2nd URSI AT-RASC, Gran Canaria, 28 May – 1 June 2018

Mitigation of Non-Narrowband Radio Frequency Interference

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

Abstract

The rapid development and implementation of wireless communication standards put increasing pressure on spectrum allocation and therefore threatens the efficacy of radio astronomy. For example, digital audio broadcasting (DAB) is a wide-bandwidth broadcast technology that is now being implemented and has spectrum allocated in the L band. We demonstrate that standard narrowband subspace subtraction methods may provide insufficient suppression of such signals. We therefore propose two algorithms that take into account the non-narrowband nature of these signals. The first proposed algorithm is based on a flat frequency response model and the second is a Taylor expansion approximation of the first. An experimental demonstration of both proposed algorithms yielded an increase of approximately a factor two in bandwidth per channel that can be processed when compared to conventional narrowband techniques (for the same attenuation of the RFI signal). The performance of the two methods is identical for LOFAR station configurations with bandwidths between 763 Hz and 195 kHz, however the Taylor expansion approximation based algorithm requires less operations (a speed-up of 1.3 was achieved). An equation is derived that gives the direction of arrival for a far-field wideband signal that causes the greatest distortion of the visibilities. Another equation is derived for the appropriate sub-band bandwidth at which to implement narrowband RFI mitigation algorithms.

1 Introduction

In the development of radio frequency interference (RFI) mitigation methods, the assumption that the RFI is narrowband, is usually made. If this is the case, spatial RFI mitigation methods such as orthogonal projection, orthogonal projection with subspace bias correction, oblique projection and subspace subtraction [1, 2, 3] can be applied. When the signal is not narrowband, the model for the array response vector becomes a function of bandwidth. The result is that the RFI will appear as an extended source that can be modelled as multiple sources, albeit with rapidly decreasing power. To evaluate the proposed RFI mitigation methods, the layout of High Band Antenna (HBA) station RS407 in the Low Frequency Array (LOFAR) [4] is used. The HBA stations in LOFAR have an operating band from 110-250 MHz, which contains many digital audio broadcasts (DABs).

To demonstrate the effect of bandwidth on a sky image a simulation was done using the layout of a LOFAR Low Band Antenna (LBA) station. The LBA station layout was used because the HBA stations have a regular array layout that is not suitable for imaging. In figure 1a the skymap is dominated by the RFI source in the top left corner. When a first order orthogonal projection filter is applied, the skymap in figure 1b is obtained. The orthogonal projector reduces the power of the RFI source by 30 dB and reveals two sources which are seperated along the radial direction due to frequency smearing [5]. Subtracting a point source model from the centroid of the RFI source results in a residual image with one peak on either side of the centroid.

Using a flat frequency response model and Zatman's approximation [6] of that model, equations are derived for the RFI covariance matrix's eigenvalues, the direction of arrival that causes the largest deviation from the narrowband point source model, the appropriate bandwidth at which to apply spatial filtering and the RFI covariance matrix's vector space. Two new subspace subtraction algorithms, that make use of the model for the RFI covariance matrix's vector space, are then presented and evaluated.

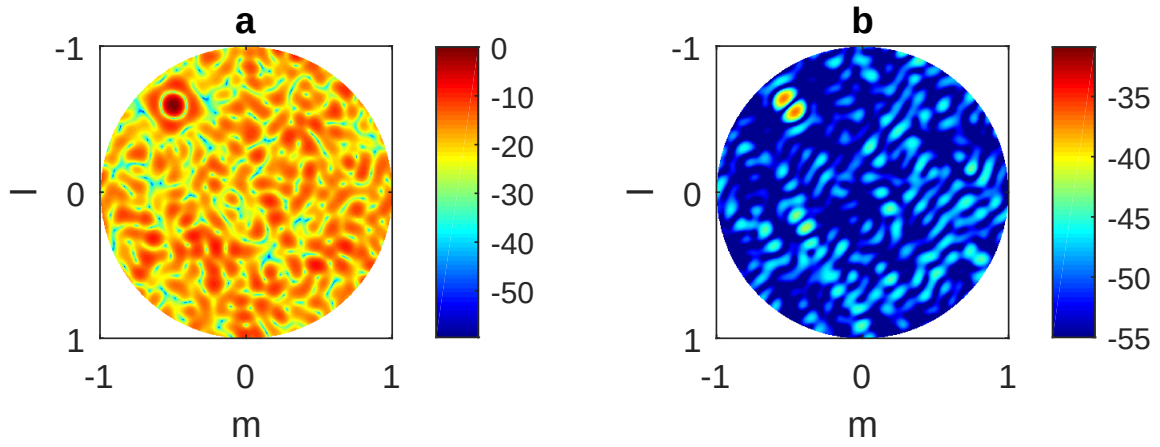

Figure 1. (a) Full skymap with wideband RFI source visible in top left corner in dB (the RFI source is the 0 dB point). **(b)** Full skymap with wideband RFI source removed using a first order orthogonal projection filter. Two weaker sources adjacent to the location of the RFI are now visible and are caused by the bandwidth of the RFI source.

