

UPPER LIMITS ON THE 21-CM EPOCH OF REIONIZATION POWER SPECTRUM FROM ONE NIGHT WITH LOFAR

A.H. PATIL¹, S. YATAWATTA^{1,2}, L.V.E. KOOPMANS^{1,†}, A.G. DE BRUYN^{2,1}, M. A. BRENTJENS², S. ZAROUBI^{1,11}, K. M. B. ASAD¹, M. HATEF¹, V. JELIĆ^{1,8,2}, M. MEVIUS^{1,2}, A. R. OFFRINGA², V.N. PANDEY¹, H. VEDANTHAM^{9,1}, F. B. ABDALLA^{7,13}, W. N. BROUW¹, E. CHAPMAN⁷, B. CIARDI⁴, B. K. GEHLOT¹, A. GHOSH¹, G. HARKER^{3,7,1}, I. T. ILIEV¹⁰, K. KAKIICHI⁴, S. MAJUMDAR¹², M. B. SILVA¹, G. MELLEMA⁵, J. SCHAYE⁶, D. VRBANEC⁴, S. J. WIJNHOLDS²

¹Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

²ASTRON, P.O.Box 2, 7990 AA Dwingeloo, The Netherlands

³Center for Astrophysics and Space Astronomy, Dept. of Astrophysics and Planetary Sciences, University of Colorado at Boulder, CO 80309, USA

⁴Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany

⁵Department of Astronomy and Oskar Klein Centre for Cosmoparticle Physics, AlbaNova, Stockholm University, SE-106 91 Stockholm, Sweden

⁶Leiden Observatory, Leiden University, PO Box 9513, 2300RA Leiden, The Netherlands

⁷Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT, London, UK

⁸Ruder Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

⁹Cahill Center for Astronomy and Astrophysics, MC 249-17, California Institute of Technology, Pasadena, CA 91125, USA

¹⁰Astronomy Centre, Department of Physics & Astronomy, Pevenssey II Building, University of Sussex, Brighton BN1 9QH, UK

¹¹Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel

¹²Department of Physics, Blackett Laboratory, Imperial College, London SW7 2AZ, UK

¹³Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa

ABSTRACT

We present the first limits on the Epoch of Reionization (EoR) 21-cm HI power spectra, in the redshift range $z = 7.9 - 10.6$, using the Low-Frequency Array (LOFAR) High-Band Antenna (HBA). In total 13.0 h of data were used from observations centred on the North Celestial Pole (NCP). After subtraction of the sky model and the noise bias, we detect a non-zero $\Delta_1^2 = (56 \pm 13 \text{ mK})^2$ ($1-\sigma$) excess variance and a best $2-\sigma$ upper limit of $\Delta_{21}^2 < (79.6 \text{ mK})^2$ at $k = 0.053 \text{ h cMpc}^{-1}$ in the range $z = 9.6 - 10.6$. The excess variance decreases when optimizing the smoothness of the direction- and frequency-dependent gain calibration, and with increasing the completeness of the sky model. It is likely caused by (i) residual side-lobe noise on calibration baselines, (ii) leverage due to non-linear effects, (iii) noise and ionosphere-induced gain errors, or a combination thereof. Further analyses of the excess variance will be discussed in forthcoming publications.

Keywords: cosmology: theory - large-scale structure of Universe - observations - diffuse radiation - methods: statistical - radio lines: general - cosmology: dark ages, reionization, first stars

1. INTRODUCTION

During the Epoch of Reionization (EoR) hydrogen gas in the universe transitioned from neutral to ionized (Madau et al. 1997). The EoR is thought to be caused by the formation of the first sources of radiation and hence its study is important for understanding the nature of these first radiating sources, the physical processes that govern them and how they influence the formation of subsequent generations of stars, the interstellar medium (ISM), intergalactic medium (IGM) and black holes; see e.g. Furlanetto et al. (2006); Morales & Wyithe (2010); Pritchard & Loeb (2012); Natarajan & Yoshida (2014); McQuinn (2015) for extensive reviews of the EoR.

Current observational constraints suggest that reionization took place in the redshift range $6 \lesssim z \lesssim 10$, with the lower

limit inferred from the Gunn-Peterson trough in high-redshift quasar spectra (Becker et al. 2001; Fan, et al. 2003, 2006), and the upper limit of the redshift range currently being set by the most recent Planck results, which yields a surprisingly low value of the optical depth for Thomson scattering, $\tau_e = 0.058 \pm 0.012$ (Planck Collaboration 2016). This small optical depth mitigates the tension that exists between the higher optical depth values obtained by the WMAP satellite (Page et al. 2007; Komatsu et al. 2011; Hinshaw et al. 2013) and the other probes. The current range can easily accommodate photo-ionisation rate measurements (Bolton & Haehnelt 2007; Calverley et al. 2011; Becker et al. 2011), Inter-Galactic Medium (IGM) temperature measurements (Theuns et al. 2002; Bolton et al. 2010; Becker & Bolton 2013), observations of high-redshift Lyman break galaxies at $7 \lesssim z \lesssim 10$ (see e.g. Oesch et al. 2010; Bouwens et al. 2010; Bunker et al. 2010; Bouwens et al. 2015; Robertson et al. 2015) and observation of Lyman- α emitters at $z = 7$ (see e.g. Schenker et al. 2014; Santos et al. 2016).

[†]koopmans@astro.rug.nl

Phase Centre (α, δ ; J2000)	$0^{\text{h}}, +90^{\circ}$	
Minimum frequency	115.039	MHz
Maximum frequency	189.062	MHz
Target bandwidth	74.249	MHz
Stations (core/remote)	48 / 13	
Raw data volume L90490	61	Tbyte
Sub-band (SB) width	195.3125	kHz
Correlator channels per SB	64	
Correlator integration time	2	s
Channels per SB after averaging	15, 3, 3, 1	
Integration time after averaging	2, 2, 10, 10	s
Data size (488 sub-bands)	50	Tbyte

Table 1. Observational and correlator set up of LOFAR-HBA observations of the North Celestial Pole (NCP).

It has been long recognized that the redshifted 21-cm emission line provides a very promising probe to observe neutral hydrogen during the EoR (see e.g. Madau et al. 1997; Shaver et al. 1999; Furlanetto et al. 2006; Pritchard & Loeb 2012; Zaroubi 2013).

To date, a number of experiments have been seeking to measure this high-redshift 21-cm emission, using LOFAR (van Haarlem et al. 2013), the GMRT (Paciga et al. 2011), the MWA (Tingay et al. 2013; Bowman et al. 2013), PAPER (Parsons et al. 2010) and the 21CMA (Zheng et al. 2016). These experiments are designed to detect the cosmological 21-cm signal through a number of statistical measures of its brightness-temperature fluctuations, such as its variance (e.g. Patil et al. 2014; Watkinson & Pritchard 2014) and its power spectrum as a function of redshift (e.g. Morales & Hewitt 2004; Bharadwaj & Ali 2005; Barkana & Loeb 2005; McQuinn et al. 2006; Bowman et al. 2006; Pritchard & Furlanetto 2007; Jelić et al. 2008; Harker et al. 2009, 2010; Pritchard & Loeb 2008).

In particular, Jelić et al. (2008), Harker et al. (2010) and more recently Chapman et al. (2013, 2016) have shown that despite the low signal-to-noise ratio and prominent Galactic and extragalactic foreground emission, the variance and power spectrum of the brightness temperature fluctuations of HI can be extracted from the data collected with LOFAR in about 600 hours of integration time on five fields, barring unknown systematic errors. Deeper integrations on fewer fields can yield similar results¹. Similar studies have been carried out for the MWA (see e.g. Geil et al. 2008, 2011; Beardsley et al. 2013) and for PAPER (see e.g. Parsons et al. 2012).

At present, a number of upper limits on the brightness-temperature power spectrum have been published. Paciga et al. (2013) have used the GMRT to set a 2- σ upper limit on the brightness temperature at $z = 8.6$ of $\Delta_{21}^2 < (248 \text{ mK})^2$ at wave number $k \approx 0.5 \text{ h cMpc}^{-1}$. Beardsley et al. (2016) provided a 2- σ limit at $z = 7.1$ of $\Delta_{21}^2 < (164 \text{ mK})^2$ at

$k \approx 0.27 \text{ h cMpc}^{-1}$ from MWA. The PAPER project provided the tightest upper limit yet of $\Delta_{21}^2 < (22.4 \text{ mK})^2$ in the wave number range $0.15 \leq k \leq 0.5 \text{ h cMpc}^{-1}$ at $z = 8.4$ (Ali et al. 2015).

Here we report the first 21-cm EoR power-spectrum limits from the LOFAR EoR Key Science Project (KSP) based on a single night of data set acquired in the first LOFAR observing cycle (i.e. Cycle-0). The approach taken in the LOFAR EoR project differs in two important aspects from those in the other experiments mentioned above. Firstly, in order to remove the chromatic response from the multitude of bright continuum sources found in a typical LOFAR observation we have developed a comprehensive sky model. This model is then used to calibrate the data in a large number of directions. We then also remove these sources and their responses from the visibility data. Secondly, we use a technique that goes by the name of Generalized Morphological Component Analysis (GMCA, hereafter GMCA) to remove the residual compact and remaining diffuse foregrounds. Both aspects, as applied to real data, have not been described in detail before. We will therefore describe these processing steps, and how we have arrived at the chosen parameters and strategy, in some detail.

This paper is organized as follows. In Sect. 2 we describe the observational set-up and the data that is being analysed. In Sect. 3 we describe the various steps in our data-processing. In Sect. 4 we describe the calibration of our data. In Sect. 5 our imaging procedures are described. The resulting power spectra are presented in Sect. 6. The paper concludes with a summary and outlook in Sect. 7. We assume the standard cosmology (Collaboration et al. 2015) and scale the Hubble constant as $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2. OBSERVATIONS

The observations conducted for the LOFAR EoR project are concentrated on two windows: the North Celestial Pole (NCP) and the bright compact radio source 3C196 (see Bernardi et al. 2010; Yatawatta et al. 2013). The results presented in this paper are based on data taken on the NCP field with the LOFAR telescope (van Haarlem et al. 2013) in the night from 2013 February 11/12. The frequency range from 115 to 189 MHz was covered using receivers in the so-called LOFAR-HBA band (where HBA refers to High Band Antenna). All 61 Dutch LOFAR-HBA stations (e.g. van Haarlem et al. 2013, and Table 1) available in early 2013, participated in the observations.

2.1. Data Sets

NCP observations are usually scheduled from “Dusk to Dawn”, and have typical durations of 12–15.5 h during the Northern hemisphere winter. The phase and pointing centre was set at RA= 0^{h} , DEC= $+90^{\circ}$ (Table 1). The NCP can be observed every night of the year making it an excellent EoR window. Currently ~ 800 h of good-quality data have been acquired during Cycles 0–5², under generally good ionospheric

¹ The power spectrum error scales inverse proportional with the integration time and with the square root of the **number of fields**, respectively. This holds in the thermal-noise dominated and low-S/N regime.

