Journal of Astronomical Instrumentation

Mitigation of Non-Narrowband Radio Frequency Interference Incorporating Array Imperfections

Jan-Willem W. Steeb[1,4], David B. Davidson[2,1] and Stefan J. Wijnholds[3,1]
[1] *Department of Electrical & Electronic Engineering, Stellenbosch University, South Africa*
[2] *International Centre for Radio Astronomy Research (ICRAR), Curtin University, Australia*
[3] *Netherlands Institute for Radio Astronomy (ASTRON), The Netherlands*

Received (to be inserted by publisher); Revised (to be inserted by publisher); Accepted (to be inserted by publisher);

In a recent paper, we presented a non-narrowband spatial radio frequency interference (RFI) mitigation algorithm for radio astronomy arrays. The algorithm constructs a second order filter by combining a first order subspace subtraction method with a non-narrowband signal model. The model is based on the assumption that the frequency response is approximately flat and that the array is calibrated. In this paper, we consider the effects of array imperfections such as unknown complex gains and mutual coupling, incorporate these into the non-narrowband signal model and extend the RFI mitigation algorithm to include a calibration step. With a calibration step and no mutual coupling, the proposed algorithm was able to process twice the bandwidth per channel when compared to conventional narrowband techniques. This performance declines to 1.6 times more bandwidth when the effect of mutual coupling is included. The evaluation of the algorithm was done using the layout of a LOFAR High Band Antenna (HBA) station and a digital audio broadcast recorded with a software defined radio.

Keywords: RFI Mitigation, Non-narrowband, Complex Gain Calibration, Mutual Coupling.

1. Introduction

Radio astronomy as a passive service competes with the telecommunication industry for radio spectrum (3 Hz to 3 THz). For example, the Low Frequency Array (LOFAR) (van Haarlem *et al.* , 2013) operates from 10 to 240 MHz and overlaps with the digital audio broadcasting (DAB) band (174 to 228 MHz) in the Netherlands. In figure 1 a plot of the power spectral density for a LOFAR antenna shows a DAB signal that is present from 182.9 MHz to 184.4 MHz and is approximately 100 dB above the noise.

When a non-narrowband signal is received by an interferometric array and the visibilities are imaged it will appear as an extended source. This phenomena is called frequency smearing (Bridle & Schwab , 1999). In this paper we address how to model this smearing if it is assumed that the frequency response of the signal is relatively flat. Furthermore, we discuss an approximation of this model which makes use of two frequency-shifted narrowband point sources.

Most radio frequency interference (RFI) that is detected by radio astronomy arrays, comprises non-narrowband signals. If the channel bandwidth of the array is sufficiently narrow, then the effects of frequency smearing will be minimal. These effectively narrowband RFI sources can be removed by using spatial RFI mitigation techniques such as orthogonal projection, orthogonal projection with subspace bias correction, oblique projection and subspace subtraction (van der Tol & van der Veen , 2005; van der Veen *et al.*, 2004; Boonstra , 2005). However, if the power of the RFI signal is far above that of the astronomical sources, then it is still possible for the frequency smearing to affect the image after a spatial RFI mitigation technique has been applied. For example, in figure 2a a simulated skymap is given of a non-narrowband RFI source with a bandwidth of 195 kHz and a center frequency of 145 MHz. For the simulation the layout of a LOFAR

[4]Corresponding author.

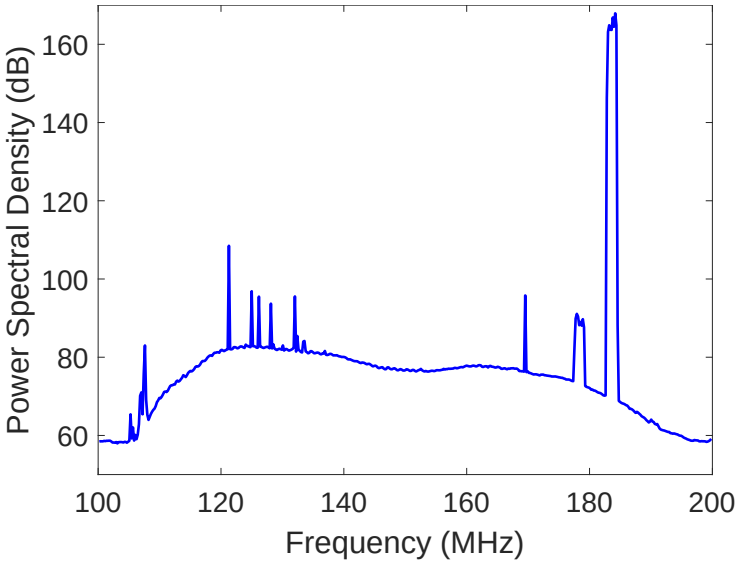

Fig. 1. Power spectral density for a LOFAR High Band Antenna (HBA) in the RS409 station. The DAB signal is clearly visible from 182.9 MHz to 184.4 MHz.

High Band Antenna (HBA) station is used. From the image it appears that the source is a narrowband point source; however, when a first order orthogonal projector is applied, this results in two sources which are separated along the radial direction (see figure 2b). These two sources are not real physical sources, but are caused by frequency smearing and, in this case, lie 40 dB below the RFI source. If the power of the astronomical sources are also in the region of -40 dB, then the effect of the RFI source cannot be fully mitigated using standard spatial RFI techniques.

The effect of array imperfections, such as unknown complex gains as well as mutual coupling, on non-narrowband signals is also considered. In figure 2c a skymap shows how unknown complex gains can distort the structure of the frequency smearing component. The mutual coupling effect is considerably more subtle for this simulation and in figure 2d only the side lobes are slightly distorted. The impact of mutual coupling will, in general, depend on the array elements and layout, as well as operating frequency.

Finally, algorithms will be presented in this paper that are able to mitigate these non-narrowband RFI signals and also take into account direction independent effects.

2. Notation

$\mathbf{A}$	Bold upper-case letters are matrices. The jk^{th} element is indicated by $\mathbf{A}_{jk}$.
$\mathbf{a}$	Bold lower-case letters are column vectors. The jth element is indicated by $\mathbf{a}_j$.
$\mathbf{I}$	Identity matrix.
$\odot$	Hadamard product.
$\ominus$	Element-wise division.
$\lvert\cdot\rvert$	Absolute value of a scalar.
$\text{Tr}(\cdot)$	Trace of a matrix.
$\text{diag}(\cdot)$	Converts a vector into a diagonal matrix.
$\angle$	Phase of a complex number.
i	Square root of -1.
c	Speed of light.
$\{\cdot\}^H$	Hermitian transpose of a matrix.
$\{\cdot\}^T$	Transpose of a matrix.
$\{\cdot\}^*$	Complex conjugate of a scalar.
$\text{sinc}(x)$	$= \sin(\pi x)/(\pi x)$, normalized sinc function.
$\lVert\cdot\rVert_F$	Frobenius norm of a matrix.
$\lvert\cdot\rvert_\odot$	Element-wise absolute value of a matrix.

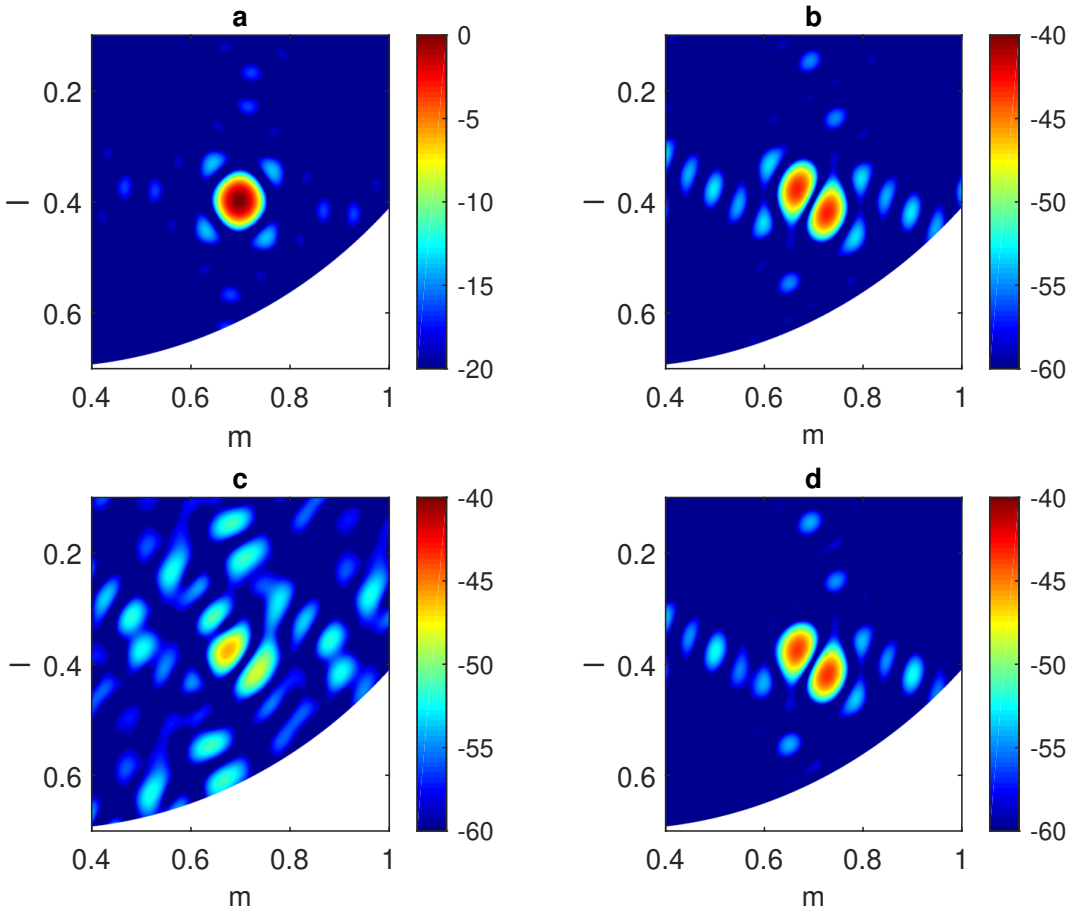

Fig. 2. The skymaps in figures **(a)** and **(b)** were originally published in Steeb *et al.* (2018). In figure **(a)** the skymap contains a non-narrowband RFI source with bandwidth of 195 kHz and the power of the RFI source is normalized to 0 dB. The result of using a first order orthogonal projection filter is shown in figure **(b)**. **(c)** Skymap with non-narrowband RFI source affected by unknown complex gains, removed using a first order orthogonal projection filter. The unknown complex gains distort the residual sources caused by the bandwidth. **(d)** Skymap with non-narrowband RFI source affected by mutual coupling, removed using a first order orthogonal projection filter. In this case, the mutual coupling slightly distorts the sidelobes of the residual sources caused by the bandwidth.