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 j^{th} element is indicated by $\mathbf{a}_j$.
$\mathbf{I}$	Identity matrix.
$\odot$	Hadamard product.
$\lvert\cdot\rvert$	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.
$\Re(\cdot)$	Real part of a complex number.
$\Im(\cdot)$	Imaginary part of a complex number.
$\text{sinc}(x)$	$= \sin(\pi x)/(\pi x)$, normalised sinc function.

3 Narrowband Signal Model

If omnidirectional antennas are used, then the normalised array response vector for an array with N_e elements and a continuous wave source with frequency ν is given by

$$\mathbf{a} = \begin{bmatrix} g_1 e^{-\text{i}2\pi\nu\tau_1} \\ \vdots \\ g_{N_\text{e}} e^{-\text{i}2\pi\nu\tau_{N_\text{e}}} \end{bmatrix}. \tag{1}$$

If the source lies in the far-field, then

$$\tau_j = -(l_\text{s} x_j + m_\text{s} y_j + n_\text{s} z_j)/c, \tag{2}$$
$$g_j = 1/\sqrt{N_\text{e}}, \tag{3}$$

where x_j, y_j, z_j are the Cartesian coordinates of the j^{th} antenna and the directional cosines are defined as

$$l_\text{s} = \sin(\theta_\text{s})\cos(\phi_\text{s}), \tag{4}$$
$$m_\text{s} = \sin(\theta_\text{s})\sin(\phi_\text{s}), \tag{5}$$
$$n_\text{s} = \cos(\theta_\text{s}), \tag{6}$$

where θ_s and ϕ_s are respectively the polar and azimuthal angles of the spherical coordinates of the source (ISO convention). The array covariance matrix for a single source without noise is given by

$$\mathbf{R} = \sigma_s^2 \mathbf{a}\mathbf{a}^H, \tag{7}$$

$$\mathbf{R}_{jk} = g_j g_k \sigma_s^2 e^{-i2\pi\tau_{jk}\nu}, \tag{8}$$

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 the 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 $\tau_j = r_{sj}/c$ and $g_j = 1/\left[r_{sj}\sqrt{\sum_{n=1}^{N_e} 1/r_{sn}^2}\right]$, where the distance between the j^{th} antenna and the source is denoted by r_{sj}.

4 Non-Narrowband Signal Model

If the channel bandwidth is not sufficiently narrow, the dependence of the array response vector on frequency becomes significant. The frequency dependent covariance matrix $\mathbf{R}(\nu)$ with only a single interferer (no noise or cosmic sources), that is modelled as a point source, can be written as

$$\mathbf{R}(\nu) = \sigma_s^2(\nu)\mathbf{a}(\nu)\mathbf{a}^H(\nu), \tag{9}$$

where $\mathbf{a}(\nu)$ is the normalised frequency dependent array response vector. When the different frequency components are uncorrelated, the total covariance matrix is found by integrating over the entire bandwidth

$$\mathbf{R} = \frac{1}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} \sigma_s^2(\nu)\mathbf{a}(\nu)\mathbf{a}^H(\nu)d\nu, \tag{10}$$

where $\Delta\nu$ is the bandwidth and ν_0 is the centre frequency. For a flat frequency response, the integral in equation (10) can be calculated and the jk^{th} element is given by

$$\mathbf{R}_{jk} = \frac{\sigma_s^2 g_j g_k}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} e^{-i2\pi\tau_{jk}\nu}d\nu = \sigma_s^2 g_j g_k \,\mathrm{sinc}(-\tau_{jk}\Delta\nu)e^{-i2\pi\tau_{jk}\nu_0}. \tag{11}$$

By taking the bandwidth into consideration (and assuming a flat frequency response) this covariance matrix model differs from the narrowband model (see equation (8)) with a sinc function that is dependent on the delay τ_{jk} and the bandwidth $\Delta\nu$. As the bandwidth increases from a single frequency, the sinc function starts to decrease from unity and the effect is that the covariances start to decorrelate. This causes the eigenvalue structure of the array covariance matrix to change. For a single frequency signal there will only be one non-zero eigenvalue. For a non-zero bandwidth signal the covariance matrix will be of full rank, since it is an infinite sum of frequencies. As the bandwidth increases, the largest eigenvalue will decrease and the other eigenvalues will increase, as can be seen in figure 2a (where even though the largest eigenvalue appears constant due to the log scale of the figure, it is decreasing). However, most of the eigenvalues will be so small relative to the cosmic sources and the noise in the system, that they can be approximated by zero. The effective rank of the RFI covariance matrix (no noise or cosmic sources) is then defined to be equal to the number of eigenvalues that are significant when compared to the eigenvalues of the covariance matrix that contains only the cosmic sources and noise.

The entire fractional bandwidth range used to generate figure 2a would be considered narrowband in the traditional signal processing sense. However, the power of some RFI sources is so high that it can be more then 70 dB above the noise and even 60 dB above a bright radio source such as Cassiopeia A. Therefore, the contribution of the RFI source to the second eigenvalue can still be significant relative to that of the cosmic sources. Figure 2b shows a plot of the three largest eigenvalues from an observation using LOFAR HBA station RS407 as a function of centre frequency. The bandwidth used for the observation is 195 kHz and the DAB signal can be clearly seen from 182.9-184.4 MHz. The second eigenvalue for the DAB signal is at approximately -40 dB and agrees with the simulated results in figure 2a. There is however no agreement between the simulated and measured third eigenvalue. This is due to calibration errors and the frequency response not being perfectly flat.

5 Approximation of RFI Eigenvalues

If a covariance matrix has an effective rank of two, it can be approximated by the sum of two discrete uncorrelated signals

$$\mathbf{R} \approx \sigma_1^2 \mathbf{a}_1 \mathbf{a}_1^H + \sigma_2^2 \mathbf{a}_2 \mathbf{a}_2^H. \tag{12}$$

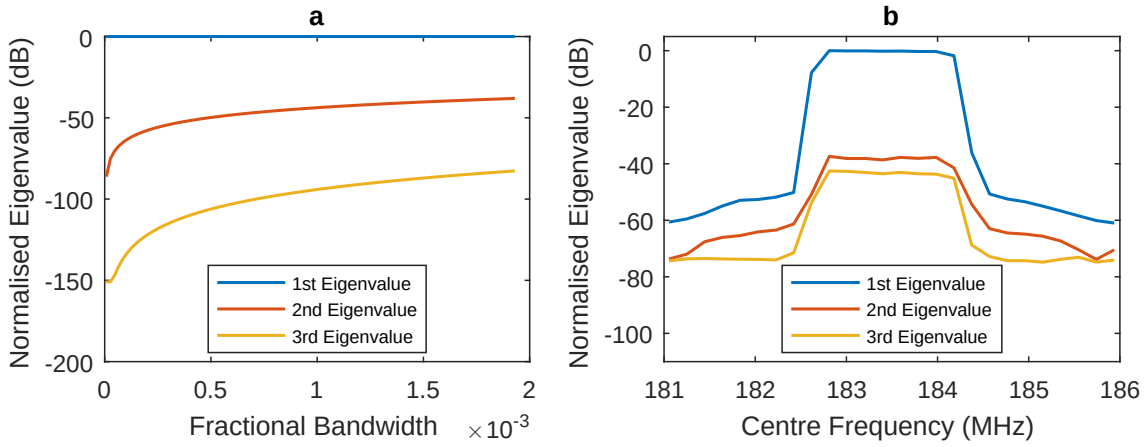

Figure 2. (a) A plot of the simulated (assuming a flat frequency response) three largest normalised eigenvalues as a function of fractional bandwidth using the layout of the LOFAR HBA station RS407 and an RFI source on the horizon (the centre frequency is 183 MHz and the bandwidth was varied from 763 Hz to 195 kHz). **(b)** The three largest normalised eigenvalues as a function of centre frequency calculated from the covariance matrices obtained from an observation using LOFAR station RS407 with a bandwidth of 195 kHz. A DAB signal can be clearly seen from 182.9-184.4 MHz.

The closed form solution for the eigenvalues of the system given in equation (12) are [3, p. 65]

$$\lambda_{1,2} = \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\left(1 \pm \sqrt{1 - 4\frac{\sigma_1^2\sigma_2^2(1 - |\mathbf{a}_1^H\mathbf{a}_2|^2)}{(\sigma_1^2 + \sigma_2^2)^2}}\right). \tag{13}$$