² <http://www.astron.nl/radio-observatory/cycles-allocations-and-observing-schedules/cycles-allocations-and-observing-schedu>

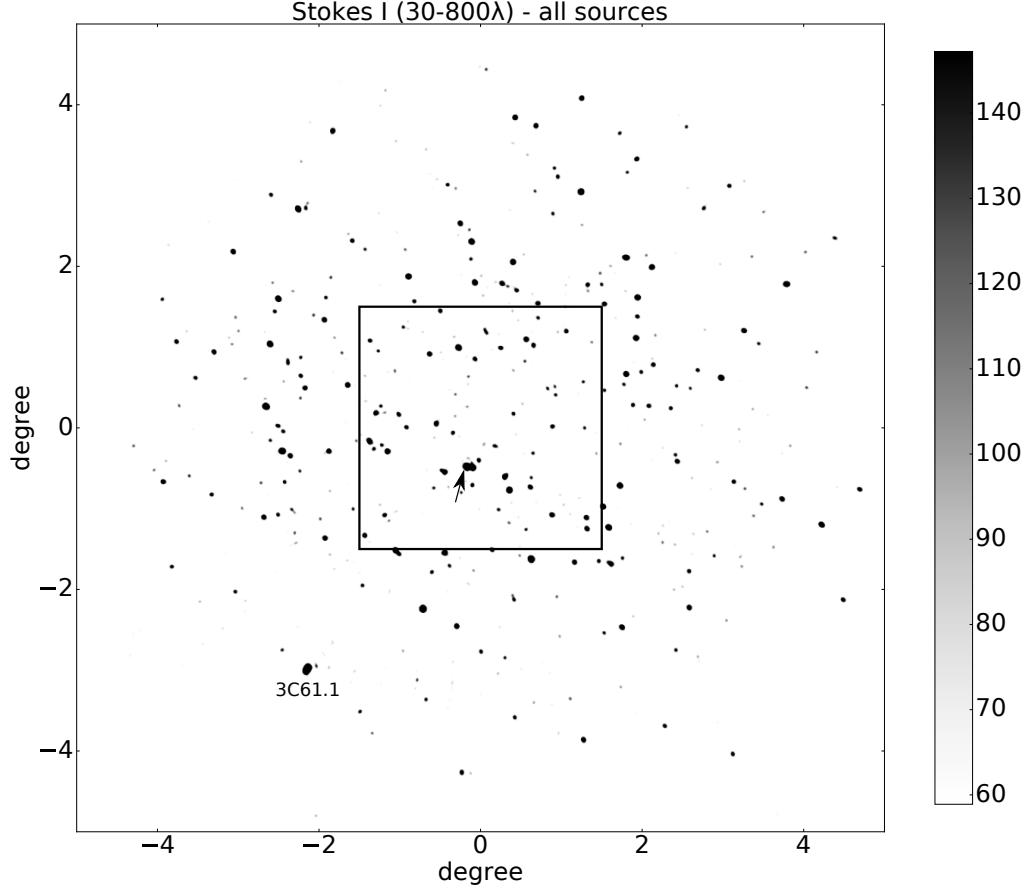


Figure 1. A relatively narrow-band continuum (134.5–137.5 MHz) LOFAR-HBA image of $10^\circ \times 10^\circ$ of the North Celestial Pole (NCP) field, centered at dec $+90.0^\circ$. Baselines between $30\text{--}800\lambda$ were included, using uniform weighting. No sources have been subtracted and the image is cleaned to a level sufficient to show the brightest few hundred sources above 60 mJy. The $3^\circ \times 3^\circ$ box delineates the area where we measure the power spectra. The bright extended source in the lower-left is 3C61.1 (J0222+8619), discussed in the text. The bright (7.2 Jy) compact source near the NCP is indicated by an arrow. The intensity units are mJy/PSF (see text). Right Ascension increases clockwise; RA=00h is towards the bottom.

conditions (see e.g. [Mevius et al. 2016](#)) and in a moderate RFI environment (e.g. [Offringa et al. 2013](#)). We refer to [Yatawatta et al. \(2013\)](#) for a detailed description of the NCP field and early LOFAR commissioning observations.

For the analyses presented in this paper, a single 13.0-hr data set (i.e. L90490) was selected from a larger set (~ 150 h of data) that was previously analysed with an earlier version of the calibration code SageCal ([Kazemi et al. 2011](#)). The data in this night is of excellent quality, based on the Stokes V rms noise, RFI levels and ionospheric conditions. We recently re-processed this dataset using an improved calibration strategy SageCal-CO (see Sect. 3; [Yatawatta 2015a, 2016](#)) yielding a more robust calibration than previously (used in e.g. [Yatawatta et al. 2013](#)).

2.2. Station Hardware and Correlation

The LOFAR array has a rather complex, hierarchical configuration. Here we give a brief summary, restricting ourselves to the HBA-band configuration in which we recorded

our data. For a more detailed description of LOFAR hardware we refer to [van Haarlem et al. \(2013\)](#).

Individual HBA-dipoles are grouped in units of 4×4 dual-polarisation dipoles. This unit is called a tile. It has a physical dimension of 5×5 m. The 16 dipole signals are combined in a summator, an analogue beam-former, the coefficients of which are regularly updated when we track a source. In the case of the NCP this is not needed. A core station (CS) consists of 24 closely packed tiles; a remote station (RS) has 48 tiles. The core stations are distributed over an area of about 2 km diameter, in co-located pairs of stations which share a receiver cabinet. The remote stations are spread over an area of about 40 km East-West and 70 km North-South. Although all remote stations have 48 tiles we only used the inner 24 tiles in the beam-former in order to give both core and remote stations the same primary beam. The receivers at a LOFAR station digitize the data at 200 MHz clock speed, fully covering the frequency range from 100–200 MHz ([van Haarlem et al. 2013](#)). This produces 512 sub-bands of each

195 kHz bandwidth. The fibre network used to bring signals from the stations to the correlator can transport a maximum of 488 of these 512 sub-bands. The correlator is located at the computing centre at the University of Groningen, about 40 km north of the LOFAR core. We therefore record a total RF bandwidth of 96 MHz (van Haarlem et al. 2013). Of this bandwidth 74 MHz, i.e. all frequencies between 115 and 189 MHz, was allocated to the target field. The remaining 22 MHz were distributed, sparsely covering the same frequency range, over a hexagonal ring of six flanking fields located at an angular distance of 3.75° from the NCP. The flanking-field data are used for calibration purposes, ionospheric studies and construction of models for sources located at the edges of the station (primary) beam. In the LOFAR EoR observations the correlator generates 64 frequency channels, each of 3.1 kHz, per sub-band and stores the visibility data at 2 s time resolution in so-called Measurement Sets (MS). Every sub-band is stored in a separate MS.

2.3. Intensity scale and Noise

The intensity scale in the data is set by the flux density of the very compact source located at RA=01h17m32s, Dec= $-89^\circ 28' 49''$ (J2000). From (unpublished) European-scale LOFAR long baseline data this source is found to have a size of about $0.3''$ and is therefore completely unresolved on the Dutch LOFAR baselines used in this work. Following calibration against 3C295 (Scaife & Heald 2012) we find the source to show a spectrally broad peak at 7.2 Jy in the range from 120–160 MHz. Note that this is its apparent flux at 31 arcmin from the pointing centre which is at DEC= 90° . However, the source bends down at frequencies below 100 MHz and above 200 MHz. We have adopted a constant flux density over the frequency range for which we show data in this paper. We estimate this value to be good to 5% on the flux scale of Scaife & Heald (2012). This flux density is about 30% larger than adopted in Yatawatta et al. (2013), where we presented the first NCP observations with LOFAR-HBA.

The thermal noise in the data is determined using the temporal statistics of the real and imaginary parts of the XY and YX visibilities in narrow 12 kHz channels. These are observed to be Gaussian distributed. The narrow-band visibility noise also correctly predicts the narrow-band image noise as determined from differences between naturally weighted images in all Stokes parameters. At this spectral resolution broad-band instrumental and ionospheric errors indeed cancel almost perfectly. The measured visibility noise implies a System Equivalent Flux Density (SEFD) of ~ 4000 Jy per station, which is close to the expected value in the direction of the NCP, after correcting for the beam gain away from the zenith (see van Haarlem et al. 2013, for the zenith SEFD values).

We note that when we quote peak flux densities of sources, or noise levels in images, we will give them as flux density per synthesized resolution element. This is what is normally called the Point Spread Function (PSF). This convention therefore differs from the terminology used in radio astronomy, which is to quote fluxes per beam. However, phased arrays, such as LOFAR, have a time-variable (primary) beam

which has often lead to confusion. So to be precise, when we refer to flux density per PSF, we refer to the flux density per solid angle as subtended by the PSF. For a Gaussian PSF, as is often used in restored images, the relevant solid angle would then be equivalent to 1.13 times the square of the Full Width at Half Maximum (FWHM) of the PSF.

3. DATA PROCESSING

3.1. Compute and Storage Resources

Processing a single 13-hr LOFAR-HBA data set is computationally expensive and currently takes ~ 50 h on a dedicated compute-cluster consisting of 124 NVIDIA K40 GPUs, called Dawn³ hereafter. Most of the processing time is needed for the calibration, specifically the direction-dependent calibration (see Sect. 4). The imaging step is computationally negligible. We are working on further optimization and automation of the calibration. All data processing on the visibilities is done on Dawn, located at the Centre for Information Technology⁴ of the University of Groningen. Petabyte-storage is distributed over Dawn, a dedicated storage cluster at ASTRON⁵ and at various locations of the LOFAR Long-Term-Archive (LTA).

The LOFAR-EoR data processing pipeline – prior to power-spectrum extraction (Sect. 6) – consists of a large number of steps: (1) Preprocessing and RFI excision, (2) data-averaging, (3) direction-independent calibration (henceforth *DI*-calibration), (4) direction-dependent calibration (henceforth *DD*-calibration) including sky-model subtraction, (5) short-baseline imaging, (6) removal of residual foregrounds. In this section we describe the hardware and software used in steps (1) and (2). The calibration of our data, steps (3) and (4), are described in detail in Sec 4. All data-processing codes are publicly available and links to the source codes and documentation are given where applicable.

3.2. Pre-processing, RFI excision and Data averaging

Standard (tabulated) corrections are applied to the raw visibilities (e.g. flagging of known bad stations or baselines) using NDPPP⁶. RFI-flagging is done on the highest-resolution data using the AOflagger⁷ (Offringa et al. 2012) and leads to a typical loss of $\sim 5\%$ of the LOFAR-HBA *uv*-data.

Several clean data products at different temporal and frequency resolutions are then created. We first flag channels 0, 1, 62, and 63 at the edges of the sub-bands to avoid low-level aliasing effects from the poly-phase filter used to provide the fine frequency resolution. The remaining 60 channels are averaged to 15 new channels each of 12 kHz. These data are

³ https://www.astron.nl/sites/astron.nl/file/cms/201500151-01_Astron_News_Winter_2015-01_%5BWeb%5D.pdf

