Array signal processing for radio astronomy [*]

Alle-Jan van der Veen[1], Amir Leshem[2] and Albert-Jan Boonstra[3]
[1] *Delft Univ. of Technology, Dept. Electrical Eng./DIMES, Mekelweg 4, 2628 CD Delft, The Netherlands;* [2] *Bar-Ilan University, Dept. Electrical Eng., Tel-Aviv, Israel; ASTRON, P.O Box 2, 7990 AA Dwingeloo, The Netherlands*

July 12, 2004

Abstract. Radio astronomy forms an interesting application area for array signal processing techniques. Current synthesis imaging telescopes consist of a small number of identical dishes, which track a fixed patch in the sky and produce estimates of the time-varying spatial covariance matrix. The observations sometimes are distorted by interference, e.g., from radio, TV, radar or satellite tranmissions. We describe some of the tools that array signal processing offers to filter out the interference, based on eigenvalue decompositions and factor analysis, a more general technique applicable to partially calibrated arrays. We consider spatial filtering techniques using projections and interference subtraction, and discuss how a reference antenna pointed at the interferer can improve the performance. We also consider image formation and its relation to beamforming. Finally, we briefly discuss some future large scale radio telescopes.

Keywords: Array signal processing, radio astronomy, interference mitigation.

1. Introduction

The future of radio astronomical discoveries depends on achieving better spatial resolution and sensitivity while maintaining immunity to terrestrial interference which is rapidly growing. The last two demands are obviously contradicting as improved sensitivity implies receiving more interfering signals. RFI detection and removal is now an important topic in radio astronomy. A promising track here is to switch to massive phased array technology, where we will gain both in terms of resolution and sensitivity while increasing the flexibility to filter out interference. The international efforts in this direction are coordinated under the framework of the Square Kilometer Array programme (SKA). The first example of a flexible massive phased array radio telescope is LOFAR (13,000 elements) which is currently under construction in The Netherlands.

The principle of interferometry has been used in radio astronomy since 1946 when Ryle and Vonberg constructed a radio interferometer using dipole antenna arrays (Ryl52). In 1962 the principle of aperture

---

synthesis using earth rotation was proposed (Ryl62), and applied for example in the five kilometer Cambridge radio telescope, the 3 km Westerbork Synthesis Radio Telescope (WSRT) in The Netherlands and the 36 km Very Large Array (VLA) in the USA.

In this paper, we present a signal processing data model (section 2) and subsequently give an overview of several problems in radio astronomy where array signal processing can make a contribution, namely calibration using factor analysis (section 3), interference removal using spatial filtering (section 5), and image formation (section 6). We also have a brief look at future radio telescope designs, in particular LOFAR (section 7).

Notation

Superscript t denotes matrix transpose, H denotes complex conjugate transpose, vec($\cdot$) denotes the stacking of the columns of a matrix in a vector, $\otimes$ the Kronecker product. $\mathbf{I}$ is the identity matrix, and $\mathbf{1}$ is a vector with all ones.

2.   Data model

2.1. RECEIVED DATA MODEL

Assume we have a telescope array with p elements. We consider the signals $x_i(t)$ received at the antennas $i = 1, \cdots, p$ in a sufficiently narrow subband. For the interference free case the array output vector $\mathbf{x}(t)$ is modeled in complex baseband form as

$$\mathbf{x}(t) = \mathbf{v}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{x}(t) = [x_1(t), \ldots, x_p(t)]t$ is the $p \times 1$ vector of telescope signals at time t, $\mathbf{v}(t)$ is the received sky signal possibly due to many astronomical sources, assumed on the time scale of (order) 10 s to be a stationary Gaussian vector with covariance matrix $\mathbf{R}_v = \mathcal{E}\{\mathbf{v}(t)\mathbf{v}(t)^H\}$ (the astronomical 'visibilities'), and $\mathbf{n}(t)$ is the $p \times 1$ Gaussian noise vector with covariance matrix $\mathbf{D}$. We assume that the noise is gaussian, and uncorrelated among the sensors, which means that $\mathbf{D}$ is diagonal. Usually identically distributed noise is assumed, for which $\mathbf{D} = \sigma^2\mathbf{I}$, but this implies accurate calibration as discussed in section 3.

Suppose there are q interfering sources, stationary only over short time intervals, with signals $s_i(t)$ for $i = 1 \cdots q$, and spatial signatures $\mathbf{a}_i$. Without loss of generality, we can absorb the unknown amplitude of $s_i(t)$ into $\mathbf{a}_i$ and thus stet the power of $s_i(t)$ to 1. Let $\mathbf{A}$ be a $p \times q$ matrix where the q columns represent the q interferer spatial signature

vectors $\mathbf{a}_i$, and let $\mathbf{s}(t)$ be a vector with the q signals $s_i(t)$. The output vector, extended with interference, is modeled as

$$\mathbf{x}(t) = \mathbf{v}(t) + \mathbf{A}(t)\mathbf{s}(t) + \mathbf{n}(t) \tag{2}$$

We assume that the processing bandwidth is sufficiently narrow, meaning that the maximal propagation delay of a signal across the telescope array is small compared to the inverse bandwidth, so that this delay can be represented by a phase shift of the signal. If the assumption is not satisfied, as for many existing telescopes, a form of subband processing has to be implemented.

2.2. COVARIANCE MODEL

Suppose that we have obtained observations $\mathbf{x}[m] := \mathbf{x}(mT_s)$, where T_s is the sampling period. We assume that $\mathbf{A}(t)$ is stationary at least over intervals of MT_s, and construct short-term covariance estimates $\widehat{\mathbf{R}}_k$,

$$\widehat{\mathbf{R}}_k = \frac{1}{M} \sum_{m=kM+1}^{(k+1)M} \mathbf{x}[m]\mathbf{x}[m]^H \tag{3}$$

where M is the number of samples per short-term average. Several filtering algorithms in this paper are based on applying operations to each $\widehat{\mathbf{R}}_k$ to remove the interference, followed by further averaging over the resulting matrices to obtain a long-term average.

Consider the $\mathbf{A}_k := \mathbf{A}(kMT_s)$ as deterministic, and denote $\mathcal{E}\{\widehat{\mathbf{R}}_k\}$ by $\mathbf{R}_k$. According to the assumptions, $\mathbf{R}_k$ has model

$$\mathbf{R}_k = \mathbf{\Psi} + \mathbf{A}_k\mathbf{A}_k^H = \mathbf{R}_v + \mathbf{D} + \mathbf{A}_k\mathbf{A}_k^H \tag{4}$$

where $\mathbf{\Psi}$ is the interference-free covariance matrix, $\mathbf{\Psi} = \mathbf{R}_v + \mathbf{D}$.

