Foreground removal using fastica

Chapman, Emma; Abdalla, Filipe B.; Harker, Geraint; Jelic, Vibor; Labropoulos, Panagiotis; Zaroubi, Saleem; Brentjens, Michiel A.; de Bruyn, A. G.; Koopmans, L. V. E.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1111/j.1365-2966.2012.21065.x

Publication date:
2012

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **423,** 2518–2532 (2012) doi:10.1111/j.1365-2966.2012.21065.x

Foreground removal using FASTICA: a showcase of LOFAR-EoR

Emma Chapman,[1⋆] Filipe B. Abdalla,[1] Geraint Harker,[2,3] Vibor Jelić,[4] Panagiotis Labropoulos,[4,5] Saleem Zaroubi,[5] Michiel A. Brentjens,[4] A. G. de Bruyn[4,5] and L. V. E. Koopmans[5]

[1]*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT*
[2]*Center for Astrophysics and Space Astronomy, University of Colorado, 389 UCB, Boulder, CO 80309-0389, USA*
[3]*NASA Lunar Science Institute, NASA Ames Research Center, Moffett Field, CA 94035, USA*
[4]*ASTRON, PO Box 2, NL-7990AA Dwingeloo, the Netherlands*
[5]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700AV Groningen, the Netherlands*

Accepted 2012 April 3. Received 2012 March 6; in original form 2011 December 16

ABSTRACT

We introduce a new implementation of the FASTICA algorithm on simulated Low Frequency Array Epoch of Reionization data with the aim of accurately removing the foregrounds and extracting the 21-cm reionization signal. We find that the method successfully removes the foregrounds with an average fitting error of 0.5 per cent and that the 2D and 3D power spectra are recovered across the frequency range. We find that for scales above several point spread function scales, the 21-cm variance is successfully recovered though there is evidence of noise leakage into the reconstructed foreground components. We find that this blind independent component analysis technique provides encouraging results without the danger of prior foreground assumptions.

Key words: methods: statistical – cosmology: theory – dark ages, reionization, first stars – diffuse radiation.

1 INTRODUCTION

Four hundred thousand years after the big bang, the recombination of electrons and protons resulted in a neutral Universe, steadily cooling with the Hubble expansion. The 'Dark Ages' followed recombination until, 400 Myr after the big bang, the first ionizing sources came into existence. This Epoch of Reionization (EoR) is one of the last unobserved eras of our Universe, but with a new generation of radio telescopes coming online [e.g. Low Frequency Array (LOFAR),[1] Giant Metrewave Radio Telescope,[2] Murchison Widefield Array,[3] Precision Array to Probe the Epoch of Reionization,[4] 21 Centimeter Array[5] and Square Kilometre Array (SKA)[6]], this is soon about to change.

It is generally accepted that the most rewarding way to probe reionization is through the 21-cm spectral line – produced by a spin flip in neutral hydrogen (van de Hulst 1945; Ewen & Purcell 1951; Muller & Oort 1951). This 21-cm radiation can be observed interferometrically at radio wavelengths as a deviation from the brightness temperature of the cosmic microwave background (CMB) (Field 1958, 1959; Madau, Meiksin & Rees 1997; Shaver et al. 1999).

Observationally, the 21-cm signal will be accompanied by system noise and Galactic and extragalactic foregrounds (e.g. Jelić et al. 2008, 2010) which are orders of magnitude larger than the 21-cm signal we wish to detect. On top of this there are systematic effects due to the ionosphere and instrument response. The foregrounds must be carefully removed using a cleaning process of high accuracy and precision as any error at this stage has the ability to destroy the underlying 21-cm signal. Foreground removal and the implications for 21-cm cosmology have been extensively researched over the past decade (e.g. Di Matteo et al. 2002; Oh & Mack 2003; Di Matteo, Ciardi & Miniati 2004; Morales & Hewitt 2004; Zaldarriaga, Furlanetto & Hernquist 2004; Santos, Cooray & Knox 2005; Bowman, Morales & Hewitt 2006; McQuinn et al. 2006; Wang et al. 2006; Gleser, Nusser & Benson 2008; Jelić et al. 2008; Harker et al. 2009b, 2010; Liu, Tegmark & Zaldarriaga 2009a; Liu et al. 2009b; Liu & Tegmark 2011, 2012; Mao 2012; Petrovic & Oh 2011). This paper constitutes only one step in the foreground removal process and assumes that bright sources have been removed, for example via a flux cut (Di Matteo et al. 2004).

There is currently no consensus on the most effective foreground removal method, though a recent method implemented by Harker et al. (2009b, 2010) has shown promising results while making only minimal assumptions. The same method highlighted that

⋆E-mail: eow@star.ucl.ac.uk
[1] http://www.lofar.org/
[2] http://gmrt.ncra.tifr.res.in/
[3] http://www.mwatelescope.org/
[4] http://astro.berkeley.edu/dbacker/eor/
[5] http://21cma.bao.ac.cn/
[6] http://www.skatelescope.org/

foreground removal techniques can be carried out both in image and visibility space – with the quality of results sometimes dependent on the choice of space. It is possible that different methods will be best suited for the extraction of different information from the data and improvements in the recovery of any one statistic, even at the expense of another, will still be useful. The statistical detection of the EoR signal is fraught with uncertainty and applying several foreground cleaning methods to the data independently will be invaluable in confirming a statistical detection.

The method presented here is based on the independent component analysis (ICA) algorithm, FASTICA (Hyvärinen, Karhunen & Oja 2001). ICA is a method originally designed to separate independent signals with minimal prior knowledge of the form of the signals. Thus ICA provides us with a foreground removal method which compensates for the fact that we do not know the form of the foregrounds at the exact resolution and frequency range of LOFAR. A non-parametric method, FASTICA, allows the foregrounds to choose their own shape instead of assuming a specific form, such as a polynomial. FASTICA is a versatile tool and has been applied recently in the field of exoplanets (Waldmann 2012) and CMB foreground removal with great success (e.g. Maino et al. 2002, 2003, 2007; Bottino, Banday & Maino 2008, 2010), motivating its implementation on other cosmological data. The results presented focus on the two main statistical aims of current EoR experiments, namely the recovery of the power spectrum and variance of the cosmological signal.

Section 2 briefly describes the FASTICA methodology and algorithm used to identify independent components (ICs) . The various methods used to simulate the 21-cm signal, foregrounds and noise are set out in Section 3. The results and sensitivity of the FASTICA method are presented in Sections 4 and 5 before a final summary and discussion in Section 6.

2 FOREGROUND REMOVAL TECHNIQUES

The statistical detection of the 21-cm reionization signal depends on an accurate and robust method for removing the foregrounds from the total signal. Since it is impossible to observe the foregrounds alone this is not a simple task.

The first attempts focused on exploiting the angular fluctuations of the 21-cm signal (e.g. Di Matteo et al. 2002, 2004; Oh & Mack 2003), but the 21-cm signal was found to be swamped by various foregrounds. The focus then moved on to the frequency correlation of the foregrounds, with the cross-correlation of pairs of maps used as a cleaning step (Zaldarriaga et al. 2004; Santos et al. 2005). While the foregrounds are expected to be highly correlated on scales of 1 MHz (e.g. Di Matteo et al. 2002; Gnedin & Shaver 2004), the cosmological signal is expected to be highly uncorrelated (Ali, Bharadwaj & Chengalur 2008) on the same frequency scales, allowing frequency correlation to differentiate the signals.

As more methods have emerged, it has become clear that different methods have different advantages and foreground subtraction has become accepted as a three-stage process. The first and last stages are bright source removal (Cooray & Furlanetto 2004; Di Matteo et al. 2004; Zaldarriaga et al. 2004; Morales, Bowman & Hewitt 2006) and residual error subtraction, respectively (Morales & Hewitt 2004), and are not dealt with in this paper. For our data we assume that the first stage has been carried out and all bright sources have been removed, which will still leave foregrounds strong enough to swamp the 21-cm signal (Di Matteo et al. 2002; Oh & Mack 2003). The second stage is known as spectral fitting, or line-of-sight fitting, and has become by far the most popular in the literature. Line-of-sight methods can be divided into subcategories of parametric and non-parametric methods. The majority of literature involves parametric methods, whereby at some point a certain form for the foregrounds is assumed, for example polynomials (e.g. Santos et al. 2005; Bowman et al. 2006; McQuinn et al. 2006; Wang et al. 2006; Gleser et al. 2008; Jelić et al. 2008; Liu et al. 2009a,b; Petrovic & Oh 2011). In contrast, non-parametric methods allow the data to determine the form of the foregrounds with many more free parameters, allowing much more freedom and not assuming a specific form. This has obvious advantages for a cosmological era so far not directly observed, but the results are often not as promising as parametric results. A possible exception is the recent method presented by Harker et al. (2009b, 2010), which preferentially considered foreground models with as few inflection points as possible. When applied to LOFAR-EoR data, this method compared very favourably with parametric methods. FASTICA is another non-parametric method which, as we will show, produces similarly promising results.

2.1 The FASTICA method

2.1.1 Background

Introduced in the early 1980s, ICA has established itself as a successful component separation technique with widespread applications. The method relies on the assumption that the multiple elements making up a mixed signal are statistically independent.

ICA methods often formulate the data model as

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

where $\boldsymbol{x}$ is a vector representing the observed signal, $\boldsymbol{s}$ is a vector the components of which are assumed mutually independent and $\mathbf{A}$ is a mixing matrix to be calculated. For our data we have signal maps of 512×512 pixels at 170 different frequencies. Equation (1) represents one line of sight where, if m ICs are assumed, the sizes of $\boldsymbol{x}$, $\mathbf{A}$ and $\boldsymbol{s}$ are [170,1], [170,m] and [m,1], respectively. Actually, FASTICA simultaneously considers all lines of sight, so $\boldsymbol{x}$ and $\boldsymbol{s}$ are in effect matrices of sizes [170, 512×512] and [m, 512×512], respectively. For clarity, we will set out the description as if only one line of sight was being considered but the reader should bear in mind that all lines of sight are simultaneously and independently treated by the algorithm, with $\mathbf{A}$ being independent of the line of sight.

It might immediately strike the reader that the model specified here is the noise-free ICA model. This is because this implementation makes no effort to model the noise component through the $\boldsymbol{x} = \mathbf{A}\boldsymbol{s} + \boldsymbol{n}$ formulation. Instead, one must appreciate that it is the way in which FASTICA is not robust to noisy components that we take advantage of here. Whereas $\boldsymbol{x}$ will represent the observed signal of foreground, noise and 21-cm signal, $\boldsymbol{s}$ is considered to be the foregrounds only. FASTICA ignores the Gaussian or non-smooth spectral components in the observed signal. When we specify m ICs, FASTICA reconstructs m ICs related to the foregrounds only.

To solve equation (1) for $\boldsymbol{s}$, we seek a linear transform:

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

where $\mathbf{W}$ is a constant weight matrix that the ICA method aims to determine assuming that the elements of $\boldsymbol{s}$ are as statistically independent as possible.

FASTICA seeks to estimate $\mathbf{W}$ using the concept of mutual information. The general philosophy behind FASTICA is outlined below, but for a full treatment refer to Hyvärinen (1999) and Hyvärinen et al. (2001).

Let us consider a single component of the signal $\boldsymbol{s}$:

$$y = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} = \sum_i w_i x_i, \tag{3}$$

where if $\boldsymbol{w}$ is one of the rows of the inverse of $\mathbf{A}$, then y is actually one of the ICs, s_i. ICA then attempts to minimize the Gaussianity of $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$. To understand why, we define a vector $\boldsymbol{z}$:

$$\boldsymbol{z} = \mathbf{A}^{\mathrm{T}}\boldsymbol{w}, \tag{4}$$

so that we have a weighted sum of the independent signal components:

$$y = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{s}. \tag{5}$$

The central limit theorem states that the greater the number of independent variables in a distribution, the more Gaussian that distribution will be. $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{s}$ is therefore always more Gaussian than any individual s_i. y will be least Gaussian when one, and only one, z_i is non-zero, and in such a case y is then one of the ICs. Thus by maximizing the non-Gaussianity of $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ we find one of the ICs.

