

Joint Calibration and Imaging for Phased Array Radio Telescopes

Liying Wei^{1*} and Stefan J. Wijnholds²

¹ *IBM Research - Netherlands, Amsterdam, The Netherlands*

² *Netherlands Institute for Radio Astronomy (ASTRON), Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, The Netherlands*

29 October 2018

ABSTRACT

We propose a novel compressed sensing based self-calibration and imaging algorithm for phased array radio telescopes, whose receiving elements are subject to the same direction dependent effects. Our approach simultaneously estimates the apparent sky brightness distribution and the direction independent receiver path gains from the observational data. The estimated source model may contain point sources and extended structures. Moreover, our algorithm allows to tune the accuracy of gain calibration and reconstructed image quality by controlling the detail level in the reconstructed image by a threshold value. We demonstrate the accuracy, robustness to extremely low Signal-to-Noise Ratio (SNR) and large sensor phase variability using Monte Carlo simulations for various array configurations. We also display the main advantages of our algorithm on actual data from a Low Band Antenna (LBA) station of the Low Frequency Array (LOFAR): firstly, our algorithm is sufficiently fast for practical applications such as LOFAR station calibration; secondly, the gain calibration accuracy improves with increasingly detailed reconstructed image; thirdly, the method is capable of handling a complex observed field that is a combination of a few dominant point sources and diffuse emission from the Galactic plane. The latter point is interesting as it poses a difficult scenario for sky source model based calibration approaches, in which the short baselines are often flagged to avoid visibility modelling issues. This feature could make our method very attractive for calibration of the Square Kilometre Array (SKA) or a future low-frequency instrument in space.

Key words: instrumentation: interferometers, methods: numerical, methods: data analysis, techniques: image processing, techniques: interferometric

1 INTRODUCTION

Modern low-frequency (< 300 MHz) radio interferometers usually consist of a (very) large number of small antennas instead of a (relatively) small number of reflector dishes. Examples of such instruments are the Low Frequency Array (LOFAR, van Haarlem et al. (2013)), the Long Wavelength Array (LWA, Ellingson et al. (2009); Eastwood & Hallinan (2017)), the Murchison Widefield Array (MWA, Tingay et al. (2013)) and the Low Frequency Aperture array (LFAA) system of the future Square Kilometre Array (SKA, Dewdney et al. (2009)). These new interferometers have a hierarchical system architecture, i.e., they consist of a large number of subarrays called stations. An attractive feature of such stations is that their individual receiving elements have an extremely large field-of-view (FoV), which poten-

tially allows for transient detection anywhere in the visible sky and formation of multiple beams spread across the sky (van Ardenne et al. 2009). The downside of this capability is that these stations are very susceptible to Radio Frequency Interference (RFI) and often have an antenna configuration that has many short baselines, for which accurate visibility modelling is known to be challenging. Typically, each station needs to form a well-defined beam into a desired direction, which requires an accurate estimation of the complex gains of each antenna element in the station. The antennas within such a station (are usually assumed to) have the same direction dependent behaviour. In this paper, we propose a blind calibration approach to deal with these challenges in the station level.

Blind calibration methods do not require a priori knowledge of the positions and powers of the sources in the observed field. These methods are therefore robust to the presence of unknown or poorly modeled RFI or visibility mod-

* E-mail: weiliyang101@gmail.com

elling issues on short baselines. As these methods are robust to the presence of unknown sources, they are also robust to transient phenomena. This can be an attractive feature for all-sky transient monitoring systems, such as the Amsterdam-ASTRON Radio Transient Facility and Analysis Centre (AARTFAAC, Prasad et al. (2014)). An extreme case, in which the observed source structure is inaccurately known, is the Orbiting Low Frequency Antennas for Radio Astronomy (OLFAR, Bentum et al. (2009)). OLFAR is envisaged to be a space-based interferometer array to observe the universe at frequencies below 30 MHz. As we do not have an accurate sky model at those frequencies, blind calibration may provide a good alternative to conventional calibration methods.

A well-known blind calibration algorithm in radio astronomy is redundancy calibration, which was first proposed by Noordam & de Bruyn (1982) to calibrate the Westerbork Synthesis Radio Telescope (WSRT), a uniform linear array (ULA). Later, this method was extended to uniform rectangular arrays (URA, Wijnholds & Noorishad (2012); Heidenreich et al. (2012)) and put to use for LOFAR station calibration (Noorishad et al. 2012; Wijnholds & Noorishad 2012). Unfortunately, redundancy calibration only works for regular array layouts while systems like LOFAR and SKA exploit a random array layout to observe more spatial frequencies instantaneously with the same number of receiver paths. If the sampling density in the aperture is high enough, the significant number of almost redundant baselines may be used to calibrate the array at the cost of reduced calibration sensitivity as demonstrated by Wijnholds (2016). Kazemi et al. (2015) proposed a Compressed Sensing (CS) based blind calibration method assuming that the observed source structure has a sparse representation. However, this method works on the raw time series data, making its application computationally infeasible for most radio interferometers, and only works when the phase variability between the elements is small. Wijnholds & Chiarucci (2016) and Chiarucci & Wijnholds (2018) proposed a blind calibration method that works on visibility data. This method iteratively solves for the direction-independent receiver path gains and the image assuming that the observed field is sparse. Compared to redundancy calibration, their method has the advantage that it works for arbitrary array geometries. As an additional bonus, their method does not only calibrate the array, but also produces a source model for the observed scene. However, this scheme assumes that the observed field only contains a few discrete point sources. As a result, they still had to flag the short baselines.

In this paper, we propose a novel compressed sensing based joint blind calibration and imaging method for phased array radio interferometers. We reformulate the problem as a general convex optimisation problem by assuming the sky brightness distribution is sparse or compressible in some sparsity basis, and present a possible solution to this generalised problem based on the Douglas-Rachford splitting algorithm (Combettes & Pesquet 2007; Wei et al. 2017). We demonstrate that our method can handle the complex visibility structure observed on short baselines by applying our method to data from a single LOFAR station observing the entire sky containing both point sources and diffuse emission from the Galactic plane. We show that the Douglas-Rachford splitting algorithm allows to tune the de-

sired detail level in the source model, which allows the user to make a tradeoff between accuracy of the source model and calibration solutions and computational load.

This paper is organised as follows. In the next section, we present the data model (or measurement equation) that we use for our analysis and formulate the problem in mathematical terms. In Section 3, we present the method proposed to solve this joint blind calibration and imaging problem. In Sections 4 and 5, we validate our method, first in simulations, then on actual data from a single LOFAR station. We conclude our paper in Section 6

Notation: $\mathcal{E}\{\cdot\}$ denotes the expectation operator, $\langle \cdot, \cdot \rangle$ denotes the inner product, $\text{diag}(\cdot)$ converts a vector to a diagonal matrix, $\text{vec}(\cdot)$ converts a matrix to a vector by stacking the columns of the matrix, $\widehat{(\cdot)}$ denotes an estimated value, $\|\cdot\|_p$ denotes p -norm, $\|\cdot\|_F$ denotes Frobenius norm, $|\cdot|$ denotes the magnitude of a variable, \angle denotes the phase or angle of a variable, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ respectively denote the real part and imaginary part of a complex number, \odot denotes the element-wise matrix multiplication (Hadamard product), \otimes denotes the Kronecker product, \circ denotes the Khatri-Rao product (a column-wise Kronecker product), $\overline{[\cdot]}$ denotes conjugation, $[\cdot]^T$ denotes transpose, and $[\cdot]^H$ denotes complex conjugate (Hermitian) transpose.

2 DATA MODEL AND PROBLEM STATEMENT

We consider a phased array radio telescope consisting of P identical antennas, and we model the sky as a collection of Q pixels with a brightness that may be zero or non-zero. In this paper, we process one frequency channel at one Short Term Integration (STI) under the following assumptions: 1) all source signals are narrow-band and far-field; 2) the noise signals of the sensors are additive and follow a zero-mean Gaussian distribution; 3) all noise and source signals are mutually independent. We consider only a single polarization for convenience of notation.

2.1 Data Model

Let the central frequency of the considered subband (satisfying the narrowband condition) be f_c with corresponding wavelength $\lambda_c = c/f_c$, where c is the speed of light. Let g_p be the complex-valued direction-independent gain of the p -th antenna located at $\mathbf{r}_p = (r_{x,p}, r_{y,p}, r_{z,p})$. Then the received signal for antenna p , as a function of time t , can be represented as

$$x_p(t) = g_p \left[\sum_{q=1}^Q a_{p,q} b_{p,q} s_q(t) \right] + n_p(t), \quad (1)$$

where $s_q(t)$ denotes signal strength from direction $\mathbf{l}_q = [l_q, m_q, n_q]$, $a_{p,q} = e^{-j\frac{2\pi}{\lambda_c} \langle \mathbf{r}_p, \mathbf{l}_q \rangle}$ describes the geometrical delay for antenna p to pixel q , $b_{p,q}$ is the direction dependent (DD) gain including antenna beam pattern to pixel q , and $n_p(t)$ is the additive zero-mean Gaussian noise. The direction coordinates, l_q , m_q and n_q , are direction cosines such that $l_q^2 + m_q^2 + n_q^2 = 1$. Under our assumption that all antennas have the same directional gain response, we have $b_{p,q} = b_q$ for all $p = 1, 2, \dots, P$.

For simplicity, we omit the time index and stack all the indexed quantities into vectors and matrices. Let $\mathbf{x} = [x_1, x_2, \dots, x_P]^T$, $\mathbf{g} = [g_1, g_2, \dots, g_P]^T$, $\mathbf{a}_p = [a_{p,1}, a_{p,2}, \dots, a_{p,Q}]^T$, $\mathbf{b} = [b_1, b_2, \dots, b_Q]^T$, $\mathbf{s} = [s_1, s_2, \dots, s_Q]^T$ and $\mathbf{n} = [n_1, n_2, \dots, n_P]^T$. We also define the direction-independent gain matrix $\mathbf{G} = \text{diag}(\mathbf{g})$, array response matrix (containing geometrical delays) $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_P]^T$ and DD gain matrix $\mathbf{B} = \text{diag}(\mathbf{b})$. We finally have our data model

$$\mathbf{x} = \mathbf{GABs} + \mathbf{n}. \quad (2)$$

Under the assumption that sources and noise are uncorrelated, the covariance matrix (or visibility matrix) of the received signals \mathbf{x} can be represented as

$$\mathbf{R} = \mathcal{E}\{\mathbf{x}\mathbf{x}^T\} = \mathbf{GAB}\Sigma_s\mathbf{B}^H\mathbf{A}^H\mathbf{G}^H + \Sigma_n, \quad (3)$$

where $\Sigma_s = \mathcal{E}\{\mathbf{s}\mathbf{s}^H\} = \text{diag}\{\sigma_s\}$ for $\sigma_s = [\sigma_{s,1}, \sigma_{s,2}, \dots, \sigma_{s,Q}]^T$ and $\Sigma_n = \mathcal{E}\{\mathbf{n}\mathbf{n}^H\}$ represent the covariance matrices associated with the source signals and the received noise power, respectively. Let $\sigma_n = [\sigma_{n,1}, \sigma_{n,2}, \dots, \sigma_{n,P}]^T$. Here, $\sigma_{s,q} = \mathcal{E}\{|s_q|^2\}$ is the brightness of the q -th pixel, $\sigma_{n,p} = \mathcal{E}\{|n_p|^2\}$ is the receiver noise power of the p -th antenna. Note that each matrix element $R_{i,j}$ represents the visibility on the baseline between antennas i and j in the array. After sampling the received signal with sample period T and stacking N time samples in a snapshot observation into $\mathbf{X} = [\mathbf{x}[1], \mathbf{x}[2], \dots, \mathbf{x}[N]]$, we have the estimated visibility matrix $\hat{\mathbf{R}} = \frac{1}{N}\mathbf{X}\mathbf{X}^H$.

Since both \mathbf{B} and Σ_s in (3) are unknown and diagonal, we can introduce $\Sigma = \mathbf{B}\Sigma_s\mathbf{B}^H = \text{diag}(\sigma)$ with apparent fluxes $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_Q]^T$ with each $\sigma_q = |b_q|^2\sigma_{s,q}$, such that the data model (3) can be simplified into

$$\mathbf{R} = \mathbf{GA}\Sigma\mathbf{A}^H\mathbf{G}^H + \Sigma_n. \quad (4)$$

Vectorizing both sides of equation (4), we get

$$\mathbf{y} = (\overline{\mathbf{GA}} \circ \mathbf{GA})\sigma + (\overline{\mathbf{I}_P} \circ \mathbf{I}_P)\sigma_n, \quad (5)$$

where $\mathbf{y} = \text{vec}(\mathbf{R})$ and \mathbf{I}_P is a $P \times P$ identity matrix. Let $\mathbf{M}_s = (\overline{\mathbf{GA}} \circ \mathbf{GA})$ and $\mathbf{M}_n = (\overline{\mathbf{I}_P} \circ \mathbf{I}_P)$. Note that the number of baselines is $P(P-1)/2$, and that the variable \mathbf{y} in eqn. (5) contains some redundancy due to the Hermitian symmetry of \mathbf{R} .

