

APPLICATION OF ROBUST CAPON BEAMFORMING TO RADIO ASTRONOMICAL IMAGING

Sebastiaan van der Tol and Alle-Jan van der Veen

Delft Univ. of Technology, Dept. Electrical Eng./DIMES, Mekelweg 4, 2628 CD Delft, The Netherlands

Radio telescopes based on large phased arrays form an interesting application area for array signal processing. LOFAR is a large low frequency (10Mz–240MHz) array consisting of 13,000 antennas grouped into 50 stations, currently under construction in the Netherlands. Data from a 60-element test station of LOFAR is available to evaluate the performance of calibration and imaging algorithms. In this paper we apply the Robust Capon Beamformer (RCB) to make images of the sky from measured data, and compare them to the classical Fourier-based images. The RCB takes uncertainty in the calibration into account. Instead of the usual spherical uncertainty sets, we have also derived a more constrained uncertainty set specifically for imaging with the RCB. The results are images with a higher dynamic range than classical or Capon beamforming. Additional simulations confirm that the images are more accurate.

1. INTRODUCTION

Future radio telescopes will be based on massive phased arrays, typically with 10,000s to 100,000s elements, dispersed over several locations. A major challenge for arrays of this size is to find efficient algorithms for calibration and imaging.

An example of such a telescope, currently under construction, is LOFAR (Low Frequency Array) [2, 3]. It will consist of 13,000 antennas grouped into 50 stations. Data is processed both at station and central level. At station level the signals are combined (beamformed) to act as a single sensor for the central level. An Initial Test Station (ITS) for LOFAR has been built consisting of 60 antennas operating in the 10MHz–40MHz range, and data is available to test algorithms for calibration, imaging and interference mitigation.

Radio astronomical imaging has always been based on what is known in array signal processing as classical (or Fourier-based) beamforming. In a previous paper, Leshem and Van der Veen have considered the use of Capon beamforming [4], but actual images have not been shown. In this paper we discuss the application of the Robust Capon Beamformer (RCB) [5–7] and test it on measured data from the ITS. RCB can use an uncertainty set to describe the calibration error. We have derived a specific uncertainty set to improve the imaging. Simulations on artificial data show that the images made by the Robust Capon Beamformer are closer to the model than the classical and the Capon beamformer.

The calibration of LOFAR at the central level consists of iteratively solving for a model including astronomical sources, the time-varying ionosphere, and instrument parameters (a generalized SELF-CAL loop). The good performance of the Robust Capon Beamformer in the presence of calibration errors makes it a candidate to be included in the LOFAR calibration either as an

initial step or as part of the loop.

2. DATA MODEL

Assume we have a telescope array with p elements. We consider the baseband signals $x_i(t)$ received at the antennas $i = 1, \dots, p$ in a sufficiently narrow subband. The $p \times 1$ array output vector $\mathbf{x}(t)$ obtained by stacking the signals $x_i(t)$ has a time-varying covariance matrix $\mathbf{R}(t) = E(\mathbf{x}\mathbf{x}^H)$, because the telescope array fixed on earth has a slowly rotating view of the sky. For imaging, the astronomer is interested only in the non-redundant off-diagonal entries of $\mathbf{R}(t)$, which contain the astronomical “visibilities”. The measurement noise covariance matrix is assumed to be diagonal and can therefore be ignored in this paper.

The telescope signals are sampled ($\mathbf{x}[m] := \mathbf{x}(mT_s)$, where T_s is the sampling period), correlated, and averaged, to obtain the short-term estimates called “snapshots”

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

Each estimate spans about 10–30 s of data and is an estimate of $\mathbf{R}(t)$ at the corresponding time instance, referred in future by the index k .

A model for \mathbf{R}_k is obtained as follows [4]. Astronomers try to estimate the intensity (brightness) $I_f(\mathbf{s})$ of the sky as a function of the unit-norm location vector \mathbf{s} and frequency f . They do this by measuring the correlation (“visibility”) V_f between identical sensors i and j with locations $\mathbf{r}_i[k]$ and $\mathbf{r}_j[k]$, corresponding to a baseline $\mathbf{r}_i[k] - \mathbf{r}_j[k]$.

Let (ℓ, m) denote normalized coordinates of the sky source ($-1 \leq \ell, m \leq 1$), and (u, v, w) the coordinates of the baseline vector between two telescopes, measured in wavelengths. Assuming a planar array, w can be removed from the equations via geometrical delay compensation. Under certain approximations, the “measurement equation” can be written as [8]

$$V(u, v) = \iint I(\ell, m) e^{-j2\pi(u\ell + vm)} d\ell dm. \quad (1)$$

It has the form of a Fourier transformation. Assume that the sky consists of a large number (d) point sources, i.e.,

$$I(\ell, m) = \sum_{l=1}^d I(\ell_l, m_l) \delta(\ell - \ell_l, m - m_l)$$

where $I(\ell_l, m_l)$ is the intensity of the source at location (ℓ_l, m_l) and $\delta(\cdot)$ is the dirac function. The relation (1) then becomes

$$V(u, v) = \sum_{l=1}^d I(\ell_l, m_l) e^{-j2\pi(u\ell_l + vm_l)} \quad (2)$$

which can also be written in more detailed notation as

$$\mathbf{R}_{ij}[k] = \sum_{l=1}^d e^{-j2\pi(u_{i0}[k]\ell_l + v_{j0}[k]m_l)} \cdot I(\ell_l, m_l) \cdot e^{j2\pi(u_{j0}[k]\ell_l + v_{i0}[k]m_l)},$$