In order to estimate and adjust $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ in such a way that its Gaussianity converges to a minimum, the methods need a robust measure of non-Gaussianity. FASTICA favours negentropy as a measure of non-Gaussianity, which is based on the idea of the entropy of a variable, $H(y)$:

$$H(y) = -\sum_i P(y = a_i)\log P(y = a_i), \tag{6}$$

where a_i are the possible values of y.

Negentropy is then defined as

$$J(y) = H(y_{\mathrm{gauss}}) - H(y), \tag{7}$$

where y_{gauss} is a random Gaussian variable with the same covariance matrix as y. Using the maximum entropy principle, one can define

$$J(y) \approx \sum_{i=1}^{n} k_i[E\{G_i(y)\} - E\{G_i(v)\}]^2, \tag{8}$$

where k_i are positive constants, v is a Gaussian variable with zero mean and unit variance, and G are non-quadratic functions. Though almost any non-quadratic function can be used, the robustness and speed of the FASTICA method depends on choosing these contrast functions well, with different contrast functions more suited to different scenarios. For this implementation we choose a non-linearity, $g(u)$, of

$$g(u) = u \times \exp\left(-\frac{u^2}{2}\right), \tag{9}$$

where $g(u) = G'(u) = \frac{\mathrm{d}G(u)}{\mathrm{d}u}$. This choice is particularly suited when robustness is important or when the components have high kurtosis.

Since $\boldsymbol{s}$ and $\mathbf{A}$ are both unknown, FASTICA cannot determine the ICs' magnitudes or order, as we can freely change the order of the components in the mixing model or multiply any of them by a scalar factor which can be balanced out by dividing out elsewhere. As such FASTICA fixes the magnitudes of the ICs by assuming that they have unit variance.

2.1.2 Algorithm

Here we summarize the fixed-point FASTICA algorithm for finding one IC.

The mixed signal is input along with a parameter representing the number of ICs we assume there to be. A typical choice in this implementation is four ICs.

These data undergo several pre-processing steps within the FASTICA program. First, the data are adjusted to be of zero mean to simplify the algorithm. Then, using a principal component analysis to estimate the eigenvalues and eigenvectors for the data, the data are whitened. This results in the vector $\boldsymbol{x}$ where the components are uncorrelated, with unit variances.

We wish to choose a unit vector $\boldsymbol{w}$ such that the non-Gaussianity of $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ is maximized. Under the assumption that the components have unit variance (which for whitened data is equivalent to assuming $||\boldsymbol{w}||^2 = 1$) these maxima occur where

$$E\{\boldsymbol{x}\,g(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x})\} - \beta\boldsymbol{w} = 0 \tag{10}$$

is satisfied. To find the roots of this equation using Newton's method we arrive at the approximate Newton iteration:

$$\boldsymbol{w}^{+} = \boldsymbol{w} - \frac{[E\{\boldsymbol{x}\,g(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x})\} - \beta\boldsymbol{w}]}{E\{g'(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x})\} - \beta}. \tag{11}$$

This iteration is carried out using the algorithm summarized in the following steps (Hyvärinen & Oja 2000).

(i) Choose an initial random weight vector $\boldsymbol{w}$.
(ii) Let $\boldsymbol{w}^{+} = E\{\boldsymbol{x}\,g(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x})\} - E\{g'(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x})\}\boldsymbol{w}$.
(iii) Let $\boldsymbol{w} = \frac{\boldsymbol{w}^{+}}{\|\boldsymbol{w}^{+}\|}$.
(iv) If the old and new values for $\boldsymbol{w}$ are not converged, then repeat the process.

Here g is the derivative and g' is the second derivative of the chosen contrast function G. The use of the contrast function derivatives comes from consideration of where the maxima of the negentropy approximation are obtained. The non-Gaussianity is maximized along the line of sight and across the map simultaneously, meaning that the method's constraining power benefits from having more pixels and more frequency maps.

To extend the algorithm to n components requires FASTICA to run simultaneously for n different weight vectors, $\boldsymbol{w}_1, ..., \boldsymbol{w}_n$, where one $\boldsymbol{w}_j$ corresponds to $\boldsymbol{w}$ in the above algorithm. To ensure that the different $\boldsymbol{w}_j\boldsymbol{x}$ converge to different maxima (i.e. the same IC is not found twice) all the outputs $\boldsymbol{w}_j\boldsymbol{x}$ must be decorrelated after every iteration. For a more detailed treatment refer to Hyvärinen (1999) and Hyvärinen et al. (2001).

2.1.3 Our implementation

We make use of the C++ implementation of FASTICA provided by the IT++ library.[7] Our foreground subtraction proceeds in the following steps.

(i) Read in the simulation data cube and specify the number of foreground ICs for FASTICA to model.
(ii) Call FASTICA to calculate the mixing matrix and ICs of the foregrounds.
(iii) Reconstruct the foregrounds by performing a multiplication of the mixing matrix, which is common to all lines of sight, with the vector of ICs for each line of sight.
(iv) Find the difference between the reconstructed foreground cube and the input cube. This residual cube should equal the 21-cm signal, noise and any foreground fitting errors.

Statistical tests can then be carried out on the residuals cube to determine if the 21-cm signal is recoverable after the foreground removal process.

[7] http://itpp.sourceforge.net/devel/fastica_8cpp.html

It is worth reiterating once more as it is such an important point: the ICs referred to in the ICA methodology as applied here are the ICs making up the foregrounds – the cosmological signal and noise are at no point modelled or even taken into account by this FASTICA implementation.

3 SIMULATED EOR DATA

We simulate 170 frequency maps between 115 and 200 MHz with spacings of 0.5 MHz. The maps consist of 512^2 pixels representing a $10^\circ \times 10^\circ$ observation window, or a resolution of 1.17 arcmin per pixel. Since an interferometer like LOFAR is insensitive to the mean value of the brightness temperature, we use mean-subtracted maps.

3.1 21-cm cosmological signal

Of the existing reionization simulation programs (e.g. Santos et al. 2010), we use the seminumeric modelling tool 21CMFAST to simulate the observable of the 21-cm radiation, the brightness temperature T_b (Mesinger & Furlanetto 2007; Mesinger, Furlanetto & Cen 2011). 21CMFAST treats physical processes with approximate methods for small realization generation times and has produced results which agree well with the most recent hydrodynamical simulations. The code was run using a standard cosmology, $(\Omega_\Lambda, \Omega_\mathrm{m}, \Omega_\mathrm{b}, n, \sigma_8, h) = (0.72,0.28,0.046,0.96,0.82,73)$ (Komatsu et al. 2011), and initialized at $z = 300$ on a 1800^3 grid. The velocity fields used to perturb the initial conditions as well as the resulting 21-cm T_b boxes were formed on a 450^3 cruder grid before being interpolated up to 512^3. We define haloes contributing ionizing photons as having a minimum virial mass of $1 \times 10^9\,\mathrm{M}_\odot$.

When a hydrogen atom undergoes a ground-state hyperfine transition from the excited triplet state, where the spins of the proton and electron are parallel, to the singlet state, where the spins are antiparallel, a photon is emitted of wavelength 21 cm or frequency $\nu_{10} = 1420$ MHz. The 21-cm spectral line is forbidden – the probability of a 21-cm transition is $2.9 \times 10^{-15}\,\mathrm{s}^{-1}$, equivalent to the triplet state lifetime of 10^7 yr (Wild 1952). Even so, the vast amount of hydrogen in the Universe leads to 21-cm observations being achievable (van de Hulst 1945). The intensity of 21-cm radiation is determined by the spin temperature, T_spin, defined through (Field 1958)

$$\frac{n_1}{n_0} = 3 \exp\left(\frac{-T^*}{T_\mathrm{spin}}\right). \tag{12}$$

The spin temperature is a fundamental measure of the number densities of the triplet and singlet states (n_1, n_0) where $T^* = \frac{h\nu_{10}}{k_\mathrm{b}} = 0.0681$ K.

T_b is detected differentially as a deviation from T_CMB and δT_b (Field 1958, 1959; Ciardi & Madau 2003):

$$\delta T_\mathrm{b} = 28\,\mathrm{mK} \times (1+\delta) x_\mathrm{HI} \left(1 - \frac{T_\mathrm{CMB}}{T_\mathrm{spin}}\right) \left(\frac{\Omega_\mathrm{b} h^2}{0.0223}\right) \times \sqrt{\left(\frac{1+z}{10}\right)\left(\frac{0.24}{\Omega_\mathrm{m}}\right)}, \tag{13}$$

where δ is the mass density contrast, h is the Hubble constant in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, x_HI is the fraction of neutral hydrogen, and Ω_b and Ω_m are the baryon and matter densities in critical density units.

3.2 Foregrounds

Though there have been foreground observations at frequencies relevant to LOFAR using the Westerbork Synthesis Radio Telescope (Bernardi et al. 2009, 2010), the foreground contamination at the frequencies and resolution of LOFAR remains poorly constrained. As a result, foreground models directly relevant for this paper rely on using constraints from observations at different frequency and resolution ranges. These constraints are used to normalize the necessary extrapolations made from observations to create a model relevant for LOFAR-EoR observations.

In general, the foreground components are modelled as power laws in 3+1 dimensions (i.e. three spatial and one frequency) such that $T_\mathrm{b} \propto \nu^\beta$ (e.g. Shaver et al. 1999; Ali et al. 2008; Jelić et al. 2008, 2010).

The foreground simulations used in this paper are obtained using the foreground models described in Jelić et al. (2008, 2010). The foreground contributions considered in these simulations are as follows.

(i) Galactic diffuse synchrotron emission (GDSE) originating from the interaction of free electrons with the Galactic magnetic field. It incorporates both the spatial and frequency variation of β by simulating in 3 spatial and 1 frequency dimension before integrating over the z-coordinate to get a series of frequency maps. Each line of sight has a slightly different power law.

(ii) Galactic localized synchrotron emission originating from supernova remnants (SNRs). Together with the GDSE, this emission makes up 70 per cent of the total foreground contamination. Two SNRs were randomly placed as discs per 5° observing window, with properties such as power-law index chosen randomly from the Green (2006) catalogue.

(iii) Galactic diffuse free–free emission due to bremsstrahlung radiation in diffuse ionized Galactic gas. This emission contributes only 1 per cent of total foreground contamination; however, it still dominates the 21-cm signal. The same method as used for the GDSE is used to obtain maps; however, the value of β is fixed to -2.15 across the map.

(iv) Extragalactic foregrounds consisting of contributions from radio galaxies and radio clusters and contributing 27 per cent of the total foreground contamination. The simulated radio galaxies assume a power law and are clustered using a random walk algorithm. The radio clusters have steep power spectra and are based on a cluster catalogue from the Virgo Consortium[8] and observed mass–luminosity and X-ray–radio luminosity relations.

Unlike Jelić et al. (2008, 2010), this paper does not consider the polarization of the foregrounds. The foregrounds simulated here are up to five orders of magnitude larger than the signal we hope to detect. Since interferometers such as LOFAR measure only fluctuations, foreground fluctuations dominate by 'only' three orders of magnitude. We will investigate polarized removal in a further analysis.

3.3 Noise

The sensitivity of a receiving system is ultimately determined by the system noise (Thompson, Moran & Swenson Jr 2001). The system noise consists of contributions from both the sky and the receivers themselves. Frequency dependence is introduced into

[8] http://www.mpa-garching.mpg.de/galform/virgo/hubble/

Table 1. LOFAR and simulation parameters.