⁴ <http://www.rug.nl/society-business/centre-for-information-technology/>

⁵ <http://www.astron.nl/>

⁶ http://www.lofar.org/operations/doku.php?id=public:user_software:ndppp

⁷ <https://sourceforge.net/p/aoflagger/wiki/Home/>

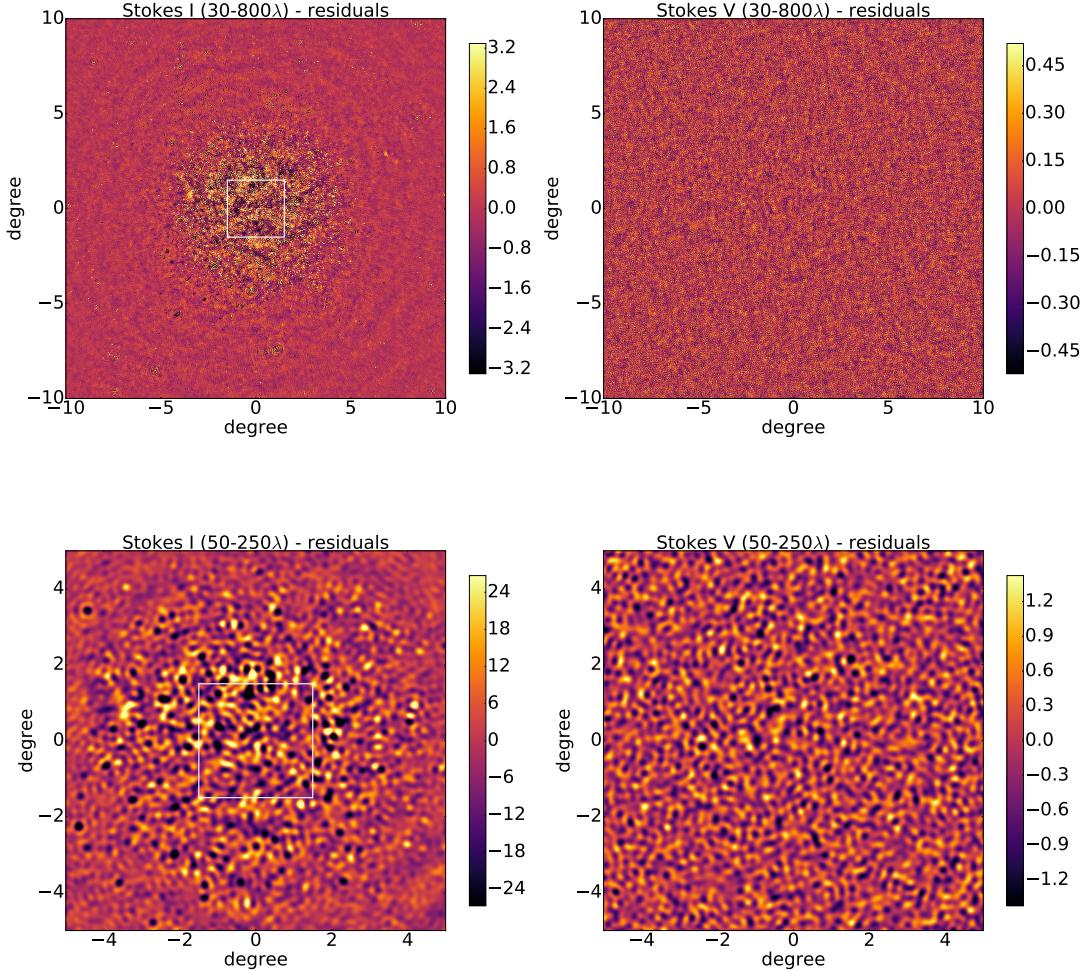


Figure 2. Stokes I and Stokes V images after sky-model subtraction for the baseline ranges $30\text{--}800\lambda$ (top panels) and $50\text{--}250\lambda$ (bottom panels). Sub-bands with frequencies between 121 and 134 MHz went into these images. Note the reduction in the displayed field-of-view from $20\times20^\circ$ to $10\times10^\circ$. Intensity units are in mJy/PSF and the scale range is set by plus and minus three times the standard deviation over the full field in all images. Note the noise-like structure in the two Stokes V images, i.e. a lack of any features. The Stokes I images, on the other hand, clearly show the LOFAR-HBA primary beam attenuation effects on the remaining diffuse emission. The level of this emission is limited by the classical confusion noise within the primary beam. The $3^\circ \times 3^\circ$ box delineates the area where we measure the power spectra.

archived for later analysis (to search for 21-cm absorption in bright sources and permit searches for fast transients). We then further average the data to 3 channels each of 61 kHz while maintaining the 2 s time resolution. At this resolution the time and frequency smearing of off-axis sources is still acceptable at the longest baselines. This is important for high-resolution source modelling (see Sect. 3.3). For initial calibration, we also formed a low-resolution product with a temporal resolution of 10 s (see Table 1). We note that in our previous analysis of the NCP field (Yatawatta et al. 2013) we used a spectral resolution of 183 kHz, i.e. a full sub-band, in the processing. Currently we conservatively flag baselines between stations that share a common electronics cabinet, to avoid any correlated spurious signals. There are 24 such baselines in the LOFAR core. These station pairs have projected baselines between about 40 and 60λ , depending on

frequency. We expect to recover most of these data in forthcoming analyses, potentially increasing the number of short baselines by up to a factor ~ 2.5 in that range.

3.3. The NCP Sky Model

The continuum foreground for EoR-experiments consist of two distinct components (Shaver et al. 1999). On very short baselines, less than about 10λ , the diffuse Galactic synchrotron emission starts to dominate the visibilities. Also the intense emission of Cas-A and Cyg-A, the two brightest radio sources in the Northern hemisphere located in or close to the Galactic plane, and very far from our EoR windows, occasionally enters a distant side-lobe and will then dominate the visibilities. The shortest baseline in LOFAR is about 35 m and corresponds to about $15\text{--}20\lambda$. This means that the diffuse Galactic component is a) hardly detectable in our data, and, b) also very difficult to model. The more problematic

Parameter	Value	Comments
Sky-model components	~20,800	Compact
Flux-limit sky model	~3 mJy	
Order P_n^S source spectra	3	Polynomial
DI-Calibration directions	2	
DD-Calibration directions	122	Source clusters
Calibration baselines	$\geq 250 \lambda$	
Order B_n^G gain regul.	3	Bernstein Polynomial
Solution interval	10 min	
uv-grid cells	$4.58 \times 4.58 \lambda$	
w-slices	128	
EoR Imaging baselines	50–250 λ	
EoR Imaging FoV	$3^\circ \times 3^\circ$	
EoR pixel size	$0.5' \times 0.5'$	
EoR Imaging Resolution	$\sim 10'$	FWHM
EoR Freq. Resolution	~ 60 kHz	
Redshift range #1	7.9 – 8.7	
Freq. range	146.8 – 159.3 MHz	
GMCA components	6/0	Stokes I/V.
Redshift range #2	8.7 – 9.6	
Freq. range	134.3 – 146.8 MHz	
GMCA components	6/2	Stokes I/V.
Redshift range #3	9.6 – 10.6	
Freq. range	121.8 – 134.3 MHz	
GMCA components	8/2	Stokes I/V.

Table 2. Calibration and sky-model parameters and settings.

component, and the one dominating our images are the extragalactic sources. Most of these have an angular size less than a few arcminutes. Source model components are determined from the highest resolution LOFAR images which have an angular resolution of ~ 6 arcsec FWHM. For some of the brightest sources we have also made use of international baselines in LOFAR, which provide a resolution down to 0.25 arcsec. The discrete source model for the NCP field has been iteratively built up over the last several years, using a program called `buildsky`⁸ (see e.g. Yatawatta et al. 2013). Fig. 1 shows a 3 arcmin resolution $10^\circ \times 10^\circ$ image of the NCP. It reveals the brightest few hundred sources down to a flux density limit of 60 mJy. Our sky model includes sources up to 19 deg distance from the NCP, excluding Cas-A and Cyg-A which are much further away. In fact all sources that are bright enough to cause (chromatic) side-lobes in the inner few degrees of the field were included in our model. We expect this model will continue to grow in the next year when we expect to go deeper. The current calibration sky model (Stokes I) consists of $\sim 20,800$ unpolarized source

components, including Cas-A and Cyg-A (see Table 2). It has components down to ~ 3 mJy, i.e. the apparent flux in our model, which are modelled either as a point-source, multiple Gaussians or shapelets. Each source has a smooth frequency model (polynomial of order 3) which is regularly updated, as data is combined and calibration improves. Although sources down to a few mJy were included in our sky model, our low resolution residual images (see Sect. 5) still show many positive and negative sources with fluxes going up to +50 and -50 mJy. These are located near the brighter sources in the field which still leave residuals following the calibration.

4. CALIBRATION

Our calibration strategy has been developed over a period of several years. In this period we have explored a wide set of processing parameters the choice of which was guided by a combination of information-theoretical arguments, end-to-end simulations, a thorough analysis of the image cubes and the effects of unmodelled structure. To give some insight into the problem we will start with an outline of our calibration strategy.

The NCP field is dominated by two bright sources (see Fig. 1). One of them (J0117+8928) is compact and has a flat spectrum (see Sect. 3.3) and is located only $31'$ from the pointing and phase centre of the observation. The other source (3C61.1; J0222+8619, an FR-II radio galaxy) is located at the edge of the primary field of view. It has a complex morphology with both intense sub-arcsecond as well as arcminute scale structure. However, the most problematic aspect of 3C61.1 is its location close to the first null of the primary beam for the highest frequencies used in this analysis. Because the LOFAR-HBA core station primary beam is much larger at 115 MHz than at 177 MHz, 3C61.1 dominates the visibilities at frequencies below 130 MHz. In fact, the source reaches an apparent flux density of ~ 14 Jy at 115 MHz. The ionospheric phase delays will therefore be dominated by those present towards 3C61.1. This frequency-dependent behaviour is exacerbated by the imperfect knowledge of the beam-gains of the 61 stations close to the edge of the primary beam. The combination of the properties of 3C61.1 forced us to depart from the normal two-step calibration of LOFAR data, which consists of a *DI*-calibration, followed by a *DD*-calibration. In essence, our *DI*-calibration is now done towards two directions simultaneously. We use SageCal-CO for both calibration steps. This is a relatively recent departure of the calibration procedure adopted in the past. The main reason is to make the direction-independent calibration solutions independent of those found towards the bright problematic source 3C61.1. However, to not unnecessarily complicate the description below we will continue to refer to this first step as *DI*-calibration. Table 2 lists the most relevant calibration parameter settings.

4.1. Direction-independent calibration

The *DI*-calibration is done at 61 kHz frequency resolution and 2 s time resolution using all baselines in the array. The sky models for the two directions consist of (i) all sources in the field, dominated by the compact 7.2 Jy source near the

⁸ Included in the SageCal-CO repository:
<https://sourceforge.net/projects/sagecal/>

centre, except 3C61.1, and (ii) the source 3C61.1 itself. In this first step the fast ionospheric phase variations towards the two brightest sources can be solved for. The S/N ratio per sub-band is sufficiently high to work at this high time resolution. We solve for the gains per sub-band of 183 kHz, but use the full frequency domain (see for details Yatawatta 2015a, 2016) to fit for the slow as well as fast variations in frequency. This *DI*-calibration will absorb the structure in the band-pass response of the stations. This structure is due to low-pass and high-pass filters in the signal chain as well as reflections in the coax-cables between tiles and receivers (see e.g. Offringa et al. 2013). In the LOFAR core stations the antennae and receivers are connected via 85 m coax-cables. These cause a 920 ns delayed signal with a relative intensity of -22dB. This causes a $\approx 1\%$ ripple in the gains with a a periodicity of 1.09 MHz. These frequency-ripples are similar for all core stations. The remote stations, on the other hand, have features at 1.09 MHz and 1.38 MHz because two sets of coax-cables with lengths of 85 m and 115 m are used. The frequency-dependent station gains and ionospheric delays found towards 3C61.1 in this first calibration step *therefore do not influence the gain solutions for the other direction*. Finally, we correct the visibilities for the gains found for the full field. Note that we do not yet remove 3C61.1 from the data in this *DI*-calibration step.

4.2. Direction-dependent calibration

We want to create a field of view – from which we want to extract the power spectra – free from as many sources and their artefacts as possible. Most of the bright sources are distributed over an area of about 8° diameter (see Fig. 1) but sources with apparent flux densities down to 3 mJy are found out to radii of at least 10° . Over such a large area the station-beam gains vary enormously and unpredictably (in detail). Also the ionospheric iso-planatic angle is expected, and indeed observed, to be typically $1\text{--}2^\circ$. To remove all these sources will therefore require *DD*-calibration. Hence *DD*-calibration is always associated with subtraction of the sky model. We do not replace these sources in our image cubes with their model (as is often done in Cleaning). We had to find a compromise between the number of directions to solve for beam and ionospheric errors, the maximum baseline to use in calibration, the time scale on which to solve for station gains and ionospheric phases, on the one hand, and the number of constraints provided by the data, on the other. Long baselines provide the most constraints. However, by using long baselines, up to a projected maximum baseline of 70 km, we are vulnerable to ionospheric and sky-model errors. Whereas *DD*-calibration is obviously important, the very large number of parameters for which we have to solve also can lead to ill-conditioning of the problem. This has led to a range of subtle and less subtle consequences, which we will describe below.

DD-calibration is an iterative process described in more detail in Yatawatta (2015b, 2016). We group the sky-model components in 122 directions, called source “clusters” (Kazemi et al. 2013). Most clusters will have a large number of components although its response might occasionally

be dominated by a single source. Clusters are typically 1–2 degrees in diameter. SageCal-CO uses an expectation maximization (EM) algorithm to solve for the four complex gains (full Stokes) in one effective Jones matrix per direction (see e.g. Hamaker et al. 1996; Smirnov 2011). This Jones matrix describes the combination of all direction-dependent effects (i.e. beam errors, ionospheric phase fluctuations, etc.) and is assumed to be the same for all sources in a cluster. We plan to relax this assumption in the future.