3. Signal Models

The standard narrowband signal model assumes that the received signal is monochromatic. In the flat frequency model, the narrowband signal model is adapted to be frequency dependent and have a uniform power spectral density. The Zatman's approximation based model approximates the flat frequency model using two monochromatic sources.

3.1. *Narrowband Signal Model*

This section contains a short description of the narrowband signal model for antenna arrays, see Johnson & Dudgeon (1993) for more detail. If perfectly calibrated omnidirectional antennas are used, then the normalized array response vector for an array with N_e elements and a continuous wave source with frequency ν is given by

$$\mathbf{a} = \begin{bmatrix} b_1 e^{-\mathrm{i}2\pi\nu\tau_1} \\ \vdots \\ b_{N_\mathrm{e}} e^{-\mathrm{i}2\pi\nu\tau_{N_\mathrm{e}}} \end{bmatrix}. \tag{1}$$

If the source lies in the far-field, then the geometrical delays and normalization constants are

$$\tau_j = -(l_\mathrm{s} x_j + m_\mathrm{s} y_j + n_\mathrm{s} z_j)/c, \tag{2}$$

$$b_j = 1/\sqrt{N_\mathrm{e}}, \tag{3}$$

where x_j, y_j, z_j are the Cartesian coordinates of the jth antenna and $l_\mathrm{s}, m_\mathrm{s}, n_\mathrm{s}$ are the directional cosines of the source. The array covariance matrix for a single source without noise is given by

$$\mathbf{R} = \sigma_\mathrm{s}^2 \mathbf{a}\mathbf{a}^H, \tag{4}$$

$$\mathbf{R}_{jk} = \sigma_\mathrm{s}^2 b_j b_k e^{-\mathrm{i}2\pi\tau_{jk}\nu}, \tag{5}$$

where σ_s^2 is the signal power, $\mathbf{R}_{jk}$ is the jk^th element in the covariance matrix and $\tau_{jk} = \tau_j - \tau_k$.

All examples in this paper make use of sources in the far-field. However, any of the presented methods can be used for near-field sources by setting $\tau_j = r_{\mathrm{s}j}/c$ and $b_j = 1/\left[r_{\mathrm{s}j}\sqrt{\sum_{n=1}^{N_\mathrm{e}} 1/r_{\mathrm{s}n}^2}\right]$, where the distance between the jth antenna and the source is denoted by $r_{\mathrm{s}j}$.

3.2. *Non-narrowband Signal*

In Compton *et al.* (1988), and most other literature (van der Tol & van der Veen , 2005; van der Veen *et al.*, 2004; Boonstra , 2005), a signal is described as narrowband if

$$2\pi\Delta\nu\tau_\mathrm{max} \ll 1, \tag{6}$$

where $\Delta\nu$ is the channel's bandwidth and τ_max is the delay given by the longest baseline (greatest distance between any two antennas). The LOFAR telescope's antennas are optimized to work from 10 to 240 MHz and typically have a channel bandwidth of 195 kHz. If the narrowband criterion in equation (6) is used, any signal detected by a LOFAR Low Band Antenna (LBA) or HBA core station (with longest baselines 107 m and 159 m, respectively) would be classified as narrowband. However, this narrowband classification is not always valid in the case of powerful RFI signals (see figure 1) with spectra that spans the channel. In the skymap in figure 2 b bandwidth related effects (for a 195 kHz channel width) are shown to be present 40 dB below the RFI point source component. Therefore, any RFI signal that is powerful enough for its bandwidth-related effects to interfere with the observed cosmic sources will be described as non-narrowband in this paper. The term wideband is avoided because it has a specific connotation in antenna and radio frequency engineering.

3.3. *Flat Frequency Model*

This section contains a short description of the flat frequency signal model for antenna arrays, see Thompson *et al.* (2004, p. 53) for more detail. For a non-narrowband signal, the total covariance response is obtained by integrating over the entire signal bandwidth $\Delta\nu$

$$\mathbf{R} = \frac{1}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} \sigma_\mathrm{s}^2(\nu)\mathbf{a}(\nu)\mathbf{a}^H(\nu)d\nu. \tag{7}$$

If the spectrum of the signal is flat, $\sigma_s^2(\nu) = \sigma_s^2$, then the jk^th element of the covariance matrix is given by

$$\mathbf{R}_{jk} = \frac{\sigma_\mathrm{s}^2 b_j b_k}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} e^{-\mathrm{i}2\pi\tau_{jk}\nu}d\nu = \sigma_\mathrm{s}^2 b_j b_k \underbrace{\mathrm{sinc}(\tau_{jk}\Delta\nu)}_{\text{Decorrelating Function}} e^{-\mathrm{i}2\pi\tau_{jk}\nu_0}. \tag{8}$$

The difference between the single frequency model in Eq. (5) and the flat frequency model in Eq. (8) is $\mathrm{sinc}(\tau_{jk}\Delta\nu)$, which causes $\mathbf{R}_{jk}$ to decorrelate as the bandwidth of the signal increases.

3.4. *Zatman's Approximation Based Model*

In the approximation of the flat frequency model proposed by Zatman (1998) two equal power sources that are frequency-shifted by $\kappa = \Delta\nu/(2\sqrt{3})$, are used

$$\mathbf{R} \approx \sigma^2 \mathbf{a}(\nu_0+\kappa)\mathbf{a}^H(\nu_0+\kappa) + \sigma^2 \mathbf{a}(\nu_0-\kappa)\mathbf{a}^H(\nu_0-\kappa) = \sigma^2(\mathbf{a}_1\mathbf{a}_1^H + \mathbf{a}_2\mathbf{a}_2^H), \tag{9}$$

where $\sigma^2 = 0.5\sigma_s^2$. Representing the two signals as Dirac deltas and integrating yields the total covariance matrix

$$\begin{aligned}\mathbf{R}_{jk} &\approx \frac{\sigma_s^2 b_j b_k}{2}\left[\int_{-\infty}^{\infty}\delta(\nu-\nu_0+\kappa)e^{-\mathrm{i}2\pi\tau_{jk}\nu}d\nu + \int_{-\infty}^{\infty}\delta(\nu-\nu_0-\kappa)e^{-\mathrm{i}2\pi\tau_{jk}\nu}d\nu\right] \\ &= \sigma_s^2 b_j b_k \underbrace{\cos\left(\pi\tau_{jk}\Delta\nu/\sqrt{3}\right)}_{\text{Decorrelating Function}} e^{-i2\pi\tau_{jk}\nu_0}.\end{aligned} \tag{10}$$

The only difference between the flat frequency model and the Zatman's approximation is the decorrelating function (see Eqs. (8) and (10)). The error between the two decorrelating functions is small for the peak around $\Delta\nu = 0$. This is shown in figure 3 using LOFAR HBA station RS407 in a simulation with a center frequency of 145 MHz and varying the signal bandwidth between 0 and 2 MHz for both models. In figures 3a and b the percentage difference between the eigenvalues and one minus the cosine similarity between the eigenvectors are respectively given between the flat frequency model and the Zatman's approximation based model as a function of fractional bandwidth. The cosine similarity is the cosine of the angle between two vectors. These plots clearly show that there is minimal error between the models even if the typical channel bandwidth of 195 kHz is increased tenfold.

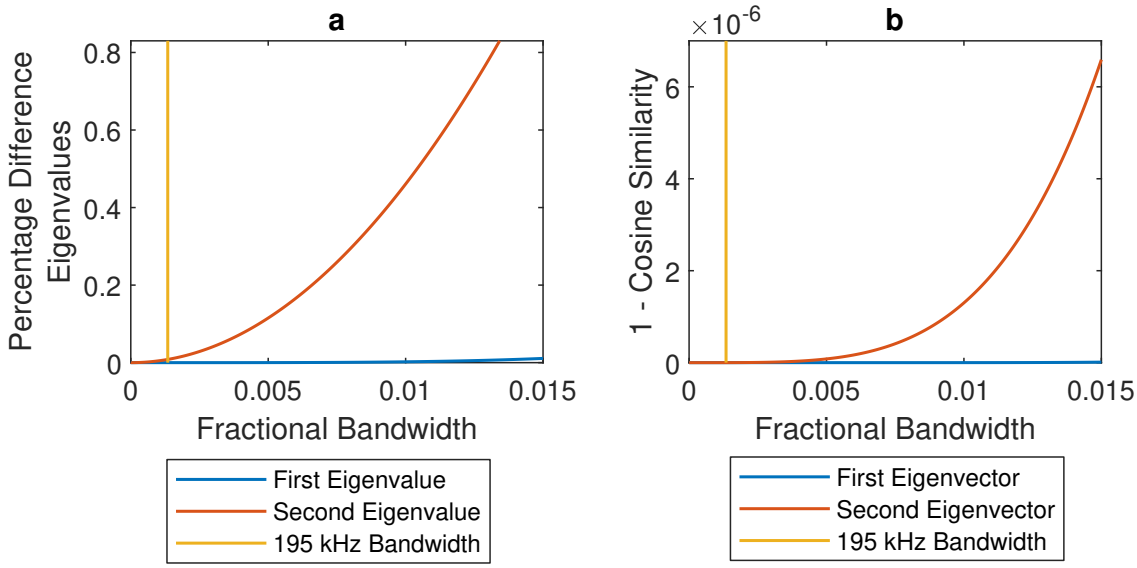

Fig. 3. Results from simulation using LOFAR HBA station RS407 at a center frequency of 145 MHz. **(a)** Percentage difference between the corresponding eigenvalues of the flat frequency model and the Zatman's approximation based model. **(b)** One minus the cosine similarity between the corresponding eigenvectors of the flat frequency model and the Zatman's approximation based model as a function of fractional bandwidth. A score of zero indicates that the eigenvectors are parallel.

The covariance matrix for the Zatman's approximation $\mathbf{R}_z$ can be written in terms of its eigenvalue decomposition

$$\mathbf{R}_z = \lambda_1\mathbf{v}_1\mathbf{v}_1^H + \lambda_2\mathbf{v}_2\mathbf{v}_2^H, \tag{11}$$

where λ_1 and λ_2 are the largest and second largest eigenvalues and $\mathbf{v}_1$ and $\mathbf{v}_2$ are the corresponding eigenvectors. Both the eigenvectors and eigenvalues can be written in terms of the Zatman's approximation

sources' power σ^2 and array response vectors $\mathbf{a}_1$ and $\mathbf{a}_2$

$$\lambda_{1,2} = \sigma^2[1 \pm |\psi|], \tag{12}$$

$$\psi = \mathbf{a}_1^H \mathbf{a}_2 = \frac{1}{N_\mathrm{e}} \sum_{p=1}^{N_\mathrm{e}} e^{i4\pi\kappa\tau_p}, \tag{13}$$

$$\mathbf{v}_{1,2} = \frac{1}{\sqrt{2(1 \pm |\psi|)}} \left[\mathbf{a}_1 \pm \frac{\psi^*}{|\psi|} \mathbf{a}_2 \right]. \tag{14}$$

A detailed derivation of these equations can be found in Steeb *et al.* (2018). The covariance matrix for each subspace is then given by

$$\mathbf{R}_1 = \mathbf{v}_1 \mathbf{v}_1^H = \frac{1}{2(1+|\psi|)} \left[\mathbf{a}_1 \mathbf{a}_1^H + \frac{\psi^*}{|\psi|} \mathbf{a}_2 \mathbf{a}_1^H + \frac{\psi}{|\psi|} \mathbf{a}_1 \mathbf{a}_2^H + \mathbf{a}_2 \mathbf{a}_2^H \right], \tag{15}$$

$$\mathbf{R}_2 = \mathbf{v}_2 \mathbf{v}_2^H = \frac{1}{2(1-|\psi|)} \left[\mathbf{a}_1 \mathbf{a}_1^H - \frac{\psi^*}{|\psi|} \mathbf{a}_2 \mathbf{a}_1^H - \frac{\psi}{|\psi|} \mathbf{a}_1 \mathbf{a}_2^H + \mathbf{a}_2 \mathbf{a}_2^H \right]. \tag{16}$$

Expanding the jk^{th} element for each matrix, using $(\mathbf{a}_1)_j = e^{-i2\pi(\nu_0+\kappa)\tau_j}/\sqrt{N_\mathrm{e}}$, $(\mathbf{a}_2)_j = e^{-i2\pi(\nu_0-\kappa)\tau_j}/\sqrt{N_\mathrm{e}}$ and $\psi = \mathbf{a}_1^H \mathbf{a}_2 = 1/N_\mathrm{e} \sum_{p=1}^{N_\mathrm{e}} e^{i4\pi\kappa\tau_p}$, yields

$$\begin{aligned}(\mathbf{R}_1)_{jk} &= \frac{1}{2(1+|\psi|)} \left[e^{-i2\pi\kappa\tau_{jk}} + \frac{\psi^*}{|\psi|} e^{i2\pi\kappa(\tau_j+\tau_k)} + \frac{\psi}{|\psi|} e^{-i2\pi\kappa(\tau_j+\tau_k)} + e^{i2\pi\kappa\tau_{jk}} \right] \frac{e^{-i2\pi\nu_0\tau_{jk}}}{N_\mathrm{e}} \\ &= \underbrace{\frac{1}{(1+|\psi|)} \left[\cos(2\pi\kappa\tau_{jk}) + \frac{1}{N_\mathrm{e}|\psi|} \sum_{p=1}^{N_\mathrm{e}} \cos(2\pi\kappa(\tau_{pj}+\tau_{pk})) \right]}_{\text{Positive Real Value if } 2\pi\kappa\tau_{jk}<\pi/2 \text{ and } 2\pi\kappa(\tau_{pj}+\tau_{pk})<\pi/2} \underbrace{\frac{e^{-i2\pi\nu_0\tau_{jk}}}{N_\mathrm{e}}}_{\text{Point Source Model}},\end{aligned} \tag{17}$$

$$\begin{aligned}(\mathbf{R}_2)_{jk} &= \frac{1}{2(1-|\psi|)} \left[e^{-i2\pi\kappa\tau_{jk}} - \frac{\psi^*}{|\psi|} e^{i2\pi\kappa(\tau_j-\tau_k)} - \frac{\psi}{|\psi|} e^{-i2\pi\kappa(\tau_j+\tau_k)} + e^{i2\pi\kappa\tau_{jk}} \right] \frac{e^{-i2\pi\nu_0\tau_{jk}}}{N_\mathrm{e}} \\ &= \underbrace{\frac{1}{(1-|\psi|)} \left[\cos(2\pi\kappa\tau_{jk}) - \frac{1}{N_\mathrm{e}|\psi|} \sum_{p=1}^{N_\mathrm{e}} \cos(2\pi\kappa(\tau_{pj}+\tau_{pk})) \right]}_{\text{Real Value}} \underbrace{\frac{e^{-i2\pi\nu_0\tau_{jk}}}{N_\mathrm{e}}}_{\text{Point Source Model}}.\end{aligned} \tag{18}$$

For both $(\mathbf{R}_1)_{jk}$ and $(\mathbf{R}_2)_{jk}$ the complex exponential part is the narrowband point source model for the signal. The real part of $(\mathbf{R}_1)_{jk}$ is positive, therefore the phase of the point source is not changed and its structure is maintained (see figure 2a). For $(\mathbf{R}_2)_{jk}$ the real part can be negative, which causes a π-rad phase shift. This phase shift in some of the elements of $\mathbf{R}_2$ causes the single point source structure to change to two adjacent sources (see figure 2b).

4. Unknown Complex Gain Model

In section 3.1 the assumption was made that there are no array imperfections or external effects on the signal. However, in reality signals are affected by direction dependent effects such as path loss, the polarization mismatch factor, the gain of the antennas and atmospheric effects. Furthermore, direction independent effects include the receiver electronics, in particular the low noise amplifiers. Direction independent and dependent effects can be combined into one unknown complex gain per antenna if only one non-narrowband RFI signal is assumed to be impinging on the array for a given band of the spectrum.

The Zatman's model will now be expanded to include unknown complex gains, $\mathbf{g} = [g_1 \dots g_{N_\mathrm{e}}]^T$. The two discrete signals from Eq. (9) are now given by

$$\mathbf{a}_{1\mathrm{g}} = \sqrt{N_\mathrm{e}}(\mathbf{a}_1 \odot \mathbf{g}), \tag{19}$$

$$\mathbf{a}_{2\mathrm{g}} = \sqrt{N_\mathrm{e}}(\mathbf{a}_2 \odot \mathbf{g}). \tag{20}$$

The complex gain vector $\mathbf{g}$ is defined as normalized and the inverse of the normalization factor is absorbed into the signal's eigenvalues (the ratio between the eigenvalues remains the same). Since the array response vectors $\mathbf{a}_{1,2}$ represent plane waves (where each element is a phasor multiplied by $1/\sqrt{N_\mathrm{e}}$) the factor $\sqrt{N_\mathrm{e}}$ is introduced to normalize both $\mathbf{a}_{1\mathrm{g}}$ and $\mathbf{a}_{2\mathrm{g}}$ in Eq. (19) and (20).

Eq. (12) which describes the relationship between the signal's power and eigenvalues, is then given by

$$\lambda_{1g,2g} = \sigma_g^2[1 \pm |\psi_g|], \tag{21}$$

where $\psi_g = \mathbf{a}_{1\mathrm{g}}^H \mathbf{a}_{2\mathrm{g}} = \sum_{p=1}^{N_\mathrm{e}} |g_p|^2 e^{i4\pi\kappa\tau_p}$ and σ_g^2 is the signal's power multiplied by the inverse of the normalization factor. The normalized covariance matrix for the first subspace (see Eq. (15)), including the effect of $\mathbf{g}$, can be expressed as

$$\begin{aligned}\mathbf{R}_{1\mathrm{g}} = \mathbf{v}_{1\mathrm{g}}\mathbf{v}_{1\mathrm{g}}^H &= \frac{1}{2(1+|\psi_\mathrm{g}|)}\left[\mathbf{a}_{1\mathrm{g}}\mathbf{a}_{1\mathrm{g}}^H + \frac{\psi_\mathrm{g}^*}{|\psi_\mathrm{g}|}\mathbf{a}_{2\mathrm{g}}\mathbf{a}_{1\mathrm{g}}^H + \frac{\psi_\mathrm{g}}{|\psi_\mathrm{g}|}\mathbf{a}_{1\mathrm{g}}\mathbf{a}_{2\mathrm{g}}^H + \mathbf{a}_{2\mathrm{g}}\mathbf{a}_{2\mathrm{g}}^H\right] \\ &= \frac{N_\mathrm{e}}{2(1+|\psi_\mathrm{g}|)}\left[\mathbf{a}_1\mathbf{a}_1^H + \frac{\psi_\mathrm{g}^*}{|\psi_\mathrm{g}|}\mathbf{a}_2\mathbf{a}_1^H + \frac{\psi_\mathrm{g}}{|\psi_\mathrm{g}|}\mathbf{a}_1\mathbf{a}_2^H + \mathbf{a}_2\mathbf{a}_2^H\right] \odot \mathbf{g}\mathbf{g}^H,\end{aligned} \tag{22}$$

where

$$(\mathbf{R}_{1\mathrm{g}})_{jk} = \underbrace{\frac{N_\mathrm{e}}{(1+|\psi_g|)}\left[\cos(2\pi\kappa\tau_{jk}) + \frac{1}{N_\mathrm{e}|\psi_g|}\sum_{p=1}^{N_\mathrm{e}}|g_p|^2\cos(2\pi\kappa(\tau_{pj}+\tau_{pk}))\right]}_{\text{Complex Gain Distorted Non-Narrowband Component}} \underbrace{g_j g_k^H}_{\text{Complex Gain}} \underbrace{\frac{e^{-i2\pi\nu_0\tau_{jk}}}{N_\mathrm{e}}}_{\text{Point Source Model}}. \tag{23}$$

The first component of $(\mathbf{R}_{1\mathrm{g}})_{jk}$ is caused by the non-narrowband nature of the signal and is distorted by the complex gains, since ψ is replaced by ψ_g. Furthermore, the complex gain distorted non-narrowband component is still real valued like the undistorted non-narrowband component (see Eq. (17)). The other two components form the model for a point source that is distorted by complex gains.

To explore the effect of the complex gains on the non-narrowband factor in $(\mathbf{R}_{1\mathrm{g}})_{jk}$, a simulation was done using the layout of LOFAR HBA station RS407 with a channel bandwidth of 195 kHz and an RFI source at the horizon. To calculate $\mathbf{g}$, a normal distribution was used to generate phases and a uniform distribution was used to generate amplitudes. Different distributions are used to reflect the distinct physical effects that lead to amplitude gains and phase errors (Brossard *et al.*, 2018; de Gasperin *et al.*, 2018). Furthermore, the uniform distribution used ensured that the generated amplitudes are positive. The normal distribution used had mean 0 and the standard deviation was varied between 0 and $\pi/3$ radians. At the final standard deviation of $\pi/3$, 99.73 % of the values lie within a π-rad band around the mean. The uniform distribution used had an upper bound fixed at 1 and the lower bound was varied from 1 to 0.0001, giving a standard deviation ranging from 0 to 0.29. The normal and uniform distributions were paired according to increasing standard deviation and for each pair 1000 realizations for each antenna were generated. Figure 4a shows plots of $1 - |\psi_g|$ where the maximum and minimum values for $|\psi_g|$ were found for the 1000 realizations as a function of the distributions' standard deviations. Even in the worst case, the distorted $|\psi_g|$ deviates only by 0.004 %. The largest percentage error simulated for the non-narrowband factor (see Eq. (23)) for each element in the covariance matrix is plotted in figure 4b. This plot was obtained by fixing the standard deviation of the normal distribution and uniform distribution to $\pi/3$ and 0.29, respectively. For the 1000 realizations used, the largest error for each element in the covariance matrix was selected for the plot. The pattern obtained is a function of the position of the RFI source. The maximum error in the plot is only 0.006 %. Therefore, for the LOFAR HBA case, the non-narrowband factor is minimally affected by the complex gains and the complex gains can be recovered by using $\mathbf{g}\mathbf{g}^H \approx \mathbf{R}_{1g} \ominus \mathbf{R}_1$ or calibration algorithms such as StEFCal (Salvini & Wijnholds , 2014).

Once the complex gains $\mathbf{g}$ have been recovered, the flat frequency model in Eq. (8) can be updated to include its effect

$$(\mathbf{R}_g)_{jk} = \frac{\sigma_\mathrm{s}^2 g_j g_k}{N_\mathrm{e}}\,\mathrm{sinc}(\tau_{jk}\Delta\nu)e^{-\mathrm{i}2\pi\tau_{jk}\nu_0}. \tag{24}$$

The Zatman's approximation in Eq. (14) can also be updated to

$$\mathbf{v}_{1g} = \frac{1}{\sqrt{2(1 \pm |\psi_g|)}} \left[\mathbf{a}_{1g} \pm \frac{\psi_g^*}{|\psi_g|} \mathbf{a}_{2g} \right]. \tag{25}$$

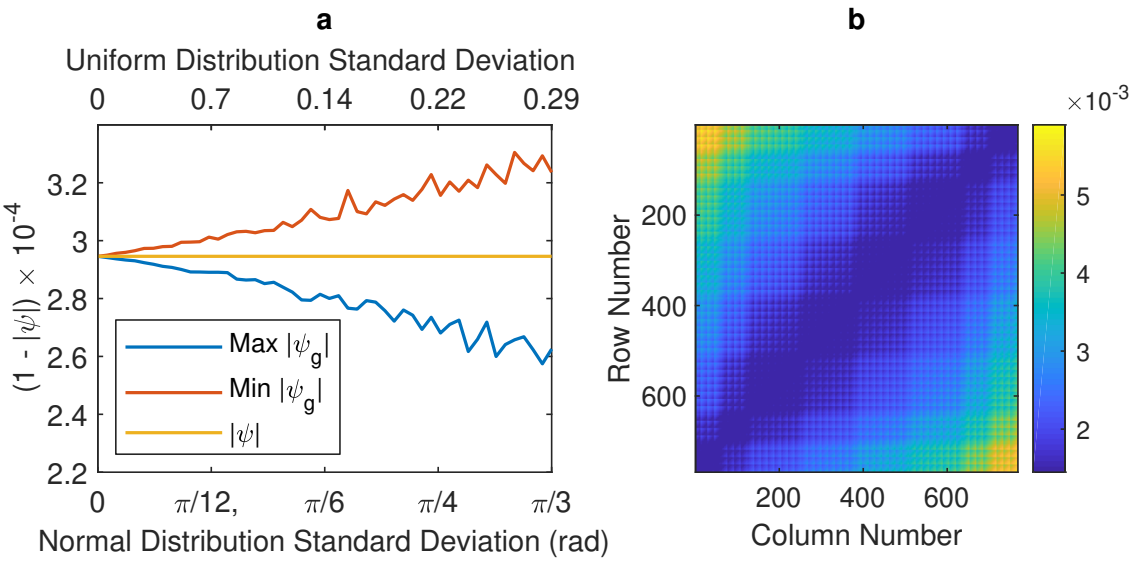

Fig. 4. **(a)** Plots of $1 - |\psi_g|$ where the maximum and minimum value of $|\psi_g|$ was found for a 1000 realizations as a function of the distributions' standard deviations, used to generate $\mathbf{g}$. **(b)** Highest percentage error of the non-narrowband factor (see Eq. (23)) for each element in the covariance matrix $\mathbf{R}_{1g}$. A 1000 realizations were generated where the normal distribution and uniform distribution had a standard deviation of $\pi/3$ and 0.29, respectively (the worst case).

5. Mutual Coupling Model

The mutual coupling in the array cannot be modelled by complex gains, since the signal in each antenna is a weighted sum of the signals in all other antennas. Doing a full electromagnetic simulation of a LOFAR HBA is computationally expensive. Therefore, the heuristic multiple input multiple output model developed in Wijnholds (2008) is used to find qualitative results. The model assumes that the signals are narrowband and that the antennas are dipoles. The heuristic model presented in this section was compared to the resonant dipole model described in Warnick *et al.* (2018). The coupling matrices of both models had a similar global structure, however, the resonant dipole model does not include reactance information. The reactance information is needed to calculate in which directions coupled signals add coherently or destructively. The heuristic model makes the reasonable assumption that phase information is strongly related to the delay between elements in the array.

The mutual coupling matrix $\mathbf{M}$ is calculated in two steps. The first is to calculate $\mathbf{M}_0$ which describes how initially each antenna signal is the weighted sum of the direct signal each antenna receives

$$(\mathbf{M}_0)_{jk} = -m_a \frac{c}{r_{jk}\nu_0}(1 - m_d|\cos(\phi_{jk})|)e^{-i2\pi r_{jk}\nu_0/c}, \tag{26}$$

$$(\mathbf{M}_0)_{jj} = 1, \tag{27}$$

where

m_a is the proportionality constant which determines the strength of the coupling,
r_{jk} is the distance between the jth and k^{th} antenna,
m_d is a directionality parameter,
ϕ_{jk} is the orientation of the line-of-sight between j^{th} and k^{th} antenna.

The initial superposition of signals will be re-radiated by each antenna, which will induce (albeit weaker)

signals in each antenna, which will in turn be again re-radiate, and so forth. To describe this iterative process the following infinite sum can be used

$$\mathbf{M} = \mathbf{I} + \sum_{i=1}^{\infty} (\mathbf{M}_0 - \mathbf{I})^i. \tag{28}$$

This series will converge if $\mathbf{M}_{0,jk} << 1$. The mutual coupling matrix $\mathbf{M}$ should be scaled so that the total power is conserved, that is $\mathrm{Tr}(\mathbf{R}) = \mathrm{Tr}(\mathbf{MRM}^H)$.

In Wijnholds (2008) values of $m_a = 0.07$ and $m_d = 0.9$ are found to give good qualitative results compared to full-EM simulations for the LOFAR HBA. Simulating the effect of m_a and m_d on the eigenvalues and eigenvectors revealed that m_d has very little effect compared to m_a. Therefore, m_d was fixed at 0.9 and m_a was varied between 0 and 0.085. The value of m_a cannot be increased above 0.085, because Eq. (28) then no longer converges in the case of a LOFAR HBA. For the simulation six different cases were considered:

- **Ideal**: no complex gain errors or mutual coupling,
- **Uncal**: the covariance matrix is uncalibrated and there is no mutual coupling,
- **MC**: mutual coupling is present and there are no complex gain errors,
- **MC Uncal**: mutual coupling is present and the covariance matrix is uncalibrated,
- **Cal**: there are complex gain errors that have been calibrated,
- **MC Cal**: mutual coupling is present and complex gain errors have been calibrated.

In figures 5a and b the effect of m_a on the first and second eigenvalues are shown. For the first eigenvalue, in figure 5, the three lines where mutual coupling is present, **MC**, **MC Uncal** and **MC Cal**, all monotonically decrease with m_a. As the first eigenvalue decreases, the second eigenvalue increases in figure 5b. In addition, for the cases where unknown complex gains are present, the first eigenvalue is smaller than the first eigenvalue of the **Ideal** case and the second eigenvalue is larger than the second eigenvalue of the **Ideal** case. Therefore, an increase in mutual coupling and unknown complex gains causes decorrelation in the signal. Even with the calibration step, the power of the first eigenvector is not completely recovered. The reason for this is that the LOFAR HBA station RS407 consists of 48 tiles with 16 antennas, each of which is analogue beamformed and only one gain solution per tile can be calculated with the recorded data.

In figure 6a the cosine similarity between the first eigenvector of the **Ideal** case and all the other cases' eigenvectors is shown as a function of m_a. The same is done for the second eigenvector in figure 6b. In figures 6a and b the cosine similarities decrease as m_a increases. The largest effect however, is caused by the unknown complex gains.

The most realistic case is **MC Cal** where there are complex gains as well as mutual coupling and a calibration step is applied. The largest error between the first eigenvalue of **Ideal** and **MC Cal** is 5×10^{-5} which is lower than all the values of the second eigenvalues. The cosine similarity between **Ideal** and **MC Cal** is one. Therefore, the effect of mutual coupling on the first eigenvalue and eigenvector is minimal. However, the effect of mutual coupling and gain calibration is relatively more pronounced on the second eigenvalue and eigenvector. This is because the gain calibration procedure outlined in section 4 provides only a rank one solution $\mathbf{gg}^H$, using only the first eigenvalue and eigenvector pair, while the mutual coupling matrix $\mathbf{M}$ has full rank. The effect of mutual coupling is therefore included in the evaluation of the proposed algorithms in section 7.

6. Proposed RFI Mitigation Algorithms

The two spatial RFI mitigation algorithms based on subspace subtraction presented in Steeb *et al.* (2018) are extended by adding a gain calibration step. These algorithms are designed for wideband RFI which is stationary, such as DAB broadcasts. The channel bandwidth should be selected such that the second eigenvalue of the sample covariance matrix $\widehat{\mathbf{R}}$ is lower or equal to the power of the cosmic sources being observed. The first algorithm is based on the flat frequency response model with unknown complex gains (see Eq. (24)) and the other on Zatman's approximation to that model (see Eq. (25)). The following preprocessing steps are required (see figure 7 for an activity diagram of the preprocessing stage):

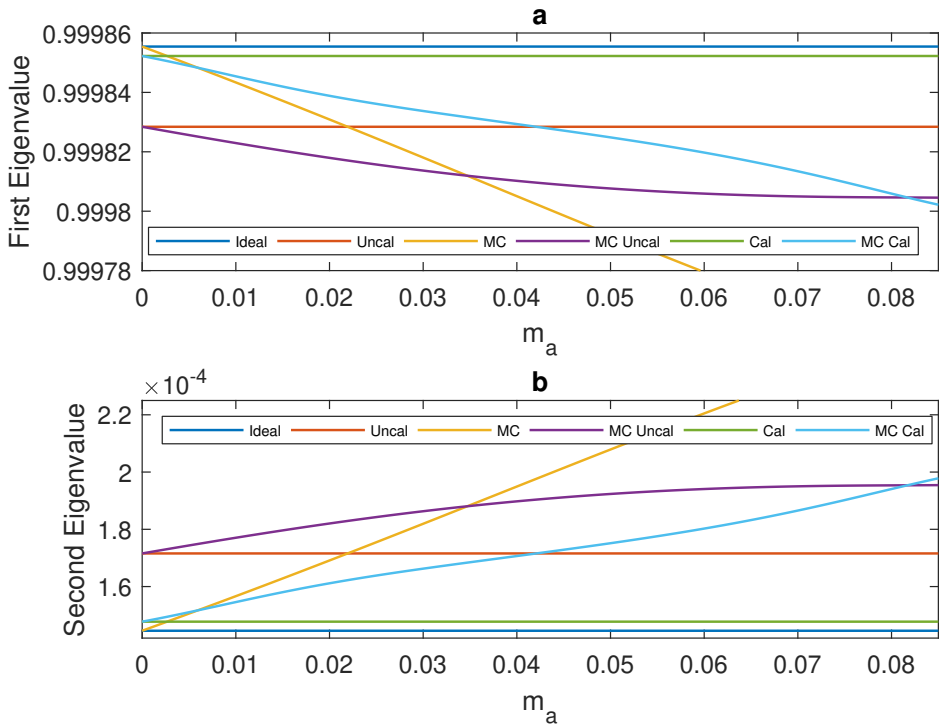

Fig. 5. **Ideal**: no complex gain errors or mutual coupling. **Uncal**: the covariance matrix is uncalibrated and there is no mutual coupling. **MC**: mutual coupling is present and there are no complex gain errors. **MC Uncal**: mutual coupling is present and the covariance matrix is uncalibrated. **Cal**: there are complex gain errors that have been calibrated. **MC Cal**: mutual coupling is present and complex gain errors have been calibrated. Figures **(a)** and **(b)** show the first and second eigenvalues, respectively, as a function of the proportionality constant m_a for different array conditions. The three horizontal straight lines **Ideal**, **Uncal** and **Cal** have no mutual coupling, therefore $m_a = 0$. All matrices have been normalized such that $\mathrm{Tr}(\mathbf{R}) = 1$.

- Obtain the location of the RFI source, (l, m, n). For example, the location of DAB towers are fixed and can be easily obtained from relevant authorities. Algorithms such as MUSIC (Balanis *et al.* , 2007, p. 80-82) or ESPRIT (Yuen *et al.* , 1998) can also be used, however this will bias the gain calibration step.
- Use the power iteration method on $\widehat{\mathbf{R}}$ to find the largest eigenvalue s_1 with the accompanying eigenvector $\mathbf{v}_1$.
- Estimates for the two largest eigenvalues of the RFI-only covariance matrix can be obtained by using the estimated location of the RFI and Eq. (12)

$$s_{\mathrm{r}1} = s_1 - \frac{\mathrm{Tr}(\widehat{\mathbf{R}}) - s_1}{N_{\mathrm{e}} - 1}, \tag{29}$$

$$s_{\mathrm{r}2} = s_{\mathrm{r}1}\left(\frac{1 - |\psi|}{1 + |\psi|}\right). \tag{30}$$

  Use these two new eigenvalue estimates to create the matrix $\mathbf{S}_{\mathrm{r}} = \mathrm{diag}([s_{\mathrm{r}1}, s_{\mathrm{r}2}]^T)$.
- An algorithm such as StEFCal can be used to calculate the complex gains in $\mathbf{g}$.

Algorithm 1: Flat frequency response model based algorithm (see figure 8a for activity diagram)

- Calculate the complex gains' distorted normalized flat frequency covariance matrix model of the RFI source $\mathbf{R}_{\mathrm{fg}}$, using Eq. (24). Note that this model's covariance matrix does not include any noise and that $\sigma_{\mathrm{s}}^2 = 1$.
- Use the power iteration method on $\mathbf{R}_{\mathrm{fg}}$ to find the second largest eigenvalue's eigenvector $\mathbf{v}_{2\mathrm{fg}}$.
- Apply subspace subtraction to obtain the flat frequency model based RFI mitigated covariance matrix

$$\widehat{\mathbf{R}}_{\mathrm{mf}} = \widehat{\mathbf{R}} - [\mathbf{v}_1, \mathbf{v}_{2\mathrm{fg}}]\mathbf{S}_{\mathrm{r}}[\mathbf{v}_1, \mathbf{v}_{2\mathrm{fg}}]^H. \tag{31}$$

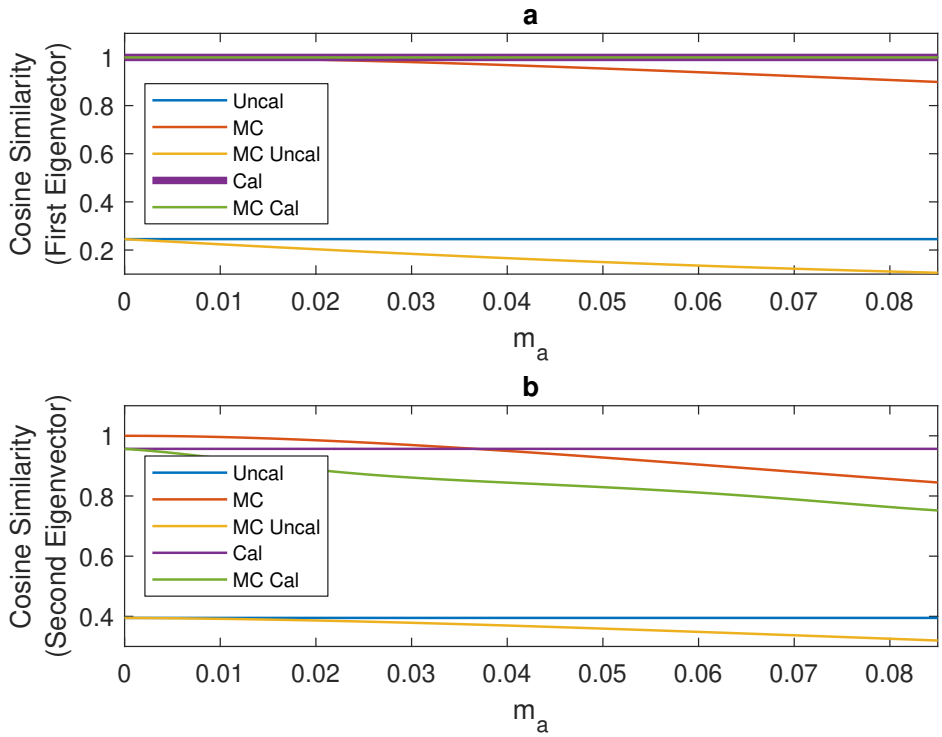

Fig. 6. See figure 5 for definitions of the legend notation. Figure **(a)** shows the cosine similarity between the first eigenvector of the **Ideal** case (no complex gain errors or mutual coupling) and different array conditions as a function of the proportionality constant m_a. Similar to figure **(a)**, figure **(b)** considers the second eigenvector. The three horizontal straight lines **Ideal**, **Uncal** and **Cal** have no mutual coupling, therefore $m_a = 0$. A cosine similarity of 1 indicates that the two vectors are parallel, while a cosine similarity of 0 indicates that the vectors are orthogonal.

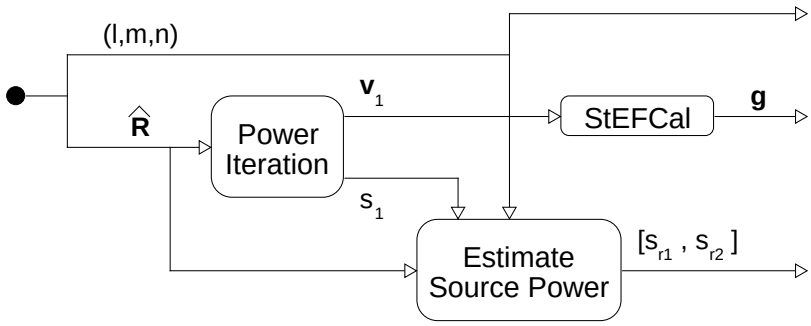

Fig. 7. Activity Diagram of the Preprocessing stage.

Algorithm 2: Zatman's approximation based algorithm (see figure 8b for activity diagram)

- Calculate the normalized Zatman's approximation based model eigenvector $\mathbf{v}_{2\mathrm{zg}}$ using Eq. (25).
- Apply subspace subtraction to obtain the Zatman's model based RFI mitigated covariance matrix

$$\widehat{\mathbf{R}}_{\mathrm{mz}} = \widehat{\mathbf{R}} - [\mathbf{v}_1, \mathbf{v}_{2\mathrm{zg}}]\mathbf{S}_{\mathrm{r}}[\mathbf{v}_1, \mathbf{v}_{2\mathrm{zg}}]^H. \tag{32}$$

7. Evaluation of RFI Mitigation Algorithms

A similar procedure described in Steeb *et al.* (2018) was used to generate estimated noise and cosmic source covariance matrices $\widehat{\mathbf{R}}_{\mathrm{cn}}$ and RFI covariance matrices $\widehat{\mathbf{R}}_{\mathrm{r}}$, with varying bandwidths (763 Hz to

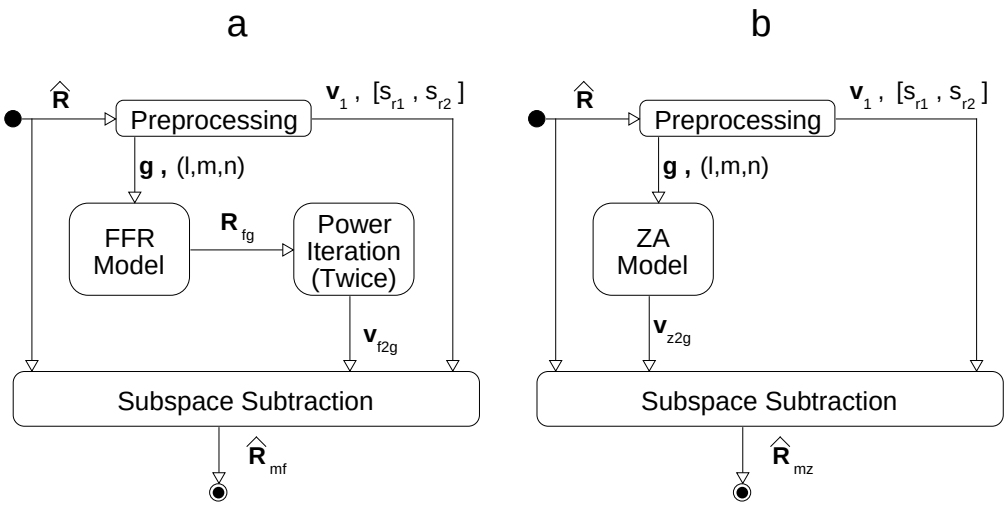

Fig. 8. **(a)** Activity diagram for algorithm 1 which uses the flat frequency model in Eq. (24). **(b)** Activity diagram for algorithm 2 which uses the Zatman's approximation in Eq. (25). The diagrams are based on those found in Steeb *et al.* (2018) and are expanded to include the unknow complex gain solutions.

218 kHz). RFI free observations from LOFAR station RS407 were used for $\widehat{\mathbf{R}}_{\mathrm{cn}}$. The $\widehat{\mathbf{R}}_{\mathrm{r}}$ matrices were generated using a DAB signal, recorded with a software defined radio. The bandwidth of the DAB signal was varied by using a finite impulse response filter. To simulate the reception of the signal by station RS407 the appropriate delay was added for each of the 768 antennas. For each bandwidth a covariance matrix was created (with an integration time of 1.5 s) and reduced from a 768×768 matrix to a 48×48 matrix by beamforming groups of 16 antennas. The effect of complex gains and mutual coupling is applied before the beamforming step.

To measure the performance of the proposed algorithms, the Covariance Matrix Distance (**CMD**, see Herdin *et al.* (2005)) and the Attenuation (**ATT**) are used

$$\mathbf{CMD} = 1 - \frac{\mathrm{Tr}(\widehat{\mathbf{R}}_{\mathrm{cn}}\widehat{\mathbf{R}}_{\mathrm{m}})}{\left\|\widehat{\mathbf{R}}_{\mathrm{cn}}\right\|_F \left\|\widehat{\mathbf{R}}_{\mathrm{m}}\right\|_F}, \tag{33}$$

$$\mathbf{ATT} = 10 \log \left[\frac{\mathrm{Tr}(|\widehat{\mathbf{R}}_{\mathrm{m}} - \widehat{\mathbf{R}}_{\mathrm{cn}}|_{\odot})}{\mathrm{Tr}(\widehat{\mathbf{R}}_{\mathrm{r}})} \right], \tag{34}$$

where $\widehat{\mathbf{R}}_{\mathrm{m}}$ is the recovered matrix. The **CMD** and **ATT** measure, respectively, how well the eigenvectors and eigenvalues of $\widehat{\mathbf{R}}_{\mathrm{cn}}$ are recovered. If the **CMD** is zero, then the matrices are equal up to a scaling factor. At negative infinity the **ATT** indicates that the total power (sum of eigenvalues) for $\widehat{\mathbf{R}}_{\mathrm{m}}$ and $\widehat{\mathbf{R}}_{\mathrm{cn}}$ are the same.

To test the effect of unknown complex gains, the generated values for $\mathbf{g}$ in section 4 were used as well as $\widehat{\mathbf{R}}_{\mathrm{r}}$ and $\widehat{\mathbf{R}}_{\mathrm{cn}}$ at a fixed bandwidth of 195 kHz. In figures 9a and b, respectively, the **CMD** and **ATT** are plotted for three different cases as a function of the normal and uniform distributions' standard deviations. The 1st order lines are the performance achieved by subtracting the estimate of the subspace associated with the largest eigenvalue $\widehat{\mathbf{R}}_{\mathrm{r}}$. Using the Flat Frequency Algorithm without a calibration step causes the performance to decline due to an increase in the distributions' standard deviations. Introducing a calibration step greatly improves the performance, but there is still a slightly increasing trend. This is caused by the beamforming stage for the LOFAR HBA station which consists of 48 tiles with 16 antennas, each of which is analogue beamformed and only one gain solution per tile can be calculated. Furthermore, the assumption made in section 4 that the complex gains non-narrowband component of Eq. (23) is minimally affected by the unknown complex gains, adds to the trend.

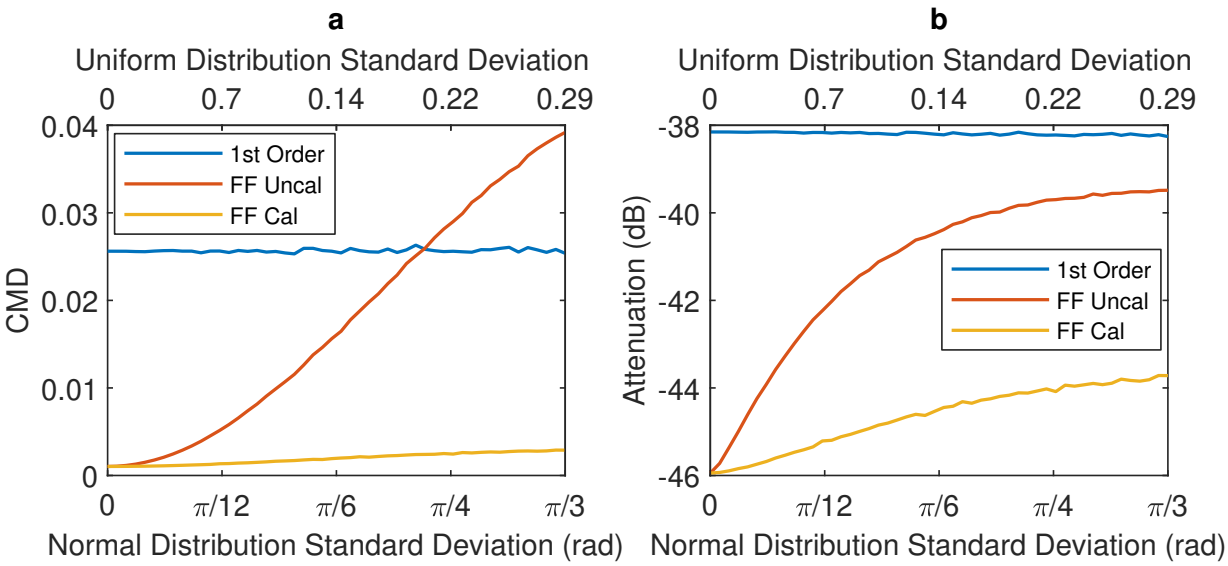

Fig. 9. **1st Order**: Subtracting the estimate of the subspace associated with the largest eigenvalue of $\widehat{\mathbf{R}}_{\mathrm{r}}$ from $\widehat{\mathbf{R}}_{\mathrm{cn}}$. **FF Uncal**: Using the Flat Frequency Algorithm without a calibration step. **FF Cal**: Using the Flat Frequency Algorithm with a calibration step. In figure **(a)** the covariance matrix distance and in figure **(b)** the attenuation is plotted for three RFI mitigation methods as a function of the distributions' standard deviations, used to generate $\mathbf{g}$.

Now the entire set of covariance matrices at different bandwidths are used. The mutual coupling matrix was generated by setting $m_a = 0.07$ and $m_d = 0.9$ and the unknown complex gains were generated using a uniform distribution (standard deviation 0.29) for the magnitude and a normal distribution (standard deviation $\pi/3$) for the phase (only one set of complex gains was used). In figures 10a and b the performances of four algorithms are given:

- **1st Order**: Subtracting the estimate of the subspace associated with the largest eigenvalue of $\widehat{\mathbf{R}}_{\mathrm{r}}$ from $\widehat{\mathbf{R}}_{\mathrm{cn}}$.
- **2nd Order**: Subtracting the estimate of the largest and second largest eigenvalue of $\widehat{\mathbf{R}}_{\mathrm{r}}$ and the accompanying eigenvectors from $\widehat{\mathbf{R}}_{\mathrm{cn}}$.
- **FF Alg**: Using the Flat Frequency Algorithm with a calibration step.
- **Zatman Alg**: Using the Zatman's Approximation Based Algorithm with a calibration step.

At close to zero fractional bandwidth, all methods have the same performance, since the estimate of the second eigenvector of $\widehat{\mathbf{R}}_{\mathrm{r}}$ is almost zero and all methods are effectively only subtracting the first subspace. The error for the **2nd Order** increases until a fractional bandwidth of approximately 8×10^{-4}. This is due to eigenvalue decomposition that cannot differentiate between the second eigenvector of $\widehat{\mathbf{R}}_{\mathrm{r}}$ and the noise. After a fractional bandwidth of 8×10^{-4}, the **2nd Order** method's performance improves, however it does not reach the same level of performance achieved at smaller fractional bandwidths. This is caused by the third eigenvalue that becomes significant due to the bandwidth and the non-perfectly flat spectrum of the signal. For all bandwidths the **FF Alg** and **Zatman Alg** methods have the same performance and both have superior **CMD** performance to the **1st Order** and **2nd Order** subtraction methods until approximately a fractional bandwidth of 11×10^{-4}. For the **ATT** performance the **FF Alg** and **Zatman Alg** methods have superior performance for the entire fractional bandwidth considered.

Channels with larger bandwidth can be processed using the **FF Alg** or **Zatman Alg**, while achieving the same level of mitigation as the 1st order method which requires channels with smaller bandwidth. This is shown in figure 11 where the bandwidth required by the **FF Alg** is given as a function of the bandwidth of the **1st Order** method. When no array imperfections are present, approximately six times as much bandwidth can be processed. When unknown complex gains are present and a calibration step is added, the performance reduces to twice as much bandwidth. Finally, if mutual coupling is also present the performance reduces further to 1.6 times more bandwidth.

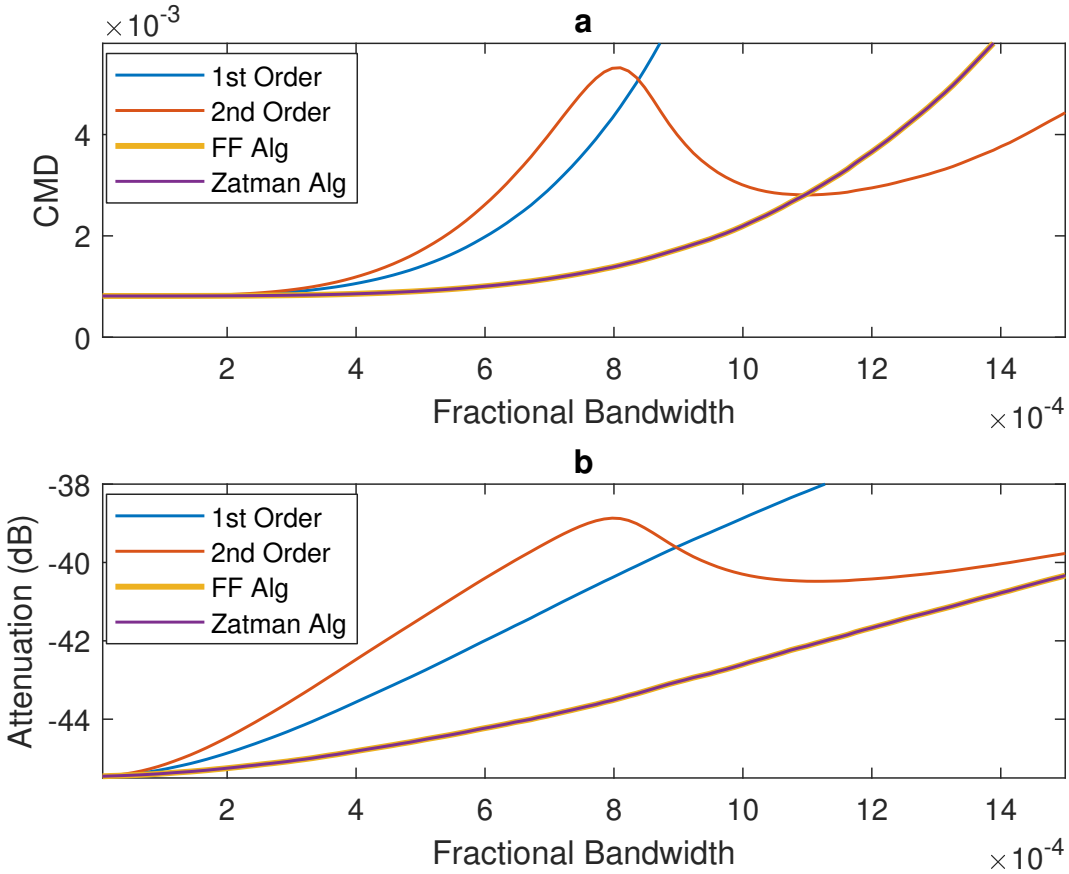

Fig. 10. **1st Order**: Subtracting the estimate of the largest eigenvalue of $\widehat{\mathbf{R}}_{\mathrm{r}}$ and its accompanying eigenvector from $\widehat{\mathbf{R}}_{\mathrm{cn}}$. **2nd Order**: Subtracting the estimate of the largest and second largest eigenvalue of $\widehat{\mathbf{R}}_{\mathrm{r}}$ and the accompanying eigenvectors from $\widehat{\mathbf{R}}_{\mathrm{cn}}$. **FF Alg**: Using the Flat Frequency Algorithm with a calibration step. **Zatman Alg**: Using the Zatman Approximation Based Algorithm with a calibration step. In figures **(a)** and **(b)** the covariance matrix distance and attenuation, respectively, for four RFI mitigation methods as a function of fractional bandwidth are plotted.

In the simulations completed, both the **FF Alg** and the **Zatman Alg** provide the same results for mitigation. However, the **Zatman Alg** is slightly computationally less expensive because, unlike the **FF Alg**, it does not have to complete the power iteration method twice (see figure 7a and b). See Steeb *et al.* (2018) for the computational complexity analysis.

8. Imaging Results

Creating an easily interpretable full sky image with the LOFAR HBA is not possible due to grating lobes. Therefore, LOFAR LBA station CS302 is used to demonstrate the effect of the **FF Alg** on the imaging plane. A simulated RFI signal at 45 MHz with 195 kHZ bandwidth was combined with real RFI free data from station CS302. The resulting skymap is shown in figure 12a and only the RFI source is visible due to the side-lobes of the array beam of the station. Using first order subspace subtraction, the point source component of the RFI is removed, but the component due to frequency smearing is still present (see figure 12b). Furthermore, two cosmic sources, Cassiopeia A and Cygnus A, are now visible as well as the structure of the Milky Way. Increasing the order of the subspace subtraction filter does not completely remove the frequency smeared component, see figure 12c. This is due to the eigenvalue decomposition not correctly isolating the frequency smeared component in the second eigenvector and value. Rather a component of the Milky Way is included and distorted (compare figure 12c with the RFI free skymap in figure 12e). The skymap obtained by the **FF Alg** is given in figure 12d where the RFI source is no longer visible and the structure of the Milky Way is recovered. The difference between the **FF Alg** skymap and

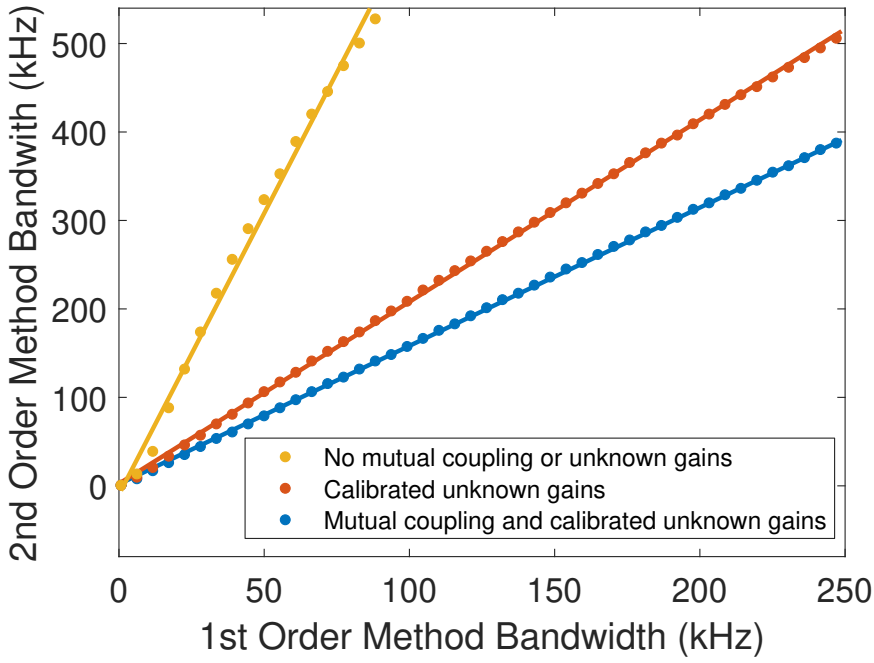

Fig. 11. A plot of the bandwidth required by the **FF Alg** as a function of the bandwidth of the **1st Order** method to achieve the same attenuation for three cases: there is no mutual coupling or unknown gains, there are unknown gains that have been calibrated for and finally unknown gains that have been calibrated for and mutual coupling.

the RFI free skymap is given in figure 12f and the largest error is 0.015 dB.

9. Conclusion

The proposed algorithm extends a previously developed algorithm that now includes a gain calibration step. The algorithm combines a first order subspace subtraction filter and a non-narrowband signal model, that may be used to mitigate powerful RFI signals for which the second eigenvalue is below the noise, but has a power that is competing with the astronomical sources. Array imperfections such as unknown complex gains and mutual coupling reduce the performance of the algorithm. It was shown that a standard gain calibration step can be applied to improve performance, since the non-narrowband factor is minimally affected by the complex gains. With a gain calibration step and no mutual coupling, the proposed algorithm was able to process twice the bandwidth per channel that can be processed when applying conventional narrowband techniques. This performance declines to 1.6 times more bandwidth when the effect of mutual coupling is included. In future work a full mutual-coupling analysis, using computational electromagnetic tools, will be completed.

10. Acknowledgements

This work was funded and supported by the South Africa Radio Astronomy Observatory, the South African Research Chairs Initiative of the Department of Science and Technology, the National Research Foundation of South Africa, and the Netherlands Organisation for Scientific Research.

References

Balanis, C. A., Ioannides, P. I. [2007] Introduction to Smart Antennas, Morgan & Claypool Publishers.

Boonstra, A. J. [2005] *Radio Frequency Interference Mitigation in Radio Astronomy*, PhD Thesis, Delft University of Technology, uuid:caa1942c-4180-4a17-88db-cec359490aad.

Bridle, A. H., Schwab, F. R. [1999] *Bandwidth and Time-Average Smearing*, In G. B. Taylor, C. L. Carilli and R. A. Perley, editors, "Synthesis Imaging in Radio Astronomy II," **180**, 18, p. 371–382, eISBN: 978-1-58381-516-8.

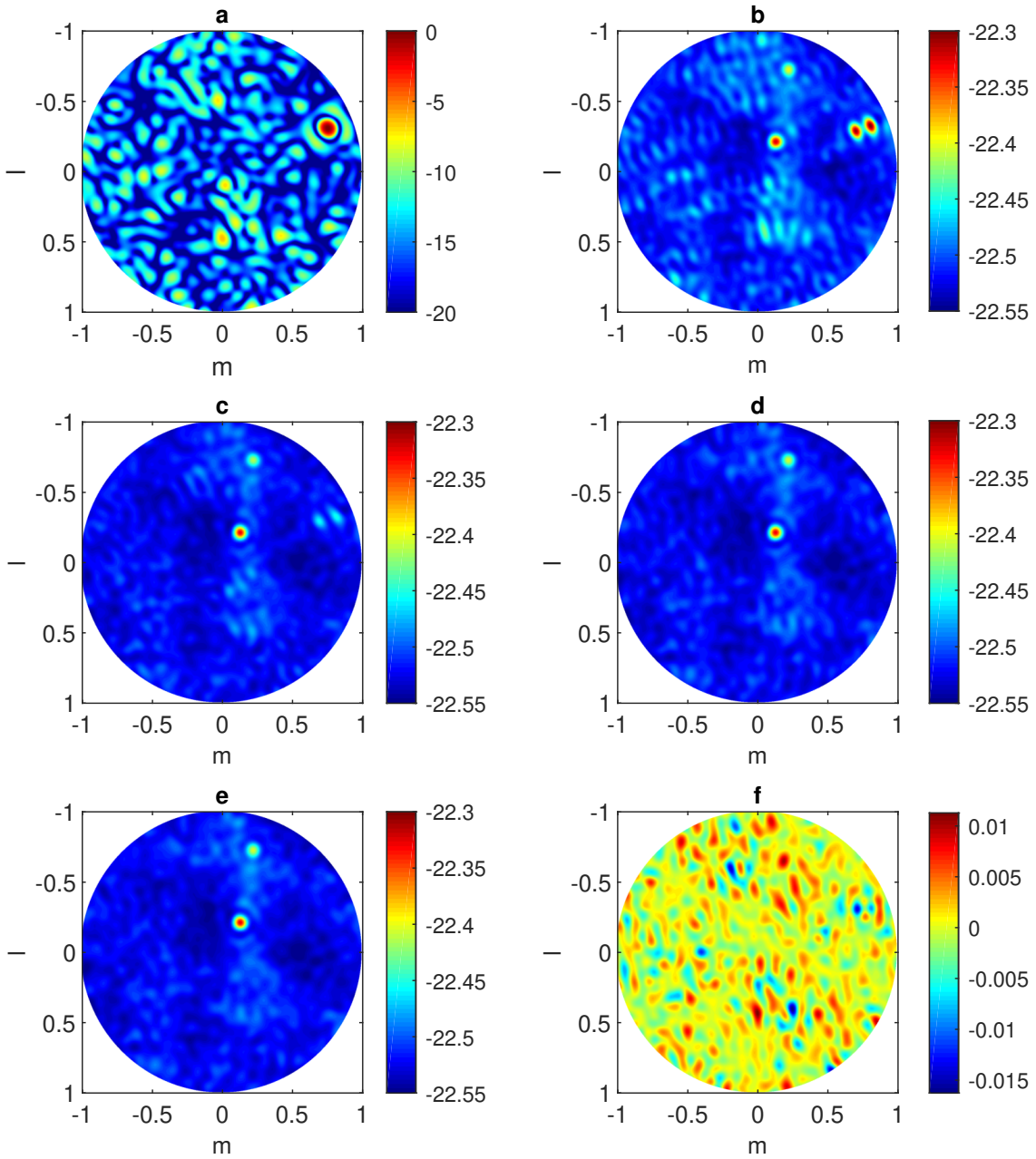

Fig. 12. **(a)** Skymap with non-narrowband RFI source in dB (the RFI source is the 0 dB point). The source has a bandwidth of 195 kHz and a center frequency of 45 MHz. The image is also distorted by unknown complex gains and mutual coupling. **(b)** Skymap with the non-narrowband RFI source removed using first order subspace subtraction. Two weaker sources adjacent to the location of the RFI are now visible and are caused by the bandwidth of the RFI source. Two unresolved cosmic sources, Cassiopeia A and Cygnus A, are also now visible and some structure of the Milky Way. **(c)** Skymap with non-narrowband RFI source removed using second order subspace subtraction. The bandwidth related component of RFI signal has been attenuated but is still visible. **(d)** Skymap with non-narrowband RFI source removed using the **FF Alg**. The bandwidth related component of RFI signal is no longer visible. **(e)** Skymap without RFI source. **(f)** Difference between RFI free skymap and skymap recovered using the **FF Alg**.

Brossard, M., El Korso, M. N., Pesaventoc, M.,Boyer, R., Larzabal, P., Wijnholds, S.J. [2018] *Parallel multi-wavelength calibration algorithm for radio astronomical arrays*, Signal Processing, **145**, p. 258–271, April 2018, doi:10.1016/j.sigpro.2017.12.014.

Compton, R. T. [1988] *Adaptive Antennas: Concepts and Performance*, Prentice-Hall.

de Gasperin, F., Mevius, M., Rafferty, D. A., Intema, H. T., Fallows, R. A. [2018] *The effect of the ionosphere on*

ultra-low-frequency radio-interferometric observations, Astronomy & Astrophysics, **615**, no. A179, July 2018, doi:10.1051/0004-6361/201833012.

Herdin, M., Czink, N., Ozcelik, H., Bonek, E. [2005] *Correlation Matrix Distance, a Meaningful Measure for Evaluation of Non-Stationary MIMO Channels*, IEEE 61st Vehicular Technology Conference, Stockholm, Sweden, doi:10.1109/VETECS.2005.1543265.

Johnson, D. H., Dudgeon, D. E. [1993] "Array Signal Processing: Concepts and Techniques," NJ: Prentice-Hall, p. 113–115.

Salvini, S., Wijnholds, S. J. [2014] *Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications*, Astronomy & Astrophysics, **571**, p. 1–14, doi: 10.1051/0004-6361/201424487.

Sardarabadi, A. M. [2016] *Covariance Matching Techniques for Radio Astronomy*, PhD Thesis, Delft University of Technology, doi:10.4233/uuid:3eb46b5b-126c-4d95-8821-fb71ce404cac.

Steeb, J. W., Davidson, D. B., Wijnholds, S. J. [2018] *Mitigation of Non-Narrowband Radio Frequency Interference*, Radio Science Bulletin, doi:10.23919/URSIRSB.2018.8572495.

Thompson, A. R., Moran, J. A., Swenson Jr., G. W. [2004]. " Interferometry and Synthesis in Radio Astronomy Second Edition," WILEY-VCH Verlag GmbH & Co. KGaA, doi:10.1002/9783527617845.

van der Tol, S., van der Veen, A. [2005] *Performance Analysis of Spatial Filtering of RF Interference in Radio Astronomy*, IEEE Transactions on Signal Processing, **53**, 3, p. 896–910, doi:10.1109/TSP.2004.842177.

van der Veen, A., Leshem, A., Boonstra, A. [2004] *Signal Processing for Radio Astronomical Arrays*, Procceedings of the Sensor Array and Multichannel Signal Processing Workshop, p. 1–10, doi:10.1109/SAM.2004.1502901.

van Haarlem, M. P., *et al.* [2013] *LOFAR: The Low Frequency Array*, Astronomy & Astrophysics, **556**, A2, p. 1–53, doi:10.1051/0004-6361/201220873.

Warnick, K., Maaskant, R., Ivashina, M., Davidson, D., Jeffs, B. [2018] *Transmitting Arrays, Network Analysis, and Pattern Overlap Integrals.*, In Phased Arrays for Radio Astronomy, Remote Sensing, and Satellite Communications, EuMA High Frequency Technologies Series, Cambridge: Cambridge University Press, p. 106–153, doi:10.1017/9781108539258.005

Wijnholds, S. J. [2008] *Mutual Coupling, Inter-tile Spacing and Inter-station Rotation in the HBA Array*, Technical Report, LOFAR-ASTRON-MEM-245.

Yuen, N., Friedlander, B. [1998] *Performance Analysis of Higher Order ESPRIT for Localization of Near-Field Sources*, IEEE Transactions on Signal Processing, **46**, 3, p. 709–719, doi:10.1049/ip-rsn:19981670.

Zatman, M. [1998] *How narrow is narrowband?*, IEE Proceedings-Radar, Sonar and Navigation, **145**, 2, p. 85–91, doi:10.1109/78.661337.