Parameter	Description	Value
n_d	Number of dipoles per tile	16
n_t	Number of tiles per station	24
n_s	Number of stations	48
η_a	Antenna efficiency	1
η_s	System efficiency	0.9
$\Delta\nu$	Frequency interval (MHz)	0.5
t_{int}	Integration time (h)	600
Ω_{arcmin}	Area of synthesized beam (arcmin^2)	4
Ω_{sr}	Area of synthesized beam (sr)	1.35×10^{-6}

the noise through the sky noise and through the frequency dependence of the effective area of the telescope. It is assumed that the noise across an image is Gaussian at any one frequency. For a detailed look at the expected noise on LOFAR measurements, see Labropoulos et al. (2009).

We have decided to reproduce the method for calculating the noise here as the noise can have a significant effect on foreground cleaning methods.

Our parameters for calculating the noise are listed in Table 1. In order to create a noise curve, Fig. 1, we use the following prescription: the system noise temperature consists of sky brightness and instrumental components. We calculate this system temperature using

$$T_{sys} = 140 + 60\left(\frac{\nu}{300\ \mathrm{MHz}}\right)^{-2.55}. \quad (14)$$

The effective area of the array is determined by multiplying the effective area of a single dipole by the number of dipoles in the array, where the effective area of a dipole is limited by the size of the tile. We calculate the effective area of the LOFAR array using

$$A_{eff} = \min\left(\frac{\lambda^2}{3}, 1.5625\right) n_d n_t. \quad (15)$$

The system equivalent flux density (SEFD) then depends on both of the quantities calculated above:

$$\mathrm{SEFD} = \frac{2T_{sys}k_b}{\eta_a A_{eff}}. \quad (16)$$

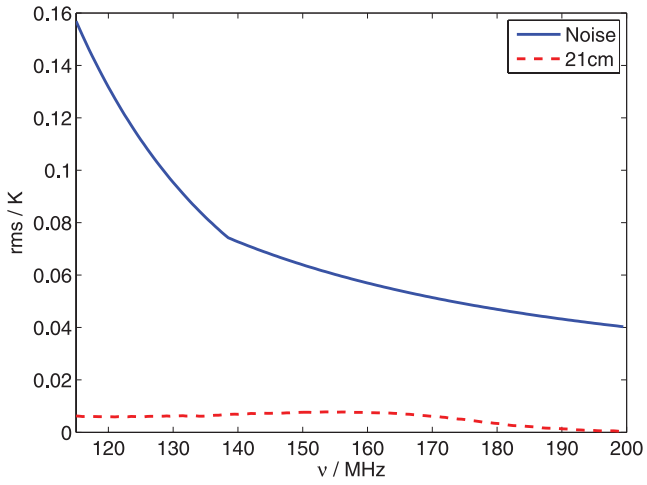

Figure 1. The rms of the simulated zero mean noise (blue; solid) and 21-cm (red; dashed) maps over frequency.

Finally we calculate the LOFAR noise sensitivity

$$\sigma = \frac{1}{\eta_s}\frac{\mathrm{SEFD}}{\sqrt{n_s(n_s-1)\Delta\nu t_{int}\Omega_{sr}}}. \quad (17)$$

Fig. 1 shows the noise curve calculated from this prescription compared to the rms of our 21-cm simulation.

For each frequency, a LOFAR measurement set was filled with Gaussian noise in the *uv* plane. This was then imaged to create a real-space image, the rms of which can be normalized to the value as given by the prescription described above. For example the noise sensitivity at 150 MHz for an integration time of 600 h and a frequency spacing of 0.5 MHz was 64 mK. The 170 noise maps were uncorrelated over frequency – i.e. a different noise realization was used to fill the measurement set for each frequency.

3.4 Dirty images

The success of an interferometer such as LOFAR is highly dependent on how *uv* space is sampled. The particular pattern of *uv* sampling forms a beam which affects how the components such as the foregrounds are seen by the interferometer. Dirty images were simulated by convolving with the point spread function (PSF) of the LOFAR setup used to simulate the noise in the previous section, Figs 2 and 3.

The PSF used for creating dirty images (and for creating the noise as described in the previous section) was chosen to be the worst in the observation bandwidth – i.e. the PSF at 115 MHz. In observations the synthesized beam decreases in size with increasing frequency, causing point-source signals to oscillate with the beam, producing a foreground signal with an oscillatory signal very much like that of the 21-cm signal. However, this mode-mixing

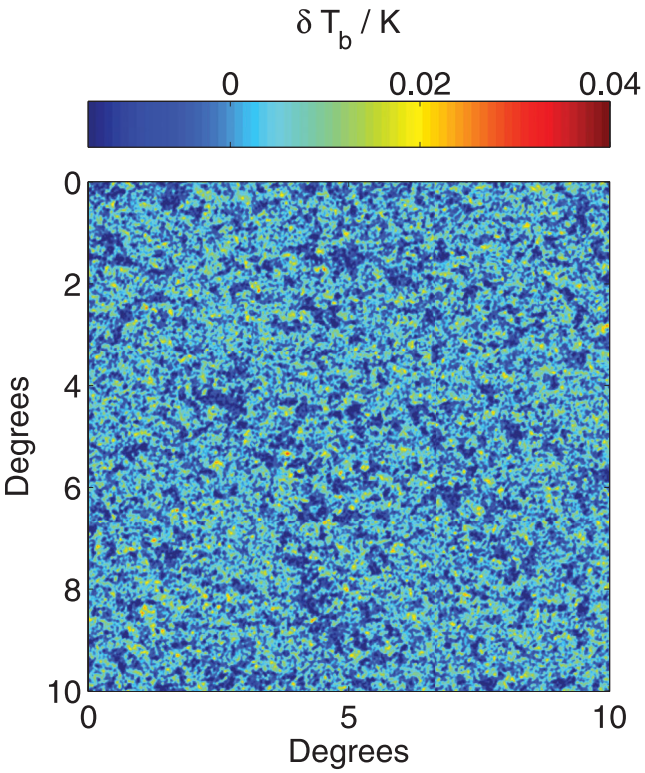

Figure 2. The 21-cm signal at 150 MHz convolved with the PSF. The signal is entirely in emission – this map has been adjusted to have a mean of zero to reflect the observations of an interferometer.

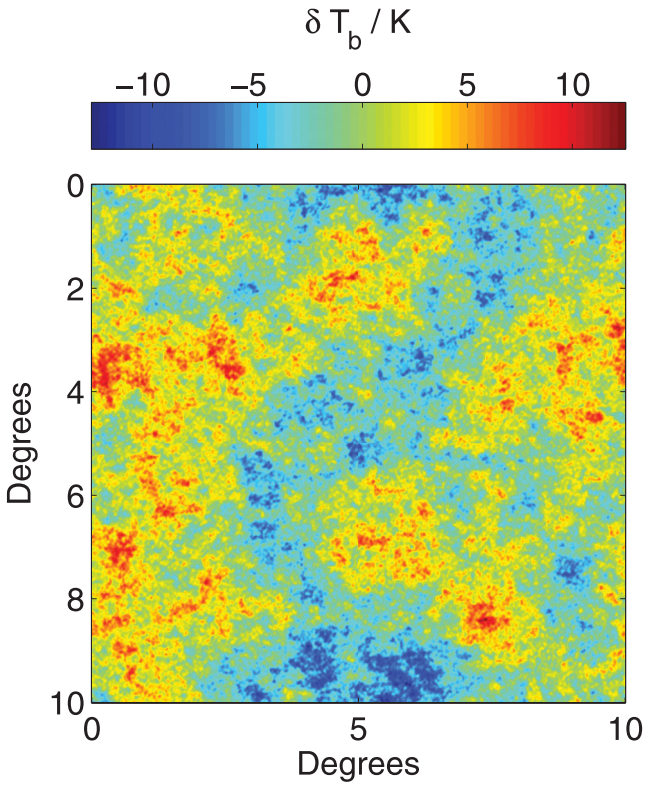

Figure 3. The total contribution of the simulated foregrounds at 150 MHz convolved with the PSF.

contribution has been found not to threaten the 21-cm recovery and has a power well below the 21-cm level (Bowman et al. 2006; Liu et al. 2009a). As such we leave the consideration of a frequency-dependent PSF to a future paper.

Once the foregrounds and 21-cm signal have been adjusted for *uv* sampling, the three component cubes are added together. The components of the total δT_b along a random line of sight are shown in Fig. 4.

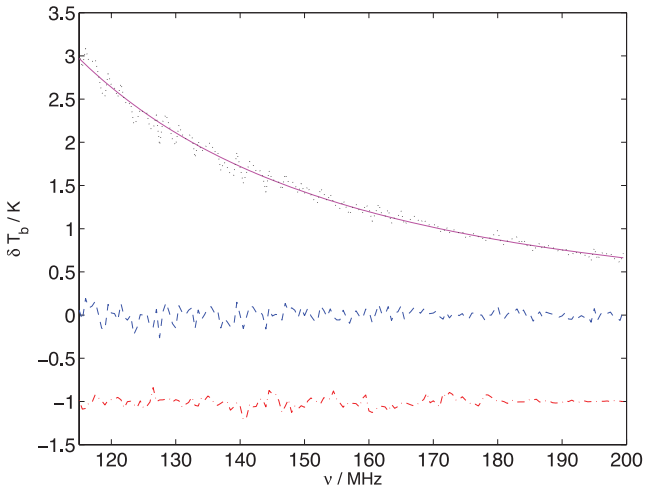

Figure 4. The redshift evolution of the simulated cosmological signal (red; dash–dotted), foregrounds (pink; solid), noise (blue; dashed) and total combined signal (black; dotted). All components have undergone the PSF convolution. Note that the 21-cm signal has been amplified by 10 and displaced by −1 K for clarity.

3.5 Fourier transformed data

The FASTICA method was implemented separately with data both in real and Fourier space. For the latter method, the fiducial image cube was 2D Fourier transformed at each frequency to create a Fourier data cube. The complete cube was then processed with FASTICA and the output reverse Fourier transformed to obtain the ICs in real space. Unless otherwise stated all results refer to real-space implementation.

4 RESULTS

4.1 The independent components

The top panel of Fig. 5 shows the four ICs found by FASTICA for a clean data cube. These ICs are the columns of the mixing matrix, **A**. For comparison we show the line of sight δT_b of the simulated foreground contributions in the bottom panel of Fig. 5.

We can see that no single component corresponds to any one single foreground contribution, even when processing a clean data cube. Instead, the components are all a mixture. While in ICs 2 and 4 the presence of Galactic synchrotron is obvious, in the other components the combination of the contributions is not so clear. It is also worth noting that while IC4 shows a significant contribution from Galactic synchrotron, it is inverted. FASTICA can only determine the ICs up to a multiplicative constant and so the sign and magnitude of the components are irrelevant.

The coefficients of the ICs are stored in the matrix **s** and are presented in Figs 6–9. We can compare these coefficients to the maps of the foreground contributions, Figs 10–12. We see that all four coefficients are a mixture of the contributions as expected.

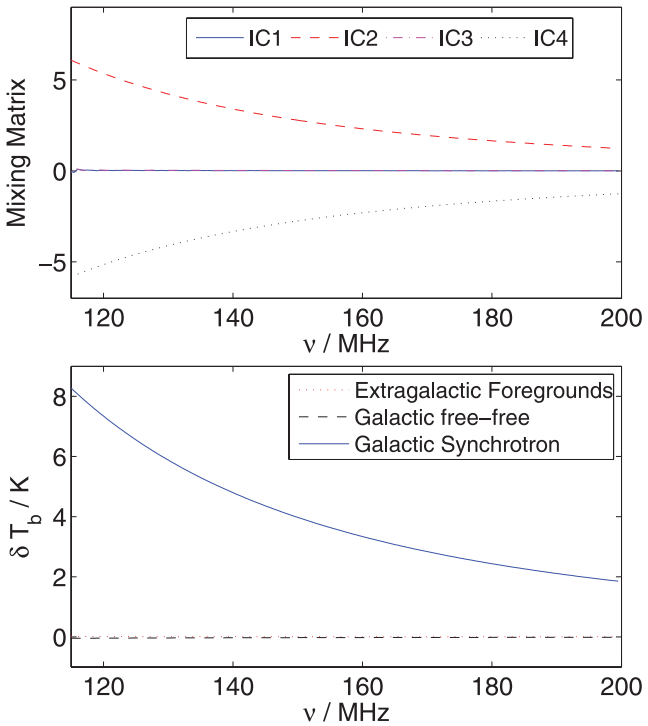

Figure 5. In the top panel we show the four columns of the mixing matrix representing the four ICs. The brightness temperatures of the foreground contributions along a random line of sight are shown in the bottom panel. We see that the ICs are each a scaled mixture of the foreground contributions.