2.2 Problem Statement

Our objective is to develop an algorithm to solve (5) for simultaneous calibration (estimation of direction-independent gains \mathbf{g}), image reconstruction (estimation of source positions \mathbf{l}_q and source powers σ_q) and noise power σ_n estimation based on the measured visibility matrix $\hat{\mathbf{R}}$. Such problems are often solved by eigen-structure methods, e.g., Weiss & Friedlander (1989); Flanagan & Bell (2001); See (1995); Schmidt (1986); Viberg & Ottersten (1991); Viberg et al. (1991). However, those approaches do not work well when the number of sources in the scene is large or when the source structures are more complicated than simple point sources.

To reduce the complexity of solving (5), we discretise the brightness distribution on the sky by $Q = n_{grid} \times n_{grid}$ pixels (hence, n_{grid} is the number of pixels in one dimension) instead of describing this brightness distribution as the superposition of a number of sources with different shapes and

sizes. Hereby, we have fixed \mathbf{A} . In radio astronomy, the actual number of sources K is usually significantly less than Q .

Assuming that the image (spatial intensity distribution) is sparse or compressible in some basis ψ (e.g., a Dirac, Fourier, or wavelet basis), i.e., $(\psi^T \sigma)$ has only K non-zero values with $K \ll Q$, we formulate the blind calibration problem as a sparse reconstruction problem:

$$\begin{aligned} \{\hat{\mathbf{g}}, \hat{\sigma}, \hat{\sigma}_n\} = \arg \min_{\{\mathbf{g}, \sigma_s, \sigma_n\}} & \|\psi^T \sigma\|_1, \\ \text{s.t.} & \|\mathbf{y} - \mathbf{M}_s \sigma - \mathbf{M}_n \sigma_n\|_2 < \epsilon, \sigma \geq \mathbf{0}, \sigma_n \geq 0, \end{aligned} \quad (6)$$

where the tolerance parameter ϵ depends on the noise level. The non-negativity constraints $\sigma \geq \mathbf{0}$ and $\sigma_n \geq 0$ are required as the intensity is always positive. The l_2 norm (data fidelity item) promotes consistency with the measurements, while the l_1 norm (regularizer) promotes sparsity. Note that our data fidelity term uses the l_2 norm instead of the l_2 norm squared as done in many other papers. This choice provides two advantages: it promotes sparsity (Donoho 2006; Tropp 2006) and it permits us to specify the required detail level in the solution. The use of a general basis ψ also extends the applicability of our blind calibration approach to more complex scenes than the scenes that can be handled by the algorithm proposed by Chiarucci & Wijnholds (2018).

As the system noise power in the individual receive paths is usually significantly higher than the source power in radio astronomy (the receiver noise instead of sky noise is dominant in the diagonal values of Σ_n), the autocorrelations in the measured visibility matrix can only be used to estimate the elements on the diagonal of the noise covariance matrix Σ_n . Following the same strategy as by Chiarucci & Wijnholds (2018) and Wei et al. (2017), we therefore mask the autocorrelation (diagonal) elements from \mathbf{R} by applying a mask matrix \mathbf{M}_{mask} with $(\mathbf{M}_{mask})_{ii} = 0$ and $(\mathbf{M}_{mask})_{ij} = 1$ for $i \neq j$. Our problem then becomes

$$\begin{aligned} \{\hat{\mathbf{g}}, \hat{\sigma}\} = \arg \min_{\{\mathbf{g}, \sigma\}} & \|\psi^T \sigma\|_1, \\ \text{s.t.} & \|(\mathbf{y} - \mathbf{M}_s \sigma) \odot \text{vec}(\mathbf{M}_{mask})\|_2 < \epsilon \quad \text{and} \quad \sigma \geq \mathbf{0}, \end{aligned} \quad (7)$$

After solving $\{\mathbf{g}, \sigma\}$ from (7) and getting a feasible solution $\hat{\sigma}$, we may estimate the system noise power from

$$\hat{\sigma}_n = \arg \min_{\{\sigma_n\}} \|\mathbf{y} - \mathbf{M}_s \hat{\sigma} - \mathbf{M}_n \sigma_n\|_2^2. \quad (8)$$

Note that noise that is not spatially white, such as electronic noise or radio frequency interference (RFI) existing in \mathbf{R} , is processed through $\|(\mathbf{y} - \mathbf{M}_s \sigma) \odot \text{vec}(\mathbf{M}_{mask})\|_2 < \epsilon$.

3 JOINT CALIBRATION AND IMAGING

In this section, we propose a Compressed Sensing (CS) based method to iteratively solve the problem formulated in (7), i.e., we jointly estimate the complex-valued direction-independent gains \mathbf{g} and the sky model σ , during which the number of true source components K , their positions \mathbf{l}_q and intensity σ_q are determined by a threshold value based on either the cumulative energy ratio C_K or the fraction of non-zero coefficients ρ_K or the known number of source components K .

Our approach is inspired by the CS based image reconstruction technique developed by Carrillo et al. (2012) and

the alternating direction implicit gain calibration (StEFCal) described by Salvini & Wijnholds (2014) that we combine into an iterative optimisation process in a similar way as done by Chiarucci & Wijnholds (2018). Compared to sky model based calibration approaches, such as those proposed by Boonstra & van der Veen (2003); Wijnholds & van der Veen (2009); Salvini & Wijnholds (2014); Brossard et al. (2016); Repetti et al. (2017), we would like to reiterate that: 1) our calibration is blind, i.e., we do not require a priori information about the brightness distribution such as position and intensity of bright sources; 2) our estimated sky model may consist of bright sources, weak sources and extended structures; and 3) our method does not only calibrate the array but also reconstructs the observed image.

3.1 Top-level Structure of the Algorithm

To handle the well known inherent phase ambiguity in gain estimation, we take the first element as phase reference. As \mathbf{g} and $\boldsymbol{\sigma}$ share a common scaling factor, we normalise the estimated gain vector either with the median of its elements or its l_2 -norm after solving \mathbf{g} in each iteration.

We propose the following algorithm to solve (7):

(i) *Initialization* Set the iteration counter $j = 1$, the maximum number of iterations J_{iter} , the threshold used as stopping criterion τ_g and the cumulative energy ratio C_K or the fraction of non-zero coefficients ρ_K or the known number of source components K . Also, specify the image size n_{grid} (and $Q = n_{grid}^2$), the initial estimate for the gains $\mathbf{G}^{[0]} = \text{diag}(\mathbf{g}^{[0]})$, for which the identity matrix is usually a good choice, the spatial signature matrix \mathbf{A} and the mask matrix \mathbf{M}_{mask} .

(ii) for each iteration j ,

(a) *Imaging step*: solve $\hat{\boldsymbol{\sigma}}^{[j]}$ with (7) for $\mathbf{g} = \hat{\mathbf{g}}^{[j-1]}$ using the robust imaging algorithm proposed by Wei et al. (2017), during which the number of source components is constrained either by the cumulative sum of the estimated ordered source solutions C_K or the fraction of non-zero coefficients ρ_K or directly given by K ;

(b) *Calibration step*: solve $\hat{\mathbf{g}}^{[j]}$ with (4) for $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma}) = \text{diag}(\hat{\boldsymbol{\sigma}}^{[j]})$ using the fast alternating direction implicit (ADI) method (StEFCal) proposed by Salvini & Wijnholds (2014);

(c) *Resolve ambiguities*: normalise the estimated gain magnitudes with the median or l_2 -norm and apply phase rotation such that the first antenna becomes the phase reference;

(d) *Check for convergence*: if the relative gain change $\rho = \|\hat{\mathbf{g}}^{[j]} - \hat{\mathbf{g}}^{[j-1]}\|_2 / \|\hat{\mathbf{g}}^{[j-1]}\|_2$ smaller than τ_g , stop; if not, update iteration index $j \leftarrow j + 1$ and repeat (a)-(d) until the maximum number of iterations is reached.

(iii) *Noise power estimation*: solve noise power $\hat{\boldsymbol{\sigma}}_n$ with (8) when $\mathbf{M}_s = \tilde{\mathbf{G}}\mathbf{A} \circ \hat{\mathbf{G}}\mathbf{A}$.

(iv) *Visibilities correction*: compute $\hat{\mathbf{G}}^{-1}(\hat{\mathbf{R}} - \hat{\boldsymbol{\Sigma}}_n)\hat{\mathbf{G}}^{-H}$ with estimated gain $\hat{\mathbf{G}}$ and noise power $\hat{\boldsymbol{\Sigma}}_n = \text{diag}(\hat{\boldsymbol{\sigma}}_n)$.

The following subsections give more details about the imaging step and the calibration step in our proposed approach.

3.2 Imaging Step

The imaging step is to solve the sky model $\hat{\boldsymbol{\sigma}}$ from (7) for given \mathbf{R} , $\mathbf{M}_s (= \overline{\mathbf{G}}\mathbf{A} \circ \hat{\mathbf{G}}\mathbf{A})$ and \mathbf{M}_{mask} . This is an image formation or deconvolution problem (Levanda & Leshem 2010) and could be solved by the state-of-the-art CLEAN and its variants (Högbom 1974; Bhatnagar & Cornwell 2004; Cornwell 2008), Least Squares Minimum Variance Imaging (LS-MVI) (Leshem & van der Veen 2000; Ben-David & Leshem 2008; Sardarabadi et al. 2016), Maximum Entropy Method (MEM) (Gull & Daniell 1978; Narayan & Nityananda 1986; Pantin & Starck 1996) or more recent Compressed Sensing (CS) algorithms (Wiaux et al. 2009; Wenger et al. 2010; Carrillo et al. 2012). Both CLEAN and MEM are limited by their slow convergence, while LS-MVI is still under development (Levanda & Leshem 2010).

Compressed Sensing (CS) techniques, which can reconstruct a signal using far fewer measurements than required by the Nyquist-Shannon theory, have been proposed in radio astronomy for some time (Wiaux et al. 2009; Wenger et al. 2010; Li et al. 2011; Carrillo et al. 2012, 2014; Dabbech et al. 2015; Garsden et al. 2015; Onose et al. 2016; Pratley et al. 2018). However, most of these algorithms have one or more drawbacks: 1) they haven't only been demonstrated on artificial visibilities obtained by random sampling a regular grid of visibilities, 2) they need an extra randomised sensing matrix, and 3) they have not been validated on observational data.

We therefore propose a Compressed Sensing (CS) based sparse image reconstruction approach based on the Douglas-Rachford splitting algorithm (Combettes & Pesquet 2007) to infer a sky source model directly from the observed visibilities that may consist of bright point sources, weak point sources and extended structures. More specifically, we reformulate the reconstruction task (7) as an convex minimisation problem (Wei et al. 2017),

$$\min_{\{\boldsymbol{\sigma}\}} \iota_{\mathbb{S}}(\boldsymbol{\sigma}) + \gamma \|\boldsymbol{\psi}^T \boldsymbol{\sigma}\|_1, \quad \text{subject to } \boldsymbol{\sigma} \geq \mathbf{0}, \quad (9)$$

where γ is a regularization parameter to control the trade-off between l_2 -norm and l_1 -norm, \mathbb{S} is an l_2 -ball with $\mathbb{S} \triangleq \{\boldsymbol{\sigma} \in \mathbb{R}^{N^2} : \|\mathbf{y}_{mask} - \mathbf{D}_{mask}\mathbf{M}_s\boldsymbol{\sigma}\|_2 < \varepsilon\}$, and $\iota_{\mathbb{S}}(\boldsymbol{\sigma})$ is the indicator function defined in the convex set \mathbb{S} with

$$(\forall \boldsymbol{\sigma}) \quad \iota_{\mathbb{S}}(\boldsymbol{\sigma}) \triangleq \begin{cases} 0, & \boldsymbol{\sigma} \in \mathbb{S} \\ +\infty, & \boldsymbol{\sigma} \notin \mathbb{S}. \end{cases}$$