*This research was supported in part by the STW under NWO-STW under the VICI programme (DTC.5893) and the ICT Delft Research Center. Parts of this paper were presented in the overview paper [1].

where $\mathbf{R}_{ij}[k]$ is entry i, j of $\mathbf{R}(t)$ at time k , equal to the visibility $V(u_{ij}[k], v_{ij}[k])$, and $(u_{i0}[k], v_{i0}[k])$ is the location of the i -th telescope at time k with reference to a stationary point. This equation can subsequently be written in matrix form as [4]

$$\mathbf{R}_k = \mathbf{A}_k \mathbf{B} \mathbf{A}_k^H$$

where $\mathbf{A}_k = [\mathbf{a}_k(\ell_1, m_1), \dots, \mathbf{a}_k(\ell_d, m_d)]$, $\mathbf{a}_k(\ell, m)$ is the array response vector,

$$\mathbf{a}_k(\ell, m) = \begin{bmatrix} e^{-j2\pi(u_{i0}[k]\ell + v_{i0}[k]m)} \\ \vdots \\ e^{-j2\pi(u_{p0}[k]\ell + v_{p0}[k]m)} \end{bmatrix} \quad (3)$$

and

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

As usual, the array response is frequency dependent. The response is slowly time-varying due to the earth rotation.

3. ASTRONOMICAL IMAGING

3.1. 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 (1). 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) is defined by

$$I_D(\ell, m) := \sum_i V(u_i, v_i) e^{j2\pi(u_i\ell + v_i m)} \quad (4)$$

Inserting the model (2), it is seen to be equal to the 2D convolution of the true image I with a point spread function known as the ‘‘dirty beam’’:

$$I_D(\ell, m) = I * B_0 = \sum_l I_l B_0(\ell - \ell_l, m - m_l)$$

$$B_0(\ell, m) := \sum_i e^{j2\pi(u_i\ell + v_i m)}$$

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, e.g., the CLEAN algorithm [9]. In this algorithm, the location of the maximum of the image is interpreted as a source location, and the corresponding value as its power. The algorithm proceeds by subtracting a small multiple of the dirty beam, centered at this location and scaled by the estimated power. It is therefore important that these locations and estimates are accurate.

We can rewrite (4) as

$$\begin{aligned} I_D(\ell, m) &= \sum_{i,j,k} V(u_{ij}[k], v_{ij}[k]) e^{j2\pi(u_{i0}[k]\ell + v_{i0}[k]m)} \\ &\quad \cdot e^{-j2\pi(u_{j0}[k]\ell + v_{j0}[k]m)} \\ &= \sum_k \mathbf{a}_k^H(\ell, m) \mathbf{R}_k \mathbf{a}_k(\ell, m). \end{aligned} \quad (5)$$

Therefore, the pixels of the dirty image can be interpreted as being the output powers of the classical beamformer.

An example of a dirty image is shown in figure 1(a). The measurement data is a 1-minute ‘‘snapshot’’ ($K = 1$) collected from

a 60-element test station for the LOFAR telescope.¹ 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. Figure 2(a) shows a similar observation but at a slightly different frequency where interference is more dominant: the sidelobes of a single interferer dominate the complete image.

It is well known in array processing that the estimated source locations of this beamformer are biased. 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., [8]).

3.2. Capon Beamformer

From now on, we assume for simplicity of notation that only a single snapshot is available, therefore we will omit the summation over k . In more general terms, we can write (5) as the output power of a beamformer:

$$\hat{I}(\ell, m) = \mathbf{w}(\ell, m)^H \mathbf{R} \mathbf{w}(\ell, m)$$

where

$$\mathbf{w}(\ell, m) = \mathbf{a}(\ell, m).$$

The weights of the classical beamformer (5) are independent of the data. It is known that image quality can be improved by using a data dependent beamformer. Capon beamforming gives a significant suppression of the sidelobes compared to classical beamforming. The Capon beamformer minimizes the output power under the constraint that the gain in the desired direction remains unity:

$$\mathbf{w}_{\text{capon}}(\ell, m) = \arg \min_{\mathbf{w}} \mathbf{w}(\ell, m)^H \mathbf{R} \mathbf{w}(\ell, m)$$

$$\text{such that } \mathbf{w}_{\text{capon}}(\ell, m)^H \mathbf{a}(\ell, m) = 1.$$