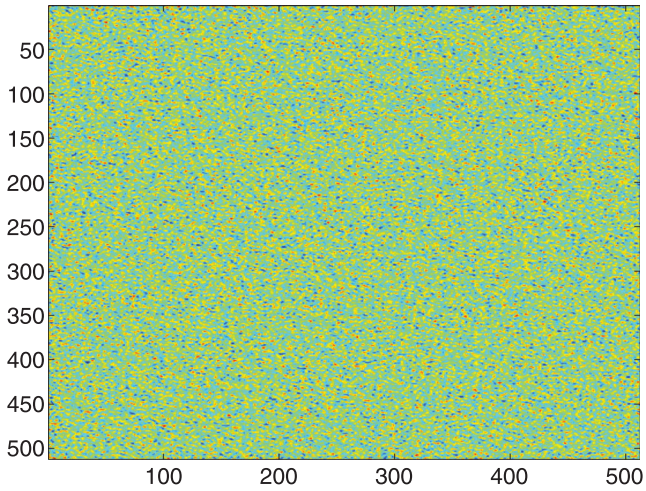

Figure 6. The first coefficient map of the ICs when FASTICA processes the clean data cube.

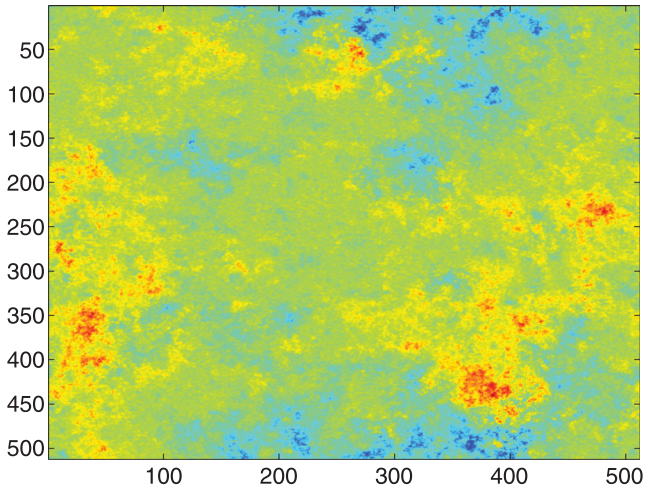

Figure 7. The second coefficient map of the ICs when FASTICA processes the clean data cube.

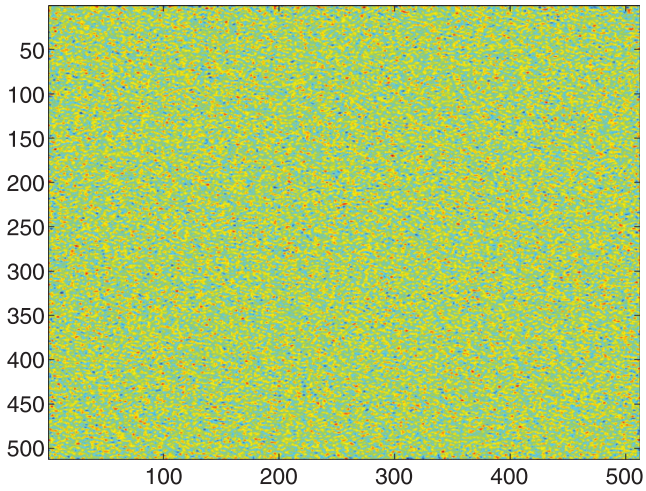

Figure 8. The third coefficient map of the ICs when FASTICA processes the clean data cube.

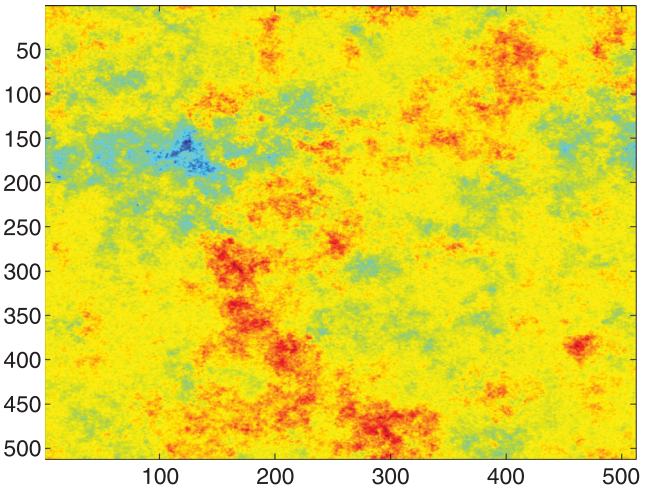

Figure 9. The fourth coefficient map of the ICs when FASTICA processes the clean data cube.

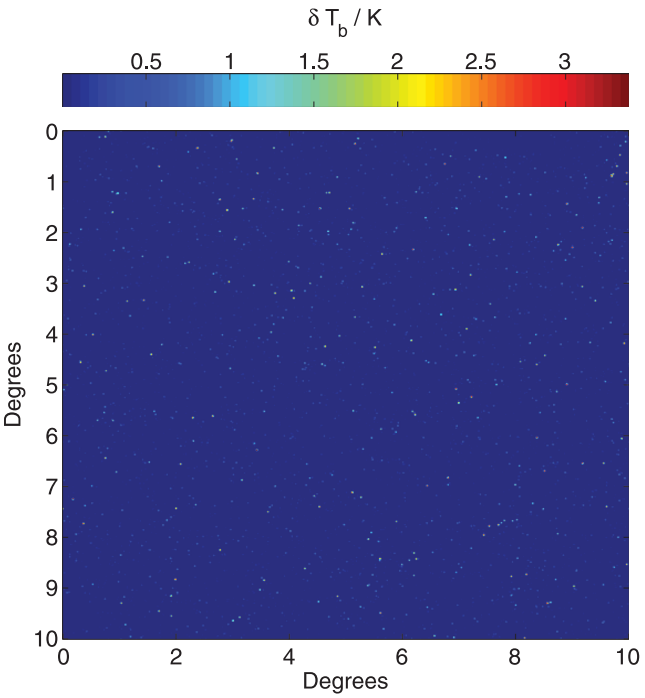

Figure 10. The simulated extragalactic foregrounds at 150 MHz.

4.2 Fitting errors and variance

We will first discuss the FASTICA results on the simulation where the data cube has been convolved with the PSF; then the data processing is carried out in real space and four ICs are assumed. The word 'simulated' is used to refer to the input maps and 'reconstructed' is used for the estimates resulting from FASTICA. The total input signal is separated into reconstructed foregrounds and residuals. The residuals are the difference between the total input signal (corrected in the online version, 2012 May 25) and reconstructed foregrounds.

To evaluate the accuracy of the foreground fitting by FASTICA, we calculated the foreground fitting error, equation (18), for each pixel:

$$\text{fitting error} = \frac{\text{fg}_{\text{reconstructed}} - \text{fg}_{\text{simulated}}}{\text{fg}_{\text{simulated}}} \times 100.0. \quad (18)$$

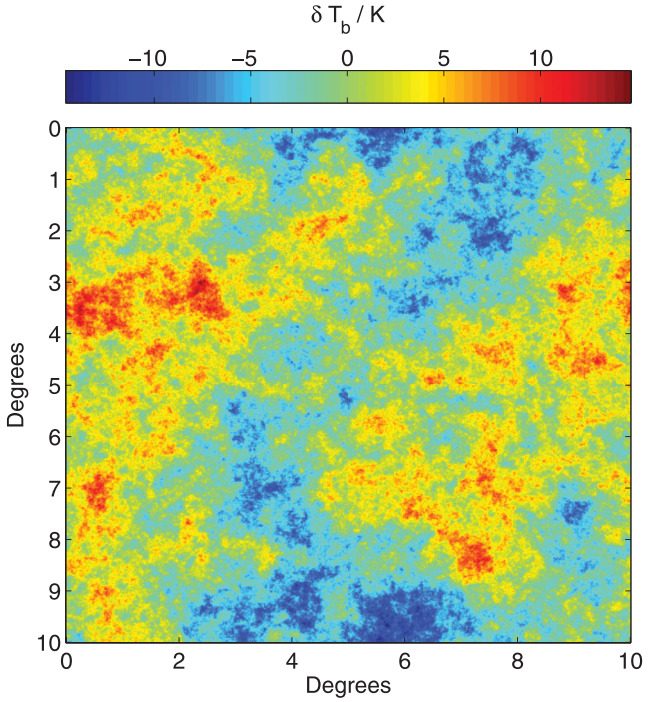

Figure 11. The simulated Galactic synchrotron foregrounds at 150 MHz.

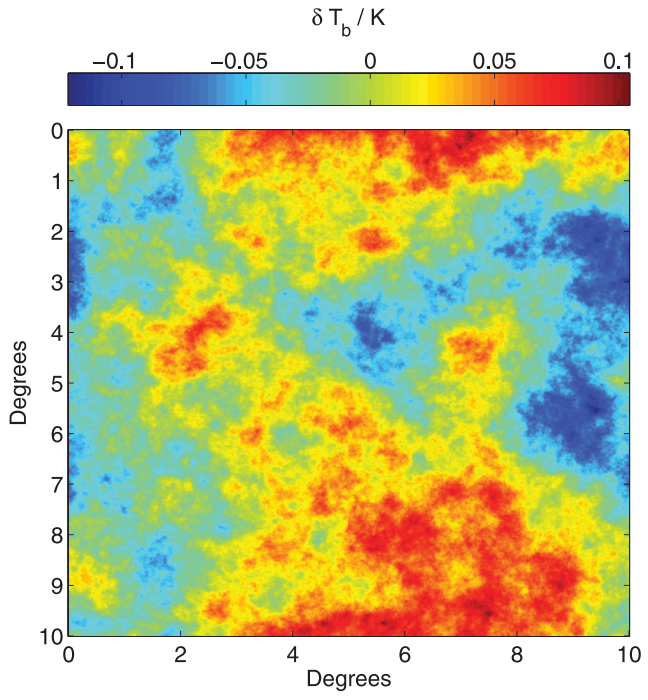

Figure 12. The simulated Galactic free–free foregrounds at 150 MHz.

In Fig. 13 we plot the Pearson correlation coefficient between the foreground fitting errors and foregrounds (top) and between the foreground fitting errors and the noise (bottom). The Pearson correlation coefficient between two data sets a and b is defined as

$$r = \frac{\sum_i (a_i - \bar{a})(b_i - \bar{b})}{\left[\sum_i (a_i - \bar{a})^2 \sum_i (b_i - \bar{b})^2\right]^{\frac{1}{2}}}, \tag{19}$$

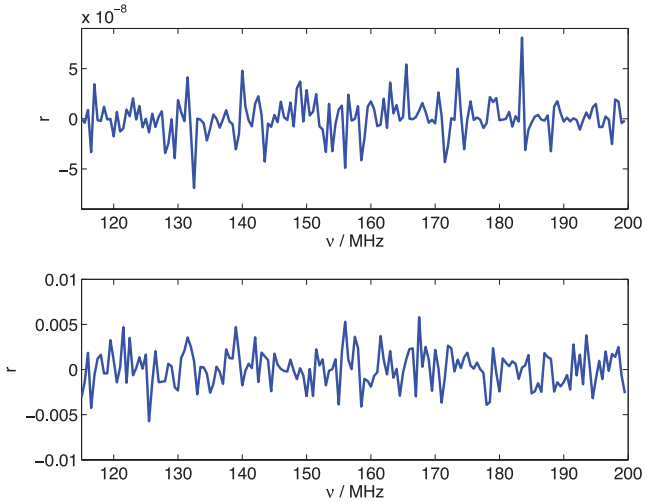

Figure 13. (a) The Pearson correlation coefficient between the foreground maps and foreground fitting errors. (b) The Pearson correlation coefficient between the noise maps and foreground fitting errors.

where $\bar{a}$ is the mean of the data set a_i, $\bar{b}$ is the mean of the data set b_i and the measure is normalized such that $r = \pm 1$ for correlation/anticorrelation.