$\iota_{\mathbb{S}}(\boldsymbol{\sigma})$ is used to ensure the residual of the data fidelity to be situated in the l_2 -ball \mathbb{S} . Our formation (9) is different from other CS based sparse image reconstruction approaches (Carrillo et al. 2014; Onose et al. 2016, 2017), in which the positivity constraint $\boldsymbol{\sigma} \geq \mathbf{0}$ was included in their objective function. Our motivation for using $\iota_{\mathbb{S}}(\boldsymbol{\sigma})$ instead of the commonly used data fidelity item $\|\mathbf{y}_{mask} - \mathbf{D}_{mask}\mathbf{M}_s\boldsymbol{\sigma}\|_2^2$ is that the $\iota_{\mathbb{S}}(\boldsymbol{\sigma})$ formulation provides the sparsest solution for under-determined systems (Donoho 2006; Tropp 2006). Our motivation for excluding the positivity constraint $\boldsymbol{\sigma} \geq \mathbf{0}$ from the objective function $\iota_{\mathbb{S}}(\boldsymbol{\sigma}) + \gamma \|\boldsymbol{\psi}^T \boldsymbol{\sigma}\|_1$ is to

- reduce the computation complexity;
- develop a computationally efficient approach to determine the number of true source components based on how

Algorithm 1 Our proposed imaging algorithm

Input: : $\hat{\mathbf{R}}, \mathbf{M}_{mask}, \mathbf{A}, n_{grid}, \mathbf{g}^{[j-1]}, \boldsymbol{\psi}, \tau, \eta, \gamma, T, C_K$ (or ρ_K or K), N_{iter}, τ_σ and N_σ

Output: : Reconstructed image $\hat{\boldsymbol{\sigma}}^{[j]}$ (and K)

- 1: Initialize: compute $Q = n_{grid}^2$, $\mathbf{y} = \text{vec}(\hat{\mathbf{R}})$, $\epsilon_0^{[1]} = \|\mathbf{y} \odot \text{vec}(\mathbf{M}_{mask})\|_2$, $\mathbf{G}^{[j-1]} = \text{diag}\{\mathbf{g}^{[j-1]}\}$ and $\mathbf{M}_s = \mathbf{G}^{[j-1]} \mathbf{A} \odot (\mathbf{G}^{[j-1]} \mathbf{A})$; let $i = 1$, $\boldsymbol{\sigma}_s^{[0]} = \mathbf{0}$, $\epsilon_*^{[1]} = 0.95 * \epsilon_0^{[1]}$
- 2: **while** $i < N_{iter}$ **do**
- 3: Compute: $\epsilon^{[i]} = \eta \epsilon_*^{[i]}$
- 4: Compute: $\hat{\boldsymbol{\sigma}}_{s0}^{[i]}$ using FISTA
- 5: Sort sources: $\hat{\boldsymbol{\sigma}}_s^{[i]} \leftarrow \hat{\boldsymbol{\sigma}}_{s0}^{[i]}$
- 6: Compute: cumulative sum $C(\hat{\boldsymbol{\sigma}}_s^{[i]})$
- 7: **if** ρ_K or K is known **then**
- 8: $\hat{\boldsymbol{\sigma}}_s^{[i]} \leftarrow \hat{\boldsymbol{\sigma}}_s^{[i]}$ with K or with $\text{round}(\rho_K \times Q)$
- 9: **else if** C_K is known **then**
- 10: $\hat{\boldsymbol{\sigma}}_s^{[i]}, K \leftarrow \hat{\boldsymbol{\sigma}}_s^{[i]}$ with $C(\hat{\boldsymbol{\sigma}}_s^{[i]})$ through C_K
- 11: Compute: $\rho_\sigma = \|\hat{\boldsymbol{\sigma}}_s^{[i]} - \hat{\boldsymbol{\sigma}}_s^{[i-1]}\|_2 / \|\hat{\boldsymbol{\sigma}}_s^{[i-1]}\|_2$
- 12: **if** $\rho_\sigma < \tau$ **then**
- 13: break
- 14: **else**
- 15: $i = i + 1$
- 16: Compute: $\epsilon_*^{[i]} = (\epsilon_0^{[1]} - \epsilon^{[\infty]}) e^{(1-i)/2} + \epsilon^{[\infty]}$
- 17: Return $\hat{\boldsymbol{\sigma}}^{[j]} = \hat{\boldsymbol{\sigma}}_s^{[i]}$ (and K)

much detail we would like to keep after we obtain the recovered solution;

- avoid over-fitting: our initial tests on real data from a LOFAR LBA station found that the solution including this positivity requirement gave an overestimated result.

We now discuss how to choose the parameters in solving (9) using the Douglas-Rachford splitting algorithm. First is the basis or dictionary $\boldsymbol{\psi}$ for a sparse representation. Currently, we focus on a Dirac basis, which consists of delta functions each representing a source of unit amplitude on a specific grid position, and orthogonal wavelets such as Daubechies wavelets. Our experiments show that there is no apparent numerical or visual difference for the reconstructed sky source model between the Dirac representation and Daubechies representation, which might indicate that the sky source intensity distribution is already sparse enough under the Dirac basis representation. Hence, we only show the results for the Dirac basis.

The step size, which affects the rate of convergence, should, in principle, be small. However, a too small step size takes more iterations towards convergence, while a large step size may diverge out or reach a wrong solution. Hence, for simplicity, we choose a constant step size $T = 10^{-2}$. The choice of the regularisation parameter γ in (9) is very important, because it determines the degree of sparsity of the sky source model. Based on our numerical simulations, the bigger γ , the higher the sensitivity for strong sources will be achieved; we found that $\gamma = 10^{-2}$ works well for the scenarios considered in this paper.

We introduce an adaptive noise tolerance parameter based on the estimated noise level (Wijnholds & Chiarucci 2016) for the data fidelity error bound ϵ to the l_2 -ball \mathbb{S} in (9), instead of the common choice (Carrillo et al. 2012): $\epsilon = \sqrt{(2P^2 + 4P)\sigma_n^2/2}$ where $\sigma_n^2/2$ is the noise variance and

P is the number of antennas. In each iteration i , we first calculate

$$\epsilon_*^{[i]} = (\epsilon^{[1]} - \epsilon^{[\infty]}) e^{(1-i)/2} + \epsilon^{[\infty]} \quad \text{for } i > 1, \quad (10)$$

where $\epsilon^{[1]} = \|\mathbf{y} \odot \text{vec}(\mathbf{M}_{mask})\|_2$ and $\epsilon^{[\infty]} = \sqrt{\frac{1}{N} \hat{\mathbf{R}} \otimes \hat{\mathbf{R}}}$. Secondly, we implement the multiplication with η for $\eta \in (0, 1.0]$, i.e., $\epsilon^{[i]} = \eta \epsilon_*^{[i]}$, such that a fast convergence is achieved. Note that η needs to be carefully chosen for different experiments.

We propose a computationally efficient approach to determine the maximum number of true sources, such that we can select the desired level of detail in the reconstructed image by introducing a stopping criterion based on the cumulative sum of the ordered source solution energy. Once the source solution $\hat{\boldsymbol{\sigma}}$ of the sky model with our proposed imaging algorithm is available, we sort it into $\hat{\boldsymbol{\sigma}}$ and then calculate its cumulative sum curve of these ordered source energy $C(\hat{\boldsymbol{\sigma}})$. Once we have decided how much detail is needed in our estimated image, we can find the number of sources K through C_K for $C_k \in (0, 1.0]$ in this curve. If K or the coefficient ratio to the total image grids ρ_K is known, we can directly use it to threshold the source solution.

The resulting algorithm is defined in Algorithm 1, where $\text{round}(\cdot)$ is a function to get the nearest integer. The main part of Algorithm 1 is to compute an new estimate for the image, $\hat{\boldsymbol{\sigma}}_{s0}^{[i]}$, which is done by the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) (Beck & Teboulle 2009). This algorithm stops when either the maximum number of iterations N_{iter} is reached or the relative change between the successive solutions $\rho_\sigma = \|\hat{\boldsymbol{\sigma}}_s^{[i]} - \hat{\boldsymbol{\sigma}}_s^{[i-1]}\|_2 / \|\hat{\boldsymbol{\sigma}}_s^{[i-1]}\|_2$ is smaller than some bound τ for $0 < \tau < 1$. In our implementation, we fix $\tau = 10^{-4}$ and initialize $\boldsymbol{\sigma}^{[0]} = \mathbf{0}$. We fix $\tau_\sigma = 10^{-4}$ and $N_\sigma = 500$.

3.3 Calibration Step

Salvini & Wijnholds (2014) proposed a fast gain calibration algorithm based on an ADI method, named StEFCal, which only needs $\mathcal{O}(P^2)$ complexity. The superior numerical efficiency of StEFCal over traditional gain calibration approaches, which use, for example, gradient based techniques like Levenberg-Marquardt (Levenberg 1944) or weighted alternating least squares (Boonstra & van der Veen 2003; Wijnholds & van der Veen 2009), has already been verified on actual LOFAR data (Salvini & Wijnholds 2014). Hence, we adapt the StEFCal algorithm in our calibration step.

Following a similar strategy as Salvini & Wijnholds (2014), i.e., masking out the dominant sky noise in auto-correlations with \mathbf{M}_{mask} , we only use the cross-correlation elements to estimate the gains,

$$\hat{\mathbf{G}} = \arg \min_{\{\mathbf{G}\}} \|(\hat{\mathbf{R}} - \mathbf{G} \mathbf{R}_0 \mathbf{G}^H) \odot \mathbf{M}_{mask}\|_F^2, \quad (11)$$

where $\mathbf{R}_0 = \mathbf{A} \hat{\boldsymbol{\Sigma}} \mathbf{A}^H$ represents the estimated visibility. Note here $\hat{\boldsymbol{\Sigma}} = \text{diag}(\hat{\boldsymbol{\sigma}})$ is the estimated sky source model from the imaging step. Different from the general sky-model based calibration, in which usually only the bright point sources are taken into account treating all other spatial signals as noise (Boonstra & van der Veen 2003; Wijnholds & van der Veen 2009; Salvini & Wijnholds 2014; Brossard et al. 2016; Repetti et al. 2017), our method provides a more

detailed model that may include bright point sources, weak point sources and extended structures. The basic idea of StEFCal is to apply the least square technique to alternately solve for \mathbf{G}^H while holding \mathbf{G} constant, and then solve for \mathbf{G} holding \mathbf{G}^H constant. As \mathbf{R}_0 is Hermitian, the above two steps are equivalent and the iteration only involves the following step:

$$\mathbf{G}^{[i]} = \arg \min_{\{\mathbf{G}\}} \|(\hat{\mathbf{R}} - \mathbf{G}^{[i-1]} \mathbf{R}_0 \mathbf{G}^H) \odot \mathbf{M}_{mask}\|_F^2. \quad (12)$$

More details about this algorithm can be found in the original paper (Salvini & Wijnholds 2014).

4 SIMULATIONS

We evaluated the proposed algorithm using Monte Carlo simulations for a Uniform Linear Array (ULA) and a Uniform Rectangular Array (URA). To evaluate the gain calibration performance, we use the Root-Mean-Square (RMS) error of the estimated gain magnitudes and phases,

$$\epsilon_m = \sqrt{\frac{1}{P} \sum_{p=1}^P (|g_p| - |\hat{g}_p|)^2}, \quad (13a)$$

$$\epsilon_{ph} = \sqrt{\frac{1}{P} \sum_{p=1}^P (\angle g_p - \angle \hat{g}_p)^2}. \quad (13b)$$

To evaluate the image reconstruction performance, we use the measure

$$\epsilon_I = \sqrt{\frac{1}{K} \sum_{q=1}^Q \|I_q - \hat{I}_q\|_2^2}, \quad (14)$$

where K is the true number of sources and Q is the total number of pixels in the image. In the equation above, K instead of Q is used for determining the accuracy, because sparse reconstruction techniques produce a zero (instead of noise) for empty pixels, which means that using the standard definition using Q would lead to an estimated RMS error that is too low as the errors are averaged with the zeros. In the following ULA case, we have $K = 4$ and $Q = 40$.

4.1 Uniform Linear Array

The ULA consists of 20 antennas separated by half a wavelength of the narrowband source signals observed at a wavelength of $\lambda_c = 2.0$ meters. The antenna locations are $r_{x,p} = -10.5 + p$ for $p = 1, 2, \dots, 20$. Four uncorrelated sources with intensities $[0.85, 0.12, 0.56, 1.0]$ are put at the one dimensional positions (direction cosines) $l = \{-1.0, -0.4, -0.2, 0.4\}$, which all coincide with grid points as the line area $[-1.0, 1.0]$ is uniformly split into 40 grid points. We assume that: 1) all antennas have an isotropic reception pattern, i.e., $\mathbf{B} = \mathbf{I}_P$, and 2) the gain magnitudes follow a Gaussian distribution $\mathcal{N}(1.0, 0.04)$ and the gain phases follow a uniform distribution $\mathcal{U}(0, \pi/2)$ in radian, i.e., the phase varies between $[0^\circ, 90^\circ]$. Without loss of generality, we use the first antenna as phase reference. Moreover, we normalise the estimated gain magnitudes with their median in our algorithm to resolve the magnitude ambiguity between the estimated gain vector and the estimated source power vector.