In the model proposed by Zatman the signals are required to have equal power ($\sigma_1^2 = \sigma_2^2 = \sigma^2$) [6]. The equal power criterion is achieved when the discrete sources are arranged in such a way that the instantaneous frequency spectrum mean and variance correspond to the mean and variance of the non-zero bandwidth signal, respectively. Consequently, the distance from the centre frequency ν_0 is given by

$$\kappa = \frac{\Delta\nu}{2\sqrt{3}}. \tag{14}$$

Thus, the model in equation (12) becomes

$$\mathbf{R} \approx \sigma^2\mathbf{a}(\nu_0 + \kappa)\mathbf{a}^H(\nu_0 + \kappa) + \sigma^2\mathbf{a}(\nu_0 - \kappa)\mathbf{a}^H(\nu_0 - \kappa) = \sigma^2(\mathbf{a}_1\mathbf{a}_1^H + \mathbf{a}_2\mathbf{a}_2^H). \tag{15}$$

Zatman's approach is now generalised from a uniform linear array to an array of any shape using the normalised response vector in equation (1) and the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Consequently, equation (13) simplifies to

$$\lambda_{1,2} = \sigma^2[1 \pm |\psi|], \tag{16}$$

where $\psi = \mathbf{a}_1^H\mathbf{a}_2$.

5.1 Worst Case Scenario

Minimising $|\psi|$ maximises λ_2, which is the worst case scenario. Also, minimising $|\psi|^2$ is the same as minimising $|\psi|$ since it is a positive function. After expanding and simplifying

$$|\psi|^2 = \sum_{p=1}^{N_e} g_p^4 + \sum_{j=1}^{N_e-1}\sum_{k=j+1}^{N_e} 2g_j^2 g_k^2 \cos\left(4\pi\tau_{jk}\kappa\right). \tag{17}$$

We will make the following assumptions:

- The array is planar (that is, $z_{jk} = 0$).
- The RFI source lies in the far-field (therefore use equations (2) and (3)).
- The argument $4\pi\tau_{jk}\kappa << \pi/2$, therefore the Taylor expansion $\cos(x) \approx 1 - \frac{x^2}{2}$ can be used.

Using these assumptions, equation (17) simplifies to

$$|\psi|^2 = 1 - \frac{16\pi^2\kappa^2}{N_e^2}\left[\sum_{j=1}^{N_e-1}\sum_{k=j+1}^{N_e}\tau_{jk}^2\right] = 1 - \frac{16\pi^2\kappa^2}{N_e^2}[\alpha_1 l^2 + \alpha_2 lm + \alpha_3 m^2] = 1 - \frac{16\pi^2\kappa^2}{N_e^2}\zeta(l,m), \tag{18}$$

where

$$\alpha_1 = \sum_{j=1}^{N_e-1}\sum_{k=j+1}^{N_e}\frac{x_{jk}^2}{c^2}, \quad \alpha_2 = 2\left(\sum_{j=1}^{N_e-1}\sum_{k=j+1}^{N_e}\frac{x_{jk}y_{jk}}{c^2}\right), \quad \alpha_3 = \sum_{j=1}^{N_e-1}\sum_{k=j+1}^{N_e}\frac{y_{jk}^2}{c^2}, \quad \zeta(l,m) = \alpha_1 l^2 + \alpha_2 lm + \alpha_3 m^2.$$

Therefore, to minimize $|\psi|^2$ the positive function $\zeta(l,m)$ must be maximised. Substituting equations (4) to (6) into ζ yields

$$\zeta(\theta,\phi) = \sin^2\theta(\alpha_1\cos^2\phi + \alpha_2\cos\phi\sin\phi + \alpha_3\sin^2\phi). \tag{19}$$

Clearly, the positive function ζ is maximized when $\theta = \pi/2$. The partial derivative of ζ with respect to ϕ is

$$\begin{aligned}\frac{\delta\zeta}{\delta\phi} &= \sin^2\theta(-2\alpha_1\cos\phi\sin\phi - \alpha_2\sin^2\phi + \alpha_2\cos^2\phi + 2\alpha_3\sin\phi\cos\phi) \\ &= \sin^2\theta[(\alpha_3 - \alpha_1)\sin(2\phi) + \alpha_2\cos(2\phi)].\end{aligned} \tag{20}$$

Setting $\theta = \pi/2$ and $\frac{\delta\zeta}{\delta\phi} = 0$, ζ is maximised by

$$\phi = \frac{1}{2}\tan^{-1}\left(\frac{\alpha_2}{\alpha_1 - \alpha_3}\right) \tag{21}$$

if $\frac{\delta^2\zeta}{\delta\phi^2} < 0$, that is

$$[(\alpha_3 - \alpha_1)\cos(2\phi) - \alpha_2\sin(2\phi)] < 0. \tag{22}$$

By maximising ζ, the arrival direction is found so that the mean delay for the RFI signal between the antennas is maximised, which then gives the largest second eigenvalue.