We see that there is very little correlation between the foreground maps and the foreground fitting errors, with around six magnitudes more correlation between the noise maps and the foreground fitting errors.

To get a representation of the foreground fitting error over an entire map, the rms error of the fitted foregrounds was calculated, Fig. 17. It should be noted that this error takes into account all scales – including those with a disproportionate error as will be seen in the power spectra. The rms difference between the simulated and reconstructed foregrounds was calculated over all 512^2 lines of sight for each frequency. Also, an rms error for each map was calculated using only 68 per cent of the pixels – with the pixels of lowest error selected first. When the outlier pixels are discounted we find that the rms error is below 10 mK for the majority of the frequency range. This is still high enough to be of concern as the 21-cm signal is itself of the order of tens of mK; however, the inclusion of all scales means that this is a worst case scenario.

For a statistical detection of the EoR, LOFAR aims to detect a non-zero variance after the noise and foregrounds have been accounted for. We begin by combining the simulated noise and simulated 21-cm signal and taking the variance of this signal. This can then be compared to the variance of the FASTICA residuals – Fig. 14. The residual variance is recovered at all but the smallest frequencies. At frequencies below 120 MHz (or $z > 10.8$) the variance is significantly underestimated, probably as a result of foreground overfitting – the leakage of noise power into the estimated foreground power. This failure at very low frequencies is hardly surprising considering that this is where the noise and foregrounds are at their strongest.

We subtract the variance of the simulated noise directly from the variance of these residuals: var(reconstructed 21-cm) = var(residuals) − var(noise). This is a fair assumption as we should be able to look at the data in narrow frequency bins and estimate the statistics (e.g. variance and power spectrum) of the noise to a very high accuracy.

We find that the recovered 21-cm variance, Fig. 15 (top-left), is not robust to small-scale power in the original signal. By removing the noise simulation maps manually from the residual maps in order

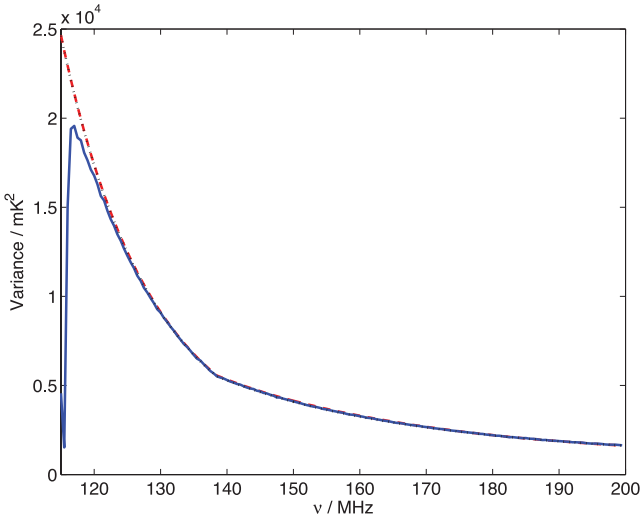

Figure 14. The variance across the combined simulated cosmological signal and noise (red, dashed), noise alone (black, dotted) and residuals (blue, solid).

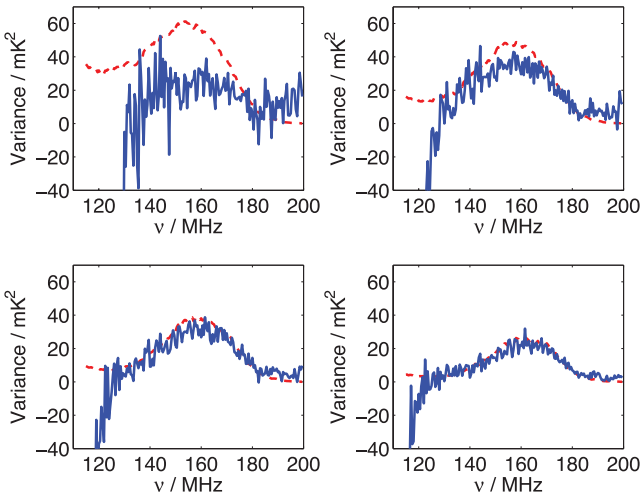

Figure 15. The variance across the simulated (red; dashed) and reconstructed 21-cm maps (blue; solid) for the fiducial data, and data which have had Fourier filtering of modes below 2, 3 and 5 PSF scales (in reading order).

to get crude maps of the recovered 21-cm signal, excess small-scale power is evident, Fig. 16. We note that we do this direct noise subtraction for a crude visual inspection only and not for any of the analytical results. The excess power is most likely due to FASTICA not being robust to the small-scale power (noise) in our data, allowing it to leak into the reconstructed foregrounds. It was found that by Fourier filtering the data to entirely remove k modes below a threshold corresponding to a multiple of the PSF scale, the variance recovery was significantly improved. A very good recovery occurs with filtering below 5 PSF scales (Fig. 15, bottom-right).

At the extremes of the frequency range the reconstructed variance increasingly diverges from that of the simulated 21-cm. Both the noise and the foregrounds are at their largest at lower frequencies meaning that both fitting errors and noise leakage are likely to be largest here, leading to less accurate 21-cm reconstruction. Equally, at the larger frequencies, the 21-cm signal is almost non-existent making an accurate reconstruction difficult when swamped with fitting errors and noise.

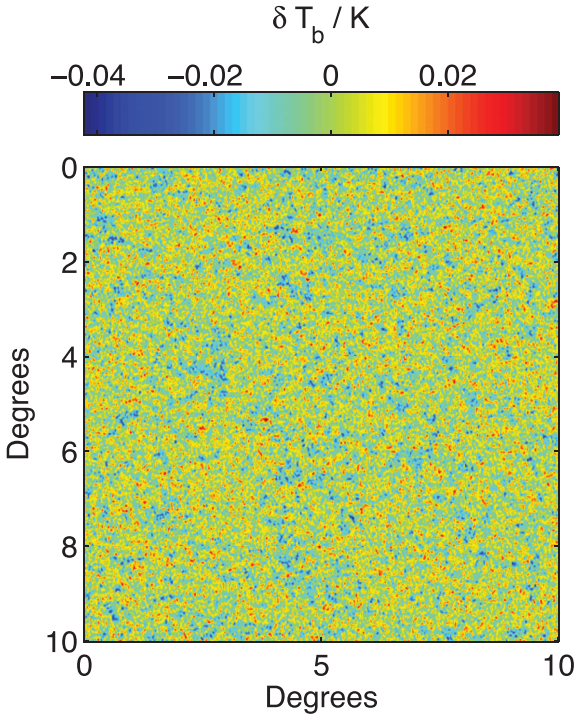

Figure 16. The reconstructed 21-cm signal at 150 MHz for dirty data. We see that while there is a strong correlation between the large-scale structure in this image and the original signal (Fig. 2), there is also a large amount of excess small-scale structure, probably due to noise leakage into the foregrounds.

This variance calculation was also carried out on a data cube where the residual and noise maps were smoothed from a 512^2 grid to a 256^2 grid before the same variance calculation was carried out above and compared to the variance of a smoothed simulated 21-cm map. The curves are, as expected, slightly smoother; however, the trend and conclusions remain the same.

4.2.1 Varying the number of ICs

The FASTICA algorithm requires specification of the number of ICs to be used in the reconstructed foreground model. Though we have modelled the various foreground contributions, it is not a trivial task to determine how these depend on each other and to what degree. To test the sensitivity of our results to the number of ICs chosen we calculate the rms error and variance recovery for IC numbers of 2, 4 and 6 in Figs 17 and 18.

We see that small variations in the number of ICs do not endanger the statistical recovery of the 21-cm signal. For the remainder of this paper, four ICs are assumed.

4.3 Power spectra

Together with the variance, EoR experiments aim to recover the power spectrum of the cosmological signal over a broad range of frequencies.

Different effects are important for modes parallel and perpendicular to the line of sight. For example, consider the scenario where the foregrounds have been under-fitted by a constant over the frequency range. This offset will not be evident in the 1D power spectrum of the residuals, however will be evident in the angular power spectrum if that constant is dependent on line of sight. Thus it has

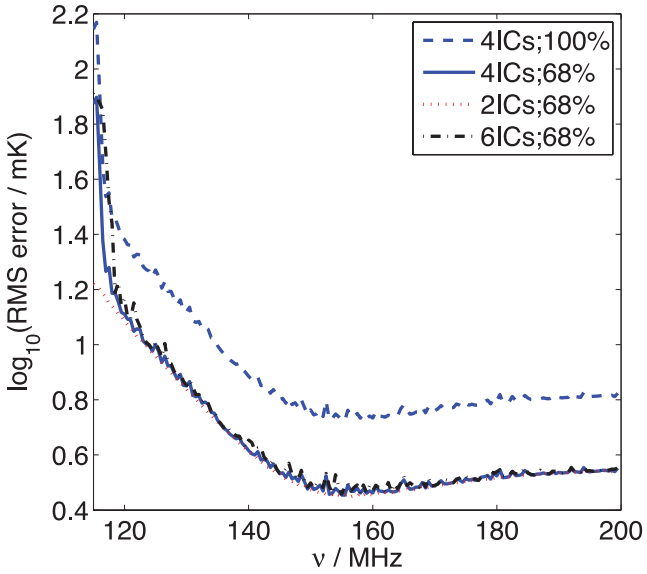

Figure 17. The rms error of the four IC reconstructed foregrounds for when all pixels are considered (blue; dashed) and when only the middle 68 per cent of the error distribution is included (blue; solid). Also, the rms errors of the reconstructed foregrounds for FASTICA applied according to models with 2 (red; dotted) and 6 (black; dash–dotted) ICs, with only the middle 68 per cent of the error distribution included.

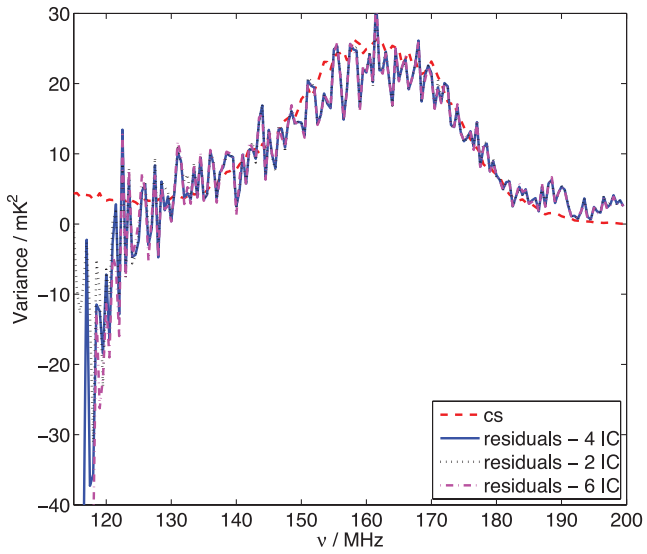

Figure 18. The variance across the simulated (red; dashed) and reconstructed maps at each frequency, for the FASTICA algorithm run with the assumption of 2 (black; dotted), 4 (blue; solid) and 6 (pink; dot–dashed) ICs. The data have been Fourier filtered at the 5 PSF scale.

been argued by Harker et al. (2010) that for LOFAR data, the separate calculation of 1D and 2D power spectra has its advantages. However, this does not consider modes either parallel or perpendicular to the line of sight and as such we calculate 2D and 3D power spectra. We note here that we performed only one simulation of the cosmological signal so the power spectrum error bars relate to this specific realization of the density field.