Zero mean Gaussian noise is added to each antenna based on the desired SNR, whose value varies in 10 logarithmically spaced steps from -20 dB to $+20$ dB. Here, the SNR is defined as the power ratio of the average power of the four sources and the average noise power. Hence, our SNR is smaller than the commonly used ‘‘peak SNR’’. The number of samples, which corresponds to the product of bandwidth and integration time for Nyquist sampled signals, for one snapshot is 10^4 , 10^5 or 10^6 . For each SNR value and each number of samples, we do 200 runs in our Monte Carlo simulation.

We have already demonstrated the excellent accuracy and robust performance of our imaging algorithm in terms of source position and intensity for different SNRs and integration intervals in our previous paper (Wei et al. 2017). In this section, we therefore only show the results for gain calibration. Fig.1 shows the averaged RMS error and bias (as fraction of the RMS difference between estimated and true gain values) obtained from our Monte Carlo simulation as a function of SNR over 100 runs for the respective number of samples. In the figures, ‘‘NS’’ means the number of samples. These results clearly demonstrate the accuracy and robustness of our algorithm: 1) the accuracy improves with the number of samples and with SNR; 2) if the SNR is greater than -10 dB, the gains are accurately estimated, especially for the important phase information where the RMS is less than 2 degrees for 10^4 samples, 0.6 degrees for 10^5 samples and 0.2 degrees for 10^6 samples; and 3) the biases are only a small fraction of the RMS, so our estimator can be considered as unbiased within the accuracy of our simulation.

4.2 Uniform Rectangular Array

In this section, we assess the performance of our algorithm when the number of source components is not a priori known in a controlled simulated experiment for a Uniform Rectangular Array (URA). To resolve the amplitude ambiguity between gain magnitudes and source powers, we normalise the estimated gain magnitude with its l_2 -norm in our algorithm.

The URA consists of $6 \times 6 = 36$ antennas separated by half a wavelength of the narrowband source signals: the wavelength is $\lambda_c = 2.0$ meters and the antenna locations are on the grid of $u_p \in [-3.0, 3.0)$ and $v_p \in [-3.0, 3.0)$. Eight uncorrelated sources whose details are provided in Table 1, are uniformly distributed over grid points of a 10×10 grid, i.e., the true number of sources equals 8. The SNR is assumed to be -20 dB and integration is done over 10^5 samples. The gain magnitudes follow a Gaussian distribution $\mathcal{N}(1.0, 0.04)$ and the gain phases follow a uniform distribution $\mathcal{U}(0, \pi)$ in radian.

Our first experiment provides a reference case in which the true number of sources, $K = 8$, is known. With our blind calibration algorithm, we achieve a gain magnitude error of $\epsilon_m = 0.0028$ and a gain phase error of $\epsilon_{ph} = 0.93$ degrees as calculated using equation (13). If we use the following formula to calculate the SNR after calibration, where \tilde{g}_p equals to $|g_p|$ or $\angle g_p$,

$$\text{SNR} = 20 \times \log_{10} \left(\frac{\sum_{p=1}^P |\tilde{g}_p|^2}{\sum_{p=1}^P |\tilde{g}_p - \hat{\tilde{g}}_p|^2} \right),$$

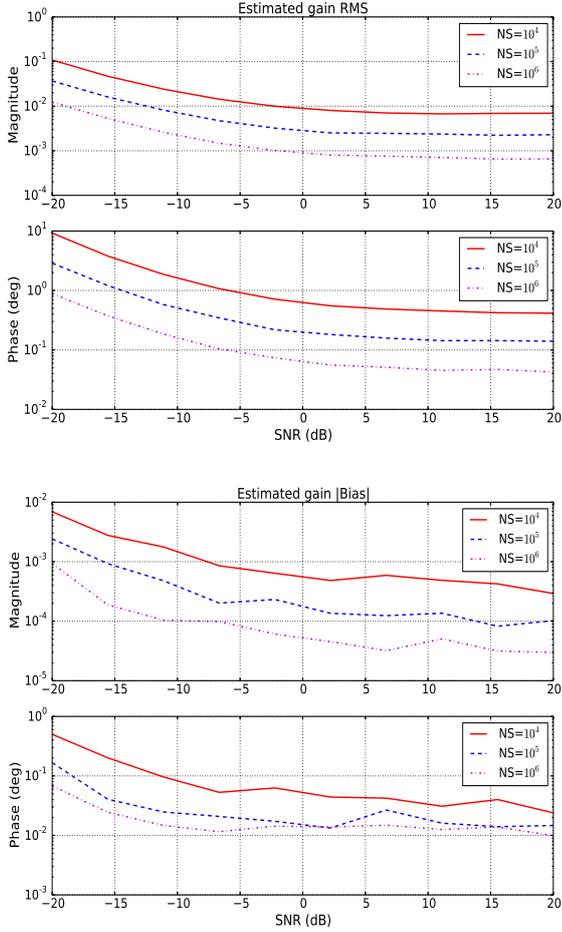


Figure 1. Monte Carlo simulation results for averaged root-mean-squared (RMS) error (top) and averaged absolute value of bias (bottom) of the estimated gain magnitudes and phases for number of samples (NS) equal to 10^4 , 10^5 and 10^6 for different SNRs.

Source Index	Intensity	Position
1	0.85	(-0.2, 0.0)
2	0.56	(0.4, -0.6)
3	0.90	(-0.8, -0.2)
4	0.30	(0.6, 0.4)
5	0.20	(-0.4, -0.4)
6	0.75	(0.2, 0.2)
7	1.00	(-0.2, 0.2)
8	0.10	(-0.6, 0.6)

Table 1. Source information for URA scenario.

we have a post-calibration SNR based on gain magnitudes and phases of 35.5 and 34.6 dB respectively. The reconstructed images and estimated gains are shown in Fig.2. Fig.2(a) shows the true image, (b) shows the normalised reconstructed image and (c) shows the difference between (a) and (b). Note that the maximum absolute value in the difference image in Fig.2(c) is less than 0.03, which tells us that all source positions are perfectly estimated and all intensities are almost correct. The most important thing is

that our algorithm correctly detects and estimates the weakest source with intensity error 0.028, whose original value is 0.10 and 10 dB below the strongest source. Fig.2(d) shows the gain estimation error between the true gain values and estimated ones. In this figure, the horizontal blue line is the mean value of the gain magnitude / phase difference. We can see that the absolute estimated error of gain magnitude is less than 0.008, while the absolute difference of gain phase is less than 3 degrees. The mean error for both gain magnitude and phase is about 0, which shows the excellent agreement between the estimated and true gain values. This clearly demonstrates the correctness of our gain estimation algorithm.

Our second experiment is to evaluate our gain calibration performance under unknown K . Therefore, we consider two cases by setting the truncation threshold value C_K of the cumulative sum of the sorted intensity estimates to $C_K = 0.65$ and $C_K = 0.90$ respectively. A low value mimics sky source model based approaches, in which only the brightest sources are used as calibrators while all other sources are treated as noise. Using our blind calibration algorithm, we obtain gain magnitude error $\epsilon_m = 0.011$ and gain phase error $\epsilon_{ph} = 4.05$ degrees for $C_K = 0.65$. For $C_K = 0.90$, we find $\epsilon_m = 0.0037$ and $\epsilon_{ph} = 1.43$ degrees. As shown in Fig.3(d) and Fig.4(d), we clearly see that: 1) the absolute gain magnitude difference is within 0.03 for $K = 3$ and 0.015 for $K = 5$; and 2) the absolute gain phase error is within 10 degrees for $K = 3$ and within 4 degrees for $K = 5$. All these errors are larger than those in the first experiment where the true number of sources was known. Furthermore, the corresponding post-calibration SNRs for gain magnitudes are 23.6 dB and 32.3 dB, respectively, while the post-calibration SNRs for gain phases are 21.9 dB and 30.9 dB, respectively. This not only shows that the post-calibration SNR improves as C_K increases, but also that the gain errors for both phase and magnitude between the estimated and true are getting smaller as C_K increases. The simulation results are shown in Fig. 3 and Fig. 4, respectively. Based on these plots, we arrive at the following conclusions:

(i) There are 3 sources for $C_K = 0.65$ and 5 sources for $C_K = 0.90$ included in the estimated images, as shown in Fig. 3(a) and Fig. 4(a). This is also reflected in Fig. 3(c) and Fig. 4(c) through the blue vertical line indicating the truncation threshold.

(ii) Although the image errors, shown as the differences between the true image in Fig.2(a) and estimated images as shown in Fig. 3(b) and Fig. 4(b), are large, the gain errors shown in Fig. 3(d) and Fig. 4(d) are relatively small. This is consistent with the practical experience that the gain estimates are most sensitive to the brightest sources in the field.

(iii) Reducing the detail level of the image source model leads to larger deviations of the estimated gain magnitudes and phase from their true values, as shown in Fig.2(d), Fig.4(d) and Fig.3(d). These figures show in increase of the maximum absolute gain magnitude estimation error from 0.008, 0.015 to 0.03 and a similar increase of the corresponding phase errors from 3.0, 4.0 to 10 degrees. We note that the gain phases are most sensitive to the structure in the image (dominated by the brightest sources), while the gain mag-

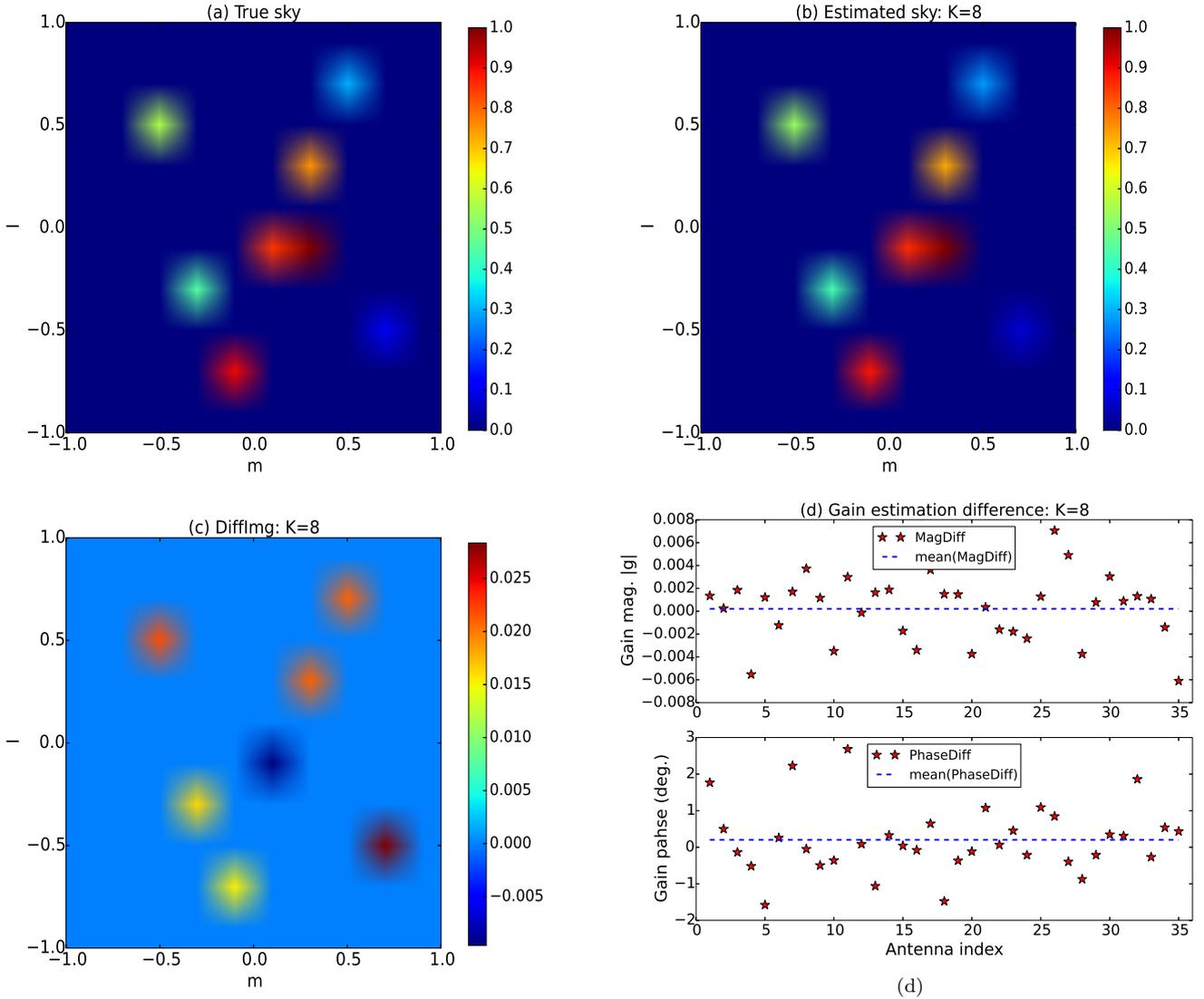


Figure 2. Source recovery for a URA for SNR=-20 dB and 10^5 samples with known source number $K = 8$. From left to right, the images are: (a) true image, (b) estimated image, (c) difference between the true and estimated images, and (d) gain estimation error between true values and estimated ones (top panel shows magnitudes and bottom panel shows phases). The l and m are the direction cosines of a point source in the source plane. The colours represent source intensity.

nitudes are most sensitive to the total power in the source model.