So far, the formalism considered only single polarization arrays. The models are easily extended to the polarization case. Let $\widetilde{\mathbf{x}}(t) \equiv (x_{1x}(t), x_{1y}(t), \cdots, x_{px}(t), x_{py}(t))^t$, where the subscript ix and iy for the ith telescope denote the two orthogonal polarizations. Then the $2p \times 2p$ polarization covariance matrix $\widetilde{\mathbf{R}}$ is defined by $\widetilde{\mathbf{R}} \equiv \mathcal{E}\{\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^t\}$. The resulting polarization data model is described in (Ham00; BV03a). Although the data model is straightforward, extending the non-polarization signal processing to polarization processing is complicated. In this overview paper we therefore focus on single polarization signal processing.

3. Subspace analysis

3.1. EIGENVALUE DECOMPOSITION

The internal structure of the covariance matrix $\mathbf{R}_k$ can be expoited for calibration purposes, for interference mitigation and imaging. Suppose that the noise covariance is equal for each sensor, $\mathbf{R}_n = \sigma_n^2 \mathbf{I}$, assume that the visibilities are much weaker than the noise powers, and assume that $q < p$. Then $\mathbf{R}_k$, dropping the index k, can be decomposed using an eigenvalue analysis as

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \mathbf{\Lambda}_s + \sigma_n^2 \mathbf{I}_q & 0 \\ 0 & \sigma_n^2 \mathbf{I}_{p-q} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{5}$$

where $\mathbf{U}_s$ is the interferer subspace. It is a $p \times q$ matrix containing the eigenvectors corresponding to the q eigenvalues in the $q \times q$ diagonal matrix $\mathbf{\Lambda}_s$. $\mathbf{U}_n$ is a $(p - q) \times (p - q)$ matrix containing the eigenvectors corresponding to the noise subspace. Note that the signal subspace and the noise subspace span the entire space, $\mathbf{U} = [\mathbf{U}_s \mathbf{U}_n]$. Note also that this technique only works for noise matrices with identical diagonal entries. An more general technique is factor analysis which is described next.

3.2. FACTOR ANALYSIS DECOMPOSITION

Factor analysis is a statistical technique with origins in psychometrics and biometrics (LM71; MKB79). It assumes a collection of data $\mathbf{X} = [\mathbf{x}(1), \cdots, \mathbf{x}(N)]$ with covariance

$$\mathbf{R} = \mathcal{E}\{\mathbf{x}(k)\mathbf{x}(k)^H\} = \mathbf{A}\mathbf{A}^H + \mathbf{D} \tag{6}$$

where $\mathbf{R} : p \times p$ Hermitian, $\mathbf{A} : p \times q$ and $\mathbf{D} : p \times p$ diagonal. The objective of factor analysis is, for given $\mathbf{R}$, to identify $\mathbf{A}$ and $\mathbf{D}$, as well as the factor dimension q. We can furthermore model $\mathbf{R}$ in terms of noise subspace $\mathbf{U}_n$ and signal subspace $\mathbf{U}_s$ (LVB00)

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}_0\mathbf{U}^H + \mathbf{D} = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \mathbf{\Lambda}_s & \\ & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} + \mathbf{D} \tag{7}$$

where $\mathbf{U} = [\mathbf{U}_s \mathbf{U}_n]$, and where $\mathbf{\Lambda}$ is the diagonal eigenvalue matrix containing the interferer powers. Here we assume $q \leq p$. Thus, the "Factor Analysis" Decomposition (FAD) can be viewed as a generalization of the eigenvalue decomposition.

This decomposition is relevant in case the noise covariance is unknown but diagonal, $\mathbf{R}_n = \mathbf{D}$, which corresponds to the noise being

uncorrelated among the sensors. In contrast, the usual eigenvalue decomposition for estimating $\mathbf{U}_s$ is only valid if the noise powers are equal among sensors ($\mathbf{R}_n = \sigma^2\mathbf{I}$), which is generally true only after accurate calibration and noise whitening.

In general we can not estimate $\mathbf{A}$ uniquely, since $\mathbf{A}$ can be replaced by $\mathbf{A}\mathbf{V}$ for an arbitrary unitary matrix $\mathbf{V}$. If the eigenvalues are not repeated and we sort them in descending order, then $\mathbf{U}_s$ and $\mathbf{\Lambda}_s$ can be uniquely determined. There are other ways to constrain $\mathbf{A}$ to be a unique factor, e.g. by taking it to be a lower-triangular rectangular Cholesky factor with positive real diagonal entries.

3.3. ESTIMATION OF THE FAD

Assume that the factor rank q is known. Given $\widehat{\mathbf{R}} = \frac{1}{N}\mathbf{X}\mathbf{X}^H$, and a sufficiently small q, we wish to estimate $\mathbf{A}^H$ and $\mathbf{D}^H$ such that $\widehat{\mathbf{R}} \approx \widehat{\mathbf{A}}\widehat{\mathbf{A}}^H + \widehat{\mathbf{D}}^H$. There are several approaches for this.

An ML estimate of the factors $\mathbf{A} : p \times q$ and $\mathbf{D}$ is dependent on the choice of q. The largest permissible value of q is that for which the number of degrees of freedom $v = (p - q)^2 - p \geq 0$, or $q \leq p - \sqrt{p}$. For larger q, there is no identifiability of $\mathbf{A}$ and $\mathbf{D}$: any sample covariance matrix $\widehat{\mathbf{R}}$ can be fitted. Even for smaller q, $\mathbf{A}$ can be identified only up to a $q \times q$ unitary transformation at the right, i.e., we can identify span($\mathbf{A}$). Luckily, this is sufficient for many applications.

For $q > 0$, there is no closed form solution to the estimation of the factors $\mathbf{A}$ and $\mathbf{D}$ in the ML estimation of $\widehat{\mathbf{R}}_q = \mathbf{A}^H\mathbf{A}^H H + \mathbf{D}^H$. There are several approaches for obtaining an estimate.

A technique known as alternating least squares, is to alternatingly minimize $\|\widehat{\mathbf{R}} - (\mathbf{A}\mathbf{A}^H + \mathbf{D})\|_F^2$ over $\mathbf{A}$ keeping $\mathbf{D}$ fixed, and over $\mathbf{D}$ keeping $\mathbf{A}$ fixed. This technique tend to converge very slowly but may be used for fine-tuning.

A fast converging technique is Gauss-Newton iterations on the original (determinant) cost function, or on the (weighted) least squares cost. This requires an accurate starting point.

Ad-hoc techniques exist for solving the least squares problem, possibly followed by a Gauss-Newton iteration. These techniques try to modify the diagonal of $\widehat{\mathbf{R}}$ such that the modified matrix is low-rank q, hence can be factored as $\mathbf{A}\mathbf{A}^H$. The case $q = 1$ was studied in more detail in (BV03b).

3.4. RADIO ASTRONOMY APPLICATION EXAMPLES

We mention two applications of the decompositions described above.

The first example is signal subspace estimation in the presence of uncorrelated noise (LV01), e.g., for the purpose of spatial filtering of
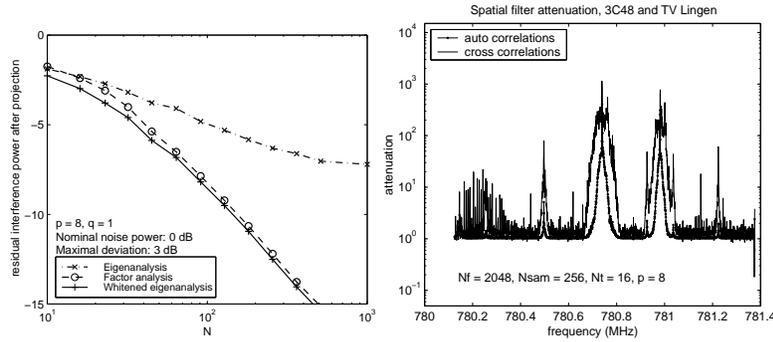
Figure 1. The left figure shows the residual interference power after projections in a simulation: $\mathbf{U}_A$ is estimated from (i) eigenvalue decomposition, (ii) Factor Analysis, and (iii) eigendecomposition after whitening by $\mathbf{D}^{-1/2}$, assuming true $\mathbf{D}$ is known. The right figure shows spatial projection filter attenuation of television sound carrier waves observed at the WSRT. The projection filter was applied after whitening by $\mathbf{D}^{-\frac{1}{2}}$; the diagonal noise term $\mathbf{D}$ was estimated by using factor analysis.

interference. An example is shown in figure 1. Here, the data model is $\mathbf{R} = \mathbf{A}\mathbf{A}^H + \mathbf{D} + \mathbf{R}_v$, where $\mathbf{A}$ corresponds to the interfering signals, $\mathbf{D}$ is the diagonal noise covariance matrix, and $\mathbf{R}_v \ll \mathbf{D}$ is the sky covariance. Using factor analysis, the number of interferers q is detected, and a basis $\mathbf{U}_A \sim \mathrm{ran}(\mathbf{A})$ is estimated, subsequently a projection $\mathbf{P}_A^\perp = \mathbf{U}_A^\perp \mathbf{U}_A^{\perp H}$ is applied to $\mathbf{R}$ to cancel the interference:

$$\widetilde{\mathbf{R}} = \mathbf{P}_A^\perp \widehat{\mathbf{R}} \mathbf{P}_A^\perp \tag{8}$$

The left figure shows $\|\widetilde{\mathbf{R}} - \mathbf{P}_A^\perp (\mathbf{D} + \mathbf{R}_v)\mathbf{P}_A^\perp\|_F$. Clearly, the solution using eigenvalue decompositions is not suitable if the noise covariance is not a multiple of the identity matrix. The right figure shows an application of factor analysis on observed data at the WSRT. It shows projection filter attenuation curves of television sound carrier waves after whitening by $\mathbf{D}^{-\frac{1}{2}}$. The diagonal noise term $\mathbf{D}$ was estimated by using factor analysis.

A second example is gain calibration (BV03b). Initially the antenna gains and noise powers of the telescopes are unknown. To estimate them, a common procedure is to point the telescopes at a strong sky source and make an observation. This produces a rank-1 factor model $\mathbf{R} = \mathbf{g}\sigma_s^2\mathbf{g}^H + \mathbf{D}$, where σ_s^2 is the source power (assumed to be known from tables), $\mathbf{g}$ is the antenna gain vector, and $\mathbf{D}$ is a diagonal matrix containing the noise powers of each antenna. These can be estimated using rank-1 factor analysis. Figure 2, left, shows the principle of a rank-1 column ratio factor analysis. Two columns $\mathbf{c}_i$ and $\mathbf{c}_j$, excluding the diagonal, are related by: $\mathbf{c}_i = \alpha_{ij}\mathbf{c}_j$. The complex "ratio" $\alpha_{ij} = \mathbf{c}_j^\dagger\mathbf{c}_i$ can
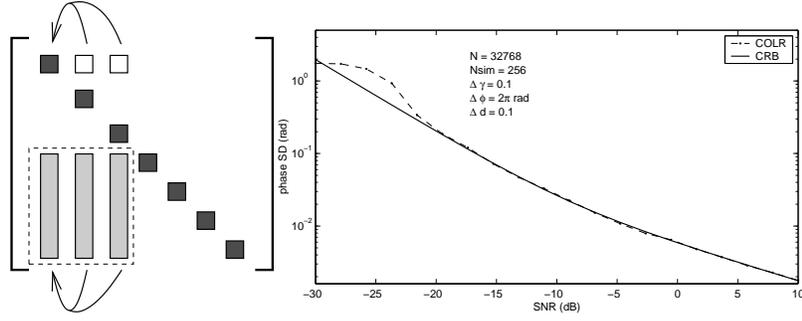
Figure 2. Column ratio factor estimation, principle (left) and estimation accuracy of the method compared to the Cramer-Rao bound (right).

then be used to estimate the diagonal terms. Applying an eigenvalue decomposition on the resulting matrix will yield the factor $\mathbf{A}$, which in this case is a vector $\mathbf{a}$. The right figure shows a simulation of a rank-1 gain estimation problem, where the phase estimation accuracy of the COLR method was compared with the theoretical Cramer-Rao bound.

4. Detection

The detection problem is given by a collection of hypotheses

$$
\begin{aligned}
\mathcal{H}_q : \ & \mathbf{x}(k) \ \sim \ \mathcal{CN}(0, \mathbf{R}_q) \\
\mathcal{H}' : \ & \mathbf{x}(k) \ \sim \ \mathcal{CN}(0, \mathbf{R}') , \qquad q = 1, 2, \cdots
\end{aligned}
\tag{9}
$$

where $\mathcal{CN}(0, \mathbf{R})$ denotes the zero-mean complex normal distribution with covariance $\mathbf{R}$, $\mathbf{R}_q$ is the covariance matrix of the model with q interferers,

$$
\mathbf{R}_q = \mathbf{A}\mathbf{A}^H + \mathbf{D} , \quad \text{where} \quad \mathbf{A} : p \times q , \quad \mathbf{D} \text{ diagonal}
\tag{10}
$$

and $\mathcal{H}'$ corresponds to a default hypothesis of an arbitrary (unstructured) positive definite matrix $\mathbf{R}'$.

The Generalized Likelihood Ratio Test (GLRT) detector for this problem tests $\mathcal{H}_q$ versus $\mathcal{H}'$, where the unknown parameters are replaced by maximum likelihood estimates under each of the hypotheses. In case the noise matrix $\mathbf{D}$ can be written as $\mathbf{D} = \sigma_n^2 \mathbf{I}$, where the noise power σ_n^2 is known, then the test statistic (Box, 1949) can be written as

$$
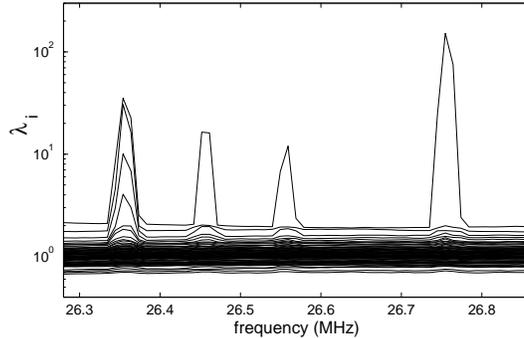T(\mathbf{X}) = -Np \log \prod_{i=1}^{p} \frac{\hat{\lambda}_i}{\sigma_n^2}
\tag{11}
$$

Figure 3. Eigenvalue distribution of the covariance matrix after whitening for an observation at the LOFAR test station (ITS). The figure shows multiple transmitters at 26.36 MHz, and three frequencies with single transmitters.

where N is the number of samples, $\hat{\lambda}_i$ is the ith eigenvalue estimate, and $\mathbf{X} = [\mathbf{x}(1), \cdots, \mathbf{x}(N)]$ is the data. The statistic $T(\mathbf{X})$ is χ^2 distributed, which allows us to select the threshold for a desired false alarm rate(LVB00).

In case the noise matrix $\mathbf{D}$ is diagonal with unknown entries, we can use a more general factor analysis approach (MKB79; LM71), resulting in a maximum likelihood test statistic (vdVLB04) given by

$$T(\mathbf{X}) = N \log |\widehat{\mathbf{R}}_q^{-1} \widehat{\mathbf{R}}| \tag{12}$$

If $\mathcal{H}_q$ is true and N is moderately large (say $N - q \geq 50$), then $2T_q(\mathbf{X})$ has approximately a χ_v^2 distribution with $v = (p - q)^2 - p$ degrees of freedom. This provides a threshold for a test of $\mathcal{H}_q$ versus $\mathcal{H}'$ corresponding to a desired probability of "false alarm" (here the probability of rejecting $\mathcal{H}_q$ when it is true).

A relatively simple to implement test is an eigenvalue threshold test based on an asymptotic formula for the largest singular value of a $p \times M$ white gaussioaan noise matrix (Ede88).

$$\gamma = \sigma_n^2 (1 + \frac{\sqrt{p}}{\sqrt{N}})^2 \tag{13}$$

Figure 3 shows an eigenvalue distribution of pre-whithed data obtained at the (60 antenna-element) LOFAR phased array test station (ITS). Clearly visible is that at three frequencies only a single transmitter can be detected; at 26.36 MHz multiple transmitters are present.

The detection theory can be applied to mitigate intermittent interference. Results concerning detection probabilities and residual interference after detection and excision can be found for example in (LVB00).

Interference detection can also improve spatial filtering approaches, by avoiding application of spatial filters (and resulting distortions) in cases when there is no interference detected.

5. Spatial filtering

Interference cancellation is becoming increasingly important in radio astronomy. Depending on the interference and the type of instrument, several kinds of RFI mitigation techniques are applicable (LVB00; FB01). For intermittent interference, the most effective techniques are based on detection and "blanking": omitting the contaminated samples from the covariance estimate, using a single sensor (Fri96; WFBD97) or multiple sensors (LVB00). For continually present interference and an array of p telescope dishes, spatial filtering is possible. The desired instrument outputs in this case are $p \times p$ correlation matrices, integrated to several seconds (eg, 10 s). Based on short-term correlation matrices (integration to e.g., 10 ms) and narrow subband processing, the array signature vector of an interferer can be estimated and subsequently projected out (RBvdV02)—we describe this technique below.

An interesting option is to utilize a reference antenna which picks up only the interference, so that adaptive cancellation techniques can be implemented (BB98; EBB01). Spatial filtering on extended arrays was first considered by Briggs et al. (BBK00) for a single dual-polarized telescope (two channels) and two reference antennas. Jeffs et al. (JKL03; JLK04) propose spatial filtering algorithms along the lines of (RBvdV02).

5.1. SPATIAL FILTERING USING PROJECTIONS

Suppose that an orthogonal basis $\mathbf{U}_k$ of the subspace spanned by interferer spatial signatures span$(\mathbf{A}_k)$ is known. We can then form a spatial projection matrix $\mathbf{P}_k := \mathbf{I} - \mathbf{U}_k\mathbf{U}_kH$ which is such that $\mathbf{P}_k\mathbf{A}_k = 0$. When this spatial filter is applied to the data covariance matrix all the energy due to the interferer will be nulled: let

$$\mathbf{Q}_k^H := \mathbf{P}_k\widehat{\mathbf{R}}_k\mathbf{P}_k \tag{14}$$

then

$$\mathcal{E}\{\mathbf{Q}_k^H\} = \mathbf{P}_k\boldsymbol{\Psi}\mathbf{P}_k \tag{15}$$

where $\boldsymbol{\Psi} := \mathbf{R}_v + \sigma^2\mathbf{I}$ is the interference-free covariance matrix. When we subsequently average the modified covariance matrices $\mathbf{Q}_k^H$, we ob-
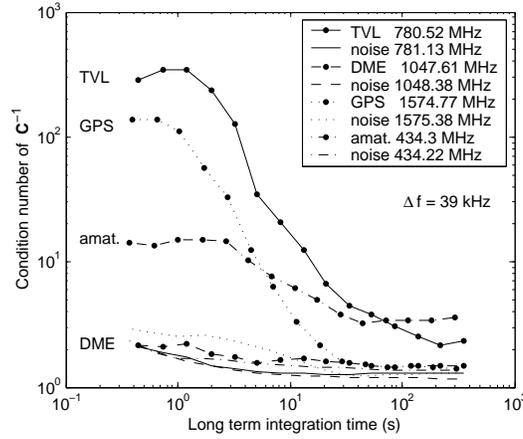
Figure 4. Observed spatial filter correction matrix condition numbers for different observed transmitters at the WSRT telescope.

tain a long-term estimate

$$\mathbf{Q}^H := \frac{1}{N} \sum_{k=1}^{N} \mathbf{Q}_k^H = \frac{1}{N} \sum_{k=1}^{N} \mathbf{P}_k \widehat{\mathbf{R}}_k \mathbf{P}_k . \tag{16}$$

$\mathbf{Q}^H$ is an estimate of $\boldsymbol{\Psi}$, but it is biased due to the projection. A bias correction matrix $\mathbf{C}$ can be derived using the relation $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^t \otimes \mathbf{A})\text{vec}(\mathbf{B})$ (RBvdV02)

$$\mathbf{C} = \frac{1}{N} \sum_{k=1}^{N} \mathbf{P}_k^t \otimes \mathbf{P}_k \tag{17}$$

leading to the following (bias-corrected) estimate of $\boldsymbol{\Psi}$:

$$\boldsymbol{\Psi} := \text{unvec}(\, \mathbf{C}^{-1}\text{vec}(\mathbf{Q}^H))\,. \tag{18}$$

 If the interference was completely projected out then $\boldsymbol{\Psi}$ is an unbiased estimate of the covariance matrix without interference. A detailed analysis of this algorithm will appear in (TV04). The main conclusion is that the variance of the estimate of $\boldsymbol{\Psi}$ is equal to $(1/N)\mathbf{C}^{-1}\sigma^4$, whereas for "clean" data it would be $(1/N)\sigma^4$. For interferers which are sufficiently moving, $\mathbf{C}^{-1}$ is well conditioned and the penalty is comparable to a loss in number of samples. Even for stationary interferers, $\mathbf{C}^{-1}$ might be well conditioned due to the motion of the telescopes, but it depends on the integration length and the location of the sky source which is being tracked. Cases where an interferer enters only on a single

telescope always lead to a singular $\mathbf{C}$ and cannot be resolved by this algorithm.

Figure 4 shows observed condition numbers of $\mathbf{C}$ for different transmitters as a function of long-term integration time. This is relevant because a high condition number implies a strong increase of the system noise due to application of the filter. For fixed location transmitters such as television (TVL) and amateur broadcasts (amat), the condition number decreases to low values (< 5) after about 100 s, as is expected from an analysis of the telescope instrumental fringe rotation. The condition number for the satellite GPS signal decreases more rapidly, because of its motion. Airplane radar (DME) transmits in bursts, the integrated covariance matrices therefore contain many short-time full-rank noise matrices. As a result, the long-term correction matrix $\mathbf{C}$ will have a low condition number.

An alternative to projection filtering is filtering by subtraction. This type of filtering will lead to comparable results. The subtraction filter however will also be biased (LVB00), and needs correction. The attenuation for both projection and subtraction filtering is limited by the spatial signature estimation accuracy, which is described in (LvdV00)

5.2. SPATIAL FILTERING WITH AN EXTENDED ARRAY

If the telescope array is extended with one or more reference antennas, we can follow the same procedure. Let p_0 be the number of primary antennas, and p be the total number of antennas. The data covariance matrix can be partitioned accordingly as

$$\mathbf{R}_k = \left[\begin{array}{cc} \mathbf{R}_{00,k} & \mathbf{R}_{01,k} \\ \mathbf{R}_{10,k} & \mathbf{R}_{11,k} \end{array} \right] . \tag{19}$$

where $\mathbf{R}_k$ has model

$$\begin{aligned} \mathbf{R}_k &= \boldsymbol{\Psi} + \mathbf{A}_k \mathbf{A}_k H = \mathbf{R}_v + \boldsymbol{\Sigma} + \mathbf{A}_k \mathbf{A}_k H \\ &= \left[\begin{array}{c|c} \mathbf{R}_{v,0} + \mathbf{A}_{0,k} \mathbf{A}_{0,k} H + \sigma_0^2 \mathbf{I} & \mathbf{A}_{0,k} \mathbf{A}_{1,k} H \\ \hline \mathbf{A}_{1,k} \mathbf{A}_{0,k} H & \mathbf{A}_{1,k} \mathbf{A}_{1,k} H + \sigma_1^2 \mathbf{I} \end{array} \right] \end{aligned} \tag{20}$$

$\boldsymbol{\Psi}$ is the interference-free covariance matrix, and $\boldsymbol{\Sigma} :=$ diag$[\sigma_0^2 \mathbf{I}, \sigma_1^2 \mathbf{I}]$ is the diagonal noise covariance matrix (assumed known). The objective is to estimate the interference-free covariance submatrix $\boldsymbol{\Psi}_{00} := \mathbf{R}_{v,0} + \sigma_0^2 \mathbf{I}$.

Following the preceding algorithm applied to $\mathbf{R}_k$, the reconstructed covariance matrix is size $p \times p$, and we can simply select the $p_0 \times p_0$ submatrix in the top-left corner, $\boldsymbol{\Psi}_{00}$ (JKL03; JLK04). An improved algorithm would not reconstruct the other blocks of $\boldsymbol{\Psi}$ (VB04). Indeed, let the projected estimates $\mathbf{Q}^H$ be as before in (16). Then (**??**) applies:
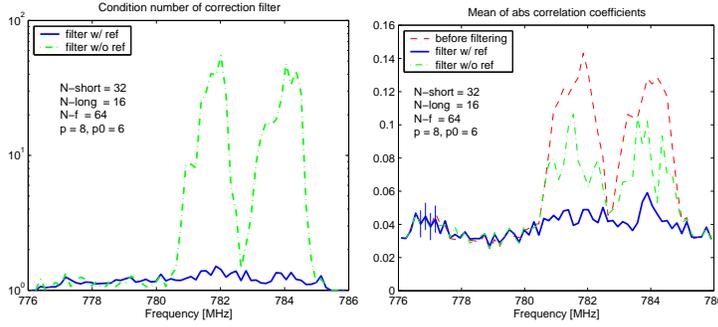
Figure 5. (*a*) Averaged autocorrelation spectrum before and after filtering, (*b*) Averaged cross-correlation spectrum

$$\mathcal{E}\mathrm{vec}(\mathbf{Q}^H) = \mathbf{C}\mathrm{vec}(\boldsymbol{\Psi}) \,. \tag{21}$$

Partition $\boldsymbol{\Psi}$ as in (20) into 4 submatrices. Since we are only interested in recovering $\boldsymbol{\Psi}_{00}$, the other submatrices in $\boldsymbol{\Psi}$ are replaced by their expected values, respectively $\boldsymbol{\Psi}_{01} = \mathbf{0}$, $\boldsymbol{\Psi}_{10} = \mathbf{0}$, $\boldsymbol{\Psi}_{11} = \sigma_1^2\mathbf{I}$. This corresponds to solving the reduced-size covariance model error minimization problem,

$$\boldsymbol{\Psi}_{00} = \arg\min_{\boldsymbol{\Psi}_{00}} \| \mathrm{vec}(\mathbf{Q}^H) - \mathbf{C}\mathrm{vec}(\left[\begin{array}{c|c} \boldsymbol{\Psi}_{00} & \mathbf{0} \\ \hline \mathbf{0} & \sigma_1^2\mathbf{I} \end{array}\right]) \|^2 \,. \tag{22}$$

The solution of this problem reduces to a standard LS problem after separating the knowns from the unknowns. Partition $\mathbf{C}$ in $\mathbf{C}_1$ (corresponding to $\mathrm{vec}(\boldsymbol{\Psi}_{00})$ and $\mathbf{C}_2$ (corresponding to $\mathrm{vec}(\sigma_1^2\mathbf{I})$, then the solution is (VB04):

$$\mathrm{vec}(\boldsymbol{\Psi}_{00}) = \mathbf{C}_1^\dagger(\mathrm{vec}(\mathbf{Q}^H) - \sigma_1^2\mathbf{C}_2\mathbf{1}) \tag{23}$$

If σ_1^2 is unknown, then it can be estimated using a straightforward modification. The advantage compared to the preceding algorithm is that $\mathbf{C}_1$ is a tall matrix, and better conditioned than $\mathbf{C}$. This improves the performance of the algorithm in cases where $\mathbf{C}$ is ill-conditioned, e.g., for stationary interferers, or an interferer entering on only a single telescope. Asymptotically for large INR of the reference array, the algorithm is seen to behave similar to the traditional subtraction technique.

5.3. EXPERIMENT

A reference signal is useful only if it has a better SNR than the primary antennas. Therefore, an omnidirectional antenna is not good enough. To be versatile, we have tested the preceding technique on a reference signal obtained from the beamformed output of a wideband phased array of 64 elements, shown in figure **??**. This system has a bandwidth of 600-1700 MHz, a baseband of 20 MHz, two digital beamforming outputs, and it is part of an envisioned "Thousand Elements Array" (THEA), developed by ASTRON. The reference signal is correlated along with the telescope signals as if it was an additional telescope, and spatial filtering algorithms can be applied to the resulting short-term integrated covariance matrices.

The test data is an observation of the strong astronomical source 3C48 contaminated by Afristar satellite signals. The primary array consists of $p_0 = 6$ of the 14 telescope dishes of the WSRT. The reference signals are $p_1 = 2$ beamforming outputs of the THEA system. One beam was pointed approximately to the satellite, the other was used for scanning. We recorded 65 kSamples at 20 MS/s, and processed these offline. After short-term windowed Fourier transforms, the data was split into 64 frequency bins, correlated, and averaged over 32 samples to obtain 16 short-term covariance matrices.

The resulting auto- and crosscorrelation spectra after filtering are shown in figure 5. The autocorrelation spectra are almost flat, and close to 1 (the whitened noise power). The cross-correlation spectra show that the spatial filtering with reference antenna has done much better to remove the interference than the case without reference antenna. The residual correlation of about 4% is known to be the SNR of the astronomical source. The lines are noisy due to the finite sample effect; the predicted standard deviation (based on number of samples averaged) are indicated for a few frequencies.

6.　Imaging

6.1. MATRIX FORMULATION

As described in more detail in (LV00), image formation is also a fruitful area for array signal processing techniques. Astronomers try to estimate the intensity (brightness) $I_f(\mathbf{s})$ of the sky as a function of the location $\mathbf{s}$ and frequency f. They do this by measuring the correlation (called the "visibility" V_f) between identical sensors i and j with locations $\mathbf{r}_i$ and $\mathbf{r}_j$, corresponding to a baseline $\mathbf{r}_i - \mathbf{r}_j$. Let (ℓ, m) denote normalized coordinates of the sky source $(-1 \leq \ell, m \leq 1)$, and (u, v, w) the

baseline vector of the antenna pair measured in wavelengths. Assuming a planar array, w can be removed from the equations via geometrical delay compensation. Under certain approximations, the "measurement equation" is given by (PSB94)

$$V_f(u, v) \;=\; \iint I_f(\ell, m) \, e^{-j2\pi(u\ell + vm)} \, d\ell dm \,. \qquad (24)$$

It has the form of a Fourier transformation.

The function $V_f(u, v)$ is sampled at various coordinates $(u_{ij}(t), v_{ij}(t))$ by first of all taking all possible sensor pairs i, j or baselines $\mathbf{r}_i - \mathbf{r}_j$, and second by realizing that the sensor locations $\mathbf{r}_i, \mathbf{r}_j$ are actually time-varying since the earth rotates. Given a sufficient number of samples in the (u, v) domain, the relation can be inverted to obtain an image (the 'map').

Assume that the sky consists of a large number (d) point sources. Equation (24) can then be written slightly differently as

$$V(u_{ij}(t), v_{ij}(t)) =$$
$$= \sum_{l=1}^{d} e^{-j2\pi(u_{i0}(t)\ell_l + v_{i0}(t)m_l)} \cdot I(\ell_l, m_l) \cdot e^{j2\pi(u_{j0}(t)\ell_l + v_{j0}(t)m_l)} \,. \qquad (25)$$

where (u_{i0}, v_{i0}) are coordinates of the ith antenna with respect to a common reference point. The connection to our previous framework is obtained by collecting the visibilities into correlation matrices $\mathbf{R}$, where $\mathbf{R}_{ij}(t) = V(u_{ij}(t), v_{ij}(t))$. The above equation can then be written as

$$\mathbf{R}_k \;=\; \mathbf{A}_k \mathbf{B} \mathbf{A}_k^H \qquad (26)$$

where $\mathbf{R}_k \equiv \mathbf{R}(t_k)$, $\mathbf{A}_k = [\mathbf{a}_k(\ell_1, m_1), \ldots, \mathbf{a}_k(\ell_d, m_d)]$, and

$$\mathbf{a}_k(\ell, m) \;=\; \begin{bmatrix} e^{-j2\pi(u_{10}(t_k)\ell + v_{10}(t_k)m)} \\ \vdots \\ e^{-j2\pi(u_{p0}(t_k)\ell + v_{p0}(t_k)m)} \end{bmatrix} \qquad (27)$$

$$\mathbf{B} \;=\; \begin{bmatrix} I(\ell_1, m_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & I(\ell_d, m_d) \end{bmatrix}$$

where $\mathbf{a}_k(\ell, m)$ is recognized as the array response vector. As usual, the array response is frequency dependent. The response is also slowly time-varying due to the earth rotation.

6.2. INVERSE FOURIER IMAGING

6.2.1. *Classical inverse Fourier imaging*

The relation between sky brightness $I(\ell, m)$ and visibilities $V(u, v)$ (where u, v are taken at frequency f) is given by the measurement equation (24). We have measured V on a discrete set of baselines $\{(u_i, v_i)\}$. The "dirty image" (a lumpy image obtained via direct Fourier inversion possibly modified with some weights c_i) is defined by

$$I_D(\ell, m) := \sum_i c_i \, V(u_i, v_i) \, e^{j2\pi(u_i\ell + v_i m)} \qquad (28)$$

It is equal to the 2D convolution of the true image $I(l, m)$ with a point spread function $B_0(l, m)$ known as the "dirty beam":

$$I_D = I * B_0, \qquad B_0(\ell, m) := \sum_i c_i \, e^{j2\pi(u_i\ell + v_i m)} \qquad (29)$$

B_0 is the dirty beam, centered at the origin. The weights $\{c_i\}$ are arbitrary coefficients designed to obtain an acceptable beam-shape, with low side lobes, in spite of the irregular sampling.

Specializing to a point source model, $I(\ell, m) = \sum_l I_l \, \delta(\ell - \ell_l, m - m_l)$ where I_l is the intensity of the source at location (ℓ_l, m_l), gives

$$V(u, v) = \sum_l I_l \, e^{-j2\pi(u\ell_l + vm_l)} \qquad (30)$$

$$I_D(\ell, m) = \sum_l I_l \, B_0(\ell - \ell_l, m - m_l) \qquad (31)$$

Thus, every point source excites the dirty beam centered at its location (ℓ_l, m_l).

From the dirty image I_D and the known dirty beam B_0, the desired image I is obtained via a deconvolution process. A popular method for doing this is the CLEAN algorithm (Hog74). The algorithm assumes that B_0 has its peak at the origin, and consists of a loop in which a candidate location (ℓ_l, m_l) is selected as the largest peak in I_D, and subsequently a small multiple of $B_0(\ell - \ell_l, m - m_l)$ is subtracted from I_D. The objective is to minimize the residual, until it converges to the noise level.

6.2.2. *Inverse Fourier imaging after projections*

If we take projections or any other linear combination $[c_{ij}]$ of the visibilities $\{V(u_i, v_i)\}$ during measurements we have instead available

$$Z(u_i, v_i) = \sum_j c_{ij} \, V(u_j, v_j) \qquad (32)$$

Suppose we compute the dirty image in the same way as before, but now from Z, then it can be shown (vdVLB04) that the dirty image is obtained via a convolution, but the dirty beam is now space-varying. Nonetheless, they are completely known if we know the linear combinations that we took during observations. Thus, the CLEAN algorithm can readily be modified to take the varying beam shapes into account: simply replace $B_0(\ell, m)$ by $B_l(\ell, m)$ everywhere in the algorithm.

6.3. IMAGING VIA BEAMFORMING TECHNIQUES

6.3.1. *CLEAN and sequential beamforming*

Using a parametric point-source model, the image deconvolution problem can be interpreted as a direction-of-arrival (DOA) estimation problem, e.g., as

$$[\{\hat{\mathbf{s}}_l\}, \mathbf{B}^H] = \arg \min_{\{\mathbf{s}_l\}, \mathbf{B}} \sum_{k=1}^{K} \| \widehat{\mathbf{R}}_k - \mathbf{A}_k(\{\mathbf{s}_l\}) \mathbf{B} \mathbf{A}_k^H(\{\mathbf{s}_l\}) - \sigma^2 \mathbf{I} \|_F \quad (33)$$

($\mathbf{B}$ is constrained to be diagonal with positive entries.) This is recognized as the same model as used for DOA estimation in array processing. Note however that the array is moving ($\mathbf{A}_k$ is time-dependent), and that there are many more sources than the dimension of each covariance matrix.

In this notation, the image formation in section 6.2.1 can be formulated as follows. If we write $I_D(\mathbf{s}) \equiv I_D(\ell, m)$ and $\mathbf{a}_k(\mathbf{s}) \equiv \mathbf{a}_k(\ell, m)$, we can rewrite the dirty image (28) as (**?**)

$$I_D(\mathbf{s}) = \sum_k \mathbf{a}_k^H(\mathbf{s}) \mathbf{R}_k \mathbf{a}_k(\mathbf{s}) \quad (34)$$

We omitted the optional weighting. Also note that, with noise, we have to replace $\mathbf{R}_k$ by $\mathbf{R}_k - \sigma^2 \mathbf{I}$. The iterative beam removing in CLEAN can now be posed as an iterative LS fitting between the sky model and the observed visibility (Sch78). Finding the brightest point $\mathbf{s}_0$ in the image is equivalent to trying to find a point source using classical Fourier beamforming, i.e., ,

$$\mathbf{s}_0^H = \arg \max_{\mathbf{s}} \sum_{k=1}^{K} \mathbf{a}_k H(\mathbf{s}) \left(\mathbf{R}_k - \sigma^2 \mathbf{I} \right) \mathbf{a}_k(\mathbf{s}). \quad (35)$$

Thus, the CLEAN algorithm can be regarded as a generalized classical sequential beamformer, where the brightest points are found one by one, and subsequently removed from $\mathbf{R}_k$ until the LS cost function (33) is minimized. An immediate consequence is that the estimated

source locations will be biased, a well known fact in array processing. When the sources are well separated the bias is negligible compared to the standard deviation, otherwise it might be significant. This gives an explanation for the poor performance of the CLEAN in imaging extended structures (see e.g., (PSB94)).

6.3.2. *Minimum variance beamforming approaches*

Once we view image formation/deconvolution as equivalent to direction-of-arrival (DOA) estimation with a moving array, we can try to adapt various other DOA estimators for handling the image formation. In particular the deflation approach used in the CLEAN algorithm can be replaced by other source parameters estimators. One approach that seems particularly relevant in this context is the Minimum-Variance Distortionless Response (MVDR) method of beamforming (Cap69). The major new aspect here is the fact that the array is moving and that there are more sources than sensors.

Instead of working with the dirty image $I_D(\mathbf{s}) = \sum_k \mathbf{a}_k H(\mathbf{s})\mathbf{R}_k\mathbf{a}_k(\mathbf{s})$, the basis for high-resolution beamforming techniques is to look at more general "pseudo-spectra"

$$I_D'(\mathbf{s}) \ := \ \sum_k \mathbf{w}_k^H(\mathbf{s})\mathbf{R}_k\mathbf{w}_k(\mathbf{s}) \tag{36}$$

Here, $\mathbf{w}_k(\mathbf{s})$ is the beamformer pointing towards direction $\mathbf{s}$, and $I_D'(\mathbf{s})$ is the output energy of that beamformer. Previously we had $\mathbf{w}_k(\mathbf{s}) = \mathbf{a}_k(\mathbf{s})$; the objective is to construct beamformers that provide better separation of close sources.

A generalized MVDR follows by defined by a minimization of a weight vector $\mathbf{w}_k$ which minimizes the output power at time k subject to the constraint that we have a fixed response towards the look direction $\mathbf{s}$ of the array. Solving this (well known) problem leads to the overall spectral estimator

$$I_D'(\mathbf{s}) \ = \ \sum_{k=1}^{K} \frac{1}{\mathbf{a}_k H(\mathbf{s})\widehat{\mathbf{R}}_k^{-1}\mathbf{a}_k(\mathbf{s})} \ . \tag{37}$$

and the locations of the strongest sources are given by the maxima of $I_D'(\mathbf{s})$. It is known that the MVDR has improved resolution compared to the classical beamformer which is the basis for the CLEAN algorithm. The algorithm is readily extended to handle the "space-varying" beamshapes that occur after spatial filtering. It is also possible to use more advanced forms of beamforming, e.g., "robust Capon beamforming" (RCB) (SWL03).
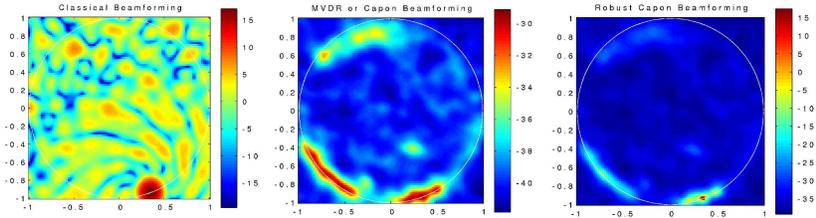
Figure 6. Examples of "dirty images" from the LOFAR test station: classical beamforming (left), MVDR image (middle), and Robost Capon beamforming (right).

Figure 6 illustrates this by comparing a dirty image produced in the classical way to the dirty image corresponding to (37) and to Robust Capon beamforming (BvdT04). The measurement data is a "snapshot" collected from a 60-element test station for the LOFAR telescope. Since this is a two-dimensional array, it does not depend on earth rotation to enable imaging. Due to the limited integration time, the sky sources are not yet observed and only interference shows up, which is visible at the horizon. All other features are due to the sidelobes of the dirty beam. An MVDR beamformer usually has sharper beams than a classical beamformer. A disadvantage of MVDR beamformers however is that due to array calibration errors, the scaling of the spectral estimator is affected. This problem is remedied in the Robust capon beamformer, also shown in figure 6.

7. Conclusion

Technological advances in the last decade have created interesting possibilities for large distributed interferometric radio telescopes with very large receiving areas and a sensitivity which is one to two orders of magnitude better than the current generation. Two proposed instruments in this context are LOFAR (wsa; Bre99) and SKA (wsb; AB02); the previously mentioned THEA system is a smallscale step-up in the design of SKA. Prominent among the challenges of designing and building these telescopes (apart from the costs) are the mitigation of radio interference, the calibration of the system, and the sheer signal processing complexity. Signal processing techniques such as discussed in this paper are vital to meet the sensitivity requirements for the next generation of radio telescopes.

8. Acknowledgements

References

A. van Ardenne and H. Butcher. Square kilometer arrray project: Concepts and technologies. *URSI GA*, 2002.

C. Barnbaum and R.F. Bradley. A new approach to interference excision in radio astronomy: Real-time adaptive cancellation. *The Astronomical Journal*, (115):2598–2614, November 1998.

F.H. Briggs, J.F. Bell, and M.J. Kesteven. Removing radio interference from contaminated astronomical spectra using an independent reference signal and closure relations. *The Astronomical Journal*, 120:3351–3361, 2000.

J.D. Bregman. Design concepts for sky noise limited low frequency array. In *Perspectives on Radio Astronomy - Technologies for Large Antenna Arrays*, pages 23–32. ASTRON, 1999.

A.J. Boonstra and A.J. van der Veen. Dual polarization gain estimation for radio telescope arrays. *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, April 2003.

A.J. Boonstra and A.J. van der Veen. Gain calibration methods for radio telescope arrays. *IEEE Transactions on Signal Processing*, 51(1):25–38, January 2003.

A.J. Boonstra and S. van der Tol. Spatial filtering of interfering signals at the initial LOFAR phased array test station. *Radio Science, special issue on large scale radio telescopes. To be submitted*, 2004.

J. Capon. High resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, pages 1408–1418, 1969.

S.W. Ellingson, J. Bunton, and J.F. Bell. Removal of the GLONASS C/A signal from OH spectral line observations using a parametric modelling. *Astrophysical Journal Supplement*, 135(1):88–93, July 2001.

A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, 9(4), 1988.

P.A. Fridman and W.A. Baan. RFI mitigation methods in radio astronomy. *Astronomy and Astrophysics*, 378:327–344, October 2001.

P.A. Fridman. A change point detection method for elimination of indiustrial interference in radio astronomy receivers. *8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Corfu, Greece*, pages 264–266, June 1996.

J.P Hamaker. Understanding radio polarimetry. IV. the full-coherency analogue of scalar self-calibration: Self-alignment, dynamic range and polarimetric fidelity. *åps*, 143:515–534, May 2000.

J.A. Hogbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. Astrophys. Suppl.*, 15:417–426, 1974.

B.D. Jeffs, K.Warwick, and L. Li. Improved interference cancellation in synthesis array radio astronomy using auxiliary antennas. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003)*, 5:77–80, April 2003.

B.D. Jeffs, L. Li, and K.Warwick. Auxiliary antenna assisted interference cancellaton for radio astronomy arrays. *Accepted for IEEE Trans. Signal Processing*, 2004.

D.N. Lawley and A.E. Maxwell. *Factor Analysis as a Statistical Method*. Butterworth & Co, London, 1971.

A. Leshem and A.J. van der Veen. Radio astronomical imaging in the presence of strong radio interference. *IEEE Tr. Information Th.*, 46(5):1730–1747, August 2000.

A. Leshem and A.J. van der Veen. Multichannel detection and spatial signature estimation with uncalibrated receivers. *Proc. 11th IEEE Workshop on Stat. Signal Proc., Singapore*, August 2001.

A. Leshem, A.J. van der Veen, and A.J. Boonstra. Multichannel interference mitigation techniques in radio astronomy. *The Astrophysical Journal Supplement Series*, 131(1):355–373, November 2000.

A. Leshem and A.J. van der Veen. The effect of adaptive interference suppression on radio astronomical image formation. *Radio Telescopes*, 4015:341–352, 2000.

K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*. Probability and Mathematical Statistics. Academic Press, London, 1979.

R.A. Perley, F.R. Schwab, and A.H. Bridle. *Synthesis Imaging in Radio Astronomy*, volume 6. Astronomical Society of the Pacific Conference Series, 1994.

J. Raza, A.J. Boonstra, and A.J. van der Veen. Spatial filtering of RF interference in radio astronomy. *IEEE Signal Processing Letters*, 9(2):64–67, February 2002.

M. Ryle. A new radio interferometer and its application to the observation of weak stars. *Proc. Royal Society A*, 211:351–375, 1952.

M. Ryle. The new Cambridge radio tlescope. *Nature*, 194:517–518, 1962.

U. J. Schwarz. Mathematical-statistical Description of the Iterative Beam Removing Technique (Method CLEAN). *AP*, 65:345+, April 1978.

P. Stoica, Z. Wang, and J. Li. Robust Capon beamforming. *IEEE Signal Processing Letters*, 10:172–175, June 2003.

S.V. van der Tol. and A.J. van der Veen. Performance analysis of spatial filtering of rf interference in radio astronomy. *IEEE Transactions on Signal Processing*, 2004.

A.J. van der Veen and A.J. Boonstra. Spatial filtering of RF interference in radio astronomy using a reference antenna. *IEEE, ICASSP 2004, Montreal, Canada*, May 2004.

A.J. van der Veen, A. Leshem, and A.J. Boonstra. Signal processing for radio astronomical arrays. *Sensor Array and Multichannel (SAM) technical committee of the IEEE Signal Processing Society, Sitges, Barcelona, Spain.*, July 2004.

R. Weber, C. Faye, F. Biraud, and J. Dansou. Spectral detector for interference time blanking using quantized correlator. *Astronomy and Astrophysics Supp.*, 126(1):161–167, November 1997.

LOFAR web site. http://www.lofar.org.

SKA web site. http://www.ska.org.