The complex gains are solved for all clusters simultaneously. We use a third-order Bernstein polynomial basis function (Yatawatta 2016) in the frequency direction as a regularisation prior on the gain solutions over the full bandwidth. Hence, although the gains are allowed to deviate from the smooth prior, this will be penalized by a quadratic regularization term (i.e. penalty function; see Yatawatta 2016). The regularization constant is optimized to minimize the mean squared error between the gain solutions per sub-band and the smooth third-order Bernstein polynomial basis function. If the regularization constant is chosen too large, the data cannot be fitted, and if chosen too small, the data are over-fitted. This fitting process is iterated typically ~ 30 times, simultaneously optimising the weights of the Bernstein polynomial basis functions and the individual gains for all 122 directions and for all sub-bands (i.e. 195 kHz). The solutions are applied to the separate narrow 61 kHz channels), until convergence is reached.

The solution time intervals are dependent on the strength of the signals in the various clusters and vary between 1 and 20 minutes. This time-scale should be sufficient to fit for the slowly varying station-beam gain variations. However, 20 minutes is too long to capture ionospheric phase variations on most baselines. The isoplanatic angle in a typical LOFAR observation in the HBA-band is typically $1\text{--}2^\circ$. Many of the relatively bright radio sources in the field, and especially those that are not dominating the cluster they are assigned to, will then be imperfectly calibrated and leave residuals. An imperfect calibration of these sources, however, will also influence the gains for the stations involved in the short baselines on which we are most sensitive to EoR signals. This could lead to baseline-dependent decorrelation effects. How these effects manifest themselves in the final residual data on the shortest baselines is still under investigation (see e.g. Vedantham & Koopmans 2016). We expect to reduce the SageCal-CO solution time in the future and also use separate solution intervals for amplitude and phase.

4.3. Suppression of Diffuse Emission

DD-calibration can remove diffuse structures (i.e. power) in Stokes I, Q and U. This has been discussed and documented in detail in Patil et al. (2016). Because our calibration sky model only consists of relatively compact sources, this removal of diffuse emission occurs because of a “conspiracy” of the direction-dependent gains – or equivalently the direction-dependent PSFs – convolving the sky model with extended low-level PSFs and removing structures in the data that are not part of the sky model. Whereas using too few calibration directions leaves artefacts around compact sources,

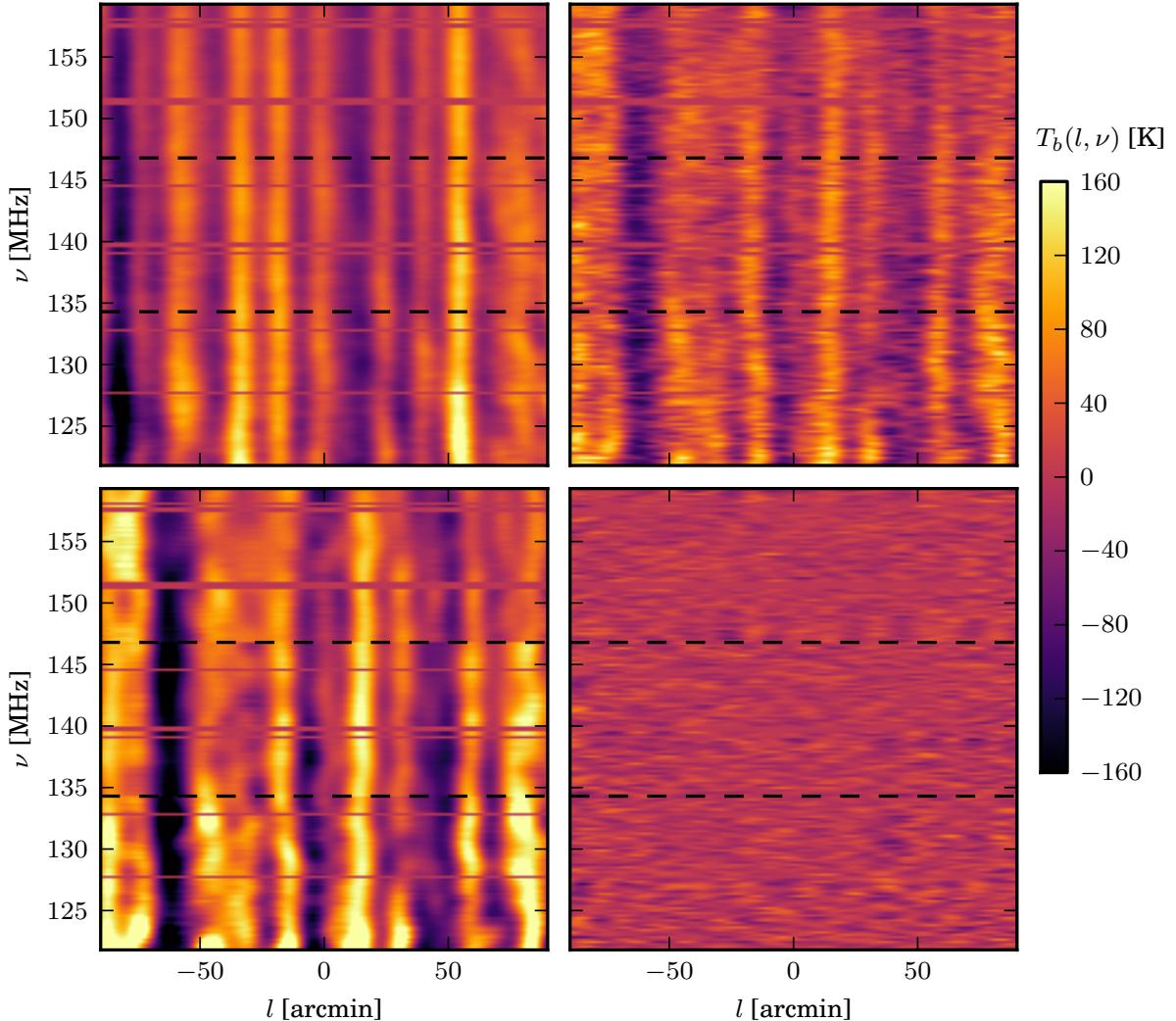


Figure 3. A slice across the centre of the $50-250\lambda$ Stokes-I data cube along the frequency direction. Top left: slice after *DI*-calibration with only 3C61.1 subtracted; the intensity scale, converted to brightness temperature, refers to this panel. Top right: after *DD*-calibration where the calibration sky model, consisting of compact sources, is subtracted with their respective direction-dependent gain solutions. The intensity scale is now multiplied by 10 for improved visualization; Bottom left: GMCA model (scale also multiplied by 10); Bottom right: GMCA residuals (scale multiplied by another factor of 20). The red horizontal bands are due to data lost due to RFI-flagging. The black dashed lines border the three redshift ranges. Note the factor ~ 200 reduction in intensity after GMCA.

using too many will remove structure (Patil et al. 2016). This is opposite (not in contradiction) to the issue noted by Barry et al. (2016), where an incomplete/inaccurate sky model in MWA data simulations causes gain errors on all baselines which then leads to excess variance in the EoR 21-cm power spectrum. To mitigate both problems we split the baseline set into non-overlapping calibration and EoR-imaging subsets, with a cut at several hundred λ , beyond which we see no evidence for diffuse emission in Stokes I, Q and U. We calibrate using the longer baselines and we analyse the EoR signal on the shorter baselines. Furthermore we use our high-resolution images to create a sky-model that reaches well below the classical confusion noise level corresponding to the resolution of the $50 - 250\lambda$ baselines (see Sect. 3.3; Fig. 2). We have tested the effects of both higher and lower cuts. The chosen cut of 250λ is the compromise adopted in our current processing. This value remains well above the baseline

lengths where (realistically speaking) LOFAR could detect an EoR signal.

We note that if diffuse emission can be included in the model, the baseline cut may not be needed. This is still under investigation and some encouraging results have already been obtained.

4.4. Excess Noise

Whereas an imposed baseline-cut largely resolves the issue of suppression of diffuse emission, it leads to excess noise on the short (imaging) baselines (see Patil et al. 2016, their Figures 11 and 12), while simultaneously decreasing the noise and unmodelled flux on long (calibration) baselines. This discontinuous change in the noise level, at the location of the uv-cut, is absent when we calibrate using *all* baselines as we did in our original calibration strategy. Extensive simulations show that this excess variance on the short baselines

that are excluded in the calibration can be caused by three effects (e.g. Patil et al. 2016):

Leverage — Leverage is an effect known in signal processing when a data set is calibrated using only a subset of the data. This leads to an increase of variance on the excluded baselines and a decrease on those that are included (see appendix in Patil et al. 2016, for a mathematical description) and is related to a bias introduced in non-linear optimization (Laurent & Cook 1992; Cook et al. 1986).

An Incomplete or Inaccurate Sky Model — Even on the long baselines, where we are not limited by classical confusion, the sky model remains incomplete and imperfect. This is partly due to our inability to determine accurate source parameters for sources with an angular size equal to the PSF. Another important source of errors in source models is related to differential ionospheric corruptions across the source clusters used in SageCal. The spectrally complex model of the brightest source (at frequencies below 130 MHz) in the field, 3C61.1 (Fig. 1), still needs improvement using sub-arcsecond structural information from the European ~ 1000 km baselines now available. The chromatic residual side-lobe noise from all these imperfectly calibrated sources, will affect the frequency-dependent gain solution on a frequency scale that depends on the distance of the source from the phase centre (see e.g. Patil et al. 2016; Barry et al. 2016).

Signal-to-Noise — Using fewer and only longer baselines increases the thermal and ionospheric speckle noise (Vedantham & Koopmans 2016), and hence the resulting gain errors. We think this effect is still the smallest of the three although it can interact or be amplified by the first two effects, especially when the optimization problem is ill-conditioned. We note, however, that SageCal-CO includes regularisation to suppress the latter (see Yatawatta 2016, for a detailed analysis).

4.5. Regularisation of complex direction-dependent gains

The three effects described in Sect. 4.4 lead to additional spectral fluctuations on short baselines (see Patil et al. 2016).

To mitigate the amplification or propagation of small (non-instrumental) gain fluctuations, we penalize irregular gain solutions via a regularisation function (Yatawatta 2015b, 2016). We use a Bernstein polynomial of third order as prior on the *DD*-gain solutions (see e.g. Farouki 2012). *DD*-calibration over the full frequency domain, splitting the calibration and imaging baselines, using a detailed sky model and regularizing the gain solutions, are all currently combined in the single framework of SageCal-CO⁹ and runs efficiently on the parallel cluster Dawn, using MPI and CUDA.

4.6. Sky-model Subtraction and Gridding

Rather than correcting the *uv*-data (or images) for direction-dependent gain errors, we subtract the sky model from the visibility data in SageCal-CO using their full-Stokes gain solutions. We use the regularised gain solutions per sub-band/channel rather than the Bernstein polynomial itself, which is purely used as a prior function for the gains (in the case of very strong regularisation these two gain solutions as function of frequency would become identical). Subtraction of the sky model also removes their polarization leakage from Stokes I to Stokes Q, U and V (see e.g. Asad et al. 2015, 2016), as well as their beam and ionospheric effects, but only on spatial scales of the cluster diameters and their respective solution time-intervals, or larger. Subsequently the *uv*-data inside the $50\text{--}250\lambda$ annulus is gridded using $4.58\lambda \times 4.58\lambda$ *uv*-cells and 128 *w*-slices, using a prolate spheroidal wavefunction kernel (see Yatawatta 2010; Noorishad & Yatawatta 2011, for details).

5. IMAGE CUBES

We make use of a GPU-enabled imager called ExCon¹⁰, which can optimize the visibility weights to minimize the spectral dependency of the PSF (Yatawatta 2014). A spectrally-independent PSF improves the performance of GMCA. We also have used WSClean (Offringa et al. 2014), and its image deconvolution features, for general verification of our images.