4.3.1 *Angular power spectra*

The angular power spectrum of a map at a single frequency is calculated by 2D Fourier transforming that map and binning the pixels according to Fourier scale, k. The power at any particular k, $\langle\delta(k)\delta^*(k)\rangle$, is the average power of all the uv cells in the bin centring on k. The error on the point for a particular bin, k_i, is calculated as $\alpha_i = \frac{\langle\delta(k_i)\delta^*(k_i)\rangle}{\sqrt{n_{k_i}}}$, where n_{k_i} is the number of uv cells that reside in that k bin. The power spectra are averaged over frequency bandwidths of 2.5 MHz and all frequencies quoted are the middle frequency of the bandwidth. The power spectrum of the reconstructed cosmological signal is calculated by subtraction of the noise power spectrum from the FASTICA residual power spectrum. The error on the simulated 21-cm power spectrum is added in quadrature with the error on the noise to reflect the error on the reconstructed 21-cm power spectrum. Note that we assume Gaussianity whereas the 21-cm signal is not Gaussian and also we calculate the error bars from the power of a single realization rather than over an ensemble of simulations. We ask the reader to bear in mind that these error bars might be considered incomplete because of this.

The quantity actually plotted is $\Delta^2_{2D}(k) = \frac{Ak^2\langle\delta(k)\delta^*(k)\rangle}{2\pi}$, where A is the area of the simulation map.

Fig. 19 shows the extent to which the FASTICA method can recover the 21-cm angular power spectrum. Overall, the 21-cm power spectrum is convincingly recovered across the redshift range. Any points where the power of the residuals is below the power of the noise are omitted, as this leads to an unrealistic negative reconstructed 21-cm power. As such, there is a lack of data at small scales indicative of noise leakage into the foregrounds.

This noise leakage could be a resolution effect (i.e. artefacts originating from correlated noise) or simply a result of FASTICA not being robust to noise, most likely the latter.

4.3.2 *3D power spectra*

To calculate the 3D power spectra we divide the cube into sub-bands of 8 MHz to avoid signal evolution effects. For each sub-band we then carry out a 3D Fourier transform and calculate the 3D power spectrum in spherical annuli in Fourier space. The frequencies attached to the plots correspond to the centre of the sub-band plotted. What we actually plot is the quantity $\Delta^2_{3D}(k) = \frac{Vk^3\langle\delta(k)\delta^*(k)\rangle}{2\pi^2}$, where V is the volume.

We find the same accurate recovery on scales above a few multiples of the PSF but with smaller errors due to the larger amount of data evaluated, Fig. 20.

4.3.3 *Cross-correlation power spectra*

To try and retrieve a more robust reconstructed 21-cm power spectrum, the cross-correlation of two data cube realizations was carried out. Two independent noise realizations were created and combined with identical foregrounds and 21-cm signals to create two data cubes with the only difference being the noise realization. FASTICA was performed on both of these cubes separately, resulting in two residual files. We carried out cross-correlations on the two reconstructed cosmological signals, the two residual files and the two 21-cm fitting error estimates (i.e. reconstructed 21-cm minus the simulated 21-cm). By cross-correlating the two residual signals consisting of two different noise realizations, we increase the amount of noise that will drop out in the noise cross terms, hopefully resulting in a more accurate power spectrum recovery when applied to real data. However, we do not expect a significant improvement

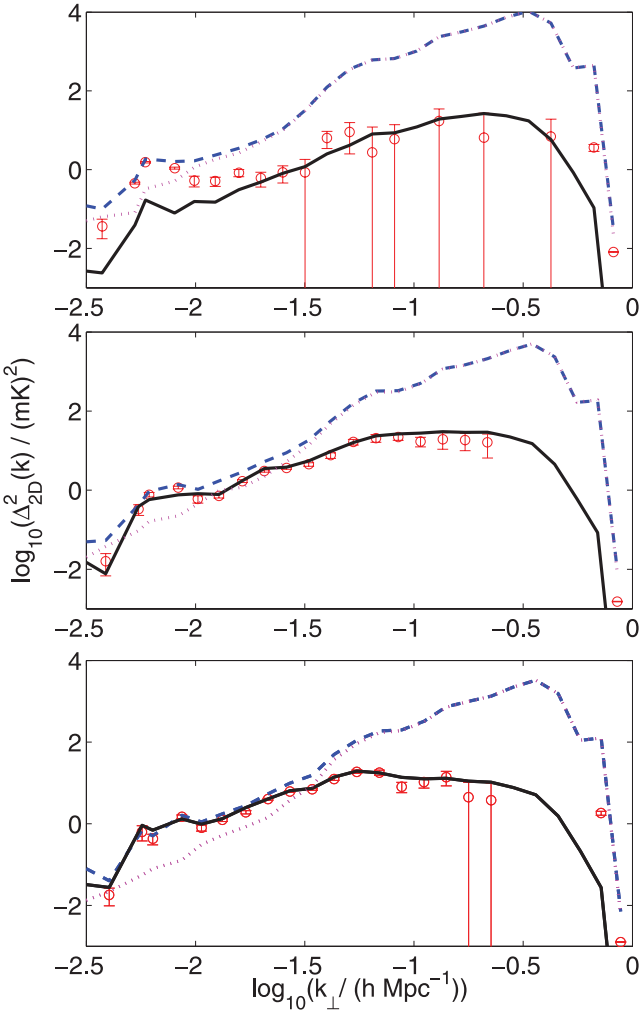

Figure 19. 2D power spectrum of the simulated 21-cm signal (black; solid), reconstructed 21-cm signal (red; points), residuals (blue; dashed) and noise (pink; dotted) at 131 MHz, or $z = 9.84$, 151 MHz, or $z = 8.40$ and 171 MHz, or $z = 7.30$ from top to bottom. Any error bars extending to below the x-axis in linear space are shown extending to negative infinity in log space.

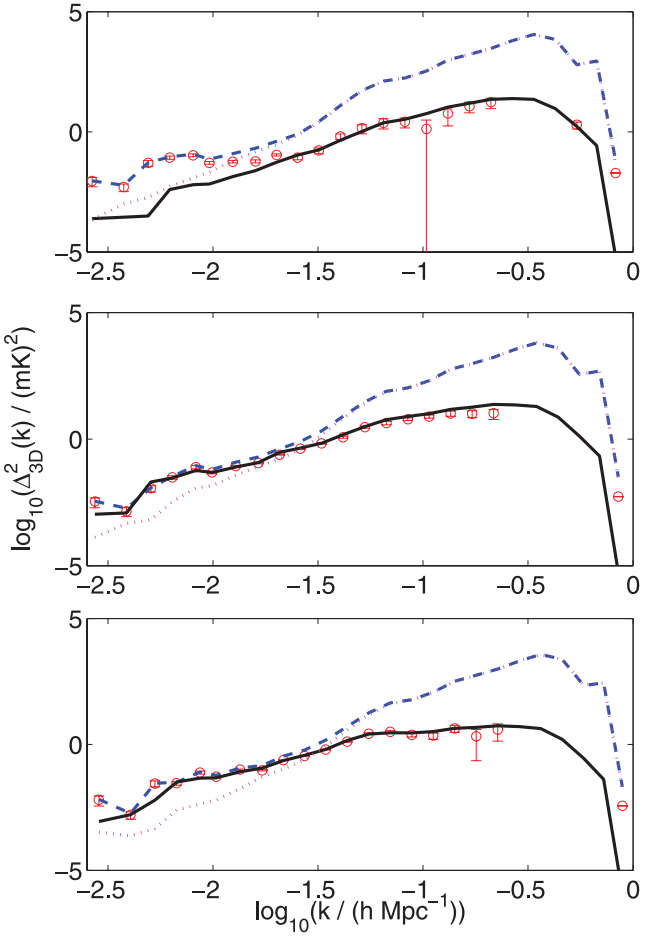

Figure 20. 3D power spectrum of the simulated 21-cm signal, reconstructed 21-cm signal, residuals and noise at 135 MHz, or $z = 9.51$, 151 MHz, or $z = 8.40$ and 175 MHz, or $z = 7.11$ over an 8 MHz sub-band. Any points where the power of the residuals is below the power of the noise are omitted, as this leads to an unrealistic negative reconstructed 21-cm power. The error bars and line styles are as described in Fig. 19.

in comparison to our 2D autospectra here as we have already assumed perfect knowledge of the noise spectrum. Instead, we do this as an example of a more robust method of power spectrum recovery for real data. Note that since correlations can be negative, it is the absolute value that is plotted. The error bars on the cross-spectra are calculated in the same way as for the autospectra, namely $\alpha_i = \frac{\langle \delta(k_i)\delta^*(k_i)\rangle}{\sqrt{n_{k_i}}}$, where n_{k_i} is the number of pixels that reside in that k bin. The power spectra recovered as a result of this process are shown in Fig. 21.

The cross-correlations were also carried out on two noise realizations which were adjusted to have roughly 10 times the signal-to-noise ratio of the LOFAR realizations (similar to what is hoped for SKA), Fig. 22. We see that with a higher signal-to-noise ratio the auto- and cross-correlation estimates are significantly improved.

4.4 Kurtosis and skewness

Skewness and, to a lesser extent, kurtosis have both been suggested as alternative statistics for the 21-cm detection due to their increased robustness to fitting errors compared to the variance (Wyithe & Morales 2007; Harker et al. 2009a). We define skewness (γ_1) and kurtosis (γ_2) in equations (20) and (21), respectively,

$$\gamma_1 = \frac{\frac{1}{N}\sum_i (T_i - \bar{T})^3}{\left(\frac{1}{N}\sum_i (T_i - \bar{T})^2\right)^{\frac{3}{2}}} \tag{20}$$

$$\gamma_2 = \frac{\frac{1}{N}\sum_i (T_i - \bar{T})^4}{\left(\frac{1}{N}\sum_i (T_i - \bar{T})^2\right)^{2}} - 3. \tag{21}$$

Kurtosis is defined in such a way that a Gaussian distribution would have a kurtosis of zero.

The structure of the 21-cm skewness and kurtosis for different source models has been discussed by Harker et al. (2009a), Wyithe & Morales (2007) and Iliev et al. (2011). There is expected to be a skewness of the 21-cm signal as the signal becomes increasingly non-Gaussian and the regions of ionized hydrogen become more numerous. Simulations also show an increase in skewness at very low redshift due to a high brightness temperature tail related to regions with some remaining neutral hydrogen.

Harker et al. (2009a) employed a Wiener filter on the dirty residual data to denoise the images, recovering the general trends of the 21-cm skew. Kurtosis recovery proved more elusive. We present

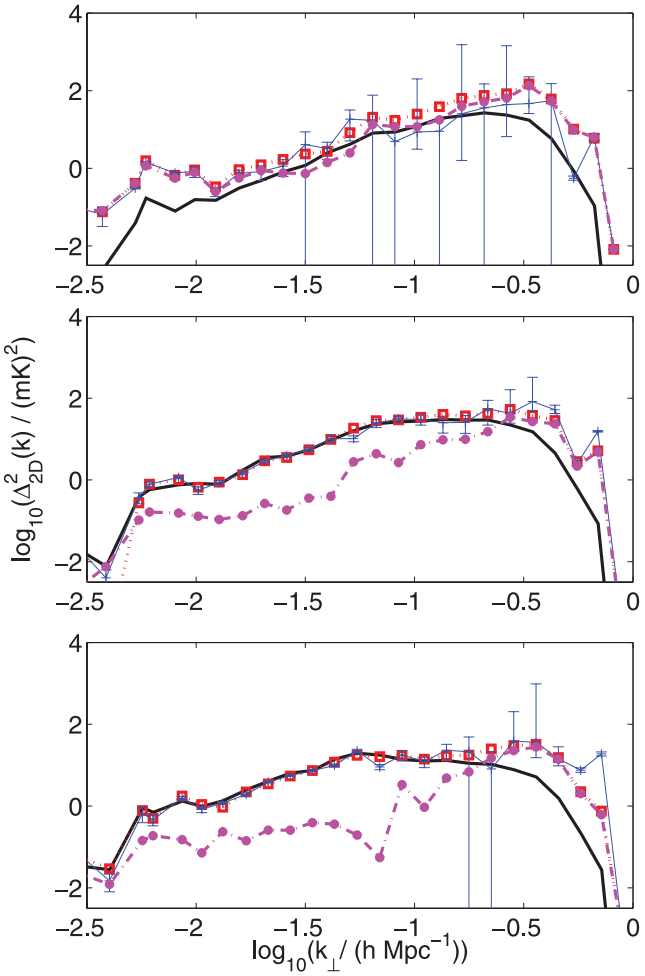

Figure 21. Cross-correlations of the two residuals (blue; cross), two reconstructed 21-cm signals (red; square), two fitting error estimates (pink; circle) and the autocorrelation of the simulated (black; solid) at 131, 151and 171 MHz. Only one set of error bars is shown for clarity.

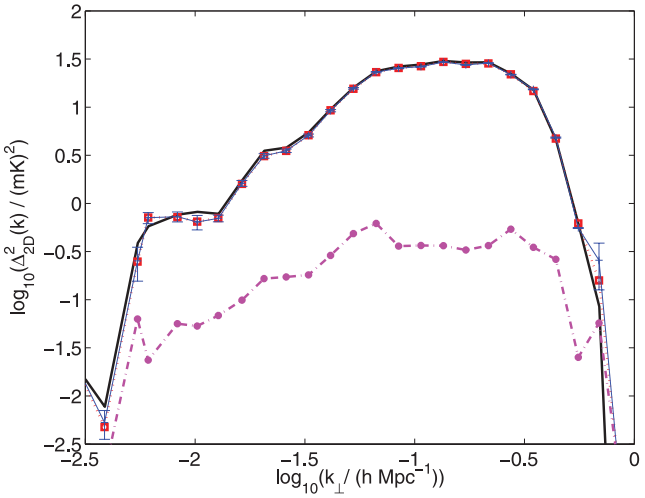

Figure 22. Cross-correlations of the two residuals (blue; cross), two reconstructed 21-cm signals (red; square), two fitting error estimates (pink; circle) and the autocorrelation of the simulated (black; solid) and reconstructed cosmological signal (for one realization) (red; circles) at 150 MHz. The noise realizations involved have been adjusted to be 10 times smaller than the LOFAR realizations.

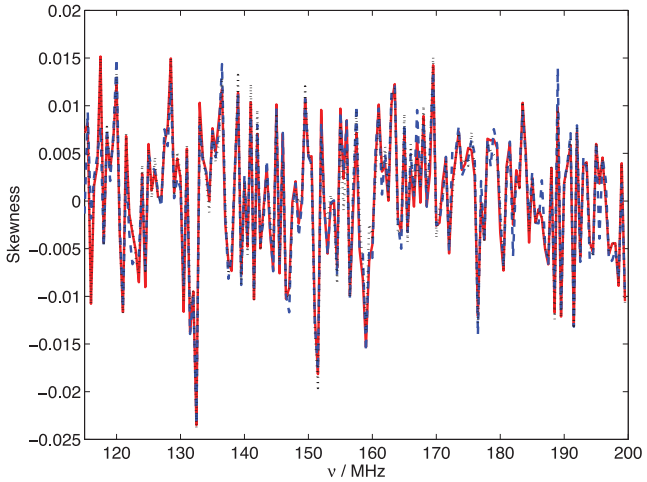

Figure 23. The skewness of the simulated 21-cm signal plus noise (red), noise alone (black dotted) and residual maps (blue dashed).

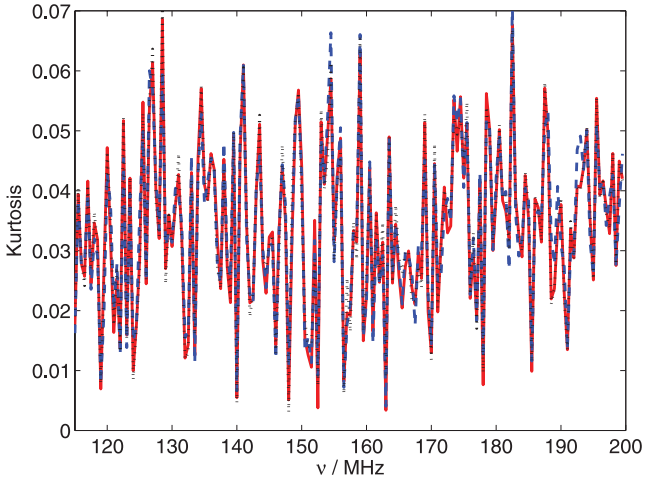

Figure 24. The kurtosis of the simulated 21-cm signal plus noise (red), noise alone (black dotted) and residual maps (blue dashed).

the skewness and kurtosis of the residual cubes, Figs 23 and 24, respectively.

The skewness and kurtosis in the residual images are recovered very well, accurately matching the simulated noise plus 21-cm signal skewness and kurtosis across the frequency range.

We now manually subtract the noise cube from the residual cube and plot the kurtosis/skewness of this reconstructed 21-cm signal, Figs 25 and 26. This amounts to assuming that we know the noise distribution perfectly which, though not viable for real data, allows us to gain an insight into the recovered signal.

We see that the skewness dip at low frequencies is only convincingly recovered with a high level of Fourier filtering. At high frequencies, where the cosmological signal is very small, the skewness is not recovered. The dip in kurtosis at frequency 165 MHz is somewhat recovered for Fourier filtering below 2 PSF scales, while it takes up to 5 PSF scales of k-space filtering before the peak centred around 140 MHz is recovered. For both statistics there is a divergence above frequencies of 180 MHz, where the cosmological signal is very small.

All of the results presented in this section are for FASTICA being implemented in real space. While an implementation was carried

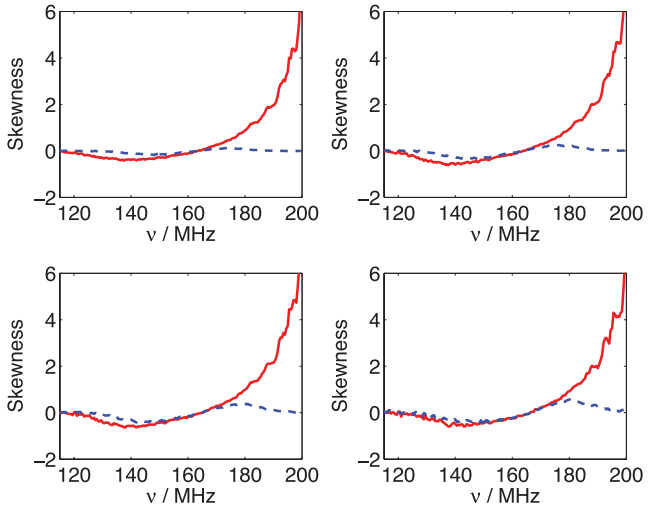

Figure 25. The skewness of the simulated 21-cm signal (red; solid) and reconstructed 21-cm maps (blue; dashed) for the fiducial signal and for different levels of Fourier filtering: 2, 3 and 5 PSF scales (in reading order).

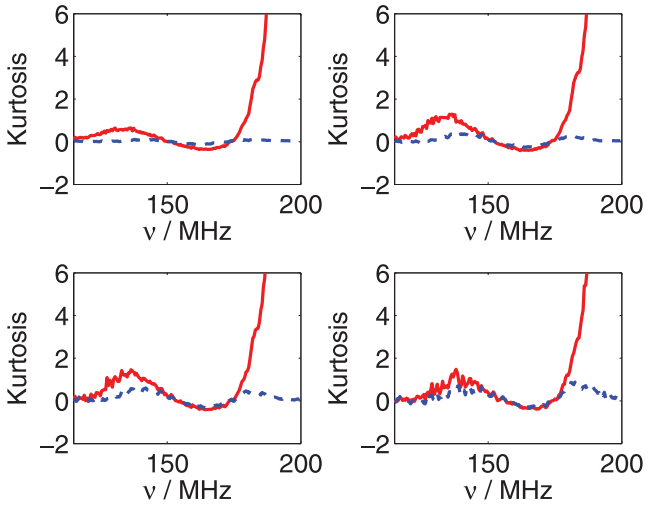

Figure 26. The kurtosis of the simulated 21-cm signal (red; solid) and reconstructed 21-cm maps (blue; dashed) for the fiducial signal and for different levels of Fourier filtering: 2, 3 and 5 PSF scales (in reading order).

out in Fourier space, the general conclusions for all results remained the same. Though there were small local variations in, for example, the recovered power spectrum points or kurtosis values, the graphs were for all intents and purposes duplications of the real-space versions and are therefore not reproduced here.

5 SENSITIVITY OF FASTICA

So far in this paper we have assumed that the full field of view and frequency bandwidth of the simulation are input to FASTICA but we must also consider whether this method will be as successful under more constrained observations.

5.1 Bandwidth of observation

First, we assess the sensitivity to bandwidth and split the dirty data cube into two smaller cubes of bandwidth 42.5 MHz, one from 115 to 157 MHz and one from 157.5 to 199.5 MHz. We perform FASTICA

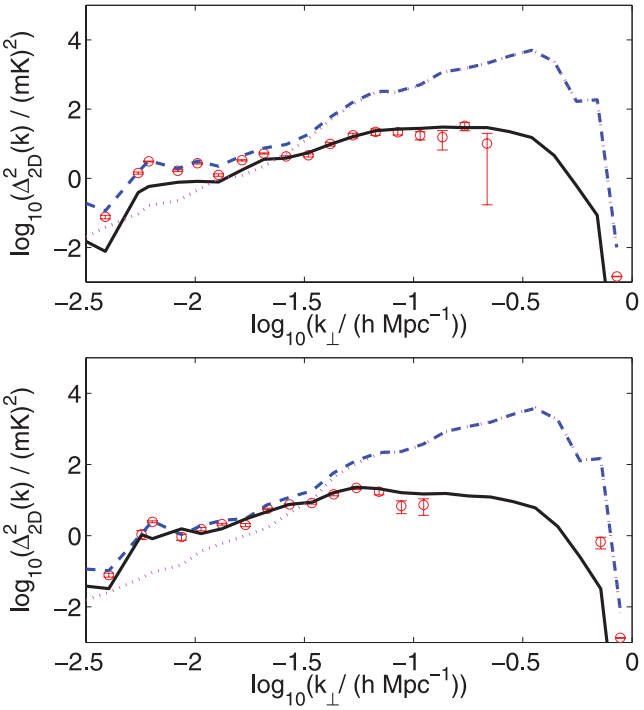

Figure 27. 2D power spectrum of the simulated 21-cm signal (black; solid), reconstructed 21-cm signal (red; points), residuals (blue; dashed) and noise (pink; dotted) at 151 MHz (top) and 171 MHz (bottom). The graphs are for data cubes of bandwidth 42.5 MHz and ranges 115–157 MHz and 157.5–199.5 MHz, respectively. Any error bars extending to below the *x*-axis in linear space are shown extending to negative infinity in log space.

on each of these separately and measure the 2D power spectrum as described previously, Fig. 27.

We can see that even for slices at the end of the cube frequency range (i.e. Fig. 27, top, shows a slice 6 MHz from the end of that cube) the 21-cm reconstruction is successful. The general degradation is not unexpected as the more data a separation technique has to fix the foregrounds, the better the reconstruction will be. We conclude that the method is not sensitive to the point of endangering the signal recovery, but larger bandwidths are preferable.

5.2 Noise

Despite the encouraging results so far, the evidence of noise leakage in the recovered maps (Fig. 16) and the variance recovery (Fig. 15) motivates us to consider the sensitivity of the 21-cm statistical recovery when there is increased noise in the observation.