This experiment clearly demonstrates the robustness of our proposed blind calibration algorithm. Finally, this experiment demonstrates that a low truncation threshold value C_K may already be sufficient in many cases.

5 APPLICATION TO LOFAR DATA

In this section, we evaluate our proposed algorithm with actual data from a LOFAR station with an irregular array configuration. We ran it on a MacBook Air with a 1.6-GHz Intel Core i5 CPU and 8 GB of 1600 MHz DDR3 RAM.

We evaluate our algorithm on data from a LOFAR Low-Band Antenna (LBA) station. This station consists of 48 dual-polarisation antennas operating between 10 and 90

MHz, as shown in Fig. 5. Note that the antennas are arranged in a randomised configuration based on rings with exponentially increasing radii. The data consisted of a series of visibility matrices captured between 21:01:29 UTC and 21:10:00 UTC on July 30, 2011, during which the frequency was swept over all 195-kHz subbands with an integration time of 1 s per subband. We selected the subbands with central frequencies of 53.125 MHz and 40.039 for our experiments.

Let the image be of size 256×256 pixels. We consider different threshold values C_K to the cumulative sum curve for the two subbands. During our optimisation, we normalise the estimated gain magnitude vector with its l_2 -norm. We evaluated with a Dirac basis and Daubechies wavelets as sparse representation, but our analysis did not show a significant difference in terms of visual appearance. Hence, we only show the results based on the Dirac basis. For the Dirac

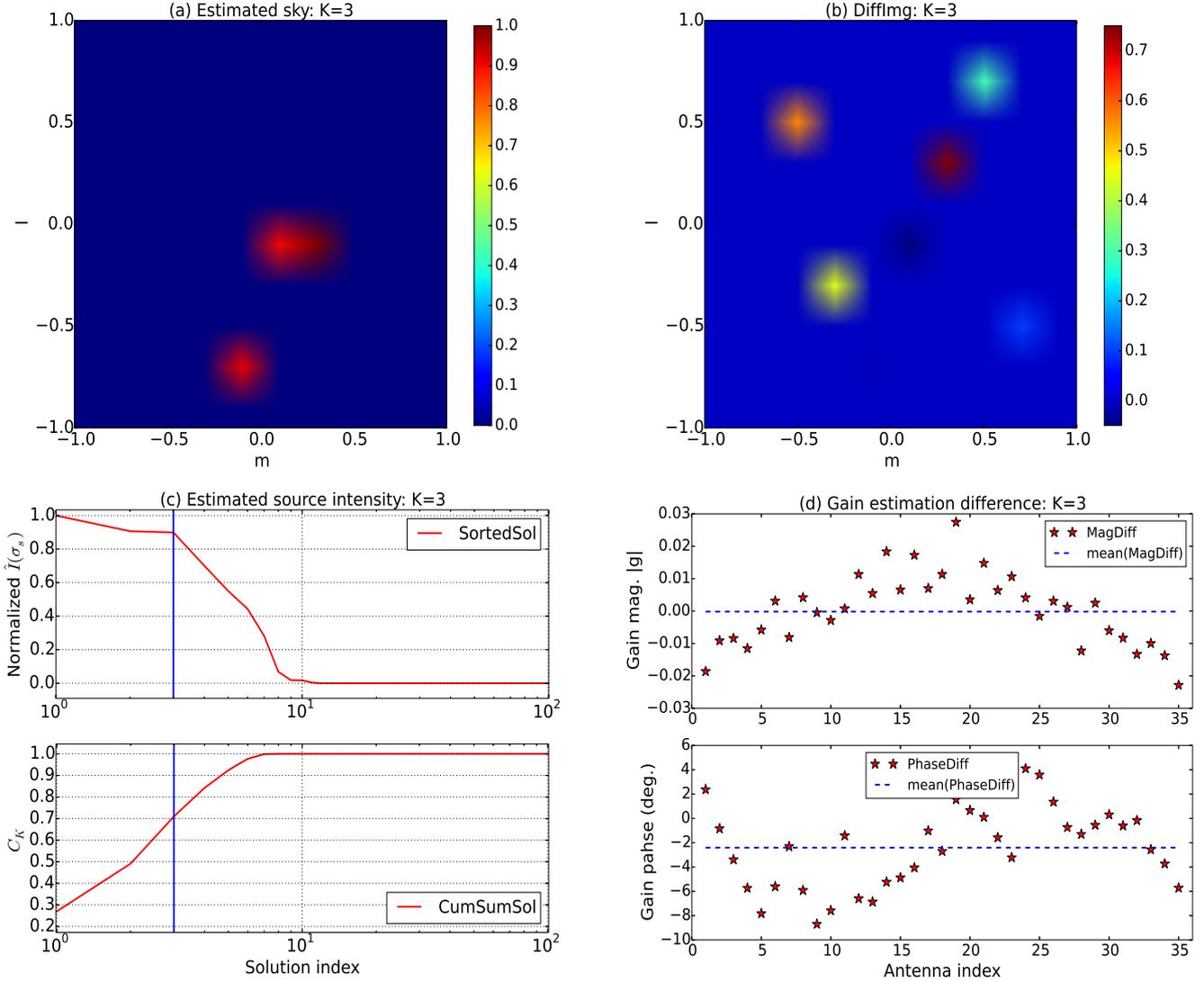


Figure 3. Source recovery for a URA for SNR=-20 dB and 10^5 samples for $C_K = 0.65$. From left to right and top to bottom, the images are: (a) estimated image, (b) difference between the true image (see Fig.2(a)) and estimated image, (c) ordered intensity estimates and cumulative sum of intensity estimates, where the vertical blue line denotes the truncation threshold value, and (d) gain estimation error between true values and estimated ones (top panel shows magnitudes and bottom panel shows phases). The l and m are the direction cosines of a point source in the source plane, while the value in the colour represents the source intensity; “solution index” refers to the rank number of sorted source solution from highest intensity to lowest, “SortedSol” refers to Sorted Solutions while “CumSumSol” stands for Cumulative Sum of Solutions.

basis and $C_K = 0.90$, our proposed algorithm took about 298 seconds with 3 iterations to stop for the 53.125-MHz subband, and about 258 seconds with 2 iterations for the 40.039-MHz subband.

First, we show the estimated sky images obtained with our proposed joint gain calibration and imaging algorithm with different thresholding values C_K in Fig. 6, while the conventional dirty images generated from the visibilities for the two subbands are shown in the top left of Fig. 9. We observe that

(i) The size of the point spread function around the bright point sources decreases as the central frequency increases as expected;

(ii) As C_K increases, more details appear in the recon-

structed images. By comparing the images for $C_K = 0.95$ for 53.125 MHz subband and $C_K = 0.90$ for 40.039 MHz with the corresponding dirty images in Fig. 9, the estimated images clearly demonstrate that our proposed imaging algorithm could reconstruct the main features: a) two bright sources, and b) the less dominant diffuse emission from the Galactic plane;

(iii) For the 53.125-MHz subband, when $C_K = 0.70$, the reconstructed image shown in Fig. 6 is dominated by two bright sources;

(iv) For the 40.039-MHz subband, when $C_K = 0.40$, the reconstructed image shown in Fig. 6 still shows a bit of diffuse emission from the Galactic plane in the bottom right besides two bright sources. This means that C_K should be decreased a little bit to keep only the two bright sources.

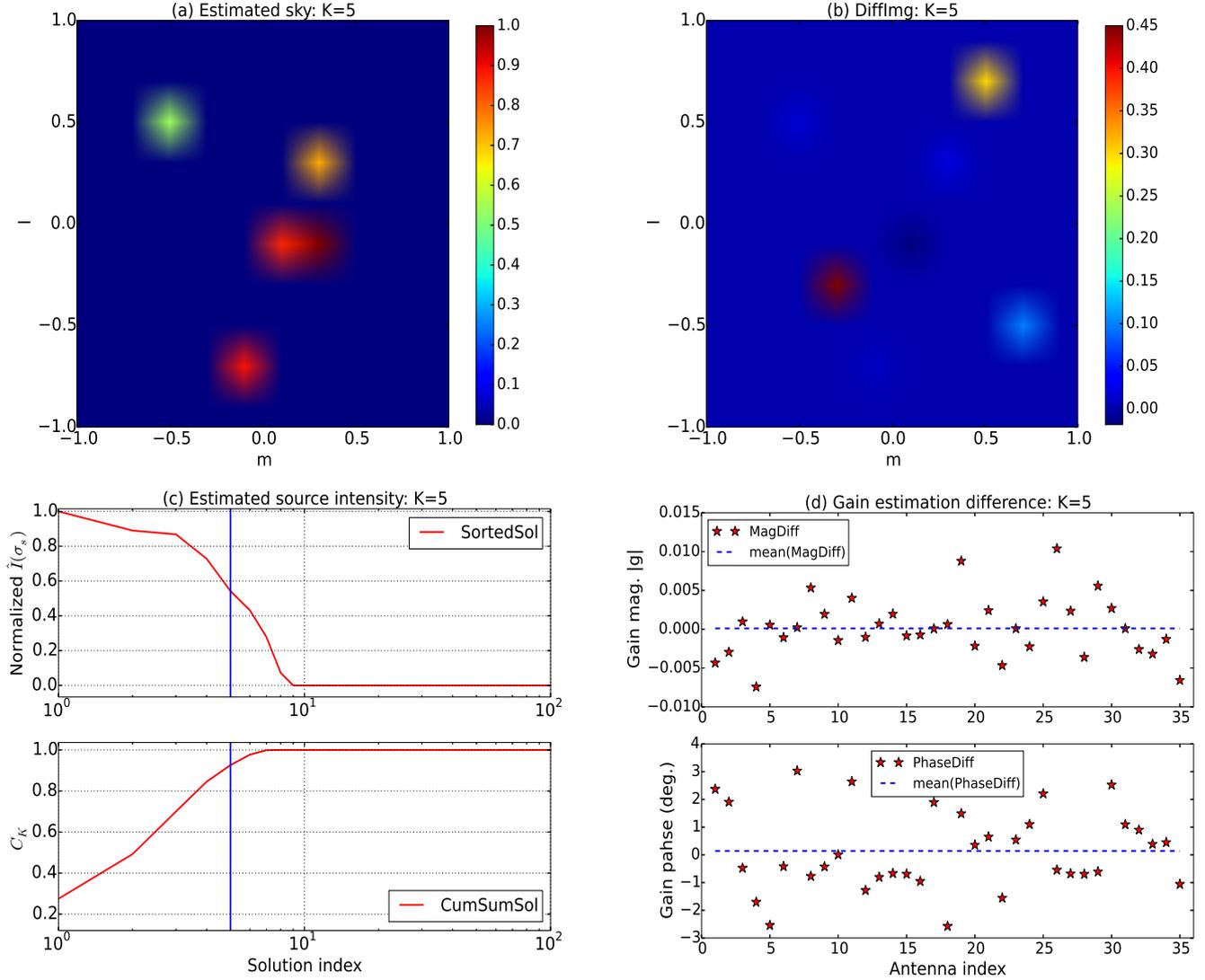


Figure 4. Source recovery for a URA for SNR=-20 dB and 10^5 samples for $C_K = 0.90$. From left to right, the images are: (a) estimated image, (b) difference between the true image (see Fig.2(a)) and estimated image, (c) ordered intensity estimates and cumulative sum of intensity estimates, where the vertical blue line denotes the truncation threshold value, and (d) gain estimation error between true values and estimated ones (top panel shows magnitudes and bottom panel shows phases). The l and m are the direction cosines of a point source in the source plane, while the value in the colour represents the source intensity; “solution index” refers to the rank number of sorted source solution from highest intensity to lowest, “SortedSol” refers to Sorted Solutions while “CumSumSol” stands for Cumulative Sum of Solutions.