6 Appropriate Bandwidth for Spatial Filtering

Spatial nulling techniques work by modifying eigenvalues in the measured covariance matrix that are associated with the RFI subspace. For example, orthogonal projection makes those eigenvalues zero. As the number of eigenvalues that are modified increases so does the loss in information [1]. Therefore, the lowest order filter that sufficiently suppresses the RFI is desired. This criterion can be met by setting the channel bandwidth so that the second eigenvalue lies sufficiently below the noise floor

$$\lambda_2 << \sigma_n^2\left(1 + \sqrt{N_e/N_t}\right)^2, \tag{23}$$

where N_t is the number of samples used to estimate the array covariance matrix [1].

Using figure 2(a) the fractional bandwidth required so that the second eigenvalue in figure 2 is sufficiently suppressed is approximately 2.6×10^{-5} (4.8 kHz). For a planar array, the required bandwidth can be calculated by substituting equation (18) into the equation for the ratio λ_2/λ_1 (see equation (16)) and solving for the bandwidth

$$\Delta\nu = 2\sqrt{3}\sqrt{\left[1 - \left(\frac{1-\lambda_2/\lambda_1}{1+\lambda_2/\lambda_1}\right)^2\right]\frac{N_e^2}{16\pi^2\zeta(l,m)}}. \tag{24}$$

A suitable value for λ_2 can be determined by using equation (23). An approximate estimate for λ_1 can be obtained from a covariance matrix, constructed from a signal with the default array bandwidth, by computing the eigenvalue decomposition and subtracting the noise power from the largest eigenvalue. Using the default bandwidth, which is usually larger than the one calculated in equation (24), causes the estimate of λ_1 to be slightly lower and has little effect on the ratio λ_2/λ_1, since $\lambda_1 >> \lambda_2$. The l and m coordinates can either be the worst case scenario (see section 5.1) or the coordinates of the RFI, which is often known in the case of a DAB broadcast.

Increasing the bandwidth so that the second eigenvalue is above the noise and then using second order spatial filtering will not sufficiently remove the RFI. The reason is that the third eigenvalue will then have a significant impact, because no array is perfectly calibrated and the frequency response is not completely flat, which causes the third eigenvalue to be substantially higher than predicted by the model (see figure 2(a) and (b)). To increase the bandwidth that can be processed or the RFI suppression, the algorithms that are proposed in section 8 construct a second order filter that does not require the second eigenvalue to be above the noise.

7 Approximating RFI Vector Space

The RFI covariance matrix in equation (15) can be rewritten in terms of its eigenvalue decomposition

$$\mathbf{R} \approx \sigma^2(\mathbf{a}_1\mathbf{a}_1^H + \mathbf{a}_2\mathbf{a}_2^H) = \lambda_1\mathbf{v}_1\mathbf{v}_1^H + \lambda_2\mathbf{v}_2\mathbf{v}_2^H, \tag{25}$$

which maximises the power in the direction of $\mathbf{v}_1$, with the remaining power contained in the direction of $\mathbf{v}_2$. Both vectors are linear combinations of $\mathbf{a}_1$ and $\mathbf{a}_2$,

$$\mathbf{v}_{1,2} = \beta_{1,2}\mathbf{a}_1 + \gamma_{1,2}\mathbf{a}_2, \tag{26}$$

with the following properties:

- $\mathbf{v}_1^H\mathbf{v}_1 = \mathbf{v}_2^H\mathbf{v}_2 = 1$ ($\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors),
- $\mathbf{v}_1^H\mathbf{v}_2 = \mathbf{v}_2^H\mathbf{v}_1 = 0$ ($\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal).

Using the definition of an eigenvector and eigenvalue

$$\mathbf{R}\mathbf{v} \approx \sigma^2(\mathbf{a}_1\mathbf{a}_1^H + \mathbf{a}_2\mathbf{a}_2^H)(\beta\mathbf{a}_1 + \gamma\mathbf{a}_2) = \sigma^2(\beta\mathbf{a}_1 + \gamma\psi\mathbf{a}_1 + \beta\psi^*\mathbf{a}_2 + \gamma\mathbf{a}_2) = \lambda\mathbf{v} = \lambda\beta\mathbf{a}_1 + \lambda\gamma\mathbf{a}_2.$$

Comparing the components of $\mathbf{a}_1$ and $\mathbf{a}_2$, respectively, yields

$$\lambda\beta = \sigma^2(\beta + \gamma\psi), \tag{27}$$
$$\lambda\gamma = \sigma^2(\beta\psi^* + \gamma). \tag{28}$$

Substituting $\lambda = \lambda_{1,2} = \sigma^2(1 \pm |\psi|)$, $\beta = \beta_{1,2}$ and $\gamma = \gamma_{1,2}$ into equations (27) and (28) yields

$$\beta_{1,2} = \pm\frac{\gamma_{1,2}\psi}{|\psi|}, \tag{29}$$
$$\gamma_{1,2} = \pm\frac{\beta_{1,2}\psi^*}{|\psi|}, \tag{30}$$
$$|\beta_{1,2}|^2 = |\gamma_{1,2}|^2. \tag{31}$$

Using the property that $\mathbf{v}$ is a unit vector gives

$$\begin{aligned}\mathbf{v}^H\mathbf{v} = 1 &= (\beta^*\mathbf{a}_1^H + \gamma^*\mathbf{a}_2^H)(\beta\mathbf{a}_1 + \gamma\mathbf{a}_2) \\ &= |\beta|^2 + \beta^*\gamma\psi + \gamma^*\beta\psi^* + |\gamma|^2,\end{aligned} \tag{32}$$

and substituting equations (30) and (31) yields

$$|\beta_{1,2}|^2 = \frac{1}{2(1 \pm |\psi|)}. \tag{33}$$

The phase of either β or γ can be arbitrarily chosen (see equations (29) to (32)). There is also no phase relationship between β_1, γ_1 and β_2, γ_2 as can be seen by expanding $\mathbf{v}_1^H\mathbf{v}_2 = 0$. The phase of $\beta_{1,2}$ is fixed to 0 and thus $\beta_{1,2} = 1/\sqrt{2(1 \pm |\psi|)}$. Substituting equation (30) into equation (26) yields

$$\mathbf{v}_1 = \frac{1}{\sqrt{2(1 + |\psi|)}}\left[\mathbf{a}_1 + \frac{\psi^*}{|\psi|}\mathbf{a}_2\right], \tag{34}$$
$$\mathbf{v}_2 = \frac{1}{\sqrt{2(1 - |\psi|)}}\left[\mathbf{a}_1 - \frac{\psi^*}{|\psi|}\mathbf{a}_2\right]. \tag{35}$$

To construct $\mathbf{v}_2$ from equation (35) the direction of arrival of the RFI (this can be obtained by using algorithms such as MUSIC and ESPRIT) as well as the signal bandwidth are required.

8 Proposed RFI Mitigation Algorithms

Two new spatial RFI mitigation algorithms based on subspace subtraction are presented in this section. These algorithms are designed for wideband RFI that is stationary, such as DAB broadcasts. The channel bandwidth should be selected so 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 and can be computed using equation (24). The first algorithm is based on the flat frequency response model (see equation (11)) and the other on Zatman's approximation to that model (see equation (35)). The following preprocessing steps are required:

- Use the power iteration method on $\widehat{\mathbf{R}}$ to find the largest eigenvalue s_1 with the accompanying eigenvector $\widehat{\mathbf{v}}_1$.
- Obtain the location of the RFI source. For example, the location of DAB towers can be easily obtained and used as the initial guess for a fast iterative algorithm such as Minimum Error Convergence [7]. If no initial guess can be made, algorithms such as MUSIC or ESPRIT can be used.
- Estimates for the two largest eigenvalues of the RFI only covariance matrix can be obtained by using the estimated location of the RFI and equation (16)

$$\widehat{\lambda}_{\mathrm{r},1} = s_1 - \frac{\mathrm{Tr}(\widehat{\mathbf{R}}) - s_1}{N_\mathrm{e} - 1}, \tag{36}$$

$$\widehat{\lambda}_{\mathrm{r},2} = \lambda_{\mathrm{r},1}\left(\frac{1-|\widehat{\psi}|}{1+|\widehat{\psi}|}\right). \tag{37}$$

  Use these two new eigenvalue estimates to create the matrix $\widehat{\mathbf{S}}_\mathrm{r} = \mathrm{diag}([\widehat{\lambda}_{\mathrm{r},1}, \widehat{\lambda}_{\mathrm{r},2}]^T)$.

Algorithm 1: Flat frequency response model based algorithm (FF algorithm)

- Calculate the normalised flat frequency covariance matrix model of the RFI source $\widehat{\mathbf{R}}_\mathrm{f}$, using equation (11). Note that this model covariance matrix does not include any noise and that $\sigma_\mathrm{s}^2 = 1$.
- Use the power iteration method on $\widehat{\mathbf{R}}_\mathrm{f}$ to find the second largest eigenvalue's eigenvector $\widehat{\mathbf{v}}_{2,\mathrm{f}}$.
- Apply subspace subtraction to obtain the flat frequency model based RFI mitigated covariance matrix

$$\widehat{\mathbf{R}}_{\mathrm{m,f}} = \widehat{\mathbf{R}} - [\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_{2,\mathrm{f}}]\widehat{\mathbf{S}}_\mathrm{r}[\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_{2,\mathrm{f}}]^H. \tag{38}$$