5.1. Residual-Image Cubes

We produce $3^\circ \times 3^\circ$ image cubes with $0.5' \times 0.5'$ pixels using the $50\text{--}250\lambda$ baselines, for the frequency ranges 121.8 – 134.3 MHz, 134.3 – 146.8 MHz and 146.8 – 159.3 MHz, respectively (see Table 2). We do not apply a correction for the slowly varying station beam in the imager. These images have a PSF of ~ 10 arcmin FWHM. The spectral resolution of the cubes, in all four Stokes parameters, is 61 kHz. We use Stokes V as a measure of the data-quality and noise level. Note that *DD*-calibration only removes the discrete source components in each source cluster using the complex gain corrections derived for that direction. That is, the residual images for all cubes processed from this point onwards have only *DI*-calibration applied to them. Table 2 lists the most relevant imaging parameter settings.

In single-night integrations we have found evidence for very faint non-celestial signals in only a dozen subbands, concentrating near the NCP. Such signals could be caused by faint stationary RFI or low-level but stable cross-talk in the system. Any stationary (w.r.t. the array) RFI sources would coherently add at the NCP (i.e. their side-lobes rotate as the sky rotates and add coherently only on the NCP). The absence of such RFI signatures is a good sign of high data fidelity. Note that strong RFI was already flagged using AOFlagger; Offringa et al. (2012). L90490 is ionospherically well-behaved with diffractive scales of 21, 12, 18 km, respectively, in consecutive ~ 4 -hr time ranges (see e.g. Mevius et al. 2016, for more details). Fig. 2 shows a panel

⁹ <https://sourceforge.net/projects/sagecal/>

¹⁰ <https://sourceforge.net/projects/exconimager/>

of Stokes I and V images of the NCP with $\sim 3'$ and $\sim 10'$ FWHM resolution, after subtraction of the sky model. The Stokes-V images appear noise-like, whereas the Stokes-I images are classical confusion noise limited.

Diffuse Stokes Q & U Emission— In the power spectra analyses (Sect. 6.1) we *only* use images made from $50\text{--}250\lambda$ baselines as motivated in Sect. 3. These short-baseline images indeed retain their diffuse Q and U power. Polarization leakage is assumed to be small (see e.g. Asad et al. 2015, 2016). In a forthcoming publication we will present the polarized structure of the NCP and its impact on the detection of the EoR signal in much deeper integrations.

Diffuse Stokes I Emission— Diffuse Stokes-I emission is harder to detect when using $50\text{--}250\lambda$ baselines, because it appears below the classical confusion noise level set by discrete sources. Images including the $10\text{--}50\lambda$ baselines clearly show diffuse emission, when averaged to lower resolution. Hence, the diffuse (EoR) emission should be retained in the images after *DD*-calibration with SageCal-CO.

5.2. Generalized Morphological Component Analysis

The remaining foreground emission inside the primary beam area (Fig. 2) should only change very slowly with frequency and thus separable from the spectrally fluctuating 21-cm EoR signal (e.g. Morales & Hewitt 2004). We use GMCA (Bobin et al. 2007b,c,a, 2008, 2013), specifically tailored to foreground removal (Chapman et al. 2013), to remove the dominant modes from the data-cubes in Stokes I and any remaining instrumental polarization leakage in Stokes V.

GMCA is a blind source separation technique introduced by Zibulevsky & Pearlmutter (2001), which uses as few assumptions about the data as possible in order to form a model of the foregrounds. The method works on the premise that the diffuse foregrounds consist of a number of statistically independent components which can be separated using the morphology of those components. An appropriate decomposition basis is sought such that the components appear sparse and in this analysis we use a wavelet decomposition. A component can then be easily separated from the other components, the cosmological signal and instrumental noise due to the components having only few significant basis coefficients which are likely to be different between components. This results in a foreground model which can be subtracted from the total data, leaving the sub-dominant cosmological signal and instrumental noise. The only user input to the default method is the number of components in the foreground model. The optimal choice for this could be led by a Bayesian model selection; however, previous analyses have shown that the foreground model is fairly robust to this choice (see Chapman et al. 2013, for details) and as such we vary this number only over a limited range in this paper.

The implementation of GMCA is the same as described in Chapman et al. (2013). No astrophysical prior information is included in the calculation. While it is possible to include spectral information about the foregrounds within the mixing matrix, we choose to implement GMCA in the blindest way possible while the data is in the early stages of being

constrained. The mixing matrix does not vary across the sky or across the wavelet scales as in more recent implementations (Bobin et al. 2013). It is possible that the variation of the mixing matrix with wavelet scale may be implemented in a later data analysis as a method of mitigating the frequency-dependent PSF. Here we instead have chosen to set our data to a common resolution through *uv*-cuts in the imaging step and careful weighting. The solutions are regularised following Eqn.13 in Bobin et al. (2013), using N_s components. The $p = 0$ formalism is not trivial to calculate and the norm is relaxed to an L^1 -norm with $p = 1$, most often the standard in GMCA implementations.

We note that GMCA does not remove most of the remaining side-lobe noise. We remove $N_s=6\text{--}8$ components in Stokes I and $N_s=0\text{--}2$ components in Stokes V. The number of components are chosen to obtain an approximately flat noise behaviour in the k_{\parallel} direction (see Table 2 for the exact numbers per redshift range). Fig. 4.2 shows a spatial-frequency slice through the Stokes-I data-cube after subtraction of the sky model. There are still spectrally smooth sources left in the data. After applying GMCA, however, the Stokes-I data cube appears noise-like. Finally, we note that whereas GMCA does not a-priori distinguishes foregrounds from the 21-cm EoR signal, extensive simulations by Chapman et al. (2013) have shown that the 21-cm power-spectrum in the current range of k -modes should not be affected significantly by the diffuse and spectrally-smooth foreground removal.

6. POWER SPECTRA

In this section we present the cylindrically and spherically averaged 21-cm power spectra. Using the former one can assess remaining systematics due to e.g. foreground residuals, side-lobe noise and frequency-coherent effects (e.g. Bowman et al. 2009; Vedantham et al. 2012). The latter achieves the highest signal-to-noise per k -mode. Given the relatively narrow LOFAR-HBA primary beam ($4.8\text{--}3.5^\circ$ at $120\text{--}160$ MHz; van Haarlem et al. 2013) and our $3^\circ \times 3^\circ$ analysis window, we can ignore sky curvature. We use the Stokes-I residual data cube, after GMCA (see e.g. Fig. 4.2), to measure the power spectra following Tegmark (1997). We use large enough cells that they can be assumed to be uncorrelated.

6.1. Power-spectrum Determination

We first transform the data cube into brightness temperature, in units of mK (see Patil 2016, for details). A Gaussian primary beam correction is applied, which is a good approximation over the $3^\circ \times 3^\circ$ analysis window (van Haarlem et al. 2013), being smaller than the FWHM of the beam (see Figs. 1 & 2). We account for *uv*-density weighting and the number of zero-valued *uv*-cells in the padded *uv*-grid¹¹ that is used to create the image data cubes. Although determining the power spectrum directly from the (ungridded) visibilities is preferable, the size of the data set of 50 Tbyte (Table 1) renders

¹¹ Due to the usual Jy PSF^{-1} convention in radio astronomy, imagers scale *uv*-visibilities such that the zero-value visibility grid-cells are properly accounted for. The scaling, however, needs to be undone when determining the power spectrum.

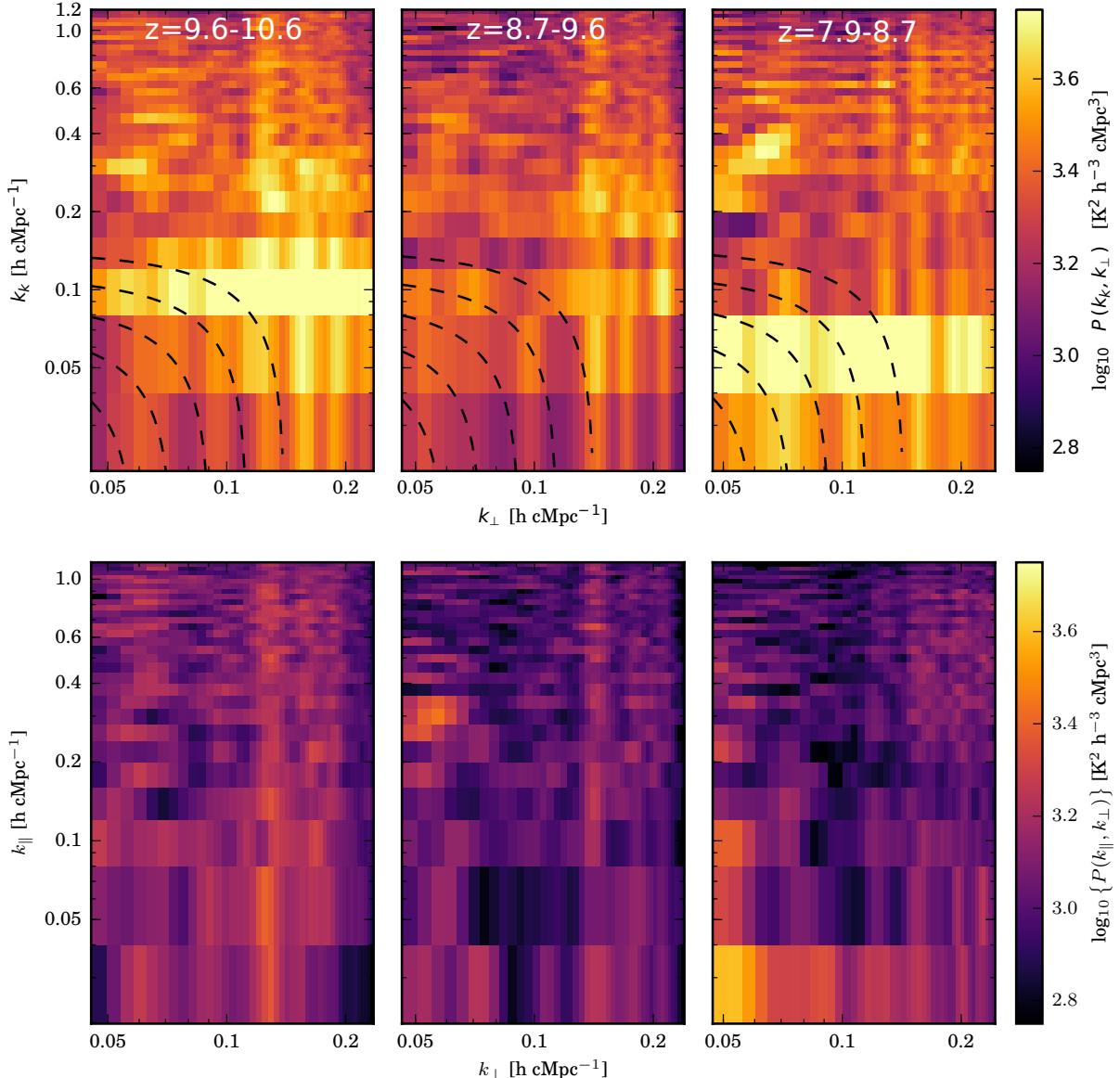


Figure 4. Stokes I (top) and V (bottom) cylindrical power spectra after sky model and GMCA-model subtraction, for L90490. From left to right are shown the redshift ranges $z = 9.6 - 10.6$, $z = 8.7 - 9.6$ and $z = 7.9 - 8.7$, respectively. The dashed curved lines in the Stokes I spectra refer to k values of 0.054, 0.067, 0.083, 0.103 and 0.128 for $z = 8.7 - 9.6$ and only slightly different values for the other redshift bins. It is along these lines that we form the spherically averaged power spectra.

this currently not feasible¹². A second reason why we do not use the visibilities is that GMCA is applied to the image cubes and not to the visibilities. The inference of the power spectrum follows Tegmark (1997); Trott et al. (2016) in part, but is adapted to the analysis of the image cube.

To determine the power spectrum, we spatially Fourier transform the cube back to the uv -domain, and use a Least Squares Spectral Analysis (LSSA) method to transform the frequency axis into a delay axis ($\nu \leftrightarrow \tau$) (see Barning 1963;

Lomb 1976; Stoica et al. 2009; Trott et al. 2016), properly accounting for the missing channels due to RFI excision (see Fig. 4.2 for the flagged channels).

We transform all axes into inverse co-moving Mpc (e.g. Morales & Hewitt 2004), using the cosmological convention of $k = 2\pi/L$. We determine power spectra $P(k)$ in units of $K^2 h^{-3} \text{cMpc}^3$ or $\Delta^2(k) = k^3/(2\pi^2)P(k)$ in units of K^2 . We also use mK units, where more conventional. Both the cylindrical and spherical power spectra are optimally weighted using the Stokes-V variance, down-weighting high noise-variance data (e.g. Tegmark 1997).

¹² Although the maximum information is retained in the un-gridded visibilities, gridding on scales substantially smaller than the inverse of the station beam ($\sim 16\lambda$) – in our case 4.58λ in the uv-domain (see Table 2) – should retain nearly all information.

6.2. Cylindrical power spectra

We present the power spectra for all redshift bins ($z = 7.9 - 8.7$, $8.7 - 9.6$ and $9.6 - 10.6$, respectively) in Fig. 5.1.0.0, for both Stokes I (left) and Stokes V (right). We note the following:

- There is some banded structure in k_{\perp} due to LOFAR-HBA uv -density variations modulating the noise variance in the Stokes-V power spectrum. No obvious structures in k_{\parallel} are seen (e.g. "wedge"; [Bowman et al. \(2009\)](#); [Vedantham et al. \(2012\)](#)). Before GMCA polarization leakage appears in Stokes V in the lowest k_{\parallel} bin, because of its broad-band nature. Because polarization leakage is also expected to be broad-band (see e.g. [Asad et al. 2015](#)), GMCA effectively removes it with at most two components (see [Chapman et al. 2013](#), for a description of GMCA components).
- The Stokes-I power spectrum appears similar to that of Stokes V after GMCA except for a residual horizontal band at $k_{\parallel} \approx 0.1 h \text{ cMpc}^{-1}$ in the $z = 9.6 - 10.6$ redshift bin and there is higher power in the $z = 7.9 - 8.7$ redshift bin around $k_{\parallel} \approx 0.05 h \text{ cMpc}^{-1}$. These are possibly caused by low-frequency structure remaining after the foreground removal with GMCA. There is at most only a mild indication in Fig. 5.1.0.0 for a wedge-like structure, suggesting that sky-model subtraction has been very effective, including the removal of side-lobes of out-of-beam sources.
- The ratios between the Stokes-I and Stokes-V power spectra for the three redshift bins is typically 2–3 in variance (see Fig. 7). Apart from the horizontal band at $k_{\parallel} \approx 0.1 h \text{ cMpc}^{-1}$, in the $z = 9.6 - 10.6$ and a similar band at $k_{\parallel} \approx 0.05 h \text{ cMpc}^{-1}$, at $z = 7.9 - 8.7$ these plots are devoid of significant features. The vertical bands have largely disappeared – in agreement with the cause of the modulation arising as a result of variations in the uv -density. It also suggests that the excess variance does not add coherently (see also Fig. 10 in [Patil et al. 2016](#)), otherwise it would not average down with the number of visibilities in the same way as thermal noise which dominates Stokes V. No evidence for signals related to cable-reflections, at their known delays (or k_{\parallel} values), is seen.

We assume that the excess variance is not the 21-cm EoR signal. It might be a mixture of side-lobe noise due to an incomplete and inaccurate sky model (Sect. 4) – causing calibration gain errors (e.g. [Barry et al. 2016](#)) –, or effects of thermal and ionospheric noise, and leverage (e.g. [Patil et al. 2016](#)). We note that the excess noise decreases as the gain solutions are regularised in the frequency direction (see Sect. 3). Because we split our baselines between calibration and imaging, and only subtract sources, but do not correct the residual visibilities after *DI*-calibration, any suppression or enhancement of Stokes-I power must have its cause in the applications of the gains to the sky model. Hence they have to come from issues relevant for the longer baselines and the most likely effects are either an incomplete/inaccurate sky model

or strong ionospheric variations. However, we have not seen evidence yet for correlations between the diffractive scale of the ionosphere and excess noise in other data sets (see Fig. 10 in [Patil et al. 2016](#)).

To illustrate the considerable impact of *DD*-calibration, we show the cylindrical power spectra for $z = 9.6 - 10.6$ before and after *DD*-calibration and sky-model subtraction in Fig. 6.2, and their ratio in Fig. 5, but before removal of the diffuse emission and residual sources in the primary beam with GMCA (Sect. 5.2).

6.3. Spherical Power Spectra

Next we determine the spherically-averaged power spectrum, optimally weighting using the Stokes-V variance, following [Tegmark \(1997\)](#); [Trott et al. \(2016\)](#), to obtain the average per k -bin. We flag two k_{\parallel} bins with show strong excess variance after running GMCA (see Fig. 7). In the $z = 9.6 - 10.6$ redshift range this corresponds to the (logarithmic) bin around $k_{\parallel} \sim 0.05 h \text{ cMpc}^{-1}$. In the $z = 7.9 - 8.7$ redshift range this corresponds to the bin around $k_{\parallel} \sim 0.1 h \text{ cMpc}^{-1}$. The integration is done along the curved lines shown in Fig. 5.1.0.0. We emphasize that we assume that the Stokes-V power-spectrum to be our best estimator of the thermal-noise power spectrum, because (i) the Stokes-V sky is by any means empty, and (ii) the thermal noise in Stokes V and I should be identical. Hence $\Delta_I^2 - \Delta_V^2$ is the noise-bias corrected residual Stokes-I power spectrum. This should in principle be consistent with the 21-cm EoR power spectrum if there were no excess variance nor other biases. Given the 13 hr integration, however, this should still be considered an upper limit on the 21-cm EoR signal. We therefore conservatively put our upper limits at 2σ on top of the excess variance and do not attempt to estimate the excess variance level itself or correct for it at present (since we have no independent estimator for it).

The resulting Stokes I, V and difference power spectra are shown in Fig. 8, up to $k = 0.2 h \text{ cMpc}^{-1}$. The errors on the power spectra are determined from the Stokes-V variance and the number of uv cells used in the integration. The errors are therefore plotted on the noise-bias-corrected powers. We note the following:

- The redshift ranges $9.6 - 10.6$ and $8.7 - 9.6$ appear power-law like¹³ in the spherically averaged power spectrum (Fig. 8). Apart from two stripes, they also have mostly featureless ratios of Stokes-I over Stokes-V power (Fig. 7),
- Whereas at all k -values the Stokes I variance exceeds the Stokes V variance, given that the EoR signal very likely is still lower than the thermal noise, we have to assume that this excess variance is due to other causes. We interpret it as a robust upper limit on the 21-cm emission power spectrum Δ_{21}^2 .

¹³ We note that such behaviour is only an approximation that would hold if $P(k)$ is roughly constant.

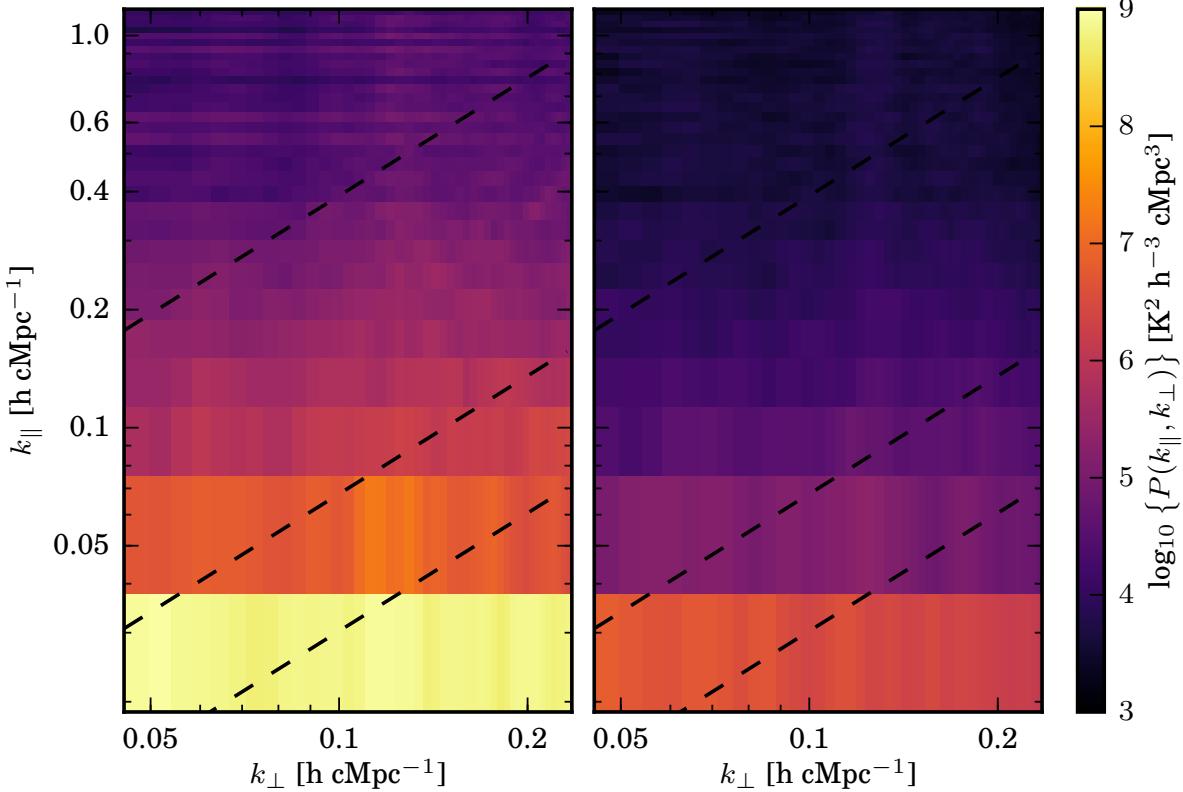


Figure 5. The Stokes-I power spectra for the redshift range $z = 9.6 - 10.6$, before (top left) and after (top right) *DD*-calibration with *SageCal*-CO, respectively. Note the large drop in power of the foregrounds at low k_{\parallel} and the removal of substantial power above the wedge as well. The dashed slanted lines indicate, from bottom to top, the location of angular distances of 4.5° and 10° from the phase centre, and the maximum delay corresponding to the horizon as seen from the zenith. The ratio between these power spectra is shown in Fig. 5

k $h \text{ cMpc}^{-1}$	$z = 7.9 - 8.7$ mK^2	$z = 8.7 - 9.6$ mK^2	$z = 9.6 - 10.6$ mK^2
0.053	$(131.5)^2$	$(86.4)^2$	$(79.6)^2$
0.067	$(242.1)^2$	$(144.2)^2$	$(108.8)^2$
0.083	$(220.9)^2$	$(184.7)^2$	$(148.6)^2$
0.103	$(337.4)^2$	$(296.1)^2$	$(224.0)^2$
0.128	$(407.7)^2$	$(342.0)^2$	$(366.1)^2$

Table 3. Δ_{21}^2 upper limits at the 2σ level.

- Up to $k_{\perp} \approx 0.2 h \text{ cMpc}^{-1}$ both the Stokes I and Stokes V power spectra follow approximate power-laws, with the power in Stokes I exceeding that in Stokes V for all k -modes and all redshift bins. At the smallest $k = 0.053 h \text{ cMpc}^{-1}$, however, these values start to approach each other with only marginal differences. This is the bin that we regard as the best upper limit in terms of mK² sensitivity yielding a 2σ upper limit of $\Delta_{21}^2 < (79.6 \text{ mK})^2$ on the 21-cm power spectrum in the range $z = 9.6 - 10.6$.

In Table 3 we summarize the 2σ upper limits for the three redshift bins for Δ_{21}^2 .