The solution to this equation is known to be

$$\mathbf{w}_{\text{capon}}(\ell, m) = \frac{\mathbf{R}^{-1} \mathbf{a}(\ell, m)}{\mathbf{a}(\ell, m)^H \mathbf{R}^{-1} \mathbf{a}(\ell, m)}$$

whereas the estimated source power is

$$\hat{I}_{\text{capon}}(\ell, m) = \mathbf{w}_{\text{capon}}(\ell, m)^H \mathbf{R} \mathbf{w}_{\text{capon}} = \frac{1}{\mathbf{a}(\ell, m)^H \mathbf{R}^{-1} \mathbf{a}(\ell, m)}.$$

Figure 1(b) shows the corresponding dirty image in the LOFAR example. It is possible to recover Cassiopeia A, which is the strongest extrasolar radio source in the sky. Figure 2(b) shows the same at a different frequency where Cas-A is not expected to be visible due to ionospheric occlusion. The interference is now seen to be much more localized. Also note that it appears as if there is a ring of interference at the horizon, whereas figure 2(a) only hinted at a single interferer.

3.3. Robust Capon Beamformer (RCB)

Power estimates (images) from an Capon beamformer have much sharper peaks compared to the classical beamformer. A drawback of this property is that calibration errors can cause the Capon beamformer to underestimate the power. Especially the higher peaks can be strongly diminished by this effect, resulting in a lower dynamic range. In the literature, methods have been proposed to

¹LOFAR (Low Frequency Array) [2, 3] is a next generation radio telescope which is currently under design. The goal of LOFAR is to enable radio astronomical observations in the 20–240 MHz band with an unprecedented high resolution and sensitivity, using $\sim 13,000$ simple antennas spread over ~ 50 stations.

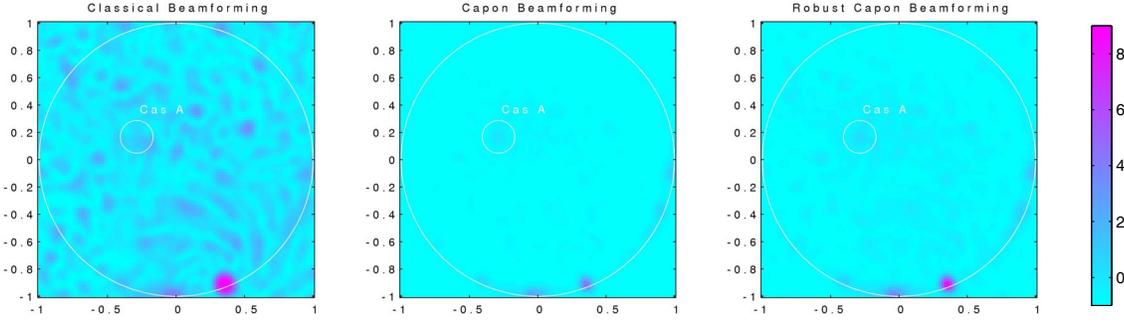


Figure 1. Skymap obtained by (a) Classical beamformer; (b) Capon beamformer; (c) Robust Capon beamformer ($\epsilon = 0.1$). Observation frequency is 18.77MHz, bandwidth is 9.77kHz.

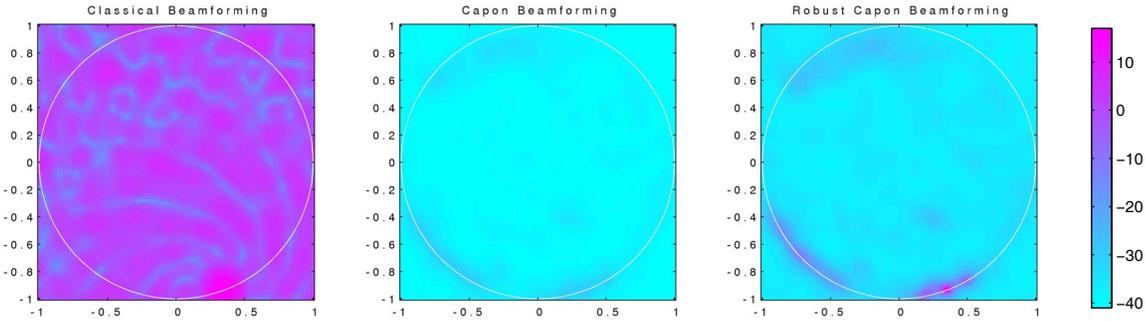


Figure 2. Similar to figure 1 but at a different observation frequency (11.87MHz).

improve the performance of the Capon beamformer for arrays with imperfect calibration [5–7]. As described in [7] the RCB searches for the maximum response of the Capon beamformer within an ellipsoid around the desired spatial signature. The problem is stated as

$$\hat{\sigma}^2, \hat{\mathbf{a}} = \arg \max_{\sigma^2, \mathbf{a}} \sigