Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **500,** 2264–2277 (2021)
Advance Access publication 2020 November 12

doi:10.1093/mnras/staa3446

Comparing foreground removal techniques for recovery of the LOFAR-EoR 21 cm power spectrum

Ian Hothi,[1★] Emma Chapman,[1] Jonathan R. Pritchard,[1] F. G. Mertens,[2,3] L. V. E. Koopmans,[2] B. Ciardi,[4] B. K. Gehlot,[5] R. Ghara,[6,7] A. Ghosh,[8] S. K. Giri,[9] I. T. Iliev,[10] V. Jelić[11] and S. Zaroubi[2,6,7]

[1] *Department of Physics, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK*
[2] *Kapteyn Astronomical Institute, University of Groningen, P. O. Box 800, NL-9700 AV Groningen, the Netherlands*
[3] *LERMA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Université, F-75014 Paris, France*
[4] *Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany*
[5] *School of Earth and Space Exploration, Arizona State University, 781 Terrace Mall, Tempe, AZ 85287, USA*
[6] *Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Ra'anana 4353701, Israel*
[7] *Department of Physics, Technion, Haifa 32000, Israel*
[8] *Department of Physics, Banwarilal Bhalotia College, GT Rd, Ushagram, Asansol, West Bengal 713303, India*
[9] *Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*
[10] *Astronomy Centre, Department of Physics and Astronomy, Pevensey II Building, University of Sussex, Brighton BN1 9QH, UK*
[11] *Ruđer Bošković Institute, Bijenička cesta 54, 10 000 Zagreb, Croatia*

Accepted 2020 November 2. Received 2020 November 2; in original form 2020 October 9

ABSTRACT

We compare various foreground removal techniques that are being utilized to remove bright foregrounds in various experiments aiming to detect the redshifted 21 cm signal of neutral hydrogen from the epoch of reionization. In this work, we test the performance of removal techniques (FastICA, GMCA, and GPR) on 10 nights of LOFAR data and investigate the possibility of recovering the latest upper limit on the 21 cm signal. Interestingly, we find that GMCA and FastICA reproduce the most recent 2σ upper limit of $\Delta^2_{21} < (73)^2\ \mathrm{mK}^2$ at $k = 0.075\ h\mathrm{cMpc}^{-1}$, which resulted from the application of GPR. We also find that FastICA and GMCA begin to deviate from the noise-limit at k-scales larger than $\sim 0.1\ h\mathrm{cMpc}^{-1}$. We then replicate the data via simulations to see the source of FastICA and GMCA's limitations, by testing them against various instrumental effects. We find that no single instrumental effect, such as primary beam effects or mode-mixing, can explain the poorer recovery by FastICA and GMCA at larger k-scales. We then test scale-independence of FastICA and GMCA, and find that lower k-scales can be modelled by a smaller number of independent components. For larger scales ($k \gtrsim 0.1\ h\mathrm{cMpc}^{-1}$), more independent components are needed to fit the foregrounds. We conclude that, the current usage of GPR by the LOFAR collaboration is the appropriate removal technique. It is both robust and less prone to overfitting, with future improvements to GPR's fitting optimization to yield deeper limits.

Key words: methods: data analysis – methods: statistical – techniques: interferometric – cosmology: observations – cosmology: theory – dark ages, reionization, first stars.

1 INTRODUCTION

The early Universe was predominantly neutral with sources of radiation yet to form and was aptly named the 'Dark Ages' (Pritchard & Loeb 2012). The first sources of ionizing radiation formed 80 Myr after the big bang, from the inhomogeneities seeded by inflation – marking the Cosmic Dawn (Bowman et al. 2018). These sources ionize their neutral surroundings producing ionized bubbles. As these bubbles grow, and more sources form, they begin to merge, slowly ionizing the neutral intergalactic medium (IGM) – the onset of the epoch of reionization (EoR) (Ciardi & Ferrara 2005; Morales & Wyithe 2010; Furlanetto 2016). The EoR was the last major phase change in the Universe's history.

Indirect observations of the EoR have provided constraints. Quasars strongly emit Ly α photons and, as they travel towards us, the photons are redshifted. Any photons that are bluewards of the Ly α photon, at a particular redshift, will be redshifted such that it corresponds to the Ly α photons. If neutral hydrogen is present at this redshift, the Ly α photon will be absorbed. This leads to the spectra of quasars having a large region of purely absorption – the Gunn–Peterson trough (Zaroubi 2013). Looking at the prominence of this feature, at progressively lower redshifts the trough starts to become incomplete and there is no longer a sharp rise that marks the end of the trough. This might suggest that the IGM is largely ionized and nearing full ionization by $z \sim 6$, leaving only a few neutral islands (Fan et al. 2006). In the spectra of high-redshift quasars, their Ly α damping absorption have been used to constrain the EoR, as having not been completed by a redshift of $z \sim 7$ (Greig et al. 2017; Greig, Mesinger & Bañados 2019).

★ E-mail: iandeepshothi@gmail.com

© 2020 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Reionization leads to an increase of free electrons in the IGM, and CMB photons will scatter off of these free electrons. The scattering of these photons by the free electrons allows us to calculate an associated optical depth to the CMB, τ (the total integrated optical depth to reionization), which indicates the amount of scattering that has occurred along the line of sight. This scattering leads to a scale-independent power suppression at scales $l \geq 10$ by a factor of $e^{2\tau}$. Planck Collaboration VI (2018) found $\tau = 0.054 \pm 0.007$. Using this, one can model the evolution of the neutral fraction and infer the bulk of reionization must have occurred at $z \lesssim 14$ (Planck Collaboration XIII 2016). Combining these indirect observations, we can loosely constrain reionization to occur between redshifts 14 to 6 (Gorce et al. 2018).

The foremost probe of the EoR is the 21 cm hyperfine forbidden line transition from neutral hydrogen (Wouthuysen 1952; Field 1958; Shaver et al. 1999). The 21 cm signal is a powerful probe, as reionization can imprint features in the 21 cm power spectrum telling us the topology of reionization which itself is sensitive to the properties of ionizing sources (Furlanetto, Oh & Briggs 2006; Furlanetto 2016).

Though it has thus far eluded detection, the EoR is likely to be directly observed for the first time by radio telescopes currently taking data, e.g. Low Frequency Array (LOFAR)[1] (van Haarlem et al. 2013); Giant Metrewave Radio Telescope (GMRT)[2] (Paciga et al. 2011, 2013); Murchison Widefield Array (MWA)[3] (Tingay et al. 2013); Precision Array to Probe the Epoch of Reionization (PAPER)[4] (Parsons et al. 2014); the Hydrogen Epoch of Reionization Array (HERA)[5] (DeBoer et al. 2017), and at lower frequencies, the new extension in Nançay upgrading LOFAR (NeNuFar)[6] (Zarka et al. 2012); the Owens Valley Long Wavelength Array (OVRO-LWA),[7] and Amsterdam-ASTRON Radio Transients Facility and Analysis Center (AARTFAAC)[8] (Prasad et al. 2016; Gehlot 2019). These are interferometric experiments that aim to measure the statistical fluctuations of the 21 cm signal.

There also exist single antenna experiments that aim to measure the sky-averaged global 21 cm signal. These include, the Global HydrOgen ReioNization Signal (BIGHORNS)[9] (Sokolowski et al. 2015); the Sonda Cosmológica de las Islas para la Detección deHidrógeno Neutro (SCI-HI) (Voytek et al. 2014); the Experiment to Detect the Global Epoch of Reionization Signature (EDGES)[10] (Bowman et al. 2018); Shaped Antenna measurement of the background RAdio Spectrum 2 (SARAS 2)[11] (Singh et al. 2018); Large-aperture Experiment to Detect the Dark Age (LEDA) [12] (Price et al. 2018), and Probing Radio Intensity at High-Z (PRIZM) (Philip et al. 2019). Bowman et al. (2018) has a tentative detection of this global signal, with the absorption depth and width, currently, unexplained. Though there have been claims that the results could be explained by exotic physics (Fialkov, Barkana & Cohen 2018), potential unaccounted for systematics (Singh & Subrahmanyan 2019), foreground polarizations (Spinelli, Bernardi & Santos 2019), or modelling (Hills et al. 2018).

The biggest obstacle preventing the detection of the EoR by radio interferometers are the foregrounds that swamp the observed signal. These foregrounds are several orders of magnitude stronger than the 21 cm signal we wish to observe (Shaver et al. 1999; Oh & Mack 2003; Di Matteo, Ciardi & Miniati 2004; Jelić et al. 2008). These foregrounds are modelled to be smooth (in the frequency domain) and this was exploited by the early foreground removal techniques, e.g. polynomial fitting (Jelić et al. 2008). This smooth nature of the foregrounds spectra is likely to deviate when met with instrumental effects, and so non-parametric techniques, such as blind source separation (BSS), are often favoured as they do not explicitly state the form of foregrounds a-priori (Chapman et al. 2012, 2013). More recently Bayesian techniques, such as Gaussian process regression (GPR) (Mertens, Ghosh & Koopmans 2018), have been looked at as an optimized fitting tool with the priors set by the various components of the observed data, such as the foregrounds and cosmological signal. Several machine learning methods have shown promising signs for the role of machine learning in further study of the EoR (Hassan et al. 2019; La Plante & Ntampaka 2019; Li et al. 2019a; Mangena, Hassan & Santos 2020).

The aforementioned experiments have already started to provide upper limits on the 21 cm power spectrum. GMRT (Paciga et al. 2013) provided an early 2σ upper limit, with 40 h worth of data, at a redshift of 8.6 with $\Delta^2 \simeq (248\,\mathrm{mK})^2$ at $k = 0.50\,h\mathrm{cMpc}^{-1}$. Here they used singular value decomposition (SVD) on each baseline, individually, with visibilities represented in a matrix by time and frequency. SVD isolates the smooth foreground modes for each baseline, removing these modes – the residual from which the power spectrum is calculated on.

MWA (Barry et al. 2019), using 21 h worth of data in Beardsley et al. (2016)'s MWA observing run to produce a limit of $\Delta^2 \simeq (62.4\,\mathrm{mK})^2$ at $k = 0.20\,h\mathrm{cMpc}^{-1}$, at a redshift of 7. Li et al. (2019b), produced a 2σ upper limit of $\Delta^2 \simeq (49\,\mathrm{mK})^2$ at $k = 0.59\,h\mathrm{cMpc}^{-1}$, at a redshift of 6.5. More recently, Trott et al. (2020) used 110 h of data to produce an upper limit of $\Delta^2 \simeq (43\,\mathrm{mK})^2$ at $k = 0.14\,h\mathrm{cMpc}^{-1}$ at a redshift of 6.5. Here they used a model for the point sources remove within their primary beam, before weighting their power spectrum estimation, to reduce the impact of foregrounds.

PAPER (Kolopanis et al. 2019), Fourier transformed each baseline spectrum into the delay domain and then filtered out foreground modes in this domain (Cheng et al. 2018). This produced the upper limit, between $0.3 < k < 0.6\,h\mathrm{cMpc}^{-1}$, of $\Delta^2 \simeq (200\,\mathrm{mK})^2$ at redshift 8.37.

LOFAR's first set of upper limits, from their high-band antenna (HBA), came in Patil et al. (2017). They produced a 2σ upper limit of $\Delta^2 \simeq (79.6\,\mathrm{mK})^2$ at $k = 0.053\,h\mathrm{cMpc}^{-1}$ in the redshift range 9.6 to 10.6. Here they produced an upper limit using Generalized Morphological Component Analysis (GMCA) (Chapman et al. 2013), as its primary removal technique. More recently, LOFAR have used GPR for upper limits, (Mertens et al. 2018; Gehlot et al. 2019), the application of which resulted in the recent 2σ upper limit of $\Delta^2 < (73\,\mathrm{mK})^2$ at $k = 0.075\,h\mathrm{cMpc}^{-1}$ at a redshift of 9.1, using 141 h worth of LOFAR data (Mertens et al. 2020). This recent upper limit has already been used to provide constraints on the intergalactic medium, excess radio background, and the neutral fraction at a redshift of 9.1 (Ghara et al. 2020; Greig et al. 2020; Mondal et al. 2020).

The goal of this paper is to assess the foreground removal techniques' pros and cons, by applying them to the 10 nights of LOFAR data as well as simulated data – with both having an injected simulated 21 cm signal. The paper is written as follows. In Section 2, we discuss the foreground removal techniques used in this paper, before moving on to Section 3, which describes the data. We then perform a full analysis of the power spectra results of the removal techniques in Section 4, followed by an investigation into the results. We summarize our results and conclude the paper with a discussion in Section 5.

---

[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] https://nenufar.obs-nancay.fr/en/homepage-en/

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

[8] http://aartfaac.org

[9] http://www.mwatelescope.org/telescope/external/bighorns

[10] https://loco.lab.asu.edu/edges/

[11] http://www.rri.res.in/DISTORTION/saras.html

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

2 FOREGROUND REMOVAL TECHNIQUES

In this section, we will provide an overview of the foreground removal techniques used in this paper, FastICA, GMCA, and GPR. We will detail how they model the foregrounds and recover the 21 cm signal.

We had also attempted to apply another non-parametric removal technique, Wp smoothing, to our data (Harker et al. 2009). Whilst it worked on the simulated data without any noise or instrumentation effects, it was unable to fit a function to the observed data. We believe that the chromaticity within the data, along with the small amount of frequencies study, result in the smoothing algorithms being unable to fit a function.

2.1 FastICA

FastICA is a non-parametric technique that aims to model the foregrounds by assuming they are statistically independent in the mixed signal we observe (Hyvärinen, Karhunen & Oja 2001). FastICA has been used in the medical field for functional magnetic resonance imaging (fMRI) analysis (Chen et al. 2003; Ge et al. 2016), as well as decomposing spinal cord signals (Tie & Sahin 2005). Early uses of FastICA for foreground removal, include its application to CMB data. Both for recovering the spatial pattern and the frequency scalings of foregrounds, as well as removing foreground to detect the CMB (Maino et al. 2002, 2003). FastICA has also been used for foreground removal in H I mapping (Wolz et al. 2015). The use of this non-parametric foreground reduction method on 21 cm data, was outlined by Chapman et al. (2012). The data are modelled as

$$\boldsymbol{x} = \mathbf{A}\boldsymbol{s} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{x}$ is the vector representation of the observed signal, $\boldsymbol{s}$ is a vector containing independent components, and $\mathbf{A}$ is the mixing matrix we wish to find. The algorithm does not attempt to fit the additive Gaussian noise term, $\boldsymbol{n}$. The number of expected independent components comprising the foregrounds, is given to the algorithm, and FastICA attempts to reconstruct this number. To solve for $\boldsymbol{s}$, we consider the following linear transform:

$$\boldsymbol{s} = \mathbf{W}\boldsymbol{x}, \tag{2}$$

with $\mathbf{W}$ being a constant weight matrix for FastICA to determine. It finds $\mathbf{W}$ by maximizing the non-Gaussianity of equation (2), quantifying non-Gaussianity by negentropy – this uses the principle that the greater the number of independent components in the set, the more Gaussian the distribution of the set will be (Central Limit Theorem).

Let us consider the example for finding a single component of $\boldsymbol{s}$. We define,

$$y = \boldsymbol{w}^T\boldsymbol{x}. \tag{3}$$

Here, $\boldsymbol{w}$ is a row from the inverse-matrix of $\mathbf{A}$. As such, y represents a single component (row) of $\boldsymbol{s}$. We first define the entropy of a variable as,

$$H(y) = -\sum_i P(y = a_i) \log \mathrm{P(y = a_i)}, \tag{4}$$

where a_i is a possible value of y. Following equation (4), we define negentropy as,

$$J(y) = H(y_{\mathrm{gauss}}) - H(y). \tag{5}$$

With y_{gauss} being a random Gaussian variable with an identical covariance matrix as y. From the central limit theorem, by maximizing the negentropy of y, and hence the non-Gaussianity, we retrieve a single independent component of $\boldsymbol{s}$.

By adjusting the weighting matrix, using negentropy as a measure of equation (2)'s non-Gaussianity, one can find the independent variables corresponding to the foregrounds.

Once $\mathbf{W}$ is found, we can find $\mathbf{A}$ as the inverse of $\mathbf{W}$ ($\mathbf{A} = \mathbf{W}^{-1}$), and $\mathbf{s}$ can be found via equation (2). Once FastICA has calculated $\boldsymbol{s}$, one can use the mixing matrix to reconstruct the foreground signal, $\mathbf{A}\boldsymbol{s}$, and deduct from $\boldsymbol{x}$. The residual of this will be a mixture of the 21 cm signal, noise, and any fitting errors.

2.2 GMCA

GMCA separates out the foregrounds from the signal (noise + EoR signal) by finding the sparsest components in an orthogonal wavelet basis that the foregrounds can be represented by (Chapman et al. 2013). It has previously been used by CMB experiments to remove Galactic foregrounds and point sources from their maps (Bobin et al. 2013b).

More recently, Carucci, Irfan & Bobin (2020) have applied GMCA to 21 cm maps and have also modelled, previously unaccounted for 'foreground', polarization leakage. GMCA also models the data with equation (1). To solve this problem, we need to estimate $\mathbf{A}$ and $\boldsymbol{s}$, where the latter represents the foreground signals which we can decompose (in the wavelet basis) as: $\boldsymbol{s} = \sum_{j=0}^{T}\boldsymbol{s}_j$. Where, $\boldsymbol{s}_j = \sum_{k=1}^{T}\alpha_j[k]\,\boldsymbol{\phi}_k$, such that $\boldsymbol{\phi}_k$ is a wavelet waveform. As we are looking for the sparsest s_j, we define it to be sparse when a few $\alpha_j[k]$ are significantly non-zero. The use of waveforms over the traditional Fourier basis functions is due to the latter being a localized basis, whereas waveforms are an infinite set in localized space. The objective of GMCA is to find a mixing matrix that yields, in the wavelet domain, the sparsest sources of $\boldsymbol{s}$. This Lagrangian optimization problem is expressed as

$$\min\left(\frac{1}{2}||\boldsymbol{x} - \mathbf{A}\alpha\Phi||_{\mathrm{F}}^2 + \lambda\sum_{\mathrm{j=0}}^{\mathrm{n}}||\alpha_{\mathrm{j}}||_{\mathrm{p}}\right). \tag{6}$$

Once we have extracted both $\mathbf{A}$ and $\boldsymbol{s}$, we can perform matrix multiplication on the two to find out foreground signal. Deducting $\mathbf{A}\boldsymbol{s}$ from $\boldsymbol{x}$ leaves us with the residual – which ideally will be the noise, EoR signal, and potentially fitting errors.

When comparing parametric to non-parametric techniques, one will often find the former outperforming the latter, on simulated data without any realistic instrumentation effects. However, this is not expected to be true for real data, as non-parametric techniques are able to model/account for the imperfections in real data as nothing is assumed a-priori about the data.

2.3 GPR

Gaussian process regression (GPR) on LOFAR was first used for foreground removal by Mertens et al. (2018),[13] and has been used to produce the upper limits within Gehlot et al. (2019) and

---

[13] https://gitlab.com/flomertens/ps_eor

Mertens et al. (2020). Within the GPR framework, one models the different components of observations, such as the astrophysical effects and instrumentation effects, and the cosmological signal, as a Gaussian process. Instrument effects can include those from the chromatic nature of the instrument, resulting in mode mixing (Hazelton, Morales & Sullivan 2013). A Gaussian process is the joint distribution of a collection of normally distributed random variables, and it is described by its mean and covariance function. This covariance function is that between pairs of points at different frequencies, and hence smoothness, in frequency, of the function. The class of covariance functions used in this paper is the Matern class:

$$\kappa_{\mathrm{Matern}}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu r}}{l} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu r}}{l} \right). \tag{7}$$

K_ν is the modified Bessel function of the second kind, and l is the hyperparameter of the kernel – the characteristic coherence-scale. Using different values of ν, we retrieve special cases of this matrix. For example, setting ν to $\frac{1}{2}$ we get the exponential kernel.

We describe the Gaussian processes as parametrized priors in GPR, and this prior is selected such that it maximizes the Bayesian evidence with the hyperparameters of the covariance function being found via Markov chain Monte Carlo (MCMC) algorithms.

GPR models the observed data $\boldsymbol{d}$ at frequency ν by a foreground, a 21 cm and a noise signal $\boldsymbol{n}$:

$$\boldsymbol{d} = \boldsymbol{f}_{\mathrm{fg}} + \boldsymbol{f}_{21} + \boldsymbol{n}. \tag{8}$$

As the 21 cm signal is expected to be uncorrelated on scales of the order of a few MHz, we can exploit this to separate from the foreground signal, which is expected to be correlated (smooth) on scales of 1 MHz. We can thus separate out the covariance function of the Gaussian process into a covariance for the foregrounds, $\mathbf{K_{fg}}$, and a covariance for the 21 cm signal, $\mathbf{K_{21}}$:

$$\mathbf{K} = \mathbf{K}_{\mathrm{fg}} + \mathbf{K}_{21}. \tag{9}$$

The foreground covariance can be decomposed into two terms: The first is the intrinsic foreground $\mathbf{K}_{\mathrm{int}}$, which is the expected smooth/coherent foreground (over large frequency scales). The second component is for the instrumental effects, such as mode-mixing, and is coherent on a smaller frequency range. As the 21-cm signal is faint compared to the foregrounds and the noise, the inclusion of the 21 cm signals covariance, $\mathbf{K}_{21}$, with appropriate priors (uniform or a more informative gamma distribution prior) in the total covariance model $\mathbf{K}$ ensures that we don't fit away the 21-cm signal part when we optimize the maximum likelihood estimation. Mertens et al. (2018) found the estimated coherence scale was significantly less biased, especially for a reference simulation with low signal to noise for 21-cm signal.

The joint probability distribution for the observed data $\boldsymbol{d}$ and function values $\boldsymbol{f}_{fg}$ of the foreground model at a given frequency ν is then given by,

$$\left[\frac{\boldsymbol{d}}{\boldsymbol{f}_{\mathrm{fg}}} \right] \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{\mathrm{fg}} + \mathbf{K}_{21} + \sigma_n^2 I & \mathbf{K}_{\mathrm{fg}} \\ \mathbf{K}_{\mathrm{fg}} & \mathbf{K}_{\mathrm{fg}} \end{bmatrix} \right), \tag{10}$$

where σ_{n}^2 is the noise variance and I is the identity matrix. Once GPR has been performed, the foreground model is retrieved via

$$E(\boldsymbol{f}_{\mathrm{fg}}) = \mathbf{K}_{\mathrm{fg}} \left[\mathbf{K}_{\mathrm{fg}} + \mathbf{K}_{21} + \mathbf{K}_{\mathrm{n}} \right]^{-1} \boldsymbol{d}, \tag{11}$$

$$\mathrm{cov}(\boldsymbol{f}_{\mathrm{fg}}) = \mathbf{K}_{\mathrm{fg}} + \mathbf{K}_{\mathrm{fg}} [\mathbf{K}_{\mathrm{fg}} + \mathbf{K}_{21} + \mathbf{K}_{\mathrm{n}}]^{-1} \mathbf{K}_{\mathrm{fg}}, \tag{12}$$

where the noise covariance, $\mathbf{K}_{\mathrm{n}}$, is the diagonal of σ_{n}^2: $\sigma_{\mathrm{n}}^2 \mathbf{I}$. Here E($\boldsymbol{f}_{fg}$) is the expectation values and cov($\boldsymbol{f}_{fg}$) is the covariance of the foregrounds. The inclusion of the $\mathbf{K}_{21}$ kernel in the definitions allows one to marginalize over all foreground models obtained by the Gaussian process – accounting for any degeneracy between the foregrounds and 21 cm signal, and providing a proper error budget. Deducting E($\boldsymbol{f}_{fg}$) from the original data gives the residual.

3 DATA

In this section we will describe the data used in this paper. The first is the 10 nights of data from LOFAR that were used for the upper limit result in Mertens et al. (2020). The second data set is simulated data that will be used to test the performance of the foreground removal techniques, and asses their pros and cons.

3.1 LOFAR 10 nights data

The LOw-Frequency ARray (LOFAR) is an interferometric array of radio antennas with stations spread across several countries in Europe (38 stations within the Netherlands and 14 international stations). LOFAR is capable of observing the frequency range 10–240 MHz. It is also capable of reaching an angular resolution of arcseconds across most of the 30–240 MHz range. Stations are made up of a mixture of low band antennas – which operate from 10 MHz (just above the ionosphere cut off) to 90 MHz (just below commercial FM) – as well as high band antennas which operate between 110 to 240 MHz (van Haarlem et al. 2013). Most of the LOFAR stations are located in the Netherlands, with the core stations being located near the town of Exloo; chosen because of its low population and (relatively) low radio frequency interference (RFI). The core has 24 dual-stations located in a 2 km radius, with six of these stations being found on the Superterp, a small island, 320 m in diameter. The stations in the Superterp make up the shortest baselines in LOFAR's array. The antennas in the core station are distributed to give optimal *uv* coverage. There are 14 stations, outside of the core, that are spaced out logarithmically, with the outer most station being at a radius of 90 km. These 14 stations are denoted as the 'remote stations'[14] (van Haarlem et al. 2013).

The observation data used here is from the LOFAR EoR project, concentrated on the North Celestial Pole (NCP) window (Bernardi et al. 2009; Yatawatta et al. 2013). The data are an accumulation of ∼140 h of observing time, coming from 10 nights worth of observations (observing nights ID: L80847, L86762, L90490, L196421, L205861, L246297, L246309, L253987, L254116, L254865) (Mertens et al. 2020). The data have undergone both direction-independent and direction-dependent calibration as well as RFI flagging (Yatawatta et al. 2008; Offringa et al. 2010; Kazemi et al. 2011). A more rigorous description of the data processing pipeline for this data set can be found in Patil et al. (2017) and Mertens et al. (2020). This data set has a total bandwidth $\Delta\nu = 11.52$ MHz, with frequency range $134.57\ \mathrm{MHz} \leq \nu \leq 146.09$ MHz. The frequency resolution of the data is 195.3 kHz, however several subbands were flagged resulting in the current 37 subband data set.

3.2 Simulated data

3.2.1 21 cm EoR signal

The 21 cm signal is simulated using `21CMFAST` (Mesinger, Furlanetto & Cen 2011). This is a semi-analytical tool which approximates

---

[14] http://old.astron.nl/radio-observatory/astronomers/users/technical-information/lofar-array-configuration/lofar-array-conf

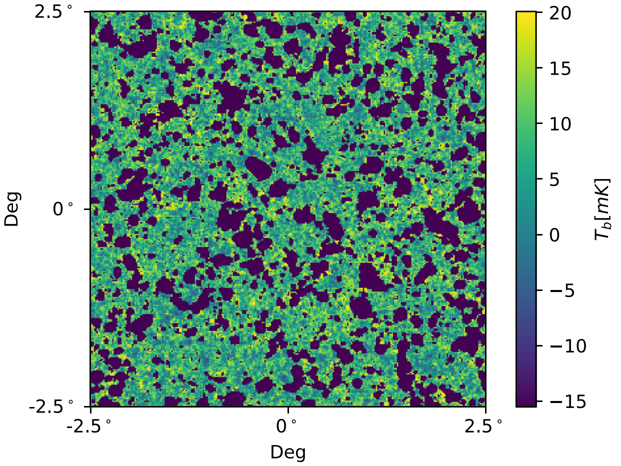


Figure 1. A frequency slice at 134 MHz ($z \sim 9.6$) of the cosmological signal realization produced by `21CMFAST`, with a neutral fraction $x_{HI} = 0.78$.

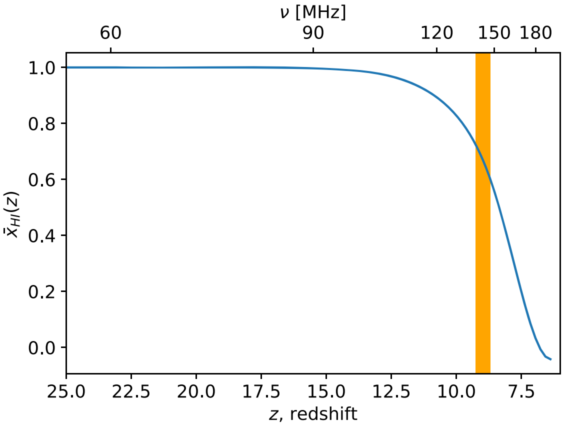


Figure 2. The neutral fraction evolution, with redshift, of the cosmological simulation produced by `21CMFAST`. The shaded region highlights the main frequency region of interest: 134.5 MHz $\leq \nu \leq$ 146 MHz.

physical processes, rather than simulating via hydrodynamic simulations, making it computationally much less expensive. The results, the evolving brightness temperature T_{b}, produced by `21CMFAST` agree with recent hydrodynamical simulations (Hutter 2018). The simulation used `21CMFAST`'s default Planck cosmology: (σ_8, h, Ω_m, Ω_Λ,n) $= (0.81,0.68,0.31,0.69,0.97)$. The simulation was initialized at $z = 25$, with a box size of 1000 Mpc, on a 2048^3 grid with subsequent evolution of the density and ionization being performed on a lower resolution 512^3 grid – with this grid being the final T_{b} box size. We have assumed saturation of the spin temperature T_s: $T_s \ggg T_{\mathrm{CMB}}$, which is valid during the redshifts of the observation; we have set the UV photons per stellar baryon to be 0.3. The output of `21CMFAST` is the δT_{b} field in the redshift range $6 \leq z \leq 25$, with these redshift slices being made into a light-cone computed with an observing angle of 5° – to match that of LOFAR. Fig. 1 shows an example slice and the evolution of neutral hydrogen for this realization is shown in Fig. 2.

With the known bandwidth of the LOFAR 10 nights data cube, 11.52 MHz, and subband width, 195.3 kHz, we simulate the cosmological signal via `21CMFAST` between 134.57 and 146.09 MHz, and then inject this cosmological signal into the data.

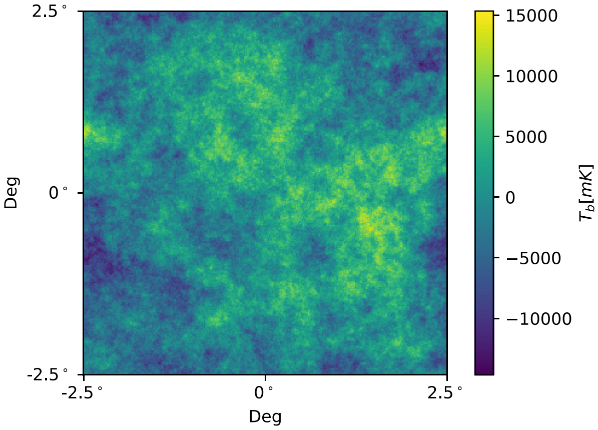


Figure 3. A frequency slice at 134 MHz of the foreground simulation produced by the lightcone outlined in Jelić et al. (2008).

3.2.2 Foregrounds

The diffuse astrophysical foregrounds were simulated using the methods described in Jelić et al. (2008). This treats the foreground emissions, within our FoV at a given frequency, as Gaussian random fields (GRF), modelling the foregrounds' emission as a power law: $T_{\mathrm{b}} \propto \nu^{-\beta}$. The foreground contributions are the following:

(i) **Galactic diffuse synchrotron emission (GDSE)**: Galactic synchrotron emission is the dominant foreground component that radio surveys encounter. This emission arises from two sources: The diffuse component arises from the interaction between relativistic free electrons from the interstellar medium (ISM) and the Galactic magnetic field. The second is due to emission from supernovae remnants (SNRs). Together these make up ~ 70 per cent of the observed foreground emission. The brightness temperature spectral index used to model this was $\beta = 2.55$ – this value was used as the mean value for a GRF with a deviation of 0.1.

(ii) **Galactic diffuse free–free emission**: This foreground is better known as bremsstrahlung emission and it only accounts for 1 per cent of the contaminating foregrounds (both Galactic and extra). Bremsstrahlung emission is due to electrons scattering off of other ions. The scattering off ions causes them to decelerate and emit radiation. At the high latitudes observed by radio arrays to detect the 21 cm signal, this emission is from diffused ionized gas which is optically thin. When simulated, it has a fixed spectral index of 2.15.

(iii) **Extragalactic foregrounds**: Extragalactic foregrounds account for 27 per cent of the foregrounds that contaminate our signal. These fall into two categories: (1) Radio galaxies, observed as point sources in the sky and the source of their observed radio signal is synchrotron emission. (2) Radio clusters, the largest virialized systems in the Universe with the emission from clusters coming from their galaxies and from the intercluster medium (ICM). Radio sources of galaxy clusters have two classifications: radio haloes and radio relics. Radio haloes have a regular morphology with a low surface brightness, typically centred on and permeating the cluster volume. Radio relics are similar to radio haloes, having low surface brightness and extending over large distances (> 1 Mpc). Jelić et al. (2008) clusters extragalactic foregrounds using a random walk algorithm.
We can see an example slice of the foregrounds in Fig. 3.

The simulated data have a field of view (FoV) of 5°×5°, with a frequency resolution of 500 kHz and pixel resolution of 35 arcsec.

Henceforth, we will refer to the data, a cube with bandwidth $\Delta\nu$ with a field of view of deg × deg, interchangeably, as a data cube or cube.

3.2.3 Simulated observations

The next step from simulating a realization of reionization is to apply observational effects on to the simulated light cone. We do this using the OSKAR visibility simulation package.[15] OSKAR produces simulated visibility data from radio telescopes containing aperture arrays. It does this by taking an input file that contains antenna placement and (permitted) beam positions. Taking these, and a sky model provided, it produces dual-polarization signals for each antenna provided before using these to beam form and hence simulate an observation. These observational effects are applied to the data provided, in our case the simulated data, to produce a measurement set. The measurement set[16] output contains visibility data, produced by OSKAR simulating an observation of our provided data. With the measurement set in hand, we next move on to producing images from these visibilities, and using WSCLEAN (Offringa et al. 2014). We use OSKAR to simulate a 120 h observation with a data cube with both the simulated foregrounds and cosmological signal. To produce an image, WSCLEAN takes the provided visibilities and grids them on to different w-layers. WSCLEAN attempts to solve the complex visibility function,

$$V(u,v,w) = \int\int \frac{A(l,m)I(l,m)}{\sqrt{1-l^2-m^2}} \times \exp\left\{2\pi i\left[ul+vm+w\left(\sqrt{1-l^2-m^2}-1\right)\right]\right\} \mathrm{d}l\,\mathrm{d}m, \qquad (13)$$

where u, v, w are the baseline coordinates in the coordinate system of the array, $A(l, m)$ is the primary-beam function, $I(l, m)$ is the sky function, and l,m are cosine sky coordinates. WSCLEAN solves this, to find $I(l, m)$ by discretizing w such that its integral becomes a summation. Once WSCLEAN solves this to produce an image, it outputs it as a Flexible Image Transport System (FITS) file. We have used natural weighting, within WSCLEAN. The visibilities produced by OSKAR are in Janskys, with WSCLEAN converting the data to Jansky per beam. We then convert to Kelvin, using the beam size for each frequency slice of the data cube. As we have used natural weighting, one must account for the PSF; see Section 3.2.1 in Mertens et al. (2020). We also note that the field of view of this image cube has been reduced to 4° × 4°, such that it is inside the primary beam.

4 RESULTS

In this section, we show the results of applying the foreground removal techniques to the aforementioned data. We attempt to replicate the recent upper limit in Mertens et al. (2020), and then look at the application of the techniques to simulated data that replicates the 10 nights of LOFAR data.

The power spectrum, on a wavenumber k, is defined as,

$$P(k) = \mathrm{V_c}|\tilde{\mathrm{T}}(k)|^2, \qquad (14)$$

where $\tilde{T}(k)$ is the Fourier transformed brightness temperature and $\mathrm{V_c}$ is the observed comoving cosmological volume, with the primary beam; spatial tapering function, and frequency tapering function delimited. We then average equation (14) in spherical shells and define the spherically averaged dimensionless power spectrum as,

$$\Delta^2(k) = \frac{k^3}{2\pi^2}P(k). \qquad (15)$$

When referring to the power spectrum, we are referring to equation (15).

We can use the Stokes I and Stokes V from the data set's observation to estimate noise. The noise estimates are made by differencing two frequency slices in the image cubes. As these slices are separated by a small frequency interval, 195.3 kHz, both the foregrounds and 21 cm should be near identical between the two – leaving only the noise. The LOFAR point spread function changes by 0.1 per cent over 0.2 MHz (Patil et al. 2016). Thus, this difference should be dominated by thermal noise, more notably after foreground removal. Once we have removed our foreground model from the data, we are left with the residual Δ^2_I. We can now subtract the spherically averaged noise power spectrum Δ^2_N, also known as noise bias removal, from the residual:

$$\Delta^2_{21} = \Delta^2_I - \Delta^2_N. \qquad (16)$$

With the associated error,

$$\Delta^2_{21,\mathrm{err}} = \sqrt{\left(\Delta^2_{I,\mathrm{err}}\right)^2 + \left(\Delta^2_{N,\mathrm{err}}\right)^2}. \qquad (17)$$

Unless stated otherwise, all power spectra have the noise bias removed.

4.1 LOFAR 10 nights

We start with the 10 nights of data (141 h) from the LOFAR EoR experiment, whose upper limit from this data can be seen in Mertens et al. (2020). In our analysis we compare FastICA, GMCA, and GPR . For FastICA and GMCA, we modelled the foregrounds with a number of different components ranging from 1 to 37; the maximum number of components is set by the dimensionality of the data set – the number of frequency channels. Here, we present only the result from the best performing number of components, which we find to be 6. We see signal suppression at component numbers higher than this. This is in line with the component number used in Patil et al. (2017)'s upper limit. For GPR, to model kernels of the data, we apply the reasoning within Mertens et al. (2020): The astrophysical foregrounds, dubbed intrinsic foregrounds, are expected to be smooth, with respect to frequency. Thus their frequency coherence-scale will be long, and in this case we have given GPR a uniform prior for the coherency range, this is between 10 and 100 MHz, from which to optimize, with a Matern kernel with $\eta_{\mathrm{mix}} = 5/2$. The Bayesian optimization finds the optimized coherency length for the intrinsic foregrounds to be 46.23 MHz. The second 'foreground', comes from instrumentation effects. We have given the frequency coherence-scale, used for a Matern kernel with $\eta_{\mathrm{mix}} = 3/2$, a range 1–10 MHz. The Bayesian optimization finds the optimized coherency length for this instrumental foreground to be 2.64 MHz. For the simulated 21 cm signal, we use an exponential kernel, with a coherency length range between 0.1 and 1.2 MHz. The Bayesian optimization finds the optimized coherency length for the 21 cm signal to be 0.73 MHz.

We can see the results of the foreground removal techniques application on the 10 nights of LOFAR data in Fig. 4, where we have removed the noise bias from the results of the removal techniques. Tables 1, 2, and 3 show the power spectrum results at a given k-bin, as well as the 2σ upper limit ($\Delta^2_{21,2\sigma}$), for GMCA, FastICA, and GPR. We see for the smaller k bins, $k <$ hcMpc^{-1}, the three removal methods perform similarly. In fact, we recover the recent upper limit presented in Mertens et al. (2020), with our use of GPR replicating the upper limit. The BSS techniques of FastICA and GMCA reproduce the upper limit at the same k-scale, $k = 0.075$ hcMpc^{-1}, with a

---

[15] https://github.com/OxfordSKA/OSKAR

[16] More information on measurement sets can be found: https://casa.nrao.edu/Memos/229.html.

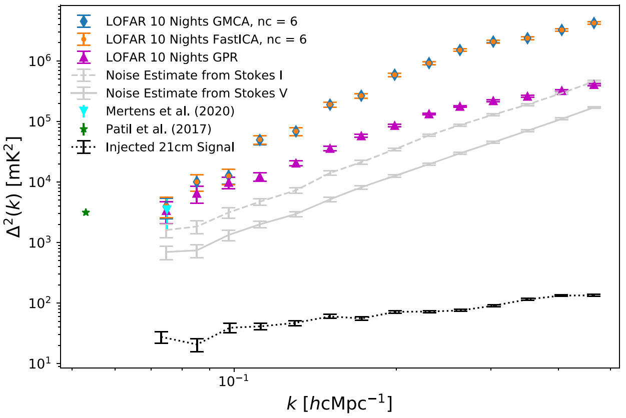


Figure 4. Comparison of the noise bias removed residual power spectra produced by the removal techniques: FastICA, GMCA, and GPR. We have also included the power spectrum of the injected cosmological signal and the noise estimates produced by Stokes I and Stokes V. One can see the comparison to recent upper limits produced by LOFAR HBA: Patil et al. (2017) and Mertens et al. (2020). Within figures involving the BSS techniques, we shall refer to the number of components used to model the foregrounds as 'nc'.

Table 1. Showing the Δ^2_{21} upper limit at 2σ ($\Delta^2_{21,2\sigma}$), and error ($\Delta^2_{21,\mathrm{err}}$), at a given k-bin for GMCA.

k (hcMpc^{-1})	Δ^2_{21} (mK^2)	$\Delta^2_{21,\mathrm{err}}$ (mK^2)	$\Delta^2_{21,2\sigma}$ (mK^2)
0.075	$(58.30)^2$	$(31.01)^2$	$(72.95)^2$
0.085	$(100.22)^2$	$(54.93)^2$	$(126.80)^2$
0.098	$(112.64)^2$	$(59.88)^2$	$(140.92)^2$
0.112	$(222.53)^2$	$(89.71)^2$	$(256.16)^2$
0.130	$(262.13)^2$	$(101.36)^2$	$(298.77)^2$
0.151	$(435.36)^2$	$(137.00)^2$	$(476.52)^2$
0.172	$(516.08)^2$	$(150.06)^2$	$(558.01)^2$
0.199	$(769.19)^2$	$(196.27)^2$	$(817.74)^2$
0.230	$(960.30)^2$	$(221.44)^2$	$(1010.07)^2$
0.263	$(1230.90)^2$	$(272.06)^2$	$(1289.63)^2$
0.303	$(1444.84)^2$	$(326.61)^2$	$(1516.87)^2$
0.351	$(1541.69)^2$	$(341.94)^2$	$(1615.75)^2$
0.406	$(1809.92)^2$	$(388.27)^2$	$(1884.69)^2$
0.466	$(2058.63)^2$	$(445.30)^2$	$(2152.80)^2$

Table 2. Showing the Δ^2_{21} upper limit at 2σ ($\Delta^2_{21,2\sigma}$), and error ($\Delta^2_{21,\mathrm{err}}$), at a given k-bin for FastICA.

k (hcMpc^{-1})	Δ^2_{21} (mK^2)	$\Delta^2_{21,\mathrm{err}}$ (mK^2)	$\Delta^2_{21,2\sigma}$ (mK^2)
0.075	$(58.24)^2$	$(31.10)^2$	$(72.98)^2$
0.085	$(100.32)^2$	$(54.98)^2$	$(126.92)^2$
0.098	$(112.00)^2$	$(59.57)^2$	$(140.14)^2$
0.112	$(224.05)^2$	$(90.30)^2$	$(257.89)^2$
0.130	$(262.41)^2$	$(101.46)^2$	$(299.07)^2$
0.151	$(436.77)^2$	$(137.43)^2$	$(478.06)^2$
0.172	$(516.43)^2$	$(150.16)^2$	$(558.39)^2$
0.199	$(769.99)^2$	$(196.21)^2$	$(817.52)^2$
0.230	$(959.09)^2$	$(221.16)^2$	$(1008.80)^2$
0.263	$(1228.46)^2$	$(271.53)^2$	$(1287.08)^2$
0.303	$(1443.43)^2$	$(326.29)^2$	$(1514.40)^2$
0.351	$(1540.83)^2$	$(341.75)^2$	$(1614.85)^2$
0.406	$(1803.51)^2$	$(388.39)^2$	$(1885.29)^2$
0.466	$(2059.14)^2$	$(445.41)^2$	$(2153.33)^2$

Table 3. Showing the Δ^2_{21} upper limit at 2σ ($\Delta^2_{21,2\sigma}$), and error ($\Delta^2_{21,\mathrm{err}}$), at a given k-bin for GPR.

k (hcMpc^{-1})	Δ^2_{21} (mK^2)	$\Delta^2_{21,\mathrm{err}}$ (mK^2)	$\Delta^2_{21,2\sigma}$ (mK^2)
0.075	$(58.06)^2$	$(31.20)^2$	$(72.92)^2$
0.085	$(80.61)^2$	$(33.45)^2$	$(93.47)^2$
0.098	$(99.22)^2$	$(35.03)^2$	$(110.90)^2$
0.112	$(110.65)^2$	$(44.04)^2$	$(126.97)^2$
0.130	$(143.23)^2$	$(44.97)^2$	$(156.72)^2$
0.151	$(189.54)^2$	$(55.69)^2$	$(205.24)^2$
0.172	$(240.49)^2$	$(57.87)^2$	$(254.03)^2$
0.199	$(293.08)^2$	$(61.88)^2$	$(305.86)^2$
0.230	$(365.19)^2$	$(65.23)^2$	$(376.66)^2$
0.263	$(421.48)^2$	$(75.85)^2$	$(434.91)^2$
0.303	$(469.14)^2$	$(88.26)^2$	$(485.46)^2$
0.351	$(510.65)^2$	$(99.14)^2$	$(529.55)^2$
0.406	$(568.37)^2$	$(112.48)^2$	$(590.21)^2$
0.466	$(638.14)^2$	$(129.66)^2$	$(663.73)^2$

2σ value of $\Delta^2_{21} < (73)^2$ mK^2. This is interesting, as neither of the two assume anything about the data a-priori, whilst GPR requires additional information about the data it is fitting. As we move to k-scales larger than hcMpc^{-1}, both FastICA and GMCA begin to deviate away from GPR, which consistently stays close to the noise limit (shown in grey). In fact, for each adjacent k-bin, at k-scales larger than hcMpc^{-1}, FastICA and GMCA produce progressively larger powers. FastICA and GMCA perform near identically, with their power over the range of k-scales, shown in Fig. 4, being near-identical.

4.2 Simulated observation

From the results in Fig. 4, we see that both FastICA and GMCA begin to diverge from the result of GPR at k-scales larger than hcMpc^{-1}. To better understand why, we further explore the performance of both GMCA and FastICA, with the performance scales $k > h$cMpc^{-1} being the motif throughout the analysis.

Within our plots we have a few graphical convention: Any k-bins where the power of the residuals, before noise bias removal, is below that of the noise are omitted. This leads to an unphysical negative power for the residual once the noise bias has been removed; any error bars extending to below the x-axis in log space are shown with equal-sized arrows to indicate this.

4.2.1 Simulated: CLEAN

We first start with how the foreground removal techniques perform on simulated data that has not gone through OSKAR nor had noise added; these data have both the cosmological signal and foregrounds, only. We can see the result of the power spectrum recovery in Fig. 5.

To find the best number of components to model the foregrounds with the BSS techniques, we applied the following technique for the goodness of fit: Taking the squared difference between the residual of our BSS techniques, being the recovered cosmological signal, and the cosmological signal we wish to recover. We sum over each pixel of the data cube's squared difference, and normalize this difference by the variance of the cosmological signal at the frequency the pixel lies in. If the residual perfectly matches the cosmological signal, the returned value is zero. The motivation for normalizing by the variance allows us to see if the squared error in the residual is below that of

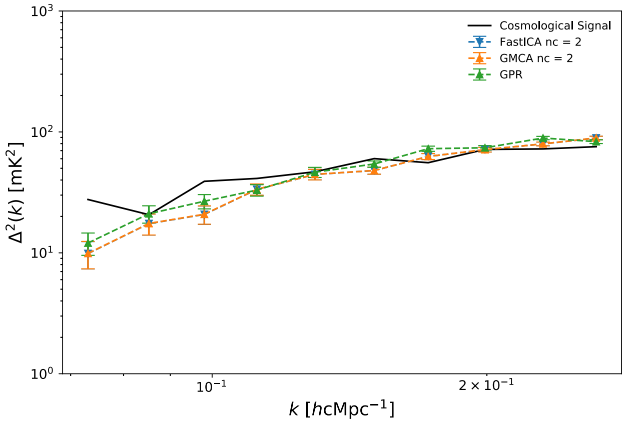


Figure 5. Comparison of FastICA, GMCA, and GPR residual power spectra as they recover the cosmological signal injected into simulated foregrounds (outlined in Jelić et al. 2008) – recovery performed over a $\Delta\nu = 30$ MHz bandwidth, 120 MHz $\leq \nu \leq$ 150 MHz. We model the foregrounds with two components for FastICA and GMCA, determined by the method outlined in Section 4.2. Due to their similar performance, the FastICA result has been overlaid by GMCA.

signals variance, giving a value below 1, and hence the cosmological signal is still detectable.

We find that modelling the foregrounds, within FastICA and GMCA, is best with two components. For GPR, we use the same covariance matrix described in Section 4.1 but without the mode-mixing kernel. All three of the techniques perform near identically, and do recover the 21 cm power spectrum fairly well, with their aforementioned technique for goodness of fit returning a value below 0.1.

4.2.2 Simulated: observation comparison

We now look to see how well FastICA and GMCA perform on simulated observations via OSKAR. We compare three simulated observations with noise included:

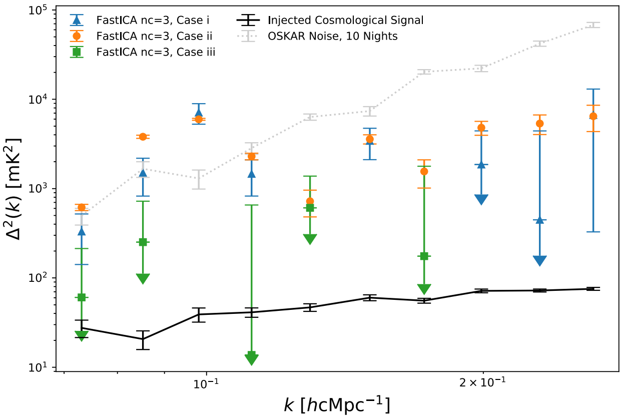
(i) 120 h observation.

(ii) 120 h observation but we have removed the primary beam from the observation.

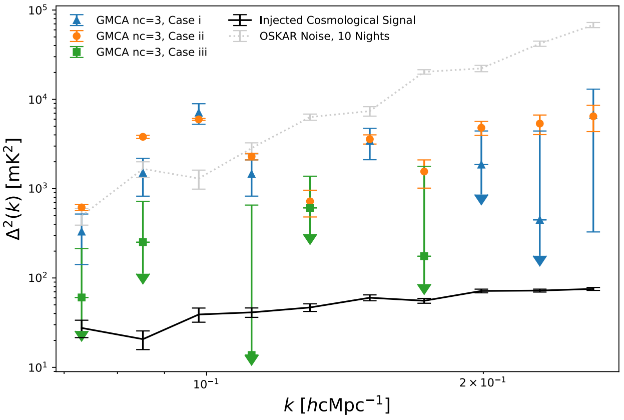
(iii) 120 h observation, but we have used the resolution from the lowest frequency in the observation's point source function (PSF). For the purposes of this paper, it is the PSF at 134.5 MHz; the primary beam is also included.

We see the results of this in Fig. 6. The reasoning for case (ii) is to see whether the primary beam effects are the limiting factor for both FastICA and GMCA (compared to GPR), and case (iii) is to see if the chromaticity of the instrument is a limiting factor for FastICA and GMCA. As mentioned in Section 4.1, GPR takes kernels for the various components in the data, one of which was $\mathbf{K}_{\mathrm{fg}}$, which is the kernel for the instrumental effects – namely, mode-mixing. Whilst GPR accounts for this, FastICA and GMCA, being BSS techniques, assume nothing of the data a-priori and solely attempt to fit the data. As it is ignorant to such effects, they may provide a hindrance to the recovery capabilities of the two techniques. For case (iii), we have removed this mode-mixing covariance from GPR's modelling as we no longer have the chromaticity of the instrument.

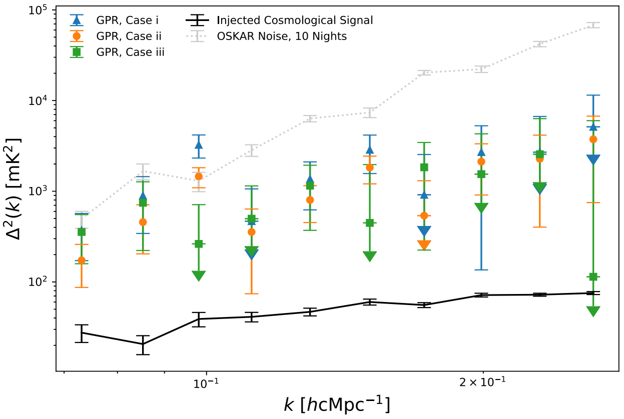
We only show the 'best' (i.e. closest to the noise limit) results. For the BSS techniques, the number of components used to define the foregrounds ranged from 1 to 37; with the number of components being limited by the dimensionality of the data set – which, here, is set by the number of frequency channels. We choose the number of components to be 3, as this is what we find to be best for case (i). We



(a) FastICA



(b) GMCA



(c) GPR

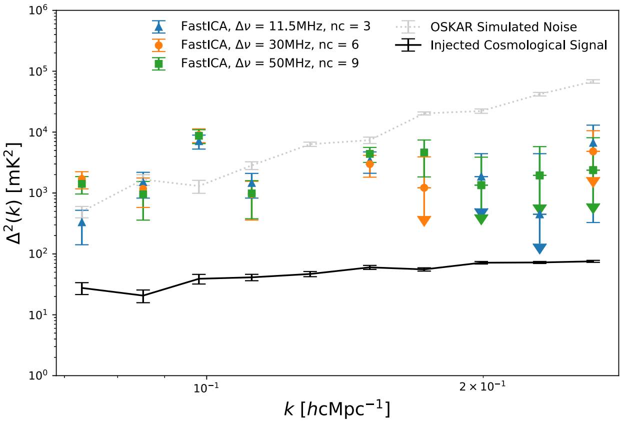
Figure 6. Comparison of FastICA, GMCA, and GPR in recovering a cosmological signal injected into simulated foregrounds (outlined in Jelić et al. 2008), with OSKAR being used to simulate observations. We then compare three cases: (i) Normal 120 h observation. (ii) 120 h observation with no primary beam included. (iii) 120 h observation with the PSF of the lowest frequency (134.5 MHz), with the primary beam included. The bandwidth of the cube is $\Delta\nu = 11.5$ MHz, in the range: 134.5 MHz $\leq \nu \leq$ 146 MHz.

perform the analysis of cases (ii) and (iii) with three components to compare how the results change. From Fig. 6, it is clear to see that case (i), which represents a normal observation, has a better result for FastICA and GMCA than that in Fig. 4. This result is due to a common downfall of simulations, as they do no reproduce residuals that arise from a normal LOFAR pipeline, such as direction-dependent and direction-independent calibration errors or RFI flagging. Offringa, Mertens & Koopmans (2019) showed that RFI flagging can lead to a flagging excess power – which was found to be mitigated by the use of a Gaussian-weighted interpolation scheme for the flagged samples along with a unitary weighting during averaging. They also showed that that GPR is able to model some of this excess at low k-scales, and so it is a reasonable question to ask if the effect seen at k-scales larger than $h\mathrm{cMpc}^{-1}$ is due to this. We tested this by taking an extreme example of $\sim\frac{1}{3}$ of the channels flagged, which is the equivalent of the gridding step with flagging. We found that, for the case (i) for GMCA and FastICA, there is excess at scales of $\sim$0.2 $h\mathrm{cMpc}^{-1}$ and larger, progressively taking the residual to the noise limit by the penultimate k-bin. We did not see any change in the GPR result, between the flagged test and Fig. 6(c). Seeing as the data used in Fig. 4 used the steps proposed by Offringa et al. (2019), we can be sure it is not this effect causing it. From Figs 6(a) and (b), we see, as expected, FastICA and GMCA perform near identically. For case (ii), one should note that we do not delimit the primary beam from the power spectrum (equation 22, Mertens et al. 2020). We see that the results for case (ii) are similar to that of case (i). This is not surprising as we reduce our FoV from 5° × 5° to 4° × 4° to be within the primary beam, and so the effects of the primary beam are less significant. For case (iii), the removal of chromaticity effects allow FastICA and GMCA to perform better on the lowest k-scales, as mode mixing introduces additional fluctuations into the observed signal – that are no longer present here. However, at larger k-scales, we see overfitting and we see more of the cosmological signal leaked into the foreground model. The removal of the primary beam and chromaticity effects reduces the number of independent components at higher k-scales, as such three components at higher k may be too many, leading to overfitting.

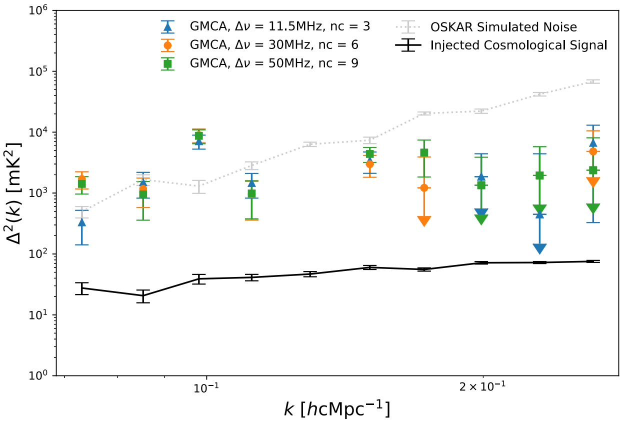
As seen for FastICA and GMCA, case (ii), unsurprisingly, with no primary beam GPR performs similarly to case (i). For case (iii), for which we do not include the mode-mixing kernel, we see that the results are similar to case (i). We conclude that instrumental effects are not the dominant effect causing the poor recovery of FastICA and GMCA, at scales $k > h\mathrm{cMpc}^{-1}$ in Fig. 4. We see instrumental effects yield no worsening of performance on scales $k > 0.1\ h\mathrm{cMpc}^{-1}$. With instrumental effects not providing any clarity on the performance of FastICA and GMCA at scales $k > h\mathrm{cMpc}^{-1}$, we extend our analysis to look at the further issues, such as limitations due to bandwidth of data and noise.

4.2.3 Simulated: not enough data

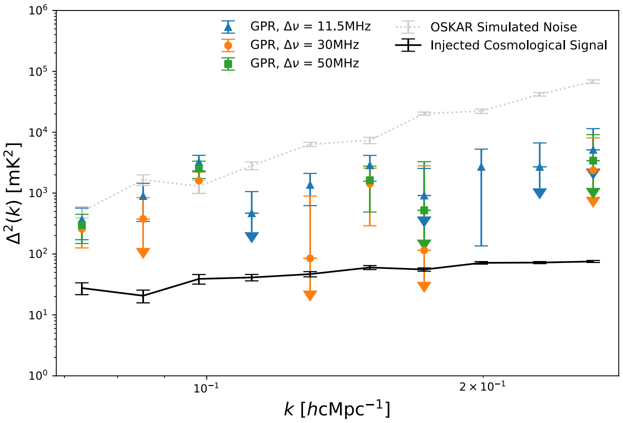
Currently, the foreground removal techniques have been applied to an image cube of bandwidth $\Delta\nu = 11.5$ MHz. The more data we give these techniques, the more information they have to fit the foregrounds and, potentially, recover the cosmological signal. We compare three cases: The first being the normal $\Delta\nu = 11.5$ MHz we have been using in our analysis thus far; the second case is a cube of bandwidth $\Delta\nu = 30$ MHz, and the final case is a cube of bandwidth $\Delta\nu = 50$ MHz. We see the results of this in Fig. 7, note that all power spectra are taken in the $\Delta\nu = 11.5$ MHz bandwidth, 134.5 MHz $\leq \nu \leq$ 146.0 MHz.



(a) FastICA



(b) GMCA



(c) GPR

Figure 7. Comparison of FastICA, GMCA, and GPR in recovering a cosmological signal injected into simulated foregrounds (outlined in Jelić et al. 2008) with OSKAR being used to simulate observations. We compare three cases: The first being the normal $\Delta\nu = 11.5$ MHz we have been using in out analysis thus far; the second case is a cube of bandwidth $\Delta\nu = 30$ MHz, and the final case is a cube of bandwidth $\Delta\nu = 50$ MHz.

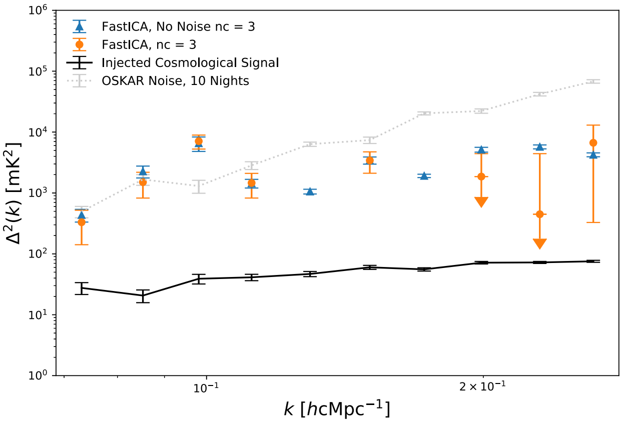
From Figs 7(a) and (b), we see there is some improvement with the larger bandwidth, with no clear improvement as we compare the $\Delta\nu = 30$ MHz and $\Delta\nu = 50$ MHz for the image cubes for GMCA and FastICA. We choose the number of components such that the cosmological signal is not suppressed, i.e. the cosmological signal doesn't leak into the foreground model. For each of the k-bins in the power spectrum, we see that there is no clear improvement of the power spectrum, at scales $k > h\mathrm{cMpc}^{-1}$, relative to that if the lower k-scales, as we increase the bandwidth. For GPR we see that, overall, the larger bandwidths perform the worse with many k-bins in the power spectrum being completely suppressed, Fig. 7(c). This result with the larger bandwidths is not surprising, as the frequency coherency of the signal evolves with redshift. Using a longer bandwidth would average this evolution, and one would expect a degradation of performance with bandwidth.

One would expect larger bandwidths to do better as there is more information for GMCA and FastICA to exploit in their decomposition. As aforementioned, GMCA has had great success in its application to CMB data. However, most applications of these BSS techniques have used large bandwidths. In Chapman et al. (2012) and Chapman et al. (2013), FastICA and GMCA were applied to a data set of bandwidth $\Delta\nu = 85$ MHz. In a most recent application of GMCA, Carucci et al. (2020), a data set of bandwidth $\Delta\nu = 400$ MHz was used; they also found that having a greater number of channels is better to perform sparsity. So it is likely, out of the realm of the motif of this analysis, that we are not giving the BSS techniques enough data, as even a bandwidth $\Delta\nu = 50$ MHz seems inadequate. The performance at longer bandwidths will not be affected by the evolution of the signal, such as that which the power spectrum suffers from as we perform it over larger bandwidths (Datta et al. 2012, 2014; Ghara, Datta & Choudhury 2015). We have seen this when looking at different bandwidth sizes for simulated data that hasn't been applied to OSKAR; as we increased the bandwidth over which removal is performed, the performance of FastICA and GMCA improved. With no clear improvement of the power spectrum, at scales $k > h\mathrm{cMpc}^{-1}$ with an increase of bandwidth, we next look at the role of instrumental noise.

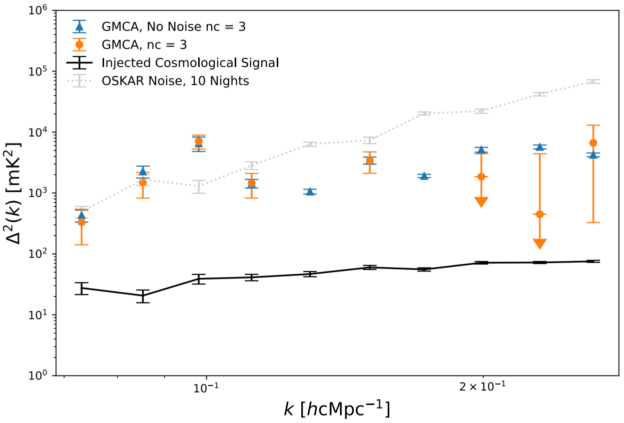
4.2.4 Simulated: noise limited

A key difference between the BSS techniques used and GPR, is that the BSS techniques are agnostic to the existence of the cosmological signal in the data set, whilst GPR not only optimizes itself for the inclusion of the 21 cm signal but also uses the Stokes V to estimate the noise. This could potentially mean that noise limitations on FastICA and GMCA are greater than GPR, as the former two cast the cosmological signal in the 'Gaussian' noise of their model. We now look at the case of no noise. In Fig. 8, we show the application of FastICA and GMCA when applied to simulated data containing a no noise. We limit the plot to a fiducial value of three components. For FastICA, Fig. 8(a), we see that the no noise case does produce a similar power spectrum to the case with noise. We find that the case with no noise for GMCA, Fig. 8(b), performs the same as the case with noise. However, without the inclusion of noise, they are still orders of magnitude above the injected cosmological signal. Whilst noise is often a limiting factor for recovery, it is not the dominant reason here as to why we see an underperformance on larger k-scales, $k > h\mathrm{cMpc}^{-1}$.

An important assumption of the blind source separation technique, as previously implemented on EoR data, is perfect noise avoidance. This is the assumption that we model the noise perfectly, and it does not leak into our foreground model. However, we see that above four components noise does begin to leak into the foreground model. This poses a problem, as we can see from Fig. 8, three components is not sufficient to model the foregrounds well enough by the residual – for Fig. 4, six components was found to be the optimal number for FastICA and GMCA.



(a) FastICA



(b) GMCA

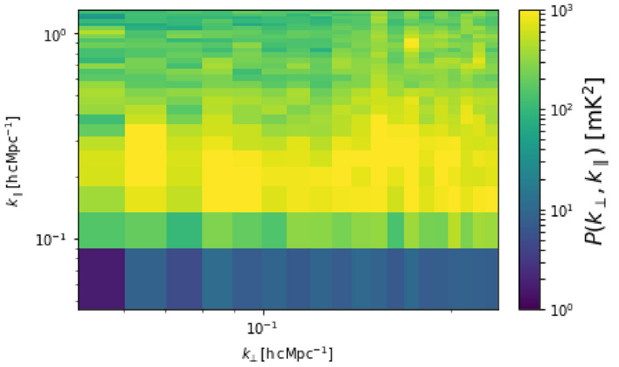
Figure 8. Comparing FastICA and GMCA by considering the case where there's no noise. The data has still gone through OSKAR to simulated 10 nights of observation. We show the results using a fiducial value of three components to model the foregrounds.

4.3 Blind source separation techniques: scale independence

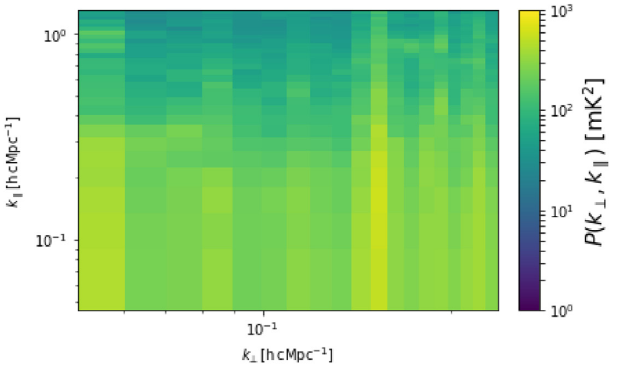
Thus far, we have only looked at the performance of the BSS techniques with respect to the data – both its constituents and the amount of data. In Fig. 4, we see that FastICA and GMCA underperform, with respect to GPR, as we increase the k-scale; with each adjacent k-bin producing a higher power than the bin that proceeded it. In Fig. 9, we show the foreground removal residuals produced by FastICA and GPR in the 2D cylindrical power spectrum form. As seen throughout this work, FastICA and GMCA perform near-identically; this similar performance was also seen by Cunnington et al. (2020). As such, for brevity, we only compare FastICA and GPR. The cylindrical power spectrum is the cylindrically averaged power spectrum and is defined as a function of angular ($k_\perp$) versus line-of-sight ($k_\shortparallel$) scales as,

$$P(k_\perp, k_\parallel) = \langle P(\mathbf{k})_{k_\perp, k_\parallel} \rangle \tag{18}$$

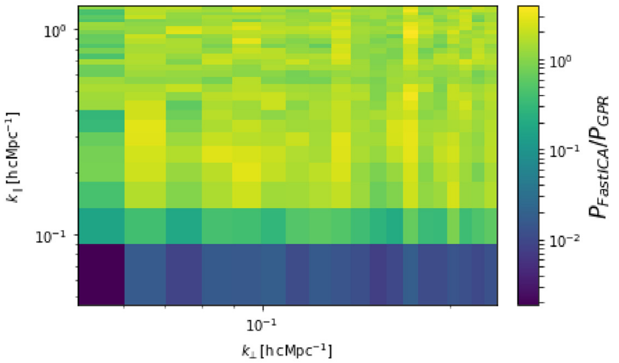
In Fig. 9(c), we see that confined to the lower $k_\shortparallel$ modes, FastICA gives a comparable removal residual to GPR. However, as we increase $k_\shortparallel$ and $k_\perp$ we see that FastICA produces a larger power for these modes – in line with what is seen in Fig. 4. At $k_\shortparallel \geq 0.8\, h\mathrm{cMpc}^{-1}$, we have artefacts introduced by the frequency Fourier transform, which arise from the discontinuities in the frequency sampling (from RFI flagging) causing spectral leakage. Interestingly, at these higher



(a) The 2D cylindrical power spectrum of the FastICA foreground removal residual produced using the LOFAR 10 Nights data, using 6 components to model the foregrounds.



(b) The 2D cylindrical power spectrum of the GPR foreground removal residual produced using the LOFAR 10 Nights data.



(c) Looking at the 2D cylindrical power spectrum for ratio between FastICA's and GPR's foreground removal residual of the LOFAR 10 Nights Data.

Figure 9. We compare the dimensionless cylindrical power spectrum of the foreground fitting residuals produced by the application of FastICA, using six components to model the foregrounds, and GPR on the 10 nights of LOFAR data.

$k_\parallel$, we see more of the spectral leakage in the FastICA residual compared to that of GPR.

BSS techniques are not scale-dependent, they apply the same mixing matrix model to all scales. On smaller k-scales (larger scales in real space), there are a smaller number of independent components and so modelling the foregrounds with six components is sufficient. As we go to larger k-scales (smaller scales in real space), the number of independent components increase, and so modelling the foregrounds with six components is no longer sufficient. As the BSS techniques do not account for this scale-dependence, one sees an increasing underperformance of the techniques with k. In Fig. 10, we apply FastICA to only our simulated foregrounds from Section 3.2.2. From Fig. 10, we see as we increase the number of components, with which we model the foregrounds, the better convergence the recovered power spectrum has with the power spectrum of the foreground we wish to model. We see that FastICA reaches convergence on lower scales, below hcMpc^{-1}, with just three components. This turnover of performance at hcMpc^{-1} is what we also observe in Fig. 4. As for larger scales in Fig. 10, as we increase the number of components, we see a progressively better fit at these scales. One can easily argue, from Fig. 10, that using a higher number of components is better as we would get a better fit to the foregrounds. However, in Fig. 11, we look at FastICA's recovery of the simulated cosmological signal. As we increase the number of components, we see that more of the simulated cosmological signal is modelled in the foreground model – as in equation (1). With the smaller k-scales being penalized more for the increase in components, than the larger k-scales. We can also potentially see this over-subtraction/signal suppression at the lower k-scales in Fig. 9(c)'s first $k_\parallel$ bin. Such suppression by these techniques on these lower k-scales is also seen by Cunnington et al. (2020).

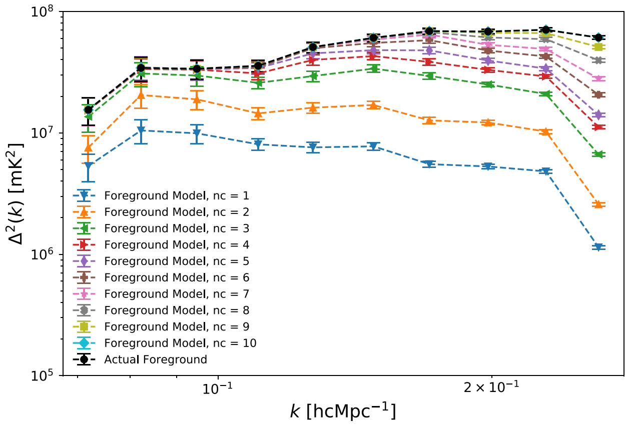


Figure 10. A comparison of FastICA's foreground model power spectra, over bandwidth $\Delta\nu = 11.5$ MHz. Using different numbers of components, ranging from 1 to 10.

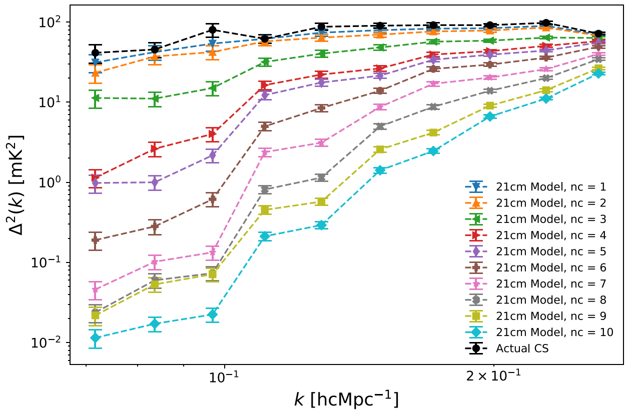


Figure 11. Comparison of power spectra of FastICA's recovered simulated cosmological signal, over bandwidth $\Delta\nu = 11.5$ MHz. Using different numbers of components, ranging from 1 to 10.

From Figs 10 and 11, we can infer that, though a higher number of components will be able to model larger k-scales, it will cause overfitting of the signal at smaller k-scales and results in the cosmological signal, on these smaller k-scales, leaking into the foreground model. Therein lies the problematic nature of the scale-independence of these BSS techniques. If one were able to fit different scales with the components best suited for it, rather than having a signal number of components across all scales, we are more likely to get a better fit. In future work, we will explore techniques that are scale-dependent and compare their performance to GPR. The version of GMCA used here is the default version used in Patil et al. (2017). There have been significant improvements and advancements of this BSS.[17] For example, local-generalized morphological component analysis (LGMCA)[18] (Bobin et al. 2013a). LGMCA, allows the mixing matrix to vary across the pixels in an image.

5 DISCUSSION AND CONCLUSION

With the recent 2σ upper limit of $\Delta^2_{21} < (73)^2\,\mathrm{mK}^2$ at $k = 0.075\,h\mathrm{cMpc}^{-1}$ by Mertens et al. (2020) from 141 h of data, and with LOFAR currently processing over 2000 h of data from their observing run of the NCP, we aimed to compare removal techniques to assess how well they perform under different conditions and their limitations. In this paper, we attempt to replicate the analysis from Mertens et al. (2020), applying FastICA, GMCA, and GPR to the data used in Mertens et al. (2020), to see which performs better. We then replicate the data via simulation to pin down the reasoning as to why certain removal techniques perform worse than others. In Section 4.1, we can see we perfectly recover the upper limit found in Mertens et al. (2020) in our use of GPR. As we know more about the signals we wish to observe, plus what constitutes it (e.g. foregrounds and mode mixing), we can have more control over the characteristics and properties of the signal we wish to model/fit. This is where GPR becomes powerful, taking the properties of these constituents (e.g. coherency length) to better fit these foregrounds and recover the cosmological signal. Interestingly, we have found that FastICA and GMCA also reproduce this 2σ upper limit. This is a powerful result as both GMCA and FastICA are blind to the data, i.e. knowing nothing a-priori about the data.

From our results in Fig. 4, we see that on scales larger than $h\mathrm{cMpc}^{-1}$ FastICA and GMCA, performing identically, begin to deviate away from the spectrum produced by GPR. We began looking at possible reasons for FastICA's and GMCA's underperformance at these scales by looking at instrumental effects. We simulate foregrounds and the cosmological signal at the same redshift and frequency range as the observed data. We use OSKAR to apply observational effects in three case, with noise included:

(i) 120 h observation.

(ii) 120 h observation, but we have removed the primary beam from the observation.

(iii) 120 h observation, but we have used the resolution from the lowest frequency in the observation's PSF. For the purposes of this paper it is the PSF at 134.5 MHz; the primary beam is also included.

From Figs 6(a) and (b), we find that the two cases (i) and (ii) give similar power spectra. As we cut our FoV to be within the primary beam in the hope of mitigating most of the effects, this result is to be expected. As for case (iii), we see that though it gives a deeper power spectrum, there are clear signs of overfitting. We use the same number of components in all three cases to give a true comparison. It is clear here that the removal of chromaticity effects removes independent components from the data, especially at higher k-scales. From this, we see it appears that no single instrumental effect is the dominant effect causing the underperformance at larger k-scales – as we see no adverse effects on scales larger than $h\mathrm{cMpc}^{-1}$ for both FastICA and GMCA. We then perform further checks that relate to the data itself. It is important to note how powerful the BSS techniques truly are. As we have said, they do not assume anything a-priori about the data. Yet, on simulated data, without OSKAR applied, they can recover the cosmological signal embedded in foregrounds.

We started our further analysis off with the amount of data we give our removal techniques. The original comparison performed in Fig. 6 was done using a $\Delta\nu = 11.5$ MHz data cube, between 134.5 and 146 MHz, to match that of the Mertens et al. (2020) data. We then looked at two additional cube frequency bandwidths: $\Delta\nu = 30$ MHz and $\Delta\nu = 50$ MHz. We see in Fig. 7, that there is an improvement in the results from FastICA and GMCA, with the $\Delta\nu = 30$ MHz and $\Delta\nu = 50$ MHz bandwidths performing better than the $\Delta\nu = 11.5$ MHz case. However, there doesn't appear to be any significant improvement between $\Delta\nu = 30$ MHz and $\Delta\nu = 50$ MHz. Though we note for most applications of these BSS techniques, the bandwidth that is used is at least double of the highest we've given, $\Delta\nu = 50$. For scales larger than $h\mathrm{cMpc}^{-1}$, there is no answer to the underperformance – as we are seeing an improvement across all scales for the larger bandwidths. In the case of GPR, increasing the bandwidth worsens its performance. Larger bandwidths will average the frequency coherence of the data, and so one would expect GPR's performance to degrade with bandwidth. We have said that typically these BSS techniques require larger bandwidths, if GPR can perform well on relatively smaller bandwidths, it could prove advantageous.

With no clear signs of degradation with the amount of data we use, we move on to see if noise was the limiting factor. We do not include noise in our OSKAR output and attempt to recover the cosmological signal. In our analysis, we limited the results to a fiducial value of three components, and did not see an improvement in the no noise case. This is interesting as the perfect noise avoidance of BSS techniques breaks down, as some noise must be in the foreground term. We see no improvement on scales larger than $h\mathrm{cMpc}^{-1}$, and hence noise is not the limiting factor here.

We next looked at the scale-independence of the BSS techniques. When applying FastICA or GMCA to data using a given number of components to model the data, it tries to model every scale of the data with that number of components. However, at smaller k-scales there are a smaller number of independent components, and so modelling these scales with a small number of components is sufficient. As we move to larger k-scales, the number of independent components also increases. Hence, the number of components that were sufficient to model the smaller k-scales, is no longer sufficient at these larger k-scales. We see this turning point at $h\mathrm{cMpc}^{-1}$, in both Figs 4 and 10. On scales below $h\mathrm{cMpc}^{-1}$, using a handful of components to model the foregrounds is sufficient, but they are not sufficient to fully capture all the independent components at larger k-scales. A simple approach would be to apply the larger number of independent components to model the foregrounds. However, all scales will be modelled with this number and we find that the 21 cm signal loss is significant on smaller scales, when using this approach. When comparing Fig. 4 to the results seen in Section 4.2, one can see that the number of components required to model the foregrounds is far greater, even on the smaller k-scales, for the real data compared to the simulations. This is due to the short comings of simulations not being able to match excess power of higher k-scales caused by, for example, RFI flagging or calibration errors. This then reduces the number of independent components seen at higher k-scales, as such the scale-independence issue is not as significant and hence we see a deeper power spectrum at larger k-scales for the simulated case versus the observed data. We also see the need of having a scale-dependent approach when looking to increase the bandwidth to improve the performance of BSS techniques. As aforementioned, applications of the BSS techniques have used larger bandwidths than that used in this paper as well as Patil et al. (2017) and Gehlot et al. (2019). The applications to simulated data have been to that without instrumental effects, as such there is a linear improvement with bandwidth. Whereas for simulated case with observational effects, we do see some improvements but it is likely that the scale-independence issue is more prominent and so the improvement isn't as clear. With real data, this is likely further exacerbated by the greater number of instrumental effects causing more independent components at high k-scales.