7. SUMMARY AND FUTURE OUTLOOK

We have presented the first upper limits on the 21-cm power spectrum (Δ_{21}^2) from the Epoch of Reionization, obtained with LOFAR-HBA, using one night of good data quality obtained toward the NCP. Our main numerical results can be summarised as follows:

- An excess variance is detected in Stokes I for all k modes and redshift ranges, leading to our best although still non-zero $\Delta_1^2 = (56 \pm 13)^2 \text{ mK}^2$ (1σ) at $k = 0.053 h \text{ cMpc}^{-1}$ in the redshift range $9.6 - 10.6$. The excess variance is seen over the entire cylindrical power spectrum range. It appears constant with no obvious outstanding features such as cable reflections.
- The most stringent 2σ upper limit of $\Delta_{21}^2 < (79.6 \text{ mK})^2$ on the 21-cm power spectrum is found at $k = 0.053 h \text{ cMpc}^{-1}$ in the range $z = 9.6 - 10.6$. For reference, in the absence of excess variance we would have reached a 2σ upper limit $\Delta_{21}^2 < (57 \text{ mK})^2$ for the same k and z ranges.
- In Table 3 we summarize the 2σ upper limits for the three redshift bins for a range of k -modes.

Currently the cause of the excess variance is still unknown. Based on simulations (see e.g. Patil et al. 2016) and data-processing tests, in particular with improved sky models and

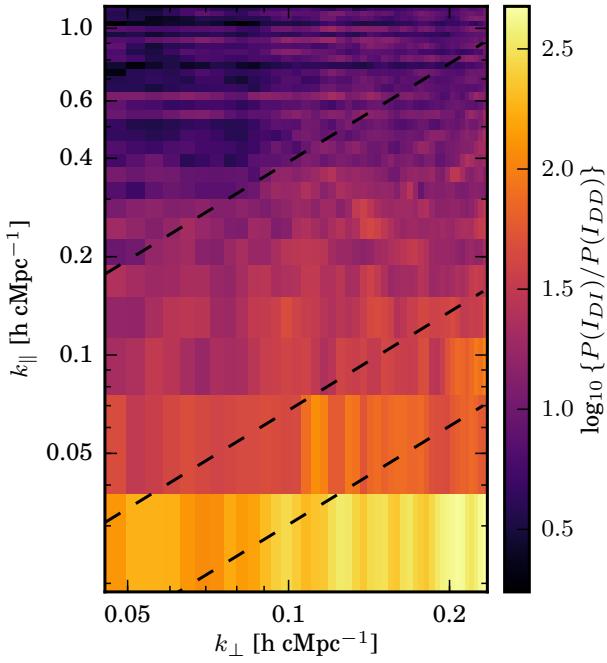


Figure 6. The ratio between the Stokes-I power before and after DD-calibration. There is a drop of two orders of magnitude in power in the foregrounds at low k_{\parallel} . The dashed slanted lines indicate, from bottom to top, the location of angular distances of 4.5° and 10° from the phase center, and the maximum delay corresponding to the horizon as seen from the zenith.

regularised gain solutions (Yatawatta 2016), it is likely due to residual side-lobe noise seen on the calibration baselines (due to an incomplete/inaccurate sky model), which affects the gain solutions on shorter baselines, as well as leverage (Patil et al. 2016). Various test are under way to find the cause, or causes.

7.1. Comparison of results

Comparing our deepest 2σ upper limit of $\Delta_{21}^2 < (79.6 \text{ mK})^2$ at $k = 0.053 \text{ h cMpc}^{-1}$ and $z = 9.6 - 10.6$, to those published by the other three teams (see Sect. 1) using the GMRT (see Paciga et al. 2013), the MWA (see Beardsley et al. 2016) and PAPER (see Ali et al. 2015), remains difficult. The reasons are the different redshift ranges and k -modes that are being quoted, as well as the considerably different integration times, being 13 h for LOFAR, 32 h for MWA, 40 h for the GMRT, and 1150 h for PAPER, respectively, as well as the use of very different instrumental configurations and post-correlation processing methods.

Currently, LOFAR-HBA reaches the highest redshift range of these experiments, with its deepest upper limits at $z = 9.6 - 10.6$ and only mildly less deep at $z = 8.7 - 9.6$ (Table 3). It also reaches considerably larger co-moving scales (i.e. smaller k -modes) compared to all other experiment, largely thanks to a strong emphasis on removal of compact sources and diffuse foreground emission from the data, allowing us to probe into the wedge region.

7.2. Lessons Learned

We have learned that a number of requirements are important in the analysis of the LOFAR-HBA EoR data (see Sections 3 and 4). We expect this to hold for other arrays as well (see e.g. Mellema et al. 2013, for earlier discussions about the SKA). Not meeting some of these requirements appears detrimental to our calibration and image quality (Sect. 4 and 3), and the resulting power spectra:

Direction-dependent calibration:— We use 122 directions, clustering sources typically in (few) degree-scale patches (see Sect. 4). This scale roughly matches that expected based on the beam forming and isoplanatic angles, but are ultimately limited in size by the signal-to-noise per baseline and the number of degrees of freedom.

Completeness and accuracy of the sky model for calibration:— We use $\sim 20,800$ source components spread over about 19° in radius from the NCP (and beyond) down to flux-density levels of ~ 3 mJy (inside/outside primary beam), below the classical confusion noise on short baselines (Sect. 3.3). Our model does not yet include diffuse emission, especially the ubiquitous diffuse polarized emission.

Diffuse-emission conservation on the short baselines:— We currently use two non-overlapping baseline sets split at 250λ (Sect. 4). Long baselines are used for calibration and short baselines for the power spectrum analyses. The fundamental reason is that DD-calibration suppresses diffuses emission in Stokes Q and U, and likely also the 21-cm EoR signal in Stokes I (Sect. 4.3).

Wide-frequency domain for calibration:— To reduce the effects of excess noise or excess variance, due to leverage, side-lobe noise, ionospheric and thermal noise, etc., highly irregular gain solutions need to be penalised if not warranted by the data (Sect. 4). We have implemented this via regularisation of the gain solutions, using third-order Bernstein polynomials.

As noted in Sect. 4, DD-calibration is necessary, but removes diffuse emission on short baselines, which is not part of the calibration model due to computational limits. Hence splitting the baselines in two sets (short and long) is necessary because diffuse emission is not measured on the longer (calibration) baselines (to our levels of sensitivity). This however leads to excess noise, which is partly mitigates by using a larger frequency domain for the gain solutions.

7.3. Future Outlook

Although the excess noise has not yet been fully eliminated, gain regularisation over a large frequency domain, as implemented in SageCal-CO (Yatawatta 2016), has considerably reduced its magnitude in recent analyses. To reduce the excess variance further, by a factor 2–3, i.e. to the level approaching Stokes-V power on all k -modes, we plan to:

- Improve the calibration sky model by including even fainter compact sources inside and outside the primary beam. With the current 20,800 component model, we still notice improvements when new sources are added.

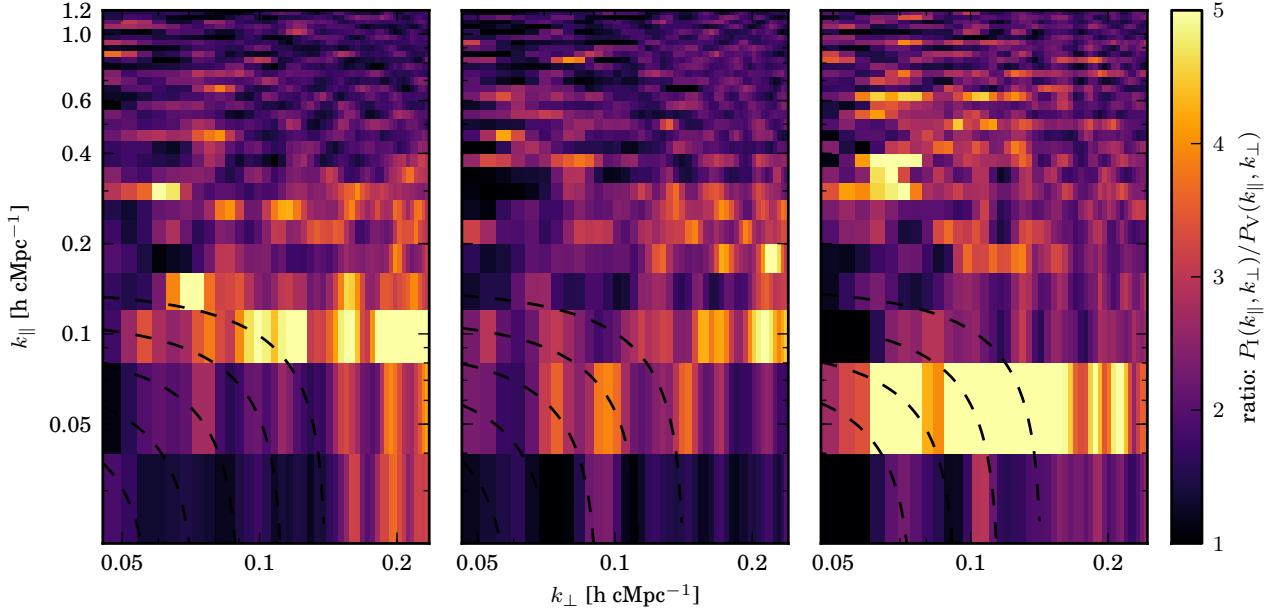


Figure 7. The Stokes I over V power spectra ratios for the redshift ranges $z = 9.6 - 10.6$, $z = 8.7 - 9.6$ and $z = 7.9 - 8.7$, respectively.

- Include diffuse Stokes Q, U and (if possible) diffuse Stokes I emission in the sky model and, if possible, avoid the split-baseline approach. This should reduce the excess variance as tests have shown, due to the elimination of leverage, while not suppressing diffuse emission.
- Improve GMCA foreground subtraction, or replace it by a spectrally-smooth diffuse foreground model and subtract it in the uv -plane on short baselines.
- Use the cross-variance between different observing epochs and assess whether the excess variance is (in)coherent. This approach avoids the need for a careful noise-power estimate and its bias correction in the Stokes-I power spectrum.
- Cross-correlate the gain solutions with data-quality metrics (e.g. diffractive scale) and sky- and calibration-model metrics to gain better insight into the nature of the excess variance.
- Include the flagged interferometers between co-located stations sharing the same electronics cabinet – with baselines in the range of $40-60\lambda$ – in the analysis. Although these baselines are the most sensitive to the 21-cm signal, they were conservatively flagged to avoid any correlated spurious signals. We have started a program to statistically analyse the signals on those baselines to quantify any non-celestial contributions and include as many of them as possible.
- Analyse the full set of data, in steps, and combine their results. If the excess variance is incoherent, and if all nights turn out to be of similar quality, we should be able to reduce the upper limits inverse proportional

with integration time (in power spectrum variance). From an earlier analysis of several nights we have indications that the excess noise is indeed only weakly correlated between nights (see Patil et al. 2016).

The results presented in this paper show that the LOFAR residual images and power spectra are still affected by low-level effects (e.g. excess variance). However, we have identified viable mitigation strategies to reduce its level. Given that the results in this paper are (i) based on only $\sim 2\%$ of the entire NCP data set in hand and (ii) still conservatively excludes some of the most sensitive short baselines, we are confident that we can reach considerably deeper limits in the near future.

