, \tau)|^2 \left(\frac{A_{\text{eff}}}{\lambda^2 \Delta \nu}\right) \left(\frac{D^2(z)\Delta D}{\Delta \nu}\right) \left(\frac{\lambda^2}{2k_B}\right)^2 \quad (13)$$

and baseline (**b**) and delay (τ) are related to $k_\perp$ and $k_\parallel$ wave numbers as

$$k_\perp = \frac{2\pi \left(\frac{|b|}{\lambda}\right)}{D(z)}, \quad k_\parallel = \frac{2\pi \nu_{21} H_0 E(z)}{c(1+z)^2} \tau, \quad (14)$$

where A_{eff} is the effective area of the antenna, λ is the wavelength of the centre frequency of the observation band, $\Delta \nu$ is the observation bandwidth, $D(z)$ is the transverse comoving distance corresponding to redshift *z*, ΔD is the comoving depth along the line of sight corresponding to $\Delta \nu$, k_B is the Boltzmann constant, ν_{21} is the rest-frame frequency of the 21 cm spin-flip transition of HI. H_0 and $E(z)$ $\equiv [\Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda]^{1/2}$ are the Hubble constant and a function of the standard cosmological parameters. We use $P(|b|, \tau)$ instead of $P(k_\perp, k_\parallel)$ and adhere to units of Jy² throughout our analysis. This is a suitable choice in this paper as we only address the severeness of the contamination effects, which are orders of magnitude ($\sim$ Kelvins in amplitude) higher than the expected 21 cm signal at the frequencies of interest. Typically, a delay spectrum is defined per visibility where the instrumental horizon (same as physical horizon) is fixed. In a phase tracking array, the instrumental horizon is no longer fixed and moves with respect to the zenith. Because of tracking, delays towards a particular source (fixed with

---

[13]In radio interferometric imaging, the (u, v) coordinates are defined in units of wavelength (λ) and are frequency invariant. Whereas, a delay spectrum is defined for baseline coordinates in physical units (metres) such that frequency dependence of baseline length is inherent to the delay transform.

respect to the phase centre) in sky will vary within a certain range, depending on the orientation of baseline and location of the phase centre. As a result of this, features due to that source in delay power spectrum produced using time-integrated image cubes will appear to be smeared over a certain range of delays. Even in drift scan arrays, a particular source appears at a certain delay only in snapshot mode with phase centre on zenith. Once the correction for earth's rotation is applied, it will appear to be smeared across several delays.

There are some differences between the delay spectrum estimation approach in Thyagarajan et al. (2015a,b) and the approach we used in our analysis. The former uses snapshot visibilities that are averaged over different observing nights (same LST) to average down the incoherent part of the visibilities. These averaged visibilities are subsequently used to estimate the delay power spectra. In our case, visibilities recorded at different times during a single observation are coherently averaged during the gridding process, ultimately averaging down their incoherent (noise) part. These gridded visibilities are Fourier transformed to produce time integrated delay power spectra. Gridding asymptotically for large numbers of visibilities leads to the delay power spectrum of average visibilities. Whereas, averaging of the individual visibility-based power spectra, as in Thyagarajan et al. (2015a,b), yields the delay power spectrum of the average visibility with the power spectra of the incoherent part of the visibility (i.e. noise and scintillation noise) added to it. We opted for the delay spectrum of gridded visibilities to (i) avoid having to separately estimate each of the incoherent power spectra and (ii) reduce computational effort since diffuse foreground subtraction is computationally prohibitive if it is done at the visibility level.

Another difference between the two approaches is that Thyagarajan et al. (2015a,b) estimate the delay power spectra directly from visibilities. Foreground subtraction in this approach will affect the power at a certain delay corresponding to the subtracted foreground source. Whereas, we determine the delay power spectra by Fourier transforming the image cubes (real-valued signal) instead of calculating it directly from the visibilities. This makes the delay transform in our case, a Hermitian transform. As an outcome of this, the resulting delay power spectrum is symmetric around $\tau = 0$ and foreground subtraction will affect the power in the same manner at positive and negative delays corresponding to the subtracted foreground source.

Estimation of $P_S(|b|, \tau)$ from image cubes requires two additional steps. (a) The image cube is Fourier transformed along the spatial axes. The spatial axes of cosine-directions (l, m) of the images are the Fourier conjugates of the baseline axes (u, v). The resulting gridded visibilities for different frequencies have fixed uv-cell size in units of wavelength (λ) causing the physical uv-cell size (in metres) to vary with frequency. (2) This physical uv-grid and corresponding visibilities are re-gridded on to a fixed grid with baseline length in metres such that the uv-coverage scales as function of frequency but the size of the physical grid remains fixed. The re-gridded visibilities are then Fourier transformed along frequency to obtain the delay spectrum. We flag several noisy sub-bands on the basis of the Stokes V rms of the images to avoid any artefacts (due to RFI etc.) in the delay power spectrum. This flagging produces image cubes and hence gridded visibilities which have irregular spacing across frequency axis, and therefore, a Fast Fourier Transform (FFT) cannot be used to FT across the frequency axis. Thus, we use an FFT to FT the image cube only across the spatial axes, whereas for the frequency axis, we use a Least Squares Spectral Analysis method (i.e. full least squares-FT matrix inversion) (see e.g. Barning 1963; Lomb 1976; Stoica, Li & He 2009; Trott et al. 2016).

The resulting cube is then squared and azimuthally averaged (annuli width $\delta b = 19.1$ m) across the spatial domain to obtain the 2D delay power spectrum. We use the I, Q, U, and V image cubes produced with 1000λ imaging scheme to determine the 2D delay power spectrum.

4.1 'Pitchfork' structure in polarized intensity

We determine the delay power spectrum from Stokes I, Q, U, V images produced after DD calibration step using 200λ cut strategy where only Cas A is subtracted. Fig. 4 shows delay power spectra for Stokes I, Q, U, V, and $\mathcal{P}$. We observe a 'pitchfork' structure in Stokes I power spectrum. A similar structure has been observed in MWA (Thyagarajan et al. 2015a,b) and PAPER (Kohn et al. 2016) observations. Moreover, we observe a similar 'pitchfork' structure in power spectrum of Stokes Q, U, and $\mathcal{P}$. Most of this polarized emission is localized on smaller baselines ($\leq$400 m) and around the delays corresponding to instrumental horizon, suggesting that the emission originates from far outside the primary beam and is diffuse in nature. This can either be caused by intrinsic diffuse polarized emission or instrumental polarization leakage from Stokes I to Q and U. One method to distinguish between intrinsic polarized emission and instrumental polarization leakage is to investigate the emission in Stokes Q, U and $\mathcal{P}$ in RM space (see the Appendix). When a polarized signal passes through an ionized medium in the presence of a net magnetic field parallel to the line of sight, the signal undergoes Faraday rotation. Due to Faraday rotation, the signal appear often at non-zero Faraday depths ($\Phi \neq 0$) in the RM space. On the contrary, any polarization leakage due to the instrument is localized around $\Phi = 0$ because the primary beam variation has a smooth but weak dependence on the frequency. In the $\mathcal{P}$ RM cubes, we do not see any polarized emission except at $\Phi = 0$. We expect that the polarized emission is depolarized by ionospheric Faraday rotation due to ionospheric Total Electron Content (TEC) varying as a function of time and position. The polarization angle $\chi \propto \lambda^2$. Thus, at low frequencies, ionospheric Faraday rotation becomes significant and depolarizes most of the intrinsic polarized signal. Fig. 4 also shows the difference $P_{\mathcal{P}} - P_V$ which represents the presence of excess polarized power over the power in Stokes V (assumed to be the noise level). The 'pitchfork' feature is also observed in the difference plot. We notice that most of the excess polarized power originates outside the primary beam and is localized around shorter baselines, i.e. $|b| < 400$ m, whereas little to no polarization power at $\tau \approx 0$ which suggests absence of intrinsic polarized emission in the field. We confirm the absence of intrinsic polarized emission also by the noise like image cubes (not shown here) in Stokes Q and U (with variance exceeding Stokes V though). We also observe a faint structure in Stokes V which appear to correlate with the 'pitchfork' structure in Stokes Q and U on certain baselines. Because there is negligible emission (circularly polarized component) in Stokes V, it is expected to have a flat power spectrum. Presence of any structure in Stokes V is another indication of instrumental polarization leakage.

4.1.1 Effect of calibration cut

To quantify the impact of DD calibration on the 'pitchfork', we used the visibilities after DD calibration step where Cas A and the in-field model has been subtracted. We compare the power spectra of visibilities from the two calibration strategies (Table 2). Fig. 5 shows Stokes I, Q, U, and V delay power spectra from the two calibration
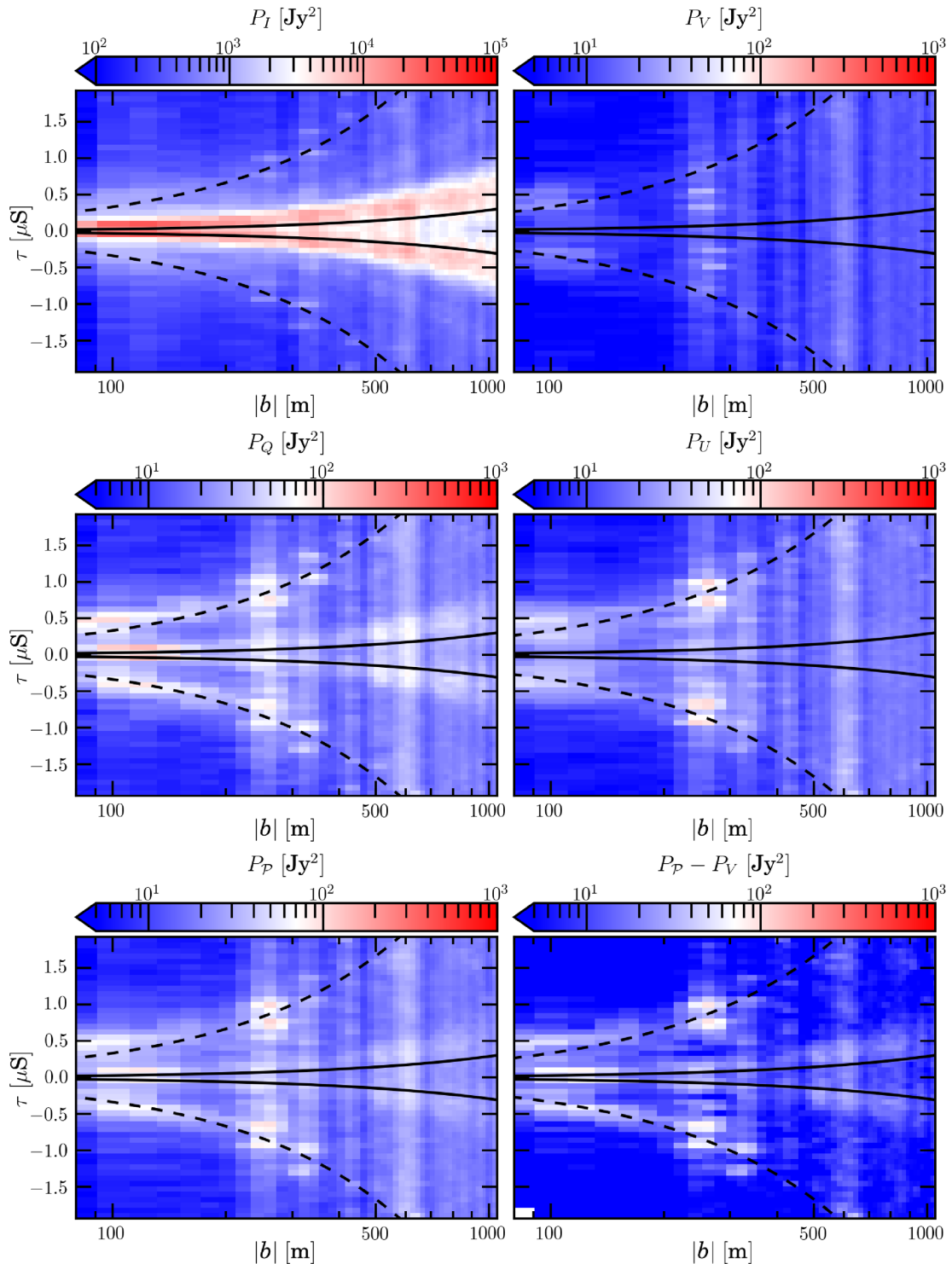
Figure 4. This figure shows the delay spectra for Stokes *I* (top-left), *V* (top-right), *Q* (middle-left), *U* (middle-right), total polarized intensity $\mathcal{P}$ (bottom-left), and the difference $P_\mathcal{P} - P_V$ (bottom-right). The black solid lines represent delays corresponding to full width at half-maximum of primary beam (see Table 1) and the dashed lines correspond to the instrumental horizon. These power spectra are computed from the image cubes produced using the 200λ DD calibrated visibilities with only Cas A subtracted. We observe a clear 'pitchfork' structure in Stokes *I*, *Q*, *U*, $\mathcal{P}$ and the difference. The difference $P_\mathcal{P} - P_V$ show the excess of polarized emission over the Stokes *V*.

strategies and their ratio ($P(200\lambda\text{cut})/P(\text{nocut})$). We notice that the power on/around the 'pitchfork' is suppressed significantly when all the baselines are used in calibration. This might happen because the unmodelled diffuse emission is absorbed in the gain solutions, hence lowering the power on smaller baselines (Patil et al. 2016).

We observe a discontinuity in the ratio at $|b| \sim 900$ m, the ratio drops for $|b| > 900$ m and continues to drop till $|b| \sim 1000$ m and becomes almost constant for $|b| > 1000$ m. The 200λ baseline cut for different frequency sub-bands lies in baseline range 900 m $< |b| <$ 1000 m. Therefore, this discontinuity around
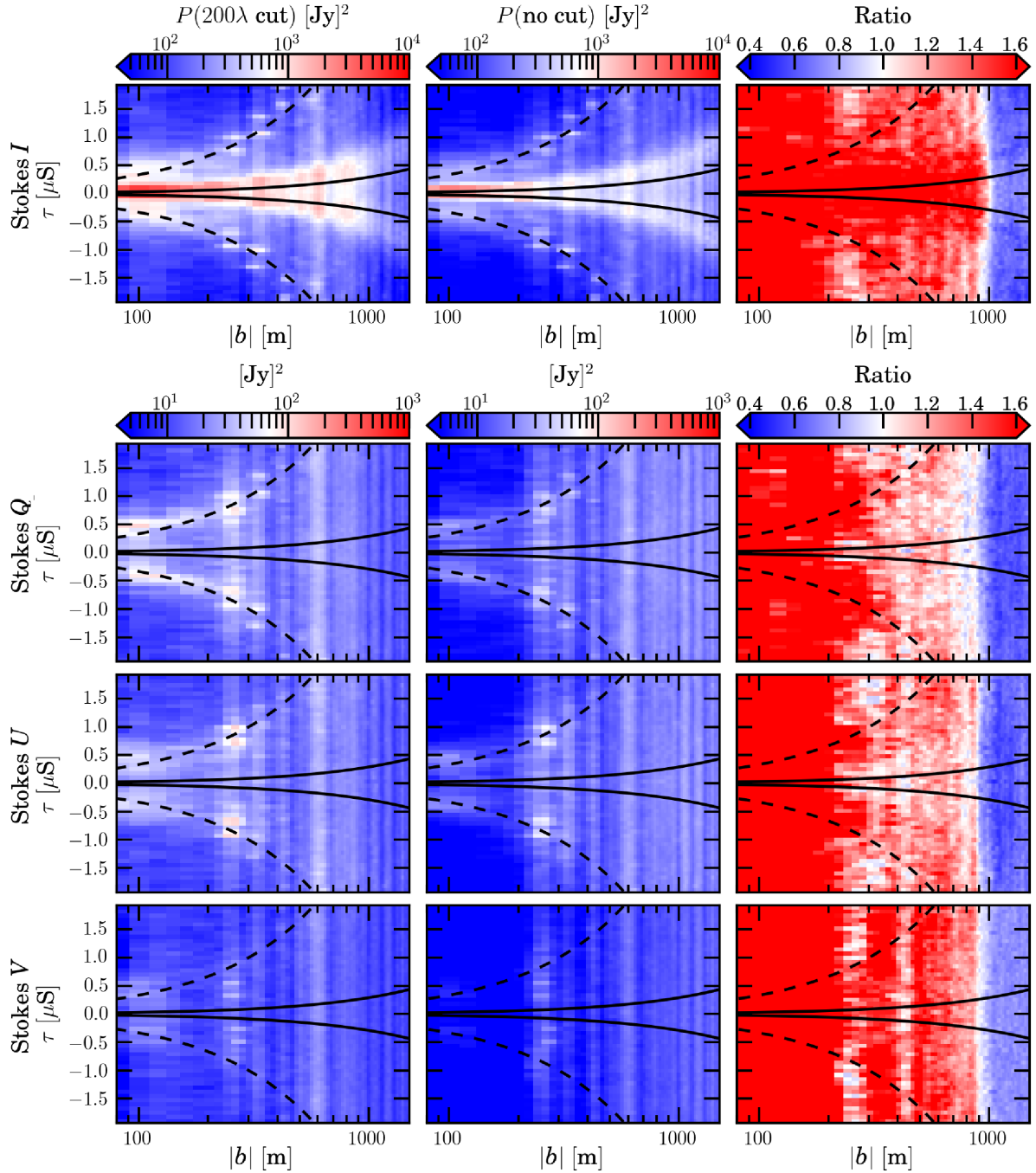
Figure 5. Stokes *I*, *Q*, *U*, and *V* delay spectra for two different DD calibration schemes and their ratio. Leftmost column corresponds to DD calibration scheme using $\geq 200\lambda$ baselines, middle column corresponds to DD calibration scheme using all baselines and rightmost column corresponds to the ratio $\frac{P(200\lambda \text{ cut})}{P(\text{no cut})}$. Different rows correspond to different Stokes parameters. Top row corresponds to Stokes *I*, second row from top corresponds to Stokes *Q*, third row from top corresponds to Stokes *U* and bottom row corresponds to Stokes *V*.

900 m < |*b*| < 1000 m corresponds to the location of baseline cut and is similar to the one in the ratio of differential power spectrum (Fig. 3, right-hand panel). This trend is observed for all the Stokes parameters. We observe that the excess power on excluded short baselines is $\gtrsim 2$ times the power on baselines included in the calibration step. A similar reasoning, as in Section 3.2, can be applied in this case as well; that power on baselines |*b*| < 200λ is enhanced because of the errors in the gain solutions (obtained solely from the longer baselines) applied to the data and to the sky-model. The source of these errors is not yet well understood, but we suspect several causes such as incomplete calibration models, ionospheric effects, and imperfect calibration (Barry et al. 2016; Patil et al. 2016).

4.1.2 Effect of source subtraction

In this section, we discuss the effect of subtraction of sources on the 'pitchfork' structure. We quantify this effect for two cases. In first

case, the in-field sky-model (sources within the primary beam) is subtracted from DI-calibrated visibilities (250λ cut) using the DD calibration (200λ cut). Note that Cas A model is already subtracted before performing the in-field model subtraction. We compare the delay power spectrum of Stokes I (P_I) and $\mathcal{P}$ ($P_\mathcal{P}$) calculated using the image cubes before and after subtracting the model. Top row of Fig. 6 shows P_I before and after in-field model subtraction and the ratio $P_I(\text{after})/P_I(\text{before})$. We observe that subtracting the sources largely within the primary beam significantly reduces the power in Stokes I within the primary beam going up till the horizon as well as above the horizon. This effect is expected as a consequence of foreground subtraction. However, the ratio on/around the 'pitchfork' remains $\sim$0.8−0.9, suggesting that the subtraction of sources within primary beam does not affect the 'pitchfork'. We observe a similar effect in comparison of $P_\mathcal{P}$ before and after in-field model subtraction. Fig. 7 (top row) shows $P_\mathcal{P}$ before and after subtracting the in-field model. We observe $\sim$30 per cent decrease in polarized power within the primary beam primarily due to subtraction of sources away from phase centre, but it does not affect the power beyond the primary beam. However, we observe an increase in ratio ($P_\mathcal{P}(\text{after})/P_\mathcal{P}(\text{before}) \gtrsim 1.0$) beyond horizon on baselines $\leq$200 m. We suspect that this increase in power is due to errors on gain solutions obtained in DD calibration step, which mainly affect the shorter baselines excluded from calibration. This effect is not visible in Stokes I, as the subtracted power is much larger than the increase in power introduced due to these gain errors. Whereas in $\mathcal{P}$, power on excluded baselines is comparable to the increase in power introduced due to errors on gain solutions and it becomes prominent in the ratio. Besides this, we do not observe any significant difference in $P_\mathcal{P}$ due to the subtraction of sources within primary beam.

In second case, we compare delay power spectra for Stokes I (P_I) and $\mathcal{P}$ ($P_\mathcal{P}$) before and after subtracting Cas A (using DD calibration) which lies outside the primary beam. In this case, we do not subtract the in-field model. In our observation, Cas A is above horizon during the whole period of observation and is $\gtrsim$40° away from the zenith ($\sim$66° away from 3C196; $\sim$31° away from NCP (North Celestial Pole)). Fig. 6 (bottom row) and Fig. 7 (bottom row) show P_I and $P_\mathcal{P}$, respectively, before and after Cas A subtraction and the ratio [$P_I(\text{after})/P_I(\text{before})$]. We observe a factor of $\sim$10 decrease in power on the 'pitchfork' in both Stokes I and $\mathcal{P}$ after Cas A subtraction. From this comparison, it is clear that subtraction of Cas A has a significant impact on the power in Stokes I and polarized intensity $\mathcal{P}$ on/around the 'pitchfork' but also on the modes within and beyond the horizon ($\sim$50 per cent decrease). Since Cas A is extremely bright at low frequencies ($\sim$21 kJy intrinsic flux at 81 MHz; Baars et al. 1977), its effects can be detected in LBA images even when it is tens of degrees away from the phase centre. LOFAR-LBA has a polarized response for angles away from zenith. For zenith angles $\approx$60°, $\mathcal{P}/I \approx 0.3$ (see e.g. Bregman 2012), causing significant fraction of the total power ($\sim$10 per cent) leak to polarized power due to the instrument. This leakage occurs from $\mathcal{P}$ to I as well. Note that most of the leaked power is on the small baselines ($<$600 m), which is probably due to the large extent of Cas A caused by ionospheric diffraction (discussed later). The leakage is reduced substantially when Cas A is subtracted using a model via DD calibration. Note that residuals after subtracting Cas A still correlate quite strongly with the power before Cas A subtraction, suggesting imperfect subtraction in DD calibration or the structure of Cas A which is harder to model. In summary, the primary cause of the 'pitchfork' structure in $\mathcal{P}$ is Cas A outside the primary beam leaking to $\mathcal{P}$ from Stokes I because of the instrumental

beam polarization. Although other sources which are spread over many directions (and delays) will also leak in to $\mathcal{P}$ as shown in Asad et al. (2016, 2018), they are unlikely to cause strong leakage. A single source as bright as Cas A, however, is clearly dominant in the power spectra.

4.1.3 Comparison with the simulations

To gain further insight on the 'pitchfork' structure, we simulate visibilities observed by LOFAR-LBA using a Stokes I only model of Cas A, with the phase centre at 3C196. We use NDPPP[14] to predict the XX, XY, YX, YY antenna correlations using the exact LOFAR-LBA station configuration for 4 h of synthesis. We chose the time and frequency resolution of the correlations to be 5 s and 183.1 kHz to save computation time. We include the LOFAR-LBA primary beam [a recent addition to LOFAR data processing pipeline (NDPPP), see footnote 12] in the prediction step in order to predict instrumental polarization leakage. We then image the predicted visibilities using WSClean using the 1000λ imaging scheme and determine the delay power spectrum for Stokes I and total polarized intensity $\mathcal{P}$. Fig. 8 shows Stokes I and $\mathcal{P}$ delay power spectrum and the ratio $P_\mathcal{P}/P_I$. We observe a clear 'pitchfork' structure in P_I and this structure appears solely due to Cas A. The structure looks nearly identical to that observed in Figs 6 and 7. If we compare P_I with $P_\mathcal{P}$, the structure in $P_\mathcal{P}$ looks exactly like that in P_I, but scaled down in power. The ratio $\frac{P_\mathcal{P}}{P_I} \approx 0.09$, which is $\sim$0.3 in amplitude. The effect of a constant ratio between Stokes I and $\mathcal{P}$ due to polarization leakage was also predicted by Asad et al. (2018) for LOFAR-HBA observations. This simulation clearly shows that the 'pitchfork' structure in $\mathcal{P}$ is indeed an artefact arising from Cas A due to instrumental polarization leakage from Stokes I to $\mathcal{P}$.

We also simulate the visibilities using a Cyg A only model to quantify the polarization leakage due to Cyg A. We used the VLSS model of Cyg A ($\sim$20 kJy at 74 MHz; Kassim et al. 2007) with the same simulation set-up as for Cas A to predict the antenna correlations. The Stokes I power spectrum (not shown here) calculated using the simulated visibilities for Cyg A shows $\sim$6–7 orders of magnitude lower power on the 'pitchfork' compared to the power due to Cas A. Although the beam model used in simulations is only approximately correct (inaccuracy of $\sim$ few percent) in the direction of Cyg A (lower elevation angles), contribution due to Cyg A is negligible and can be ignored for any practical purpose in observations with LOFAR-LBA centred on 3C196, at the current level of accuracy.

When the model of Cas A is subtracted from the visibilities during the DD calibration step, the 'pitchfork' structure due to Cas A should in principle (if the model is accurate) disappear. However, we still observe some residual power on/around the pitchfork. The residual power on small baselines ($\leq$400 m) is $\sim$10 per cent of the power before Cas A subtraction. These residuals can be caused by other factors such as unmodelled sources, diffuse emission, an inaccurate Cas A model, imperfect calibration and ionospheric effects. For example, Cas A is 3 arcmin in extent, which should be resolved only on the baselines $>$800λ. Cas A should therefore appear approximately as a compact source on baselines $\leq$100λ. Thus, inaccuracy in the Cas A model should not cause such significant residuals on these baselines. Ionospheric turbulence, on the other hand, can cause Cas A to scintillate significantly and visibilities to decorrelate

---

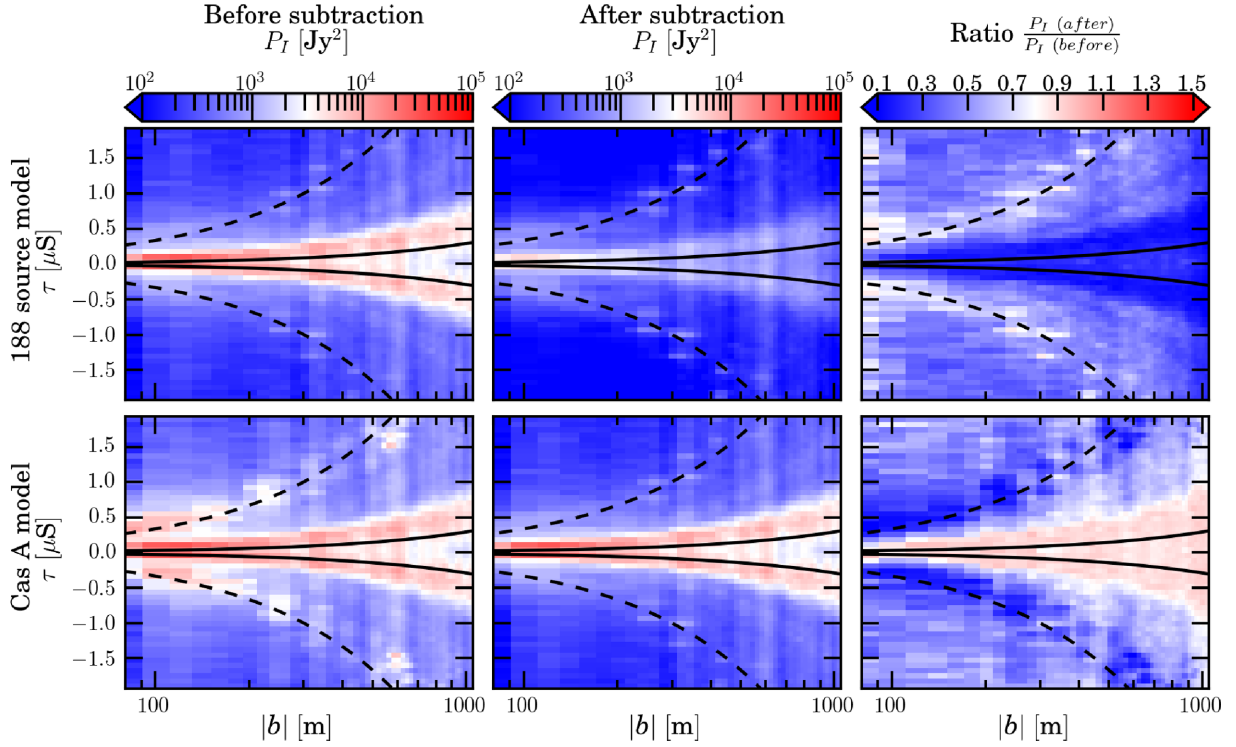[14]http://www.lofar.org/operations/doku.php?id=public:user_software:ndppp

Figure 6. Stokes *I* delay power spectra before (left column) and after (centre column) model subtraction and their ratio [P_I(after)/P_I(before)] (right column). The top row shows P_I before and after in-field model subtraction and their ratio. The bottom row shows P_I before and after Cas A model subtraction and their ratio.
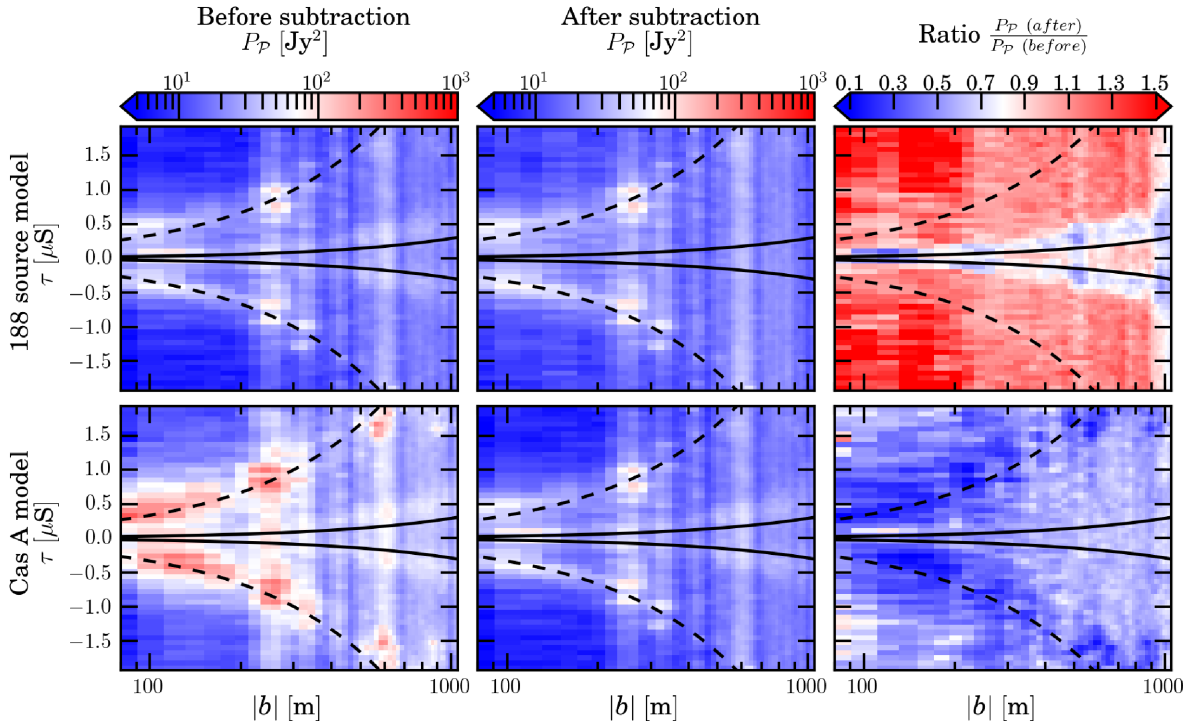

Figure 7. Polarized intensity $\mathcal{P}$ delay power spectra before (left column) and after (centre column) model subtraction and their ratio (right column). The top row shows $P_{\mathcal{P}}$ before and after in-field model subtraction and their ratio. The bottom row shows $P_{\mathcal{P}}$ before and after Cas A model subtraction and their ratio.
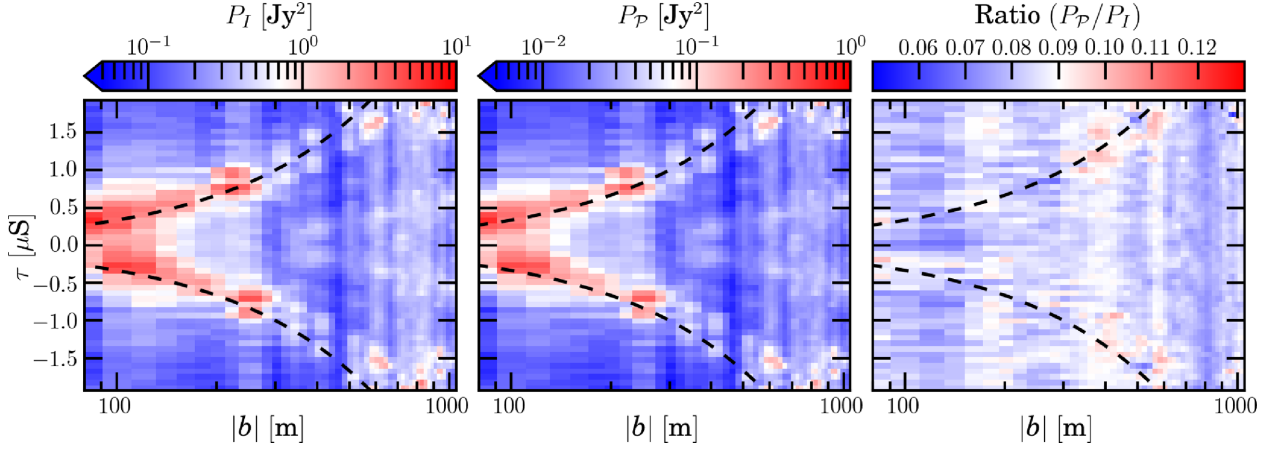
Figure 8. Stokes *I* (left-hand panel) and Polarized intensity $\mathcal{P}$ (middle panel) delay power spectrum determined using simulated visibilities (see Section 4.1.3) and the ratio $P_{\mathcal{P}}/P_I$ (right-hand panel). Note that the amplitude scales are apparent and are only meant to compare the power between *I* and $\mathcal{P}$.
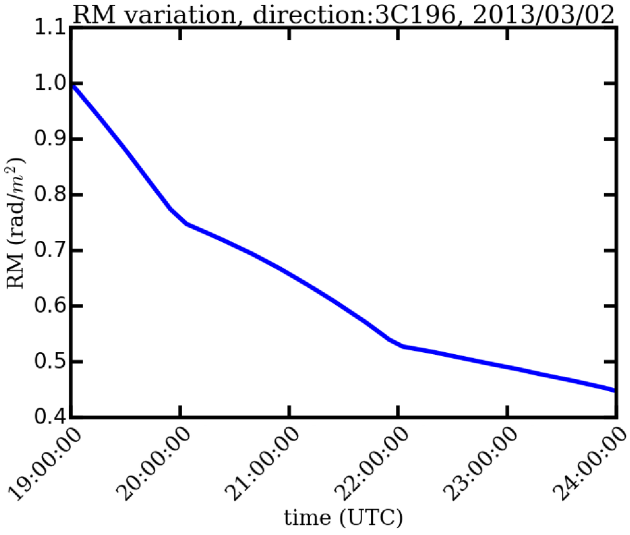

Figure 9. Ionospheric RM variation in direction of 3C196 as a function of time on 2013 March 2. This variation is calculated using `RMextract` developed by Maaijke Mevius. `RMextract` uses the GPS data to extract RM in particular direction.

Figure 10. This figure shows $\tau = 0$ slice of Stokes *I* (solid-red curve) and $\mathcal{P}$ (solid-blue curve) delay power spectra determined using the DI calibrated visibilities (250λ cut) which are phase rotated towards Cas A. The dashed curves show power spectrum for pure Kolmogorov turbulence (equation 22) with best-fitting values listed in Table 4 for Stokes *I* (red) and $\mathcal{P}$ (blue).

within the DD calibration solution interval. On baselines <5 km, the ionosphere decorrelates on time-scales of less than a minute, which is shorter than the solution interval in the DD calibration (5 min). Therefore, this 'scintillation noise' (see e.g. Vedantham & Koopmans 2015, 2016) might lead to imperfect calibration causing residual flux. We discuss this effect in the next section.

5 IONOSPHERIC SCINTILLATION

In previous analysis, we observed that the $\mathcal{P}$ delay power spectrum has more power concentrated on the 'pitchfork' on smaller baselines (<400 m). This feature is present in delay power spectra for both calibration strategies and is associated with polarization leakage from Cas A. There is residual flux in $\mathcal{P}$ delay power spectrum even after subtraction of Cas A. Given the low frequency and large angle away from zenith (i.e. large vTEC, see e.g. Fig. 9), we expect Cas A to be strongly affected by the ionosphere. The ionospheric turbulence is usually carried along with the bulk motion of ionospheric
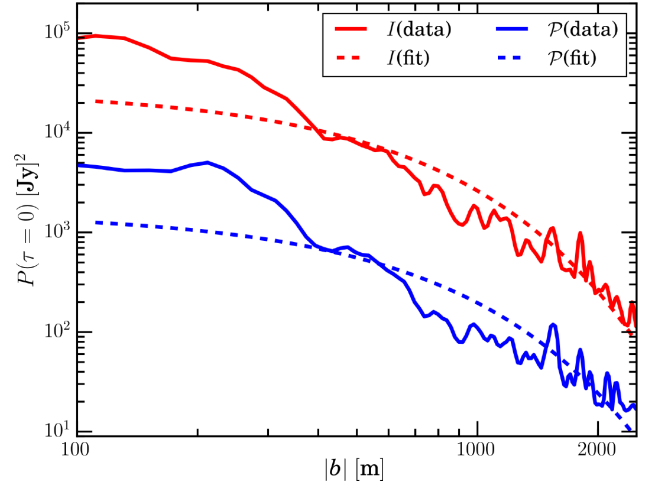
plasma, which has typical speeds between 100 and 500 km h^{-1}. Turbulent plasma in the ionosphere introduces time, frequency, and position-dependent phase shifts to the propagating wave. Under the phase-screen approximation, the phase shift φ introduced due to the wave propagation through ionospheric plasma is

$$\phi = \int \frac{2\pi\eta(z)}{\lambda} \mathrm{d}z, \tag{15}$$

where z is the distance along the direction of propagating wave. η (refractive index of non-magnetized plasma) is given by

$$\eta = \sqrt{1 - \frac{\nu_p^2}{\nu^2}} \approx 1 - \frac{1}{2}\frac{\nu_p^2}{\nu^2} \quad \text{for } \nu_p << \nu, \tag{16}$$

where ν_p is the plasma frequency (order of few MHz) and ν is the frequency of the propagating wave. By combining equation (15 and 16), φ can be written as

$$\phi = \int \frac{2\pi\eta(z)}{\lambda} \mathrm{d}z = \int \frac{2\pi\nu}{c} \mathrm{d}z - \frac{1}{2}\int \frac{2\pi\nu_p^2}{c\nu} \mathrm{d}z. \tag{17}$$
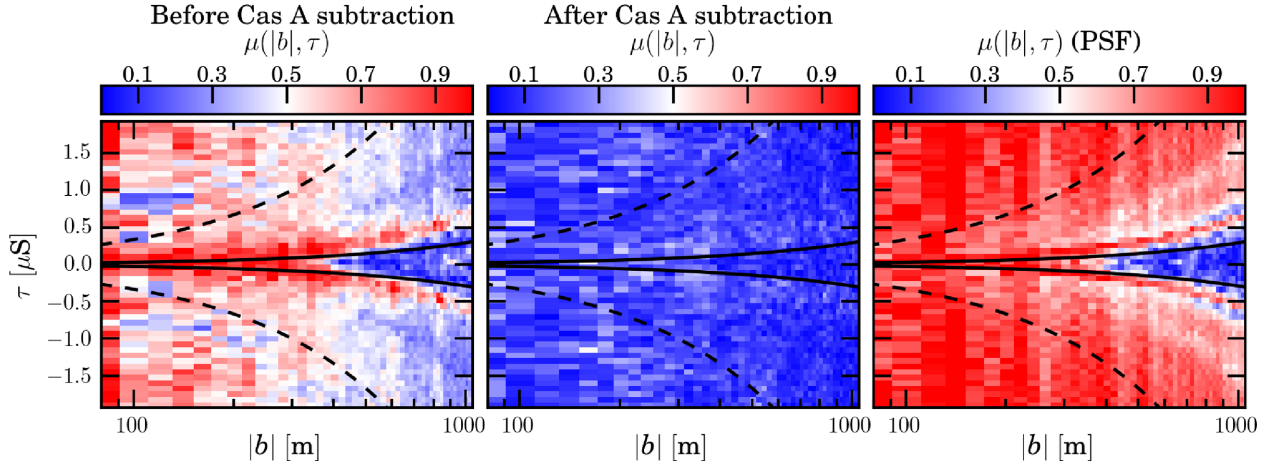
Figure 11. Stokes I cross-coherence ($\mu(|b|, \tau)$) in delay baseline space before (left-hand panel) and after (middle panel) subtracting Cas A. Right-hand panel shows the PSF cross-coherence μ_{PSF}.

The first term in equation (17) is a geometric delay term which is generally absorbed in the interferometer measurement equation. The second term in equation (17) is inversely proportional to the frequency of the propagating wave (ν):

$$\phi(\nu) \propto \frac{\nu_p^2}{\nu}, \text{ where } \nu_p = \frac{1}{2\pi} \sqrt{\frac{n_e q_e^2}{m_e \epsilon_\circ}}, \quad (18)$$

n_e is plasma density, q_e is the electron charge, m_e is the mass of electron, and $\epsilon_\circ$ is the permittivity of free space. The spatial variations in n_e can be described by Kolmogorov-type turbulence (Rufenach 1972; Singleton 1974; Koopmans 2010; Vedantham & Koopmans 2015). The power spectrum for Kolmogorov-type turbulence is represented by a $-11/3$ index power law. Since $\varphi \propto n_e$ (follows from equation 18), the phase fluctuations are also described by a Gaussian random field with power spectrum (assuming isotropy) given by

$$|\tilde{\phi}(k)|^2 \propto k^{-11/3}, \ k_o < k < k_i, \quad (19)$$

where k is the length of the spatial wavenumber vector $\mathbf{k}$, k_o is the wavenumber corresponding to the outer scale or the energy injection scale, and k_i corresponds to the inner scale or energy dissipation scale. If the visibility of a source in absence of ionospheric effects is given by $\mathcal{V}_S(b)$, then the expectation value of visibilities corrupted by the ionospheric phase fluctuations (assuming that the calibration solution interval significantly exceeds the time-scale on which the phases fluctuate) is given by (see e.g. Vedantham & Koopmans 2015, 2016):

$$\langle \mathcal{V}_C(b) \rangle = \mathcal{V}_S(b) \exp\left(-\frac{1}{2}\mathcal{D}(b)\right), \quad (20)$$

where $\mathcal{V}_C(b)$ are the time-averaged visibilities. $\mathcal{D}(b)$ is the phase structure function and is defined as

$$\mathcal{D}(b) = \left(\frac{b}{r_{diff}}\right)^{5/3}, \quad (21)$$

where r_{diff} is the diffractive scale. The power spectrum of the visibilities corrupted by the ionosphere is given by

$$P_C(b) = |\mathcal{V}_S(b)|^2 \, e^{-\mathcal{D}(b)} = P_S \exp\left[-\left(\frac{b}{r_{diff}}\right)^{5/3}\right]. \quad (22)$$

The power spectrum of an unresolved source as a function of baselines is constant in absence of ionospheric effects, whereas if the

Table 4. Best-fitting parameters.

Fit parameters	$P_I(|b|, \tau = 0)$ slice	$P_{\mathcal{P}}(|b|, \tau = 0)$ slice
P_S (Jy^2)	26775.1 ± 3010.6	1580.5 ± 209.1
r_{diff} (in λ)	429.8 ± 13.3	479.3 ± 19.3

source is affected by the ionospheric phase fluctuations, it will take the form of $P_C(b)$. To determine $P(|b|, \tau = 0)$, we selected DI calibrated visibilities with 250λ cut strategy. We phase rotate these visibilities towards Cas A and image them with the 1000λ scheme. We used the resulting image cubes to obtain $P_I(|b|, \tau)$ and $P_{\mathcal{P}}(|b|, \tau)$. We then choose the $\tau = 0$ slice from each $P_I(|b|, \tau)$ and $P_{\mathcal{P}}(|b|, \tau)$, which are expected to be dominated by the power due to Cas A, and fit them with $P_C(b)$ in equation (22) using P_S and r_{diff} as free parameters. We use $100 \, m \le |b| \le 2500 \, m$ baselines for fitting. Because Cas A exhibits a 3 arcmin structure and is only resolved on baselines $>800\lambda$ ($\sim$4 km at 60 MHz), means that the intrinsic power spectrum for Cas A is flat for selected baselines. Fig. 10 shows the $P_I(|b|, \tau = 0)$ and $P_{\mathcal{P}}(|b|, \tau = 0)$ slices fitted with equation (22).

We can see that equation (22) fits $P_I(|b|, 0)$ and $P_{\mathcal{P}}(|b|, 0)$ for over three orders of magnitude in power. The best-fitting values for P_S and r_{diff} for both power spectra are listed in Table 4. We find a diffractive scale r_{diff} towards Cas A of order $\sim$430 m for P_I and $\sim$480 m $P_{\mathcal{P}}$. Estimated r_{diff} values for P_I and $P_{\mathcal{P}}$ agree with each other within 10 per cent error. Typical values of r_{diff} at zenith vary between 3 and 20 km at 150 MHz and scale with frequency as $r_{diff} \propto \nu^{6/5}$ (Mevius et al. 2016; Vedantham & Koopmans 2016) and varies between 1 and 10 km for zenith at 60 MHz. Therefore, the diffractive scales we have measured are the smallest scales ever measured at $\sim$60 MHz. $P_{\mathcal{P}}$ has the same r_{diff} (within the errors) as P_I but is scaled down by one order of magnitude in power. This is additional evidence of instrumental polarization leakage from Stokes I to $\mathcal{P}$. The ratio of the power $P_{\mathcal{P}}/P_I \sim 0.1$ is approximately same as the estimate of the polarization leakage for Cas A obtained in simulation results shown in Section 4.1.3. We also observe that P_I and $P_{\mathcal{P}}$ deviate from the fit for $|b| < 400 \, m$, which corresponds to the Fresnel scale ($r_F \sim 400$ m at 60 MHz, see e.g. Vedantham & Koopmans 2015, 2016 for more details). Baselines below Fresnel scale are dominated by amplitude scintillation, whereas the baselines above Fresnel scale are dominated by phase scintillation producing a better fit on $|b| > 400 \, m$ compared to $|b| < 400 \, m$.

For time-scales of 5 min which correspond to the solution interval of DD calibration, ionospheric scintillation is expected to decorrelate on baselines $<1000\lambda$. We calculate the cross-coherence ($\mu(|b|, \tau)$) to quantify the decorrelation of this scintillation. To determine $\mu(|b|, \tau)$, we select visibilities before Cas A subtraction, and visibilities after Cas A subtraction. We arrange each visibility set in subsets of 5 min duration, such that each subset corresponds to a different DD calibration solution. Next, we divide the visibility set in two nonoverlapping consecutive subsets such that one subset consists of visibilities ($\mathcal{V}_{odd}$) corresponding to `odd` numbered calibration solutions and the other subset consists of visibilities ($\mathcal{V}_{even}$) corresponding to `even` numbered calibration solutions. The resulting visibility subsets $\mathcal{V}_{odd}$ and $\mathcal{V}_{even}$ are interleaved in time. We phase rotate the visibilities towards Cas A and image them to get the corresponding image cubes I_{odd} and I_{even}. We calculate the cross-coherence ($\mu(|b|, \tau)$), i.e. the normalized cross power spectrum as

$$\mu(|b|, \tau) = \frac{|\mathcal{V}_{even}\mathcal{V}_{odd}^*|}{\sqrt{|\mathcal{V}_{even}|^2|\mathcal{V}_{odd}|^2}} = \frac{|\tilde{I}_{even}\tilde{I}_{odd}^*|}{\sqrt{|\tilde{I}_{even}|^2|\tilde{I}_{odd}|^2}}. \tag{23}$$

To determine $\mu(|b|, \tau)$, we calculate the 3D power spectra $|\tilde{I}_{even}|^2$ and $|\tilde{I}_{odd}|^2$ and the cross-power spectrum $|\tilde{I}_{even}\tilde{I}_{odd}^*|$. We perform azimuthal averaging to obtain the corresponding delay power spectra. Finally, we use these delay power spectra to calculate $\mu(|b|, \tau)$ in delay-baseline space. Fig. 11 shows the cross-coherence in the direction of Cas A before and after subtracting Cas A. Left-hand panel of Fig. 11 shows the cross-coherence between I_{even} and I_{odd} in direction of Cas A. We observe that $\mu(|b|, \tau) \sim 0.8$–1.0 for $|b| \lesssim 400$ m and drops afterwards. The middle panel of Fig. 11 shows $\mu(|b|, \tau)$ after the subtraction of Cas A with its DD gain solutions. We notice that effectively all correlation disappears, suggesting that most of the Cas A residuals seen in Fig. 7 are incoherent over 5 min intervals as expected for ionospheric scintillation noise.

Fig. 11 (right-hand panel) shows $\mu_{PSF}(|b|, \tau)$ in delay baseline space. Note that the incoherent structure inside primary beam delay line on $|b| > 400$ m in $\mu(|b|, \tau)$, before Cas A subtraction, correlates with the structure at the same location in $\mu_{PSF}(|b|, \tau)$. We can attribute this structure to the migration of baselines from one *uv*-cell to another in 5 min time-scale. In the *uv*-plane, a typical baseline vector $\mathbf{b}(u, v)$ with small baseline length will traverse a smaller distance in a given time compared to a baseline vector with larger baseline length. This migration of baseline vector across the *uv*-plane mixes with the frequency dependence of the baseline vector to produce this incoherence effect in delay-baseline space. This effect is purely a *uv*-plane sampling effect and appears in cross-coherence between PSF$_{even}$ and PSF$_{odd}$.

6 CONCLUSIONS AND SUMMARY

The LOFAR-EoR project aims at statistical detection of HI signal from redshifts $z = 7$–12 and to measure the 21 cm power spectrum as a function of redshift (Patil et al. 2017). LOFAR also operates at frequencies corresponding to the CD, making it in principle possible to measure or set limits on the CD power spectrum using the LOFAR-LBA system. Several contamination effects such as foreground contamination, instrumental polarization, ionospheric effects, and calibration effects make the detection of redshifted 21 cm emission from neutral hydrogen at high redshifts an extremely challenging task. These contamination effects are orders of magnitude stronger than the expected signal in terms of the brightness temperature. Therefore, understanding the nature of these contaminants and how they corrupt the 21 cm power spectrum becomes a crucial step

in the calibration and signal extraction process. In this paper, we use several techniques such as the differential power spectrum, delay power spectrum and cross-coherence to study various contamination effects in LOFAR-LBA data at low frequencies (56–70 MHz). The main results of the paper are summarized below:

(i) We find that the excess power in the differential power spectrum of Stokes I is $\sim$10 times larger than that of Stokes V. A similar behaviour has been observed in HBA observations, but it is by far not as severe as we observe in our analysis. This ratio is almost flat and does not change between the two calibration strategies with or without a baseline cut (i.e. using $|b| \geq 200\lambda$ or using all baselines in calibration), even though the power spectra themselves change. The reasons for this excess power might be incomplete sky-model, ionospheric effects and/or imperfect calibration.

(ii) Introducing a baseline cut in calibration decreases the power on baselines outside the cut and increases it on the baselines inside the cut similar to Patil et al. (2016). However, the power in Stokes I, when using all baselines in calibration, seems to decrease to smaller scales. Some decrease in power might occur when a diffuse sky-model is not included in the calibration and is calibrated away.

(iii) The discontinuity in the ratio of differential Stokes I and V power spectra for two calibration strategies appears at the location of calibration cut. We suggest that this effect is purely an artefact of the calibration cut. If the gains estimated during the calibration process using only a subset of baselines are erroneous, then the errors on gain estimates might transfer to smaller baselines, which are excluded in the calibration process. This enhances the excess noise on the excluded baselines (e.g. Barry et al. 2016; Patil et al. 2016). These errors on gains might occur due to incomplete sky-model and/or ionospheric scintillations.

(iv) We observe a 'pitchfork' structure in the delay power spectrum of total polarized intensity ($\mathcal{P}$). The 'pitchfork' structure appears due to bright sources (Cas A in our case) leaking from Stokes I to $\mathcal{P}$ due to instrumental polarization. Most of the power on and around this structure disappears when Cas A is subtracted (using DD calibration). The residual power after Cas A subtraction correlates strongly with the power before Cas A subtraction, suggesting inaccurate Cas A model and/or imperfect source subtraction during DD calibration. Subtraction of sources within the primary beam does not affect the 'pitchfork'.

(v) Inclusion of short baselines in the calibration scheme suppresses the residual power around the 'pitchfork' compared to the scheme where the short baselines are excluded from the calibration step. We expect that any unmodelled flux (diffuse) outside the primary beam gets absorbed in the gains when short baselines are used in the calibration step suppressing the power around the 'pitchfork'. We show that the delay spectrum of $\mathcal{P}$ is a scaled down version of the Stokes I delay spectrum.

(vi) Ionospheric scintillations are dominant at low frequencies. The power spectrum of Cas A at small baselines (where Cas A can be treated as compact source) takes the form of a compact source corrupted by Kolmogorov-type turbulence. We observe extremely small ionospheric diffractive scales $r_{diff} \sim 400$ m towards Cas A. To our knowledge, these are the smallest scales ever measured at 60 MHz. The power spectrum of $\mathcal{P}$ in direction of Cas A fits very well with the Kolmogorov-type turbulence and appears to be a scaled down version of the Stokes I power spectrum, which is another confirmation of the strong instrumental polarization leakage in LBA.

(vii) Cross-coherence between two residuals images of Cas A, when rotated to the phase centre disappears on 5 min intervals. This

suggests that the residuals after Cas A subtraction are incoherent as expected for ionospheric scintillation noise, even though the coherent part of the source should be nearly constant in time for a source in the phase centre. This also point towards strong ionospheric activity during observation. The ionosphere typically decorrelates on time-scales of ∼10 s at lower frequencies. Solving for ionospheric effects in direction-dependent calibration step requires solutions intervals ≤10 s. This requires effectively calibrating each visibility snapshot and also requires higher SNR (to achieve better quality solutions) than afforded by the current LOFAR-LBA data.

The contamination effects which we discussed in this work, although in part identified in LOFAR-HBA data at frequencies around 150 MHz, appear much stronger in LOFAR LBA data. This can in part be not only due to the small diffractive scale of the ionosphere but also due to the calibration process and the incomplete sky-model. The level of these effects we have observed in our study is a clear indication that these and other far-field effects (such as scintillation of Cas A) pose much more severe concerns in current/upcoming CD experiments compared to the EoR experiments. These effects need to be accounted for before the thermal noise (or Stokes *V* rms) level can be reached at frequencies relevant for 21 cm CD observations. In upcoming CD experiments, such as with SKA-low, NENUFAR and LEDA, which will observe in the frequency range of 30–80MHz, and will probe the same short baselines as studied here, these effects have to be mitigated to an accuracy of ∼0.01 per cent or be incoherent and below the thermal noise such that they average down in time in order to get a detection. This study will prove to be helpful in understanding the behaviour of these contamination effects at low frequencies and mitigating them.

ACKNOWLEDGEMENTS

REFERENCES

Asad K. M. B. et al., 2015, MNRAS, 451, 3709
Asad K. M. B. et al., 2016, MNRAS, 462, 4482
Asad K. M. B., Koopmans L. V. E., Jelić V., de Bruyn A. G., Pandey V. N., Gehlot B. K., 2018, MNRAS, 476, 3051
Asgekar A. et al., 2013, A&A, 551, L11
Aslanian A. M., Dagkesamanskii R. D., Kozhukhov V. N., Malumian V. G., Sanamian V. A., 1968, Astrofizika, 4
Baars J. W. M., Genzel R., Pauliny-Toth I. I. K., Witzel A., 1977, A&A, 61, 99
Barning F. J. M., 1963, Bull. Astron. Inst. Netherlands, 17, 22
Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, MNRAS, 461, 3135
Becker R. H. et al., 2001, AJ, 122, 2850
Bernardi G. et al., 2010, A&A, 522, A67
Bolton J. S., Becker G. D., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2010, MNRAS, 406, 612
Bowman J. D. et al., 2013, PASA, 30, e031
Bregman J., 2012, System Design and Wide-field Imaging Aspects of Synthesis Arrays with Phased Array Stations, PhD thesis, University of Groningen
Brentjens M. A., de Bruyn A. G., 2005, A&A, 441, 1217
DeBoer D. R. et al., 2017, PASP, 129, 045001
Ewall-Wice A. et al., 2016, MNRAS, 460, 4320
Ewall-Wice A., Dillon J. S., Liu A., Hewitt J., 2017, MNRAS, 470, 1849
Fan X. et al., 2003, AJ, 125, 1649
Fan X. et al., 2006, AJ, 131, 1203
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Hamaker J. P., Bregman J. D., Sault R. J., 1996, A&AS, 117, 137
Hinshaw G. et al., 2013, ApJS, 208, 19
Jelić V. et al., 2015, A&A, 583, A137
Kassim N. E. et al., 2007, ApJS, 172, 686
Kazemi S., Yatawatta S., 2013, MNRAS, 435, 597
Kazemi S., Yatawatta S., Zaroubi S., Lampropoulos P., de Bruyn A. G., Koopmans L. V. E., Noordam J., 2011, MNRAS, 414, 1656
Kazemi S., Yatawatta S., Zaroubi S., 2013, MNRAS, 430, 1457
Kohn S. A. et al., 2016, ApJ, 823, 88
Komatsu E. et al., 2011, ApJS, 192, 18
Koopmans L. V. E., 2010, ApJ, 718, 963
Koopmans L. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 1
Lomb N. R., 1976, Ap&SS, 39, 447
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Mellema G. et al., 2013, Exp. Astron., 36, 235
Mevius M. et al., 2016, Radio Sci., 51, 927
Mohan N., Rafferty D., 2015, Astrophysics Source Code Library, record ascl:1502.007
Offringa A. R., de Bruyn A. G., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010, MNRAS, 405, 155
Offringa A. R., van de Gronde J. J., Roerdink J. B. T. M., 2012, A&A, 539, A95
Offringa A. R. et al., 2013a, MNRAS, 435, 584
Offringa A. R. et al., 2013b, A&A, 549, A11
Offringa A. R. et al., 2014, MNRAS, 444, 606
Ono Y. et al., 2012, ApJ, 744, 83
Paciga G. et al., 2011, MNRAS, 413, 1174
Page L. et al., 2007, ApJS, 170, 335
Pandey V. N., van Zwieten J. E., de Bruyn A. G., Nijboer R., 2009, in Saikia D. J., Green D. A., Gupta Y., Venturi T., eds, ASP Conf. Ser. Vol. 407, The Low-Frequency Radio Universe, Astron. Soc. Pac., San Francisco. p. 384
Parsons A. R., Backer D. C., 2009, AJ, 138, 219
Parsons A. R. et al., 2010, AJ, 139, 1468
Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012, ApJ, 756, 165
Patil A. H. et al., 2016, MNRAS, 463, 4317
Patil A. H. et al., 2017, ApJ, 838, 65
Pentericci L. et al., 2011, ApJ, 743, 132
Planck Collaboration XLVII, 2016, A&A, 596, A108
Price D. C. et al., 2018, MNRAS, sty1244
Pritchard J. R., Furlanetto S. R., 2007, MNRAS, 376, 1680
Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901
Rufenach C. L., 1972, J. Geophys. Res., 77, 4761
Schenker M. A., Stark D. P., Ellis R. S., Robertson B. E., Dunlop J. S., McLure R. J., Kneib J.-P., Richard J., 2012, ApJ, 744, 179
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Singleton D. G., 1974, J. Atmos. Terr. Phys., 36, 113
Smirnov O. M., 2011a, A&A, 527, A106
Smirnov O. M., 2011b, A&A, 527, A107
Stoica P., Li J., He H., 2009, IEEE Trans. Signal Process., 57, 843
Theuns T., Schaye J., Zaroubi S., Kim T.-S., Tzanavaris P., Carswell B., 2002, ApJ, 567, L103
Thyagarajan N. et al., 2015a, ApJ, 804, 14
Thyagarajan N. et al., 2015b, ApJ, 807, L28
Tingay S. J. et al., 2013, PASA, 30, e007
Trott C. M. et al., 2016, ApJ, 818, 139
van Haarlem M. P. et al., 2013, A&A, 556, A2
van Straten W., 2009, ApJ, 694, 1413
Vedantham H. K., Koopmans L. V. E., 2015, MNRAS, 453, 925
Vedantham H. K., Koopmans L. V. E., 2016, MNRAS, 458, 3099

Wang F. Y., 2013, A&A, 556, A90

Yatawatta S., 2013, Exp. Astron., 35, 469

Yatawatta S., 2015, MNRAS, 449, 4506

Yatawatta S., 2016, 24th European Signal Processing Conference (EU-SIPCO), EURASIP , Hungary p. 265

Zahn O. et al., 2012, ApJ, 756, 65

Zarka P., Girard J. N., Tagger M., Denis L., 2012, in Boissier S., de Laverny P., Nardetto N., Samadi R., Valls-Gabaud D., Wozniak H., eds, SF2A-2012: Proceedings of the Annual Meeting of the French Society of Astronomy and Astrophysics, SF2A, France. p. 687

Zaroubi S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library Vol. 396, The First Galaxies, Springer, Germany. p. 45 (arXiv:1206.0267)

Zheng Q., Wu X.-P., Johnston-Hollitt M., Gu J.-h., Xu H., 2016, ApJ, 832, 190

APPENDIX: RM SYNTHESIS

The rotation of the polarization angle (χ) of an electromagnetic wave, while propagating through magnetized plasma is called Faraday rotation. The value of χ depends on the frequency (ν) of the wave, electron density (n_e), and magnetic field component parallel to the line of sight ($B_{\parallel}$). For a single Faraday screen, χ can be written as $\chi = \chi_{\circ} + \Phi\lambda^2$, where $\chi_{\circ}$ is the intrinsic polarization angle of the wave and Φ is the Faraday depth, which can be expressed as

$$\frac{\Phi}{[\text{rad}\,\tilde{}\,\text{m}^{-2}]} = 0.81 \int_{\text{source}}^{\text{observer}} \frac{n_e}{[\text{cm}^{-3}]} \frac{B_{\parallel}}{[\mu\text{G}]} \frac{\text{d}l}{[\text{pc}]}. \tag{A1}$$

The RM is defined as the slope of $\chi(\lambda^2)$:

$$\text{RM} = \frac{\text{d}\chi(\lambda^2)}{\text{d}\lambda^2}, \quad \text{where } \chi = \frac{1}{2}\tan^{-1}\left(\frac{U}{Q}\right), \tag{A2}$$

and Q and U are the Stokes parameters. The RM synthesis technique (Brentjens & de Bruyn 2005) takes the advantage of the λ^2 dependence of the complex polarized emission $\mathcal{P}(\lambda^2) = Q(\lambda^2) + iU(\lambda^2)$. Using this, the complex Faraday dispersion function $F(\Phi)$ (which measures the Faraday rotation) can be defined as

$$F(\Phi) = R(\Phi) * \int_{-\infty}^{\infty} \mathcal{P}(\lambda^2)e^{-2i\Phi\lambda^2}\text{d}\lambda^2, \tag{A3}$$

where $R(\Phi)$ is the Fourier transform of the wavelength sampling function $W(\lambda^2)$ and is known as the rotation measure spread function (RMSF). Rotation measure (RM) synthesis can be used to distinguish between intrinsic and instrumental polarization by examining the polarized emission in RM space. Before performing the RM synthesis, the time varying ionospheric Faraday rotation has to be corrected. This correction is a global Faraday rotation correction (or de-rotation) with low time resolution ($\sim$15 min in our case) and applies a single correction (corresponding to the phase centre) to the entire field. It does not correct for any differential Faraday rotation with variations on shorter time-scales and as a function of position. The de-rotation can be performed after the DI calibration. We use the RM estimates as a function of time from the GPS data (see e.g. Fig. 9) to perform de-rotation using the BBS package. We have produced RM cubes (we do not show any RM cubes in the paper) before and after applying de-rotation. We used Stokes Q and U images produced with 200λ imaging scheme (see Table 3) for RM synthesis, because diffuse polarized emission is significant only on small baselines. We observed that before de-rotation, $\mathcal{P}$ in RM space appears noise-like except at $\Phi = 0$, which is dominated by the instrumental polarization leakage. The de-rotation causes the emission due to polarization leakage at $\Phi = 0$ to move to some other Faraday depth whose value depends on the integrated RM over the duration of observation. Apart from this shift, $\mathcal{P}$ RM cubes appear similar (noise-like) to RM cubes before de-rotation.

Amount of depolarization due to time varying ionospheric RM can be estimated using RM values as a function of time (t). For time-dependent RM (RM (t)), $\chi(t)$ can be written as

$$\chi(t) = \chi_{\circ} + \text{RM}(t) \times \lambda^2 , \tag{A4}$$

Since $\mathcal{P} = Q + iU$, the variation in Stokes Q and U due to $\chi(t)$ is

$$Q(t) = \mathcal{P}_{\circ}\cos\chi(t) \text{ and } U(t) = \mathcal{P}_{\circ}\sin\chi(t) . \tag{A5}$$

The remaining total polarized intensity $|\mathcal{P}|$ (after ionospheric depolarization) after integrating over the observation time is given by

$$|\mathcal{P}| = \sqrt{\langle Q \rangle_t^2 + \langle Q \rangle_t^2}. \tag{A6}$$

Assuming $\chi_{\circ} = 0$ and $\mathcal{P}_{\circ} = 1$ gives fractional polarized intensity after the depolarization. We observed that time-varying ionospheric RM produces depolarization of $\sim$75–80 per cent for 56–70MHz frequency range. Jelić et al. (2015) observed bright polarized emission ($\sim$10 K) in 3C196 field at 150 MHz using LOFAR-HBA observations. Assuming a spectral index of -2.55, we get polarized emission of $\sim$100 K at 60 MHz. After taking ionospheric depolarization into account, we expect polarized emission of $\sim$20 K at 60 MHz. Since, we do not observe any polarized emission in 3C196 field at LBA frequencies, it means either Galactic polarized emission at low frequencies is depolarized more than 5 per cent by intervening magneto-ionic medium because Faraday rotation scales as λ^2; or the differential Faraday rotation due to the ionosphere is significant, since we only correct for the phase centre. Thus, a combination of both Galactic and ionospheric depolarization might be the cause of the absence of any polarized emission at low frequencies.

The resolution in Faraday depth space is $\delta\Phi = 2\sqrt{3}/(\lambda_{\text{min}}^2 - \lambda_{\text{max}}^2)$, which corresponds to the full width at half-maximum of the RMSF. For the frequency range of 56–70MHz, $\delta\Phi \approx 0.33$ rad m^{-2}, while the largest structure that can be resolved is only $\Delta\Phi_{\text{max}} = \pi/\lambda_{\text{min}}^2 \approx 0.17$ rad m^{-2}. Whereas at 150 MHz, $\delta\Phi \sim 1.75$ rad m^{-2} and $\Delta\Phi_{\text{max}} \sim 1.15$ rad m^{-2} which is almost an order of magnitude larger compared to lower frequencies. It is possible that the polarized structures observed in 3C196 field at 150 MHz are Faraday thick ($\lambda^2\Delta\Phi >> 1$) at lower frequencies and therefore cannot be observed in LBA. This is similar to Faraday thick structures in 3C196 field at 150 MHz which are not observed with LOFAR-HBA but have been detected at 350 MHz with WSRT (see section 6 in Jelić et al. 2015).

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