---

[17] http://www.cosmostat.org/software/gmcalab

[18] https://www.cosmostat.org/software/lgmca

If one were able to perform BSS on a scale-dependent approach, with a mixing matrix for different scales being modelled with different number of components, we believe we would likely see a better fit to the data and remove the divergence seen in Fig. 4. There is also no way to determine the best number of components to recover the signal, for each of these scales. One could, in future, aim to make these techniques less blind. As we have a grasp on the properties of the data, one could add a Bayesian backend to these BSS techniques to optimize the number of components, on each scale, on which sparsity is performed.

Comparatively, GPR only takes into account the variance and optimizes its fit with the knowledge of the constituents of the observed signal: that there is a weaker signal present, as it has a kernel modelling the signal, in the foregrounds and mode-mixing effects. It is less likely to overfit and remove power for the recovered signal, as it is accounted for, and is also able to model the mode mixing effect in the foregrounds as we optimize for its inclusion. GPR can also output the optimized covariance matrix, so one can see how it chose the fit – something neither FastICA or GMCA can do. We also see clear improvements in the performance of GPR as we increase the bandwidth over which we apply GPR, likely showing its accounting of mode-mixing makes instrumental effects less detrimental.

GPR also has scope for improvement. GPR, currently, fits visibility by visibility, equivalent to one line of sight at a time, and so can be further improved by utilizing all of the lines of sight together, to include spatial correlations, to better improve the fitting. As mentioned in Section 2.3, the covariance parameters for mode-mixing were based on treating mode-mixing as a Gaussian process rather than simulating the effect (Mertens et al. 2018; Gehlot et al. 2019). If it were to be updated and tested on a more realistic mode-mixing, the covariance $\mathbf{K}_{\mathbf{fg}}$ is more accurately tested and parameters refined, it will likely fit the data even better improving the recovery. GPR is also not scale-dependent, as the same covariance model for all visibilities. Improving this, is likely to yield a better recovery of the cosmological signal as well.

With LOFAR still processing the observed data from its observing run of the NCP, we conclude that the foreground removal technique that is best suited, out of the three tested methods, to recover the 21 cm power spectrum by LOFAR is GPR, as used in Mertens et al. (2020). Through the testing we have done in this paper that GPR has shown to be the most rigorous and still has scope to be improved on. Its application on the Mertens et al. (2020) data, looks to be close to noise-limited and as the full observational data of the NCP it processed, will likely detect the signal in this low-noise regime.

There is also need for further exploration of these BSS techniques in the production of upper limits. On k-scales below hcMpc^{-1}, we have seen on these scales, on actual LOFAR data, they perform as well as GPR. As aforementioned, BSS techniques are very powerful as they do not assume anything about the data a-priori, and there have been many further iterations of GMCA. Application of a scale-dependent approach is something to be looked at in future works. Machine learning is likely going to play a prominent role in foreground mitigation in the near future, with a framework being set up for LOFAR-EoR. Sooknunan et al. (in preparation), are developing a machine learning framework for LOFAR-EoR foreground mitigation techniques.

We have also seen that Mertens et al. (2020) compared to Patil et al. (2017), as well as using a different foreground removal technique, has improved the direction-dependent calibration of the data. Once LOFAR has fully processed the data from its observing run of the NCP, it is likely to have significantly improved its calibration, as well as improving the current sky model and the inclusion of phase errors introduced by the ionosphere. This will yield deeper power spectra limits as well as increasing the likelihood of detecting the cosmological signal with the fully processed data.

ACKNOWLEDGEMENTS

EC acknowledges the support of a Royal Society Dorothy Hodgkin Fellowship and a Royal Society Enhancement Award. JRP acknowledges financial support from the European Research Council under ERC grant number 638743-FIRSTDAWN. FGM and LVEK would like to acknowledge support from a SKA-NL Roadmap grant from the Dutch ministry of OCW. ITI was supported by the Science and Technology Facilities Council [grant numbers ST/I000976/1 and ST/T000473/1] and the Southeast Physics Network (SEP-Net). VJ acknowledges support by the Croatian Science Foundation for the project IP-2018-01-2889 (LowFreqCRO). BKG would like to acknowledge the National Science Foundation grant AST-1836019. We also thank Catherine A. Watkinson for her useful discussions and comments. This research made use of astropy, a community-developed core Python package for astronomy (Astropy Collaboration 2018); SCIPY, a Python-based ecosystem of open-source software for mathematics, science, and engineering (Virtanen et al. 2020) – including numpy (Walt, Colbert & Varoquaux 2011); MATPLOTLIB, a Python library for publication quality graphics (Hunter 2007); SCIKIT-LEARN, a free software machine learning library for the Python programming language, for the use of FastICA (Pedregosa et al. 2011).

DATA AVAILABILITY

The data underlying this article were provided by LOFAR under licence / by permission. Data will be shared on request to the corresponding author with permission of LOFAR.

REFERENCES

Astropy Collaboration, 2018, AJ, 156, 123
Barry N. et al., 2019, ApJ, 884, 1
Beardsley A. P. et al., 2016, ApJ, 833, 102
Bernardi G. et al., 2009, A&A, 500, 965
Bobin J., Starck J. L., Sureau F., Basak S., 2013a, A&A, 550, A73
Bobin J., Sureau F., Paykari P., Rassat A., Basak S., Starck J. L., 2013b, A&A, 553, L4

Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
Carucci I. P., Irfan M. O., Bobin J., 2020, MNRAS, 499, 304
Chapman E. et al., 2012, MNRAS, 423, 2518
Chapman E. et al., 2013, MNRAS, 429, 165
Chen H., Yao D., Zhuo Y., Chen L., 2003, Brain Topography, 15, 223
Cheng C. et al., 2018, ApJ, 868, 26
Ciardi B., Ferrara A., 2005, Space Sci. Rev., 116, 625
Cunnington S., Irfan M. O., Carucci I. P., Pourtsidou A., Bobin J., 2020, preprint (arXiv:2010.02907)
Datta K. K., Mellema G., Mao Y., Iliev I. T., Shapiro P. R., Ahn K., 2012, MNRAS, 424, 1877
Datta K. K., Jensen H., Majumdar S., Mellema G., Iliev I. T., Mao Y., Shapiro P. R., Ahn K., 2014, MNRAS, 442, 1491
DeBoer D. R. et al., 2017, PASP, 129, 045001
Di Matteo T., Ciardi B., Miniati F., 2004, MNRAS, 355, 1053
Fan X. et al., 2006, AJ, 132, 117
Fialkov A., Barkana R., Cohen A., 2018, Phys. Rev. Lett., 121, 011101
Field G. B., 1958, Proc. IRE, 46, 240
Furlanetto S. R., 2016, in Mesinger A., ed., The 21-cm Line as a Probe of Reionization. Springer International Publishing, p. 247
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Ge R., Wang Y., Zhang J., Yao L., Zhang H., Long Z., 2016, J. Neurosci. Method., 263, 103
Gehlot B. K. et al., 2019, MNRAS, 488, 4271
Gehlot B., 2019, PhD thesis, University of Groningen
Ghara R., Datta K. K., Choudhury T. R., 2015, MNRAS, 453, 3143
Ghara R. et al., 2020, MNRAS, 493, 4728
Gorce A., Douspis M., Aghanim N., Langer M., 2018, A&A, 616, A113
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
Greig B., Mesinger A., Bañados E., 2019, MNRAS, 484, 5094
Greig B. et al., 2020, preprint (arXiv:2006.03203)
Harker G. et al., 2009, MNRAS, 397, 1138
Hassan S., Liu A., Kohn S., La Plante P., 2019, MNRAS, 483, 2524
Hazelton B. J., Morales M. F., Sullivan I. S., 2013, ApJ, 770, 156
Hills R., Kulkarni G., Meerburg P. D., Puchwein E., 2018, Nature, 564, E32
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Hutter A., 2018, MNRAS, 477, 1549
Hyvärinen A., Karhunen J., Oja E., 2001, Independent Component Analysis. John Wiley and Sons, Hoboken, NJ
Jelić V. et al., 2008, MNRAS, 389, 1319
Kazemi S., Yatawatta S., Zaroubi S., Lampropoulos P., de Bruyn A. G., Koopmans L. V. E., Noordam J., 2011, MNRAS, 414, 1656
Kolopanis M. et al., 2019, ApJ, 883, 133
La Plante P., Ntampaka M., 2019, ApJ, 880, 110
Li W. et al., 2019a, MNRAS, 485, 2628
Li W. et al., 2019b, ApJ, 887, 141
Maino D. et al., 2002, MNRAS, 334, 53
Maino D., Banday A. J., Baccigalupi C., Perrotta F., Górski K. M., 2003, MNRAS, 344, 544
Mangena T., Hassan S., Santos M. G., 2020, MNRAS, 494, 600
Mertens F. G., Ghosh A., Koopmans L. V. E., 2018, MNRAS, 478, 3640
Mertens F. G. et al., 2020, MNRAS, 493, 1662
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Mondal R. et al., 2020, MNRAS, 498, 4178
Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127
Offringa A. R., de Bruyn A. G., Zaroubi S., Biehl M., 2010, preprint (arXiv:1007.2089)
Offringa A. R. et al., 2014, MNRAS, 444, 606
Offringa A. R., Mertens F., Koopmans L. V. E., 2019, MNRAS, 484, 2866
Oh S. P., Mack K. J., 2003, MNRAS, 346, 871
Paciga G. et al., 2011, MNRAS, 413, 1174
Paciga G. et al., 2013, MNRAS, 433, 639
Parsons A. R. et al., 2014, ApJ, 788, 106
Patil A. H. et al., 2016, MNRAS, 463, 4317
Patil A. H. et al., 2017, ApJ, 838, 65
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Philip L. et al., 2019, J. Astron. Instrum., 8, 1950004
Planck Collaboration XIII, 2016, A&A, 596, A108
Planck Collaboration VI, 2020, A&A, 641, A6
Prasad P. et al., 2016, J. Astron. Instrum., 5, 1641008
Price D. C. et al., 2018, MNRAS, 478, 4193
Pritchard J. R., Loeb A., 2012, Rep. Progr. Phys., 75, 086901
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Singh S. et al., 2018, ApJ, 858, 54
Singh S., Subrahmanyan R., 2019, ApJ, 880, 26
Sokolowski M. et al., 2015, Publ. Astron. Soc. Austr., 32, e004
Spinelli M., Bernardi G., Santos M. G., 2019, MNRAS, 489, 4007
Tie Y., Sahin M., 2005, J. Eng., 2, 90
Tingay S. J. et al., 2013, Publ. Astron. Soc. Austr., 30, e007
Trott C. M. et al., 2020, MNRAS, 493, 4711
van Haarlem M. P. et al., 2013, A&A, 556, A2
Virtanen P. et al., 2020, Nature Methods, 17, 261
Voytek T. C., Natarajan A., Jáuregui García J. M., Peterson J. B., López-Cruz O., 2014, ApJ, 782, L9
Walt S. v. d., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Wolz L. et al., 2015, MNRAS, 464, 4938
Wouthuysen S. A., 1952, AJ, 57, 31
Yatawatta S., Zaroubi S., de Bruyn G., Koopmans L., Noordam J., 2008, MNRAS, 414, 1656
Yatawatta S. et al., 2013, A&A, 550, A136
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. LSS/NenuFAR: The LOFAR Super Station project in Nançay, p. 687
Zaroubi S., 2013, The First Galaxies, Astrophysics and Space Science Library, 396, 45

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