LOFAR, the Low Frequency Array designed and constructed by ASTRON, has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the International LOFAR Telescope (ILT) foundation under a joint scientific policy. SZ and AP would like to thank the Netherlands Organisation for Scientific Research (NWO) VICI grant for financial support. LVEK, HV, AG, SD, BKG thank the European Research Council Starting Grant (639.043.308) for support. AGdB, ARO, SBY, VNP, MM, HV and MH, acknowledge support from the ERC (grant 339743, LOFARCORE). MM, VNP and SBY also acknowledge support from the NWO TOP grant (614.001.005). VJ would like to thank the Netherlands Foundation for Scientific Research (NWO) for financial support through VENI grant 639.041.336. ITI was supported by the Science and Technology Facilities Council [grant number ST/I000976/1]. SM would like to acknowledge the financial assistance from the European Research Council under ERC grant number 638743-FIRSTDAWN and from the European Unions Seventh Framework Programme

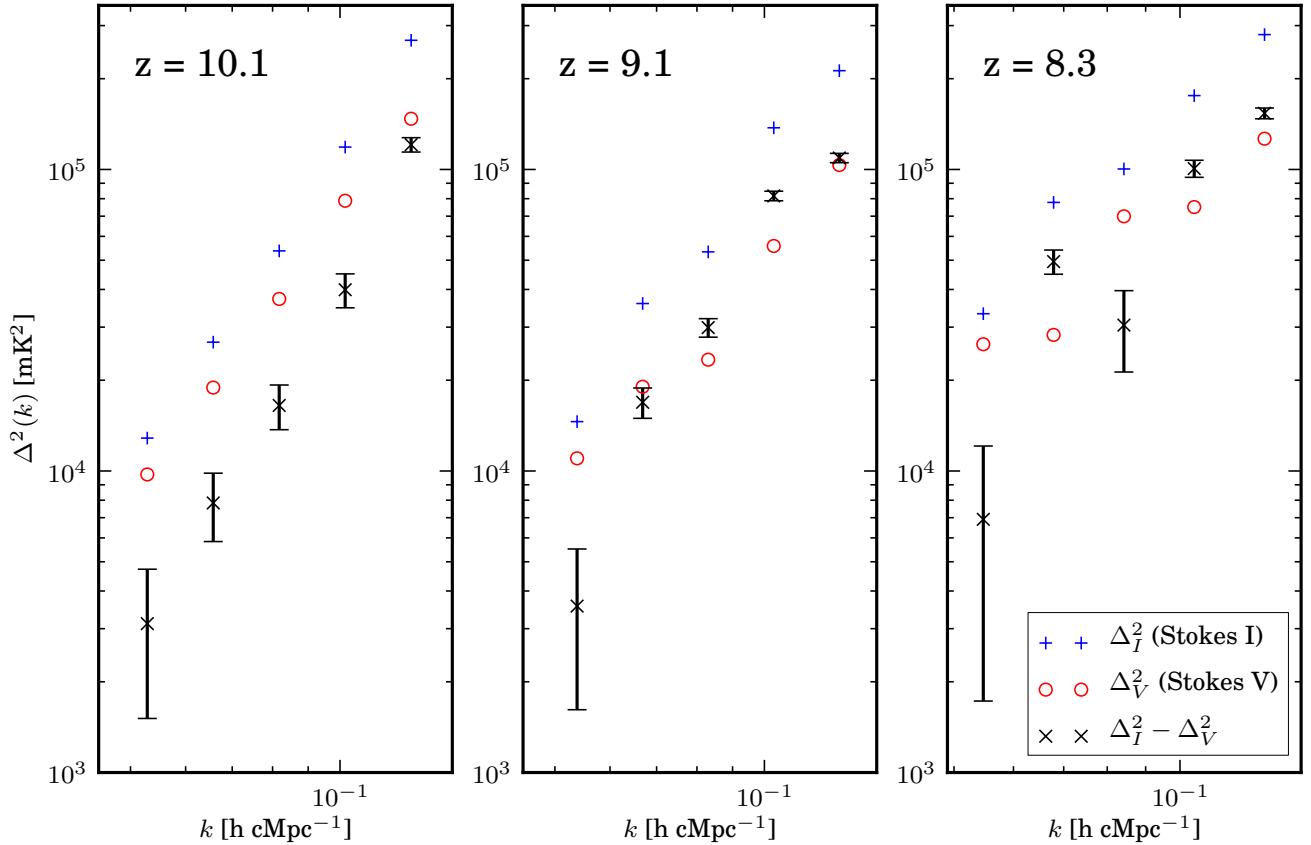


Figure 8. The spherically averaged Stokes I and V power spectra after GMCA for L90490; From left to right are shown the redshift ranges $z = 9.6 - 10.6$, $z = 8.7 - 9.6$ and $z = 7.9 - 8.7$ from left to right, respectively. The mean redshifts are indicated in the panels.

FP7-PEOPLE-2012-CIG grant number 321933-21ALPHA.

REFERENCES

- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, ApJ, 809, 61
 Asad, K. M. B., Koopmans, L. V. E., Jelić, V., et al. 2015, MNRAS, 451, 3709
 Asad, K. M. B., Koopmans, L. V. E., , et al. 2016, arXiv.org, arXiv:1604.04534
 Barkana, R., & Loeb, A. 2005, ApJL, 624, L65
 Barnard, F. J. M. 1963, BAN, 17, 22
 Barry, N., Hazelton, B., Sullivan, I., Morales, M. F., & Pober, J. C. 2016, MNRAS, 461, 3135
 Beardsley, A. P., Hazelton, B. J., Morales, M. F., et al. 2013, MNRAS: Letters, 429, L5
 Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, arXiv.org, 1608.06281v1
 Becker, G. D., & Bolton, J. S. 2013, MNRAS, 436, 1023
 Becker, G. D., Bolton, J. S., Haehnelt, M. G., & Sargent, W. L. W. 2011, MNRAS, 410, 1096
 Becker, R., Fan, X., White, R., et al. 2001, AJ, 122, 2850
 Bernardi, G., de Bruyn, A. G., Harker, G., et al. 2010, A&A, 522, 67
 Bharadwaj, S., & Ali, S. S. 2005, MNRAS, 356, 1519
 Bobin, J., Fadili, J., Moudden, Y., & Starck, J.-L. 2007a, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6701, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
 Bobin, J., Moudden, Y., Starck, J.-L., Fadili, J., & Aghanim, N. 2008, Statistical Methodology, 5, 307
 Bobin, J., Starck, J., Sureau, F., & Basak, S. 2013, Astronomy and Astrophysics, 550, A73
 Bobin, J., Starck, J.-L., Fadili, J., & Moudden, Y. 2007b, IEEE Transactions on Image Processing, 16, 2662
 Bobin, J., Starck, J.-L., Fadili, J. M., Moudden, Y., & Donoho, D. L. 2007c, IEEE Transactions on Image Processing, 16, 2675
 Bobin, J., Starck, J.-L., Sureau, F., & Basak, S. 2013, A&A, 550, A73
 Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., & Sargent, W. L. W. 2010, MNRAS, 771
 Bolton, J. S., & Haehnelt, M. G. 2007, MNRAS, 382, 325
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, ApJL, 709, L133
 ——. 2015, ApJ, 803, 34
 Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, ApJ, 638, 20
 Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183
 Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, e031
 Bunker, A. J., Wilkins, S., Ellis, R. S., et al. 2010, MNRAS, 409, 855
 Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, MNRAS, 412, 2543
 Chapman, E., Zaroubi, S., Abdalla, F. B., et al. 2016, MNRAS, 458, 2928
 Chapman, E., Abdalla, F. B., Bobin, J., et al. 2013, MNRAS, 429, 165
 Chapman, E., Abdalla, F. B., Bobin, J., et al. 2013, MNRAS, 429, 165
 Collaboration, P., Ade, P. A. R., Aghanim, N., et al. 2015, arXiv.org, arXiv:1502.01589
 Cook, R. D., Tsai, C. L., & Wei, B. C. 1986, Biometrika, 73, 615
 Fan, et al., X. 2003, AJ, 125, 1649

- . 2006, AJ, 131, 1203
- Farouki, R. T. 2012, Computer Aided Geometric Design, 29, 379
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhR, 433, 181
- Furlanetto, S. R., Peng Oh, S., & Briggs, F. H. 2006, Physics Reports, 433, 181
- Geil, P. M., Gaensler, B. M., & Wyithe, J. S. B. 2011, MNRAS, 1416
- Geil, P. M., Wyithe, J. S. B., Petrovic, N., & Oh, S. P. 2008, MNRAS, 390, 1496
- Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137
- Harker, G., Zaroubi, S., Bernardi, G., et al. 2009, MNRAS, 397, 1138
- . 2010, MNRAS, 405, 2492
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
- Jelić, V., Zaroubi, S., Labropoulos, P., et al. 2008, MNRAS, 389, 1319
- Kazemi, S., Yatawatta, S., & Zaroubi, S. 2013, MNRAS, 430, 1457
- Kazemi, S., Yatawatta, S., Zaroubi, S., et al. 2011, MNRAS, 414, 1656
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
- Laurent, R. T. S., & Cook, R. D. 1992, J. Am. Stat. Assoc., 87, 985
- Lomb, N. R. 1976, Ap&SS, 39, 447
- Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
- Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
- McQuinn, M. 2015, arXiv.org, 1512.00086v1
- McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, ApJ, 653, 815
- Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2013, Experimental Astronomy, 36, 235
- Mevius, M., Tol, S., Pandey, V. N., et al. 2016, Radio Science, 51, 927
- Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7
- Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7
- Morales, M. F., & Wyithe, J. S. B. 2010, Annual Review of A&A, 48, 127
- Natarajan, A., & Yoshida, N. 2014, Progress of Theoretical and Experimental Physics, 2014, 06B112
- Noorishad, P., & Yatawatta, S. 2011, in 2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) (IEEE), 326–330
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, ApJL, 709, L16
- Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, MNRAS, 444, 606
- Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. 2012, A&A, 539, A95
- Offringa, A. R., de Bruyn, A. G., Zaroubi, S., et al. 2013, A&A, 549, A11
- Paciga, G., Chang, T.-C., Gupta, Y., et al. 2011, MNRAS, 413, 1174
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639
- Page, L., Hinshaw, G., Komatsu, E., et al. 2007, ApJS, 170, 335
- Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J. 2012, ApJ, 753, 81
- Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, The Astronomical Journal, 139, 1468
- Patil, A. 2016, PhD thesis
- Patil, A. H., Zaroubi, S., Chapman, E., et al. 2014, MNRAS, 443, 1113
- Patil, A. H., Yatawatta, S., Zaroubi, S., et al. 2016, MNRAS, 463, 4317
- Planck Collaboration. 2016, arXiv:1605.03507, arXiv:1605.03507
- Pritchard, J. R., & Furlanetto, S. R. 2007, MNRAS, 376, 1680
- Pritchard, J. R., & Loeb, A. 2008, PhRvD, 78, 103511
- Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
- Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19
- Santos, S., Sobral, D., & Matthee, J. 2016, arXiv.org, 1606.07435v1
- Scaife, A. M. M., & Heald, G. H. 2012, MNRAS: Letters, 423, L30
- Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2014, ApJ, 795, 20
- Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, A&A, 345, 380
- Smirnov, O. 2011, A&A, 527, A106
- Stoica, P., Li, J., & He, H. 2009, IEEE Transactions on Signal Processing, 57, 843
- Tegmark, M. 1997, Physical Review D, 55, 5895
- Theuns, T., Schaye, J., Zaroubi, S., et al. 2002, ApJL, 567, L103
- Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, Publications of the Astronomical Society of Australia, 30
- Trott, C. M., Pindor, B., Procopio, P., et al. 2016, ApJ, 818, 139
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2
- van Haarlem, M. P., Wise, M. W., Gunst, A., et al. 2013, A&A, 556, A2
- Vedantham, H., Udaya-Shankar, N., & Subrahmanyam, R. 2012, ApJ, 745, 176
- Vedantham, H. K., & Koopmans, L. V. E. 2016, Monthly Notices of the Royal Astronomical Society, 458, 3099
- Watkinson, C. A., & Pritchard, J. R. 2014, MNRAS, 443, 3090
- Yatawatta, S. 2010, ArXiv e-prints, arXiv:1008.1892
- Yatawatta, S. 2014, in 2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS) (IEEE), 1–4
- Yatawatta, S. 2015a, MNRAS, 449, 4506
- . 2015b, MNRAS, 449, 4506
- . 2016, arXiv.org, 1605.09219
- Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, A&A, 550, 136
- Zaroubi, S. 2013, in Astrophysics and Space Science Library, Vol. 396, Astrophysics and Space Science Library, ed. T. Wiklind, B. Mobasher, & V. Bromm, 45
- Zheng, Q., Wu, X.-P., Johnston-Hollitt, M., Gu, J.-H., & Xu, H. 2016, eprint arXiv:1602.06624, 1602.06624
- Zibulevsky, M., & Pearlmutter, B. A. 2001, Neural Computation, 13, 863