This experiment clearly shows that the proposed algorithm is able to reconstruct not only those bright point sources, but also the extended structures. As demonstrated by Chiarucci & Wijnholds (2018), one advantage of blind calibration techniques, such as our proposed method, over sky source model based methods is that they will be able to detect and reconstruct an unexpected bright transient signal and / or a RFI signal at the moment that it appears, while commonly used source model based calibration may fail in such cases as these unexpected sources cannot be included in the required source model.

Next, we show the ordered intensity distribution and its cumulative sum of energy for different threshold values C_K in Fig.7. The decaying curve of $\hat{I}(\hat{\sigma})$ in this figure clearly shows three regimes: a few strong source components, hun-

dreds of intermediate source components and ten thousands of weak source components. Taking $C_K = 0.90$ as an example, for the 53.125-MHz subband, there are $K = 3996$ sources kept, which is $(3996/256^2) \times 100\% = 6.10\%$ of the total grid points. For the 40.039-MHz subband, there are 10075 sources that occupy 15.37% of the total grid points. We also note that the curves of $\hat{I}(\hat{\sigma})$ show an obvious turning point whose value corresponds to $C_K = 0.70$ for the 53.125-MHz subband and to $C_K = 0.40$ for the 40.039-MHz subband, which corresponds to the strong point sources in the sky image as shown in Fig.6. Moreover, the estimated source intensity values $\hat{I}(\hat{\sigma})$ generally get smaller as C_K increases.

Next, we show the estimated direction independent antenna gain magnitudes and phases in Fig.8, where the left

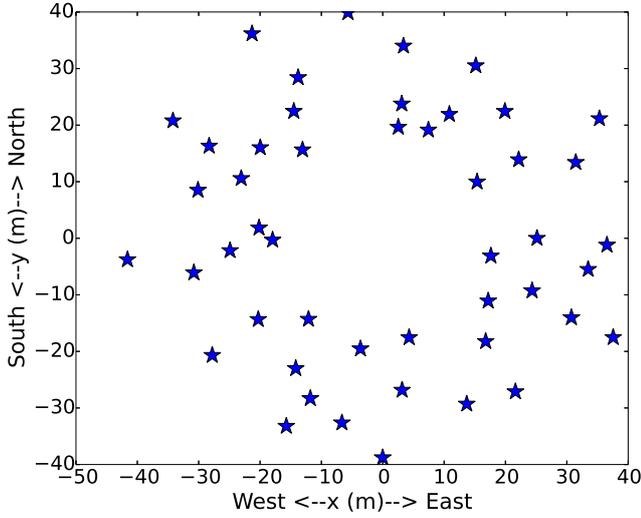


Figure 5. LOFAR station antenna layout (48 antennas).

panels are for the 53.125-MHz subband and the right panels for the 40.039-MHz subband. Observing this figure, we find that

(i) For the higher frequency subband, the estimated gain magnitudes vary a lot for different C_K while the estimated gain phases follow almost the same trend for different C_K , which demonstrates that a sky source model only including the brightest point sources is acceptable to get reasonable estimates of the gain phases. However, the maximal absolute gain phase difference between $C_K = 0.70$ and $C_K = 0.95$ is about 28.5 degrees and the RMS difference between these two threshold values is about 10.9 degrees.

(ii) For the lower frequency subband, the gain magnitudes still have large variation, but, in this case, the gain phases also vary significantly with the choice for C_K . The maximal absolute gain phase difference between $C_K = 0.40$ and $C_K = 0.90$ is about 45.8 degrees and the RMS difference between them is about 16.6 degrees. This can be explained by the fact that extended structures, such as the Galactic plane, start to play an important role as can be seen from the ordered estimated intensity distribution and its cumulative sum of energy as shown in Fig. 7. As phased array stations are getting more sensitive to the diffuse structure at lower frequencies, the gain calibration requires a more accurate sky source model with (much) more source components.

Based on our simulation results for the URA in Subsection 4.2, where we showed that the accuracy of the gain estimation improves when more sources are included in the model, we can conclude that the gain estimates for $C_K = 0.40$ and $C_K = 0.70$ in the respective subbands are less accurate than that for $C_K = 0.90$ and $C_K = 0.95$ respectively. The result suggests that a larger threshold value is required when the central frequency of the subband is lower. Moreover, this experiment indicates that our algorithm is able to calibrate a complex scene where the extended structures play a significant role. This is an attractive feature as commonly used sky source model based calibration schemes

usually struggle to construct an accurate visibility model for the short baselines.

Next, we look into the corrected visibilities $\hat{\mathbf{G}}^{-1}(\hat{\mathbf{R}} - \hat{\Sigma}_n)\hat{\mathbf{G}}^{-H}$ using the estimated gains $\hat{\mathbf{G}}$ and noise powers $\hat{\Sigma}_n = \text{diag}(\hat{\sigma}_n)$ obtained by the proposed blind calibration algorithm and obtained by sky source model based calibration, where the latter only contains two bright sources whose positions and intensities are known. Fig. 9 shows the conventional dirty images generated from the visibilities before and after gain calibration, during which all autocorrelation elements are removed. The images in Fig. 9 (a) are for the 53.125-MHz subband, while those in Fig. 9 (b) are for the 40.039-MHz subband. Let us first focus on Fig. 9(a). Comparing the dirty image in the top left, which is generated from the uncorrected visibilities, with the dirty image obtained after model based calibration (see Fig. 9 “SkymodelCal”) or after calibration using our proposed method for different threshold values clearly show that: 1) the two bright point sources stand out, 2) the point spread function around the bright sources appears less distorted; and 3) the background and side-lobes are greatly reduced. All these phenomena indicate that both sky model based calibration and our blind calibration method are successful. However, the differences between the calibrated images are quite small. The main reason is that the pixel values in these dirty images are an average over all calibrated visibilities such that the effect of residual calibration errors is smoothed. This is consistent with our simulation results presented in Subsection 4.2 and is expected given the good phase solutions as the phase solutions are most important for reconstruct the structure in the image. Similar conclusions could be drawn from dirty images in Fig. 9(b).

Finally, we show that the proposed algorithm has a reasonably fast computational speed while achieving an acceptable calibration result. Fig.10 shows the estimated source intensities and Fig.11 shows the estimated gain phases and magnitudes difference between different image resolutions $Q = n_{grid} \times n_{grid}$ for $n_{grid} = \{64, 128, 256\}$ using our proposed blind calibration with threshold value $C_K = 0.90$. The computational times and gain phase errors for the two subbands are:

- subband at 53.1 MHz:
 - (i) the running times are 48 s, 87 s and 426 s, respectively;
 - (ii) the maximum absolute difference of estimated gain phase is 3.6 degrees between $n_{grid} = 64$ and $n_{grid} = 256$ and 0.9 degrees between $n_{grid} = 128$ and $n_{grid} = 256$.
- subband at 40.039 MHz:
 - (i) the running times are 32 s, 60 s and 299 s, respectively;
 - (ii) the maximum absolute difference of estimated gain phase is 3.1 degrees between $n_{grid} = 64$ and $n_{grid} = 256$ and 0.9 degrees between $n_{grid} = 128$ and $n_{grid} = 256$.

As we scaled the gain magnitudes, there is no physical reason to discuss its exact values. However, the difference between different n_{grid} are very small as shown in the bottom of Fig. 11. The improvement in the phase solutions between $n_{grid} = 64$ and $n_{grid} = 256$ is less than 4 degrees for the *maximum* error. Given the fact that the coherence loss in a beamformer for an *RMS* phase error of 4 degrees is less than

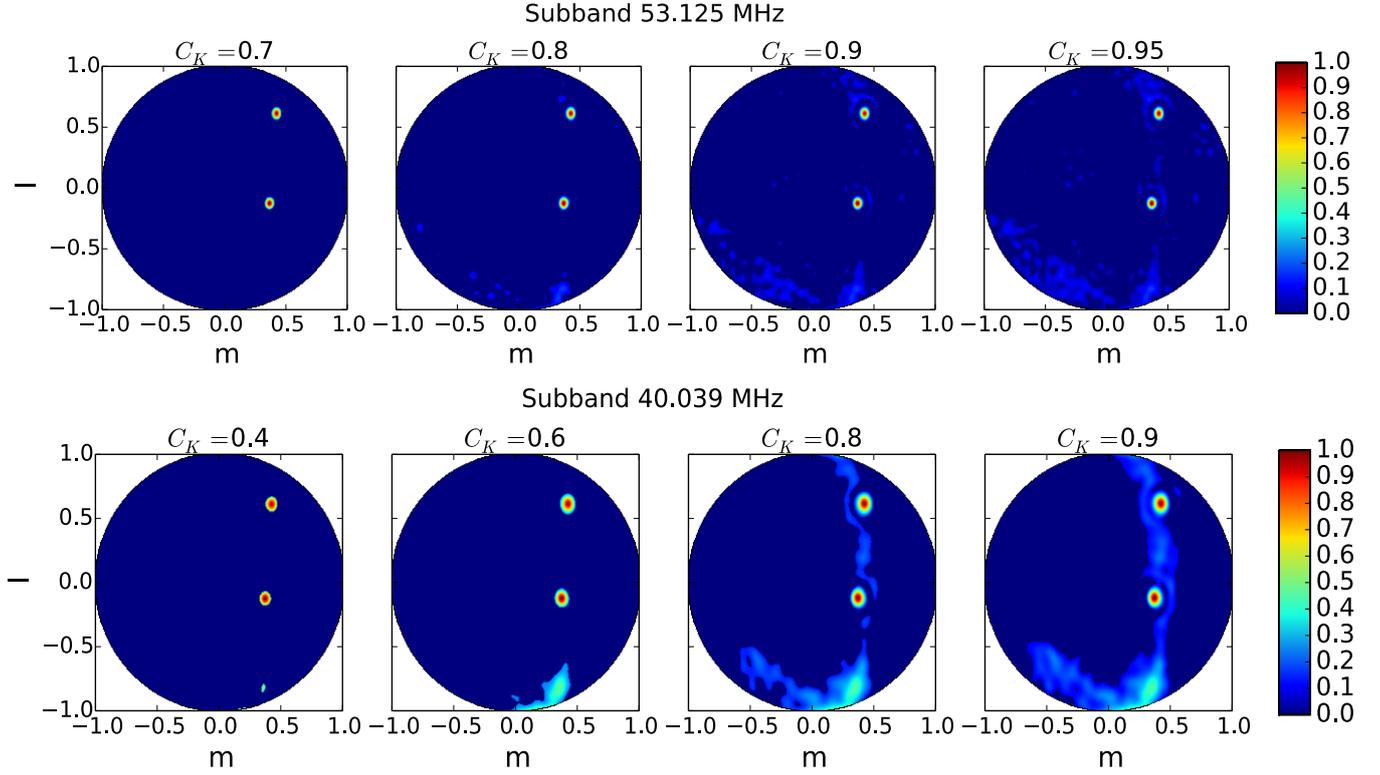


Figure 6. Estimated sky images obtained by the proposed blind calibration algorithm for two subbands with different threshold values C_K : (top) subband with central frequency 53.1 MHz; (bottom) subband with central frequency 40.039 MHz. The l and m are the direction cosines of a point source in the source plane, while the value in the colour represents the normalized intensity by the maximum intensity in each figure. All panels in each row share the same y -axis.

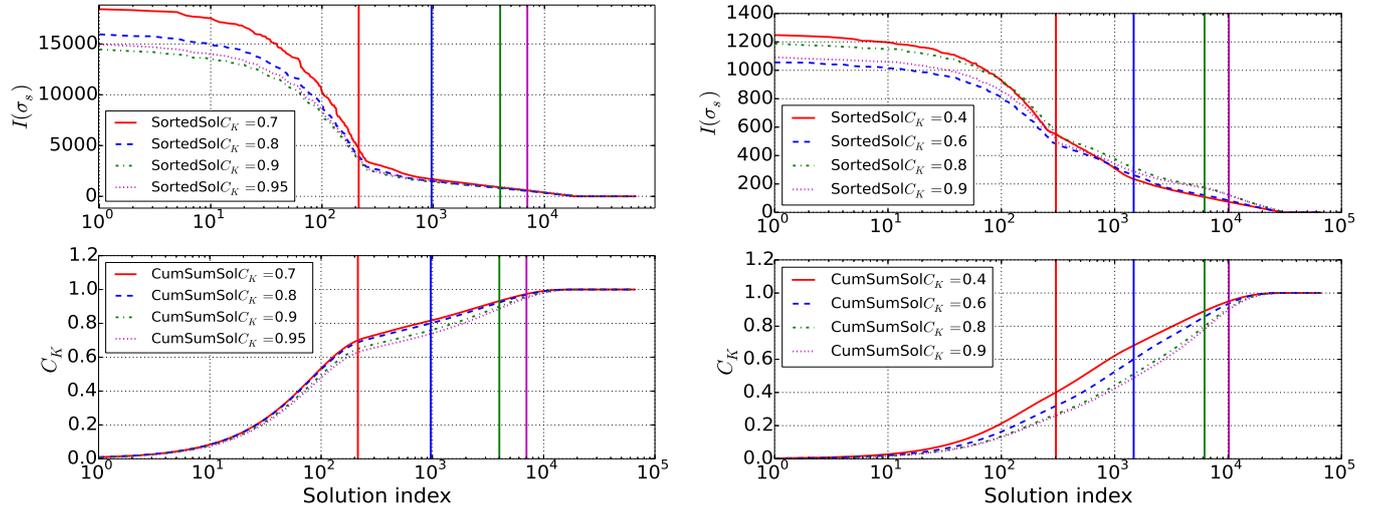


Figure 7. Estimated ordered source power solution distribution (top panels) and its cumulative sum of energy (bottom panels) for two subbands with different threshold values C_K . The vertical lines with different colour indicate truncation values to the cumulative sum curves with the same colour. Here, $I(\sigma_s)$ is actually corresponding to $I(\hat{\sigma})$, and “solution index” refers to the rank number of sorted source solution from highest intensity to lowest.