Algorithm 2: Zatman's approximation based algorithm (ZA algorithm)

- Calculate the normalised Zatman's approximation based model eigenvector $\widehat{\mathbf{v}}_{2,\mathrm{z}}$ using equation (35).
- Apply subspace subtraction to obtain the Zatman model based RFI mitigated covariance matrix

$$\widehat{\mathbf{R}}_{\mathrm{m,z}} = \widehat{\mathbf{R}} - [\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_{2,\mathrm{z}}]\widehat{\mathbf{S}}_\mathrm{r}[\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_{2,\mathrm{z}}]^H. \tag{39}$$

9 Evaluation of RFI Mitigation Algorithms

To evaluate the performance of both proposed algorithms, an estimated covariance matrix was created by adding an estimated noise and cosmic source covariance matrix $\widehat{\mathbf{R}}_\mathrm{nc}$ to an estimated RFI covariance matrix $\widehat{\mathbf{R}}_\mathrm{r}$. The matrix $\widehat{\mathbf{R}}_\mathrm{nc}$ was obtained from a real observation done with LOFAR HBA station RS407 where there is no RFI present. A software defined radio was used to record a DAB signal which the power spectrum can be seen in figure 3(a) which has a reasonably flat frequency spectrum. Finite impulse response filters were used to produce 70 signals with bandwidths ranging from 763 Hz to 195 kHz (typical values for LOFAR). Each filtered signal was upsampled and frequency shifted to 145 MHz (the same centre frequency as $\widehat{\mathbf{R}}_\mathrm{nc}$). Covariance matrices were created for each of the 70 processed signals by adding an appropriate delay for each antenna and correlated.

To measure the performance of the proposed algorithms, the Frobenius norm of the difference between the recovered covariance matrix $\widehat{\mathbf{R}}_{\mathrm{m}}$ and matrix $\widehat{\mathbf{R}}_{\mathrm{nc}}$ is used

$$\mathrm{FN} = \sqrt{\sum_{j=1}^{N_{\mathrm{e}}} \sum_{k=1}^{N_{\mathrm{e}}} \left| \widehat{\mathbf{R}}_{\mathrm{nc}}(j,k) - \widehat{\mathbf{R}}_{\mathrm{m}}(j,k) \right|^2} \tag{40}$$

In figure 3b a plot is given of the Frobenius norm as a function of fractional bandwidth. The difference in performance of the FF and the ZA algorithms is less than 10^{-13} and both are represented by the proposed 2nd order line. The 1st order line is the performance achieved by using single frequency subspace subtraction. Close to zero fractional bandwidth, the performance of the 1st order method and that of the 2nd order methods are the same. As the bandwidth increases so does the Frobenius norm for the 1st order method, since the second eigenvalue becomes significant, however the FN for the 2nd order methods shows very little increase. The minimum achievable FN is just under 0.2 and is due to the estimation errors in $\widehat{\mathbf{v}}_1$ and $\widehat{\lambda}_{\mathrm{r},1}$. Channels with larger bandwidth can be processed using the 2nd order methods, while achieving the same level of mitigation as the 1st order method which requires channels with smaller bandwidth. This is shown in figure 3(c) where the bandwidth required by the second order methods is given as a function of the bandwidth of the first order method. Using a fitted straight line indicates an increase by approximately a factor two in bandwidth that can be processed.

The FF algorithm and the ZA algorithm have the same performance for the bandwidths selected. Both algorithms have computational complexity $\mathscr{O}(N_{\mathrm{e}}^2)$ (which includes the preprocessing step making use of the Minimum Error Convergence algorithm). However, the FF algorithm requires an additional calculation of an eigenvector and the computational complexity of the FF algorithm's equation (11) is $\mathscr{O}(N_{\mathrm{e}}^2)$ compared to the ZA algorithm's equation (35) which has computational complexity $\mathscr{O}(N_{\mathrm{e}})$. The quadratic growth of computational time as a function of the number of antennas is shown in figure 3d. The ZA algorithm achieved a median speed-up of 1.3.