We have seen that by much reducing the expected LOFAR noise to expected SKA levels, the 21-cm cross-correlation power spectrum recovery is extremely accurate, Fig. 22. For completeness, here we set up some 'worst case' scenarios, whereby we measure the recovered power spectra in the presence of twice, three times and five times the expected LOFAR noise, Fig. 28.

As expected, the more noise present, the less accurately the 21-cm power spectrum is recovered. For twice the expected level of noise we see the larger scales beginning to be overestimated, the extent of which worsens for three times the expected noise. For five times the expected amount of noise the power spectrum is significantly overestimated across the scale range. However, we must stress that the fact that the 21-cm power spectrum is recovered across a

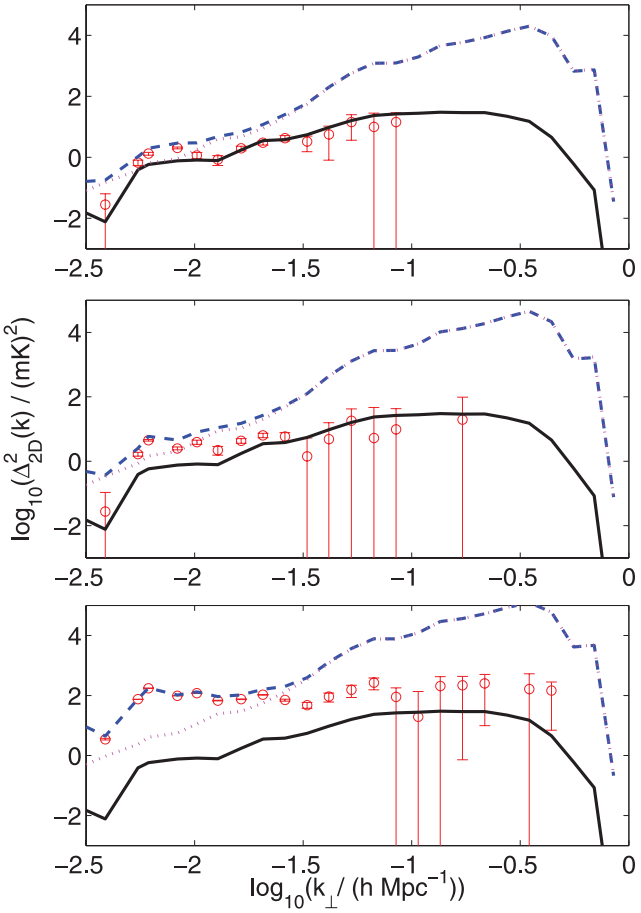

Figure 28. 2D power spectrum of the simulated 21-cm signal (black; solid), reconstructed 21-cm signal (red; points), residuals (blue; dashed) and noise (pink; dotted) at 151 MHz, or $z = 7.30$. From top to bottom, the noise simulation is set at twice, three times and five times the expected LOFAR noise, respectively. Any error bars extending to below the x-axis in linear space are shown extending to negative infinity in log space.

wide-scale range, even in the presence of twice the noise levels expected, can only be seen as extremely promising.

5.3 Field of view

In this paper we have assumed a $10^\circ \times 10^\circ$ field of view, which is at the upper limit of what we can expect for LOFAR observations. To explore the sensitivity of the analysis to the field of view, we now process a $2.^\circ5 \times 2.^\circ5$ data cube. If we had kept the noise and the resolution the same, analysing such a data cube would be plagued with noise as we would have reduced the number of pixels that we are analysing. Hence, we can choose to analyse a smaller patch in the sky with a higher resolution and same noise or decrease the noise and have similar resolution in order to have similar constraining power as the fiducial analysis and establish the effect of the sky area coverage. In actual observations a decrease in field of view and an increase in resolution would be related to the size of the stations and distribution of the stations, respectively. If we had changed the resolution we would no longer correspond, strictly speaking, to a LOFAR case scenario. We therefore decide to decrease the field of view by a factor of 4 and enhance the signal-to-noise ratio by a factor of 16. We see that the residuals are actually lower than

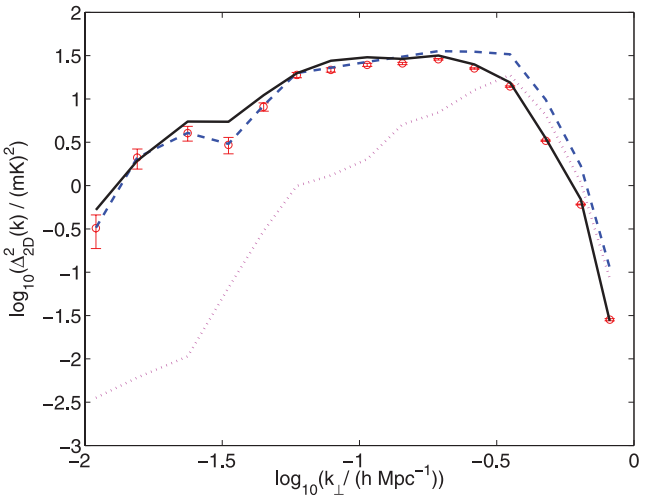

Figure 29. 2D power spectrum of the simulated 21-cm signal (black; solid), reconstructed 21-cm signal (red; points), residuals (blue; dashed) and noise (pink; dotted) at 151 MHz, or $z = 7.30$, for a $2.^\circ5 \times 2.^\circ5$ field of view. Any error bars extending to below the x-axis in linear space are shown extending to negative infinity in log space.

the original 21-cm signal at the larger scales, Fig. 29; however, the 21-cm power spectrum is still well recovered at the smaller scales. We interpret this as evidence that the 21-cm signal has been mixed into the other signals by FASTICA, potentially because FASTICA did not have as many lines of sight to remove the foregrounds, making the reconstruction less accurate though still successful.

6 CONCLUSIONS

We have presented new implementation of a non-parametric foreground cleaning method using the FASTICA algorithm. FASTICA is an ICA technique which uses negentropy as a measure of non-Gaussianity. By maximizing the non-Gaussianity of a signal mixture, the ICs of the foregrounds can be separated. FASTICA can then reconstruct the foregrounds, with any data not considered to be part of the foregrounds forming the residuals. The residuals consist of the 21-cm signal, system noise and fitting errors.

The success of using the FASTICA method to obtain an EoR signature was tested by attempting extraction of the two main statistical markers of the EoR, the 21-cm power spectrum and variance. The rms foreground fitting error is bounded below 10 mK across almost all of the frequency range when pixels with disproportionate errors due to unusually small foreground values are discarded.

Once the variance of the noise has been subtracted from the variance of the residuals, an excess variance is recovered. To accurately recover the 21-cm variance it was necessary to Fourier filter the data up to about five times the PSF scale. In this case the excess variance accurately recovers the order and shape of the simulated 21-cm variance across the majority of the frequency range, failing only where the signal-to-noise ratio is extremely low.

The 21-cm angular power spectrum and 3D power spectrum are recovered very well across a wide frequency range.

Performing the ICA in Fourier space provides no particular advantages or disadvantages according to the statistical tests carried out in this paper. This is in contrast to other methods which have shown preference towards processing in Fourier space (Harker et al. 2009b).

The FASTICA method has not proved to be robust in the presence of large amount of noise. Though impressive results are obtained at large scales even for twice the expected levels of noise, levels above this endanger the recovery.

We have shown that FASTICA can be a competitive foreground removal technique for EoR data, though for a full treatment of the LOFAR-EoR data, the polarization of the simulated data and a more accurate frequency-dependent PSF model needs to be considered in future work.

ACKNOWLEDGMENTS

FBA acknowledges the support of the Royal Society via a University Research Fellowship.

GH is a member of the LUNAR Consortium, which is funded by the NASA Lunar Science Institute (via Cooperative Agreement NNA09DB30A) to investigate concepts for astrophysical observatories on the Moon.

The authors would like to acknowledge Jacques Delabrouille and Mario Santos for useful discussion.

REFERENCES

Ali S. S., Bharadwaj S., Chengalur J. N., 2008, MNRAS, 385, 2166
Bernardi G. et al., 2009, A&A, 500, 965
Bernardi G. et al., 2010, A&A, 522, A67
Bottino M., Banday A. J., Maino D., 2008, MNRAS, 389, 1190
Bottino M., Banday A. J., Maino D., 2010, MNRAS, 402, 207
Bowman J. D., Morales M. F., Hewitt J. N., 2006, ApJ, 638, 20
Ciardi B., Madau P., 2003, ApJ, 596, 1
Cooray A., Furlanetto S. R., 2004, ApJ., 606, L5
Di Matteo T., Perna R., Abel T., Rees M. J., 2002, ApJ, 564, 576
Di Matteo T., Ciardi B., Miniati F., 2004, MNRAS, 355, 1053
Ewen H., Purcell E., 1951, Nat, 168, 356
Field G., 1958, Proc. IRE, 46, 240
Field G., 1959, ApJ, 129, 536
Gleser L., Nusser A., Benson A. J., 2008, MNRAS, 391, 383
Gnedin N. Y., Shaver P. A., 2004, ApJ, 608, 611
Green D., 2006, A Catalogue of Galactic Supernova Remnants (2006 April version) available at http://www.mrao.cam.ac.uk/surveys/snrs
Harker G. J. A. et al., 2009a, MNRAS, 393, 1449
Harker G. et al., 2009b, MNRAS, 397, 1138
Harker G. et al., 2010, MNRAS, 405, 2492
Hyvärinen A., 1999, IEEE Trans. Neural Netw., 10, 626
Hyvärinen A., Oja E., 2000, Neural Netw., 13, 411
Hyvärinen A., Karhunen J., Oja E., 2001, Independent Component Analysis. John Wiley and Sons, New York
Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., Mao Y., Koda J., Ahn K., 2011, preprint (astro-ph/1107.4772)
Jelić V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E., 2010, MNRAS, 409, 1647
Jelić V. et al., 2008, MNRAS, 389, 1319
Komatsu E. et al., 2011, ApJS, 192, 18
Labropoulos P. et al., 2009, preprint (astro-ph/0901.3359v1)
Liu A., Tegmark M., 2011, Phys. Rev. D, 83, 103006
Liu A., Tegmark M., 2012, MNRAS, 419, 3491
Liu A., Tegmark M., Zaldarriaga M., 2009a, MNRAS, 394, 1575
Liu A., Tegmark M., Bowman J., Hewitt J., Zaldarriaga M., 2009b, MNRAS, 398, 401
McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Maino D., Banday A. J., Baccigalupi C., Perrotta F., Grski K. M., 2003, MNRAS, 344, 544
Maino D., Donzelli S., Banday A. J., Stivoli F., Baccigalupi C., 2007, MNRAS, 374, 1207
Maino D. et al., 2002, MNRAS, 334, 53
Mao X.-C., 2012, ApJ, 744, 29
Mesinger A., Furlanetto S., 2007, ApJ, 669, 663
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Morales M. F., Hewitt J., 2004, ApJ, 615, 7
Morales M. F., Bowman J. D., Hewitt J. N., 2006, ApJ, 648, 767
Muller C., Oort J., 1951, Nat, 168, 357
Oh S. P., Mack K. J., 2003, MNRAS, 346, 871
Petrovic N., Oh S. P., 2011, MNRAS, 413, 2103
Santos M. G., Cooray A., Knox L., 2005, ApJ, 625, 575
Santos M. G., Ferramacho L., Silva M. B., Amblard A., Cooray A., 2010, MNRAS, 406, 2421
Shaver P. A., Windhorst R. A., Madau P., de Bruyn A. G., 1999, A&A, 345, 380
Thompson A. R., Moran J. M., Swenson G. W. Jr, 2001, Interferometry and Synthesis in Radio Astronomy. John Wiley and Sons, New York
van de Hulst H., 1945, Ned. Tijdschr. Natuurkd, 11, 201
Waldmann I. P., 2012, ApJ, 747, 12
Wang X., Tegmark M., Santos M. G., Knox L., 2006, ApJ, 650, 529
Wild J., 1952, ApJ, 115, 206
Wyithe J. S. B., Morales M. F., 2007, MNRAS, 379, 1647
Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622

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