0.3%, this indicates that $n_{grid} = 64$ is likely to be sufficient in most cases.

6 CONCLUSIONS

We proposed an iterative joint blind calibration and imaging algorithm for phased array radio telescopes. In the imaging step, we estimate a sky image that may consist of bright

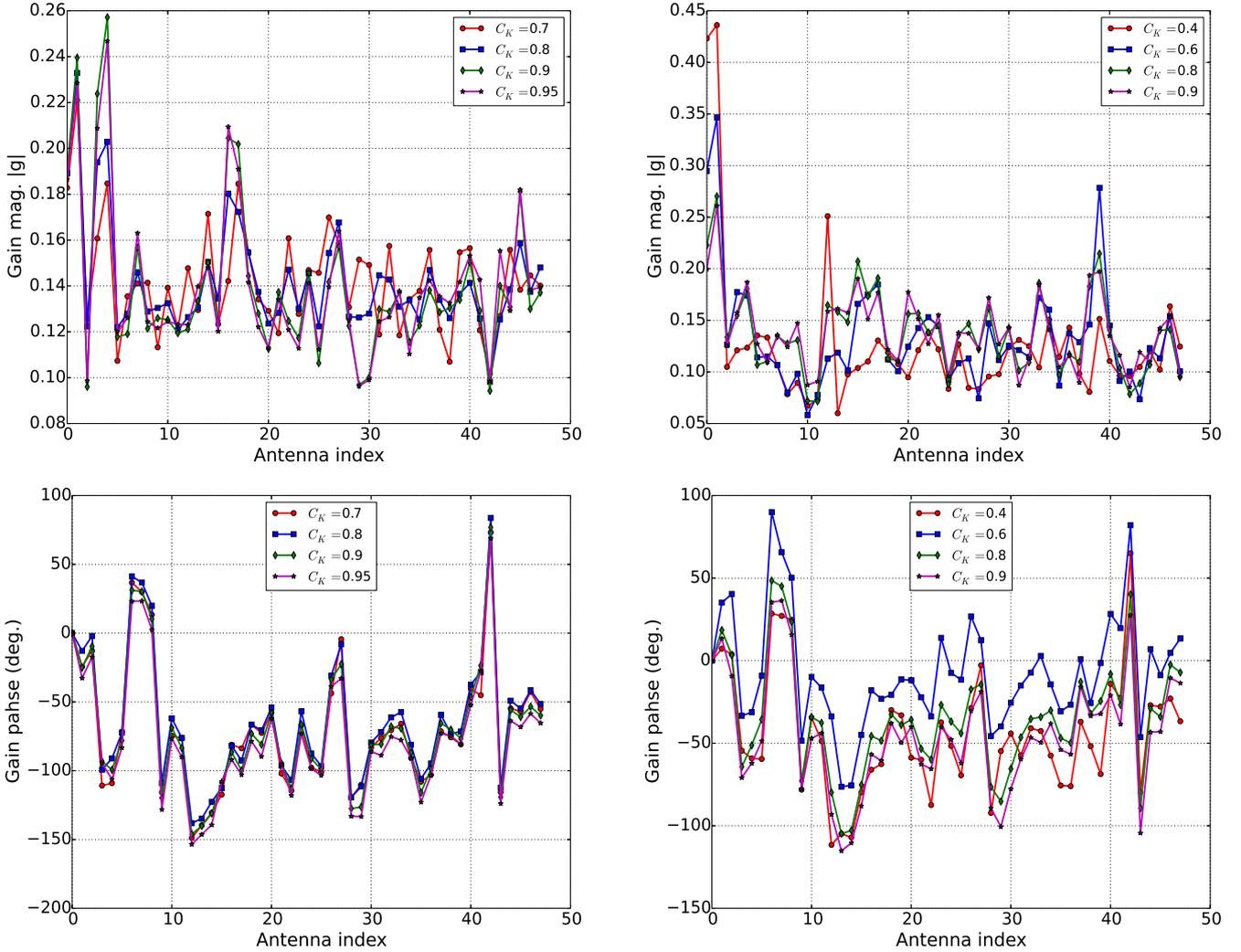


Figure 8. Estimated gain phase and magnitude obtained using the proposed blind calibration algorithm for the subbands at 53.125 MHz (left) and at 40.039 MHz (right) with different threshold values C_K : (top) estimated gain magnitudes, (bottom) estimated gain phases.

point sources, weak point sources and extended structures under the assumption that the observed scene is sparse in some suitably chosen basis. We use a Douglas-Rachford splitting algorithm to handle this Compressed Sensing (CS) based image reconstruction problem. In the calibration step, we estimate the complex-valued direction independent antenna gains using an alternating direction implicit method. We verified the excellent performance of our proposed algorithm in terms of accuracy and robustness with phase variations up to 0.9 rad RMS in simulations for regular arrays. We demonstrated its excellent performance in processing complex scenes containing extended structures with real data from an irregular array - a LOFAR LBA station. Our future work is to extend our algorithm with GPU implementation in order to fulfil the needs for on-line calibration.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Paul Hurley from the IBM Zurich Research Laboratory for inspiring discussions.

This work was supported by IBM, ASTRON, the Dutch Ministry of Economic Affairs and the Province of Drenthe as well as by the Netherlands Organisation for Scientific Research.

REFERENCES

- Beck A., Teboulle M., 2009, *SIAM Journal Imag. Sci.*, 2, 183
- Ben-David C., Leshem A., 2008, *IEEE Journal of Selected Topics in Signal Processing*, 2, 670
- Bentum M. J., Verhoeven C. J. M., Boonstra A. J., Gill E. K. A., van der Veen A. J., 2009, in *60th International Astronautical Congress A Novel Astronomical Application for Formation Flying Small Satellites*
- Bhatnagar S., Cornwell T. J., 2004, *Astronomy & Astrophysics*, 426, 747
- Boonstra A. J., van der Veen A. J., 2003, *IEEE Transactions on Signal Processing*, 51, 25
- Brossard M., El Korso M. N., Pesavento M., Boyer R.,

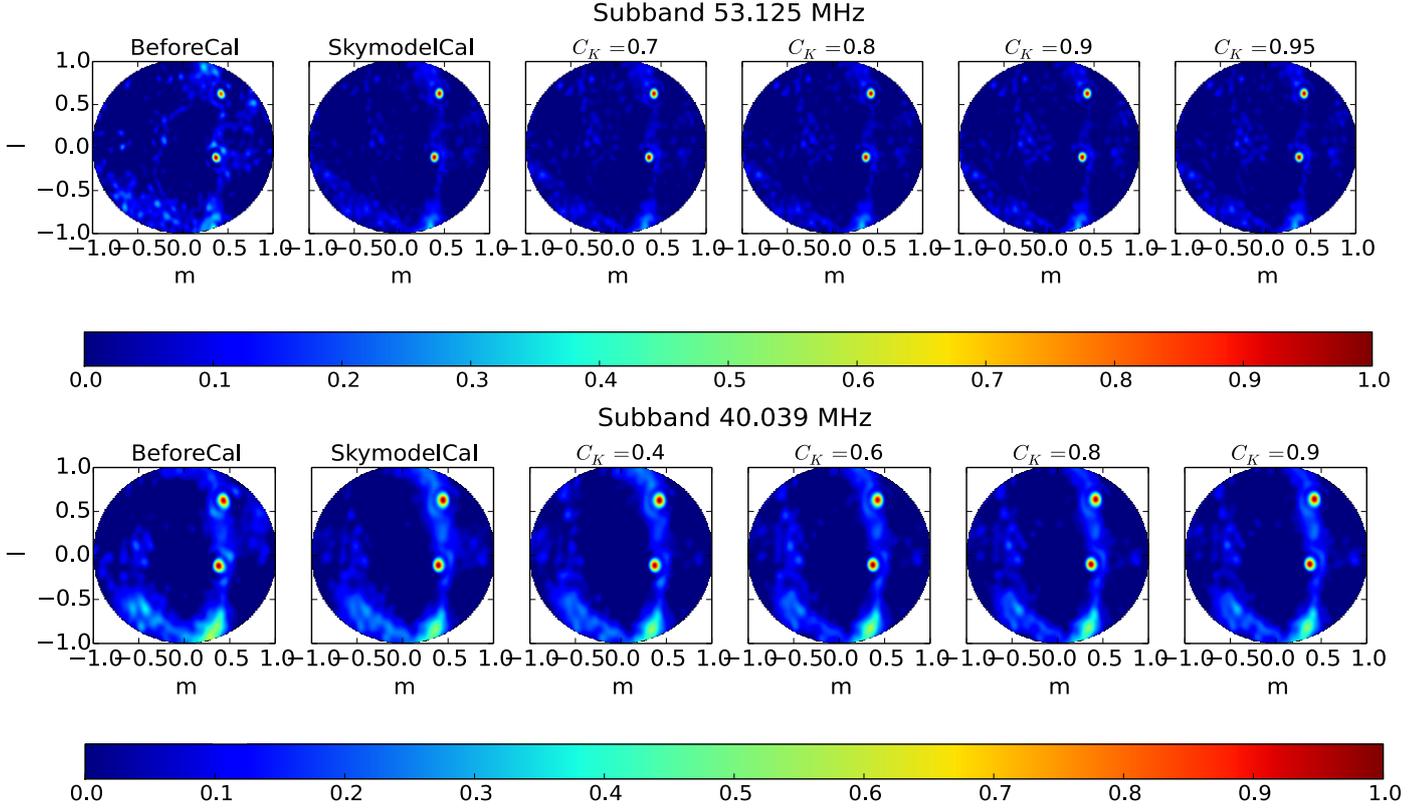


Figure 9. All-sky normalised dirty images for a LOFAR LBA station with 48 antennas for two subbands with different threshold values C_K . Note that all of the autocorrelation elements in the visibility matrix are removed before imaging. The l and m are the direction cosines of a point source in the source plane, while the value in the colour represents the normalized intensity by the maximum intensity in each figure. All panels in each row share the same y-axis.