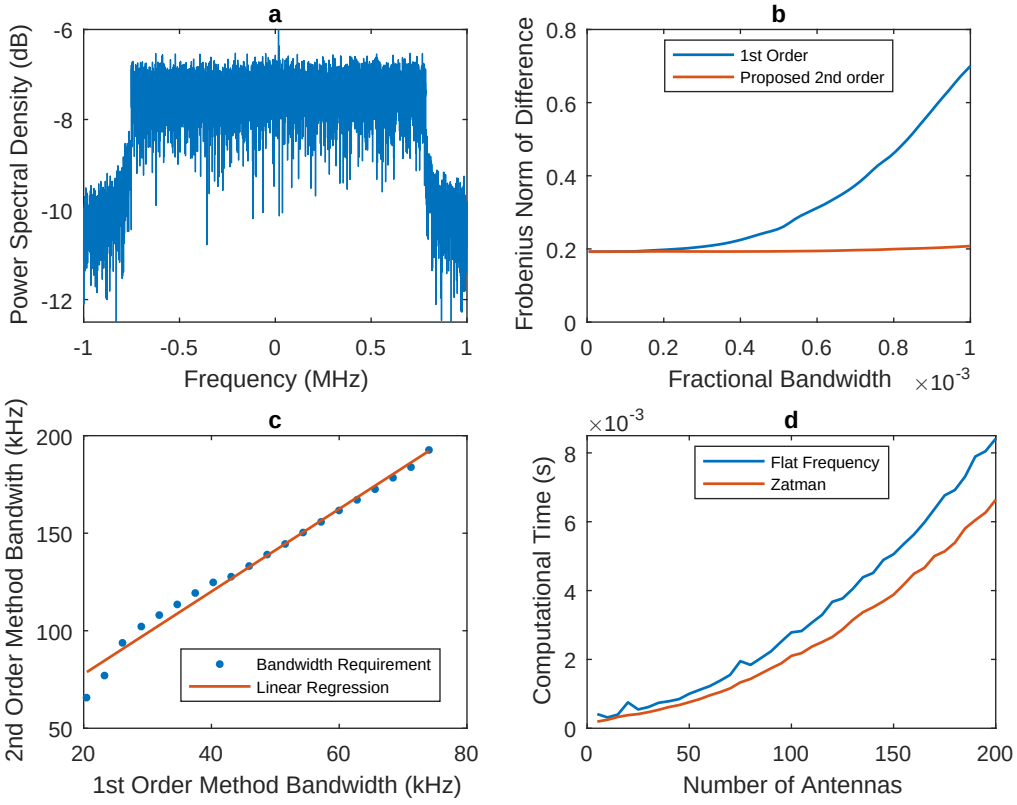

Figure 3. **(a)** Power spectrum of a DAB signal measured with a software defined radio. **(b)** Frobenius norm of the difference between the recovered matrices using RFI mitigation methods and the noise and cosmic source covariance matrix as a function of fractional bandwidth. The 1st order mitigation method makes use of single frequency subspace subtraction. The performance of the FF algorithm and the ZA algorithm are within 10^{-13} of each other and are both represented by the proposed 2nd order line. **(c)** A plot of the bandwidth of the 2nd order methods as a function of the bandwith of the 1st order method to achieve the same attenuation. **(d)** The computational time of both proposed mitigation algorithms as a function of the number of antennas.

10 Conclusion

Strong wideband RFI cannot be modelled as a single point source, but rather as an infinite sum of sources that rapidly decrease in power. For traditional spatial filtering to work on powerful wideband RFI, it must be filtered into sub-bands which are sufficiently narrow so that for each sub-band, the RFI source is a single point source. This greatly increases the computational cost. To reduce this cost, the FF and ZA algorithms are presented that combine a wideband signal model with a subspace subtraction method and, in so doing, decreases the number of sub-bands that must be processed. The proposed algorithms are able to process approximately twice more bandwidth than conventional spatial filtering methods. For bandwidths between 763 Hz and 195 kHz and a LOFAR HBA station layout, the performance of the proposed methods is similar. However, the ZA algorithm showed a speed-up of 1.3 relative to the FF algorithm.

11 Acknowledgements

This work is funded and supported by IBM, ASTRON, the Dutch Ministry of Economic Affairs and the Province of Drenthe, SKA South Africa, 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

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

[2] A. van der Veen, A. Leshem and A. Boonstra, "Signal Processing for Radio Astronomical Arrays," *Processing Workshop Proceedings, Sensor Array and Multichannel Signal*, July 2004, pp. 1–10, doi:10.1109/SAM.2004.1502901.

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

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

[5] A. H. Bridle and F. R. Schwab, "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, 1999, pp. 371–382, eISBN: 978-1-58381-516-8.

[6] M. Zatman, "How narrow is narrowband?," *IEE Proceedings-Radar, Sonar and Navigation*, **145**, 2, April 1998, pp. 85–91, doi:10.1049/ip-rsn:19981670.

[7] J. W. Steeb, D. B. Davidson, and S. J. Wijnholds, "Computationally efficient near-field radio frequency source localisation," *2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, August 2017, pp. 1–4, doi: 10.23919/URSIGASS.2017.8104498.