Larzabal P., 2016, in Proc. 24th European Signal Processing Conference (EuSIPCo2016) Calibration of radio interferometers using a sparse DoA estimation framework. Budapest, Hungary, pp 275–279

Carrillo R. E., McEwen J. D., Wiaux Y., 2012, Monthly Notices of the Royal Astronomical Society, 426, 1223

Carrillo R. E., McEwen J. D., Wiaux Y., 2014, Monthly Notices of the Royal Astronomical Society, 439, 3591

Chiarucci S., Wijnholds S. J., 2018, Monthly Notices of the Royal Astronomical Society, 474, 1028

Combettes P. L., Pesquet J. C., 2007, IEEE Journal of Selected Topics on Signal Processing, 1, 1

Cornwell T. J., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793

Dabbech A., Ferrari C., Mary D., Slezak E., Smirnov O., Kenyon J. S., 2015, Astronomy & Astrophysics, 576, 1

Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, Proceedings of the IEEE, 97, 1482

Donoho D. L., 2006, Comm. Pure Appl. Math., 59, 797

Eastwood M. W., Hallinan G., 2017, in General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS) Full-Sky Maps of the VHF Radio Sky with the Owens Valley Long Wavelength Array. Montreal, Canada

Ellingson S. W., et al., 2009, Proceedings of the IEEE, 97,

1421

Flanagan B. P., Bell K. L., 2001, Signal Processing, 81, 2201

Garsden H., et al., 2015, Astronomy & Astrophysics, 575, 1

Gull S. F., Daniell G. J., 1978, Nature, 272, 686

Heidenreich P., Zoubir A. M., Rubsam M., 2012, IEEE Transactions on Signal Processing, 60, 4683

Högbom J. A., 1974, Astronomy & Astrophysics Supplement, 15, 417

Kazemi S., et al., 2015, in Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa) Blind calibration for radio interferometry using convex optimization. Pisa (Italy), pp 164–168

Leshem A., van der Veen A. J., 2000, IEEE Transactions on Information Theory, 46, 1730

Levanda R., Leshem A., 2010, IEEE Signal Processing Magazine, 27, 14

Levenberg K., 1944, The Quarterly J. of App. Math., 11, 431

Li F., Cornwell T. J., de Hoog F., 2011, Astronomy & Astrophysics, 528, 1

Narayan R., Nityananda R., 1986, Annual Review of Astronomy & Astrophysics, pp 127–170

Noordam J. E., de Bruyn A. G., 1982, Nature, 299, 597

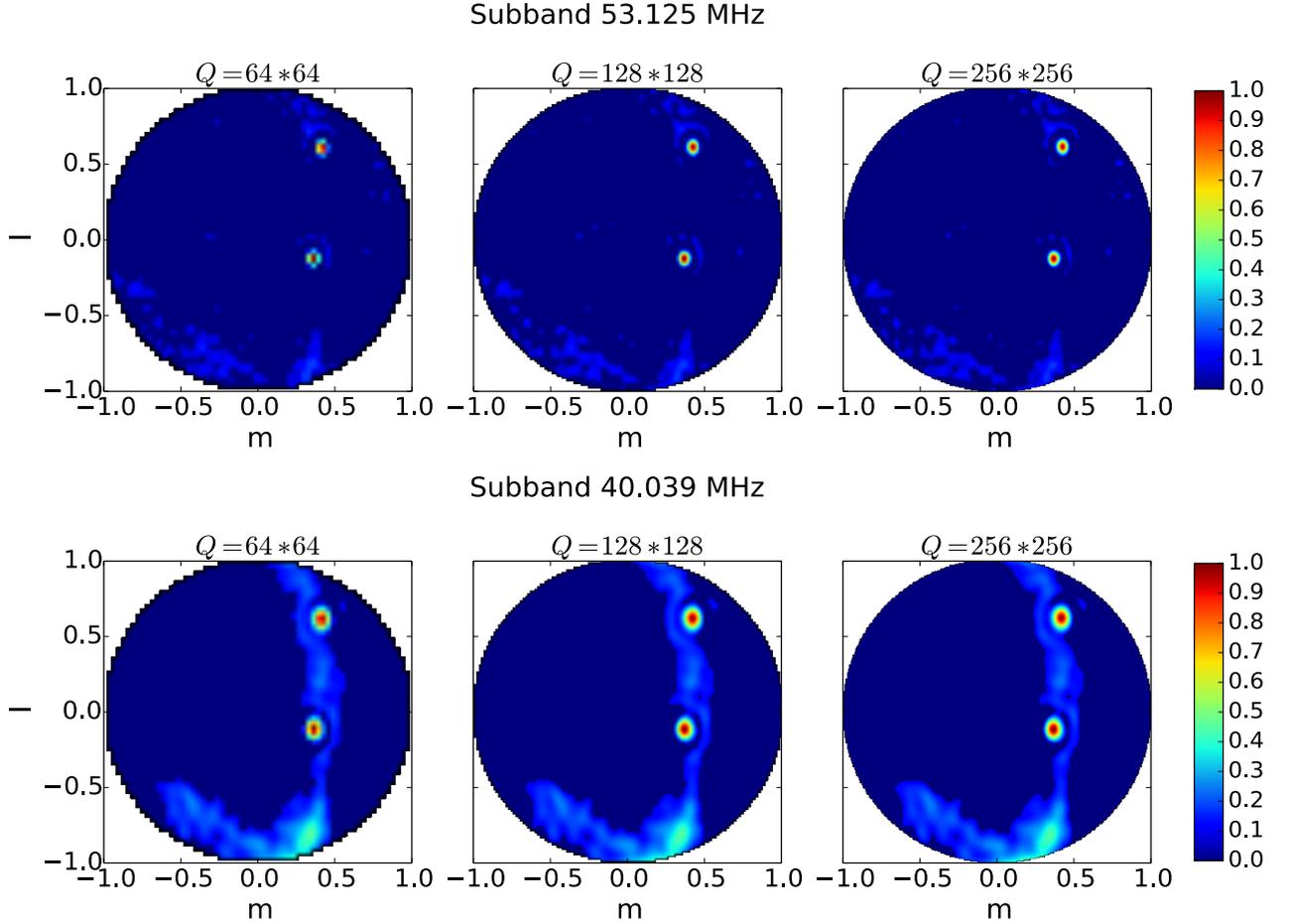


Figure 10. Estimated source intensities obtained using our proposed blind calibration algorithm for two different subbands with threshold value $C_K = 0.90$: (top) for subband 53.125 MHz, (bottom) for subband 40.039 MHz. The l and m are the direction cosines of a point source in the source plane, while the value in the colour represents the normalized intensity by the maximum intensity in each figure. Each panel in each row figure shares the same y-axis.

Noorishad P., Wijnholds S. J., van Ardenne A., van der Hulst J. M., 2012, *Astronomy & Astrophysics*, 545, 1

Onose A., Carrillo R. E., A., McEwen J. D., Thiran J. P., Pesquet J. C., Wiaux Y., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 4314

Onose A., Dabbech A., Wiaux Y., 2017, *Monthly Notices of the Royal Astronomical Society*, 469, 938

Pantin E., Starck J. L., 1996, *Astronomy & Astrophysics Supplement*, 118, 575

Prasad P., Wijnholds S. J., Huizinga F., Wijers R. A. M. J., 2014, *Astronomy & Astrophysics*, 568, 1

Pratley L., McEwen J. D., d’Avezac M., Carrillo R. E., Onose A., Wiaux Y., 2018, *Monthly Notices of the Royal Astronomical Society*, 473, 1038

Repetti A., Birdi J., Dabbech A., Wiaux Y., 2017, *Monthly Notices of the Royal Astronomical Society*, 470, 3981

Salvini S., Wijnholds S. J., 2014, *Astronomy & Astrophysics*, 571, 1

Sardarabadi A. M., Leshem A., van der Veen A. J., 2016, *Astronomy & Astrophysics*, 588

Schmidt R. O., 1986, *IEEE Transactions on Antennas and Propagation*, AP-34, 276

See C., 1995, *IEE Proc. Radar, Sonar and Navig.*, 142

Tingay S. J., et al., 2013, *Publications of the Astronomical Society of Australia*, 30, 1

Tropp J. A., 2006, *IEEE Transactions on Information Theory*, 52, 1030

van Ardenne A., Bregman J. D., van Cappellen W. A., Kant G. W., bij de Vaate J. G., 2009, *Proceedings of the IEEE*, 97, 1531

van Haarlem M. P., et al., 2013, *Astronomy & Astrophysics*, 556, 1

Viberg M., Ottersten B., 1991, *IEEE Transactions on Signal Processing*, 39, 1110

Viberg M., Ottersten B., Kailath T., 1991, *IEEE Transactions on Signal Processing*, 39, 2436

Wei L., Wijnholds S. J., Hurley P., 2017, in *IEEE International Conference on Image Processing (ICIP) Robust recovery for aperture synthesis imaging*. Beijing, China

Weiss A., Friedlander B., 1989, *IEEE Trans. Acoustics, Speech and Signal Processing*, 37

Wenger S., Magnor M., Pihlström Y., Bhatnagar S., Rau U., 2010, *Publications of the Astronomical Society of the Pacific*, 122, 1367

Wiaux Y., Jacques L., Puy G., Scaife A. M. M., Vanderghenst P., 2009, *Monthly Notices of the Royal As-*

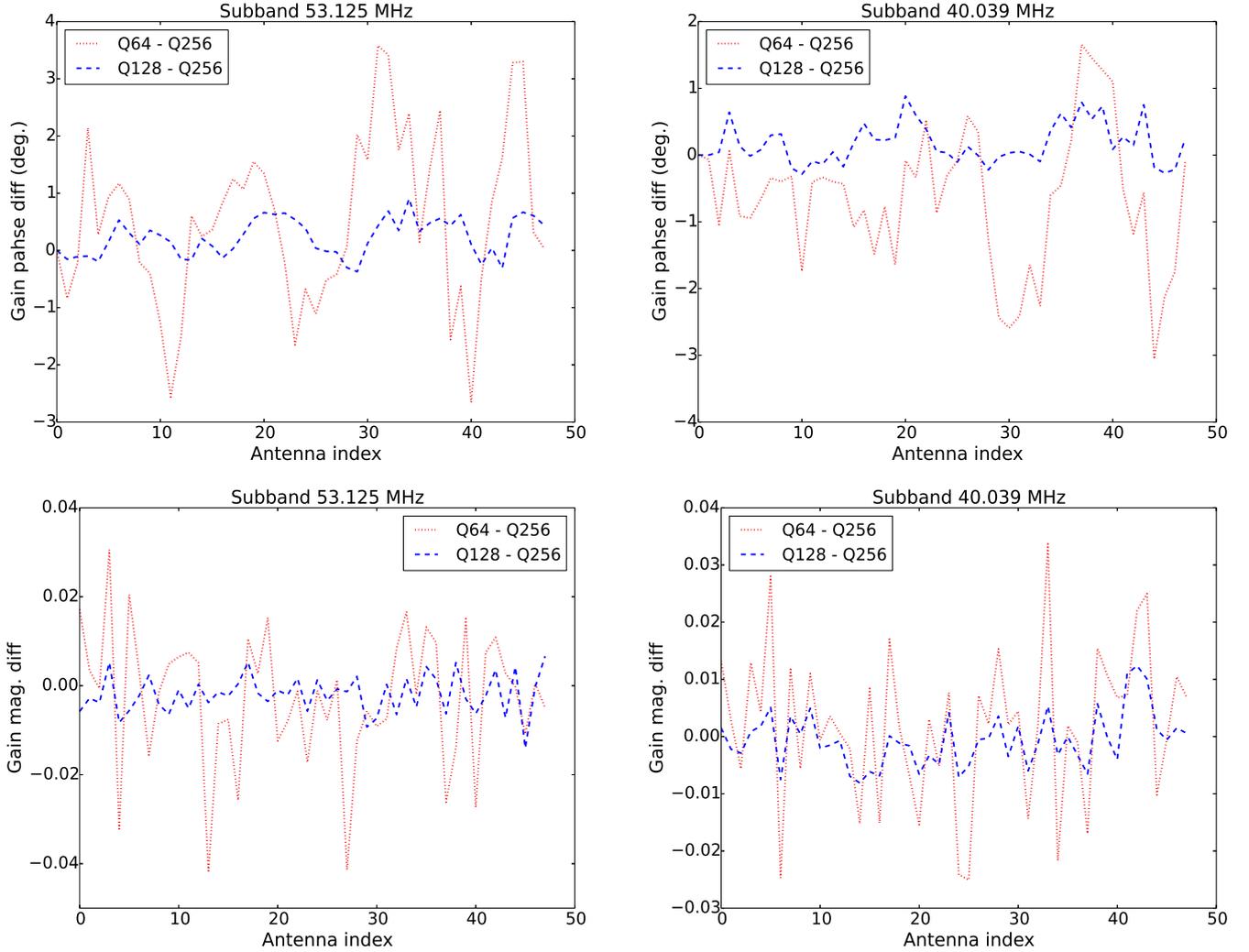


Figure 11. Estimated gain phases and magnitudes difference between different reconstructed image size Q obtained using our proposed blind calibration algorithm for two different subbands with threshold value $C_K = 0.90$: (top) estimated gain phases, (bottom) estimated gain magnitudes. The left column is for subband 53.125 MHz while the right is for 40.039 MHz. In the figures, the red curve is the difference between $Q = 64 \times 64$ and $Q = 256 \times 256$, while the blue one is for between $Q = 128 \times 128$ and $Q = 256 \times 256$.

tronomical Society, 395, 1733

Wijnholds S. J., 2016, in IEEE International Symposium on Antennas and Propagation (APS) Blind calibration of dense irregular arrays by exploiting multiply-measured nearly-identical spatial frequencies. Fajardo (Puerto Rico)

Wijnholds S. J., Chiarucci S., 2016, in 24th European Signal Processing Conference (EuSiPCo) Blind calibration of phased arrays using sparsity constraints on the signal model. Budapest (Hungary)

Wijnholds S. J., Noorishad P., 2012, in European Signal Processing Conference (EUSIPCO) Statistically Optimal Self-Calibration of Regular Imaging Arrays. Bucharest (Romania), pp 1304–1308

Wijnholds S. J., van der Veen A. J., 2009, IEEE Transactions on Signal Processing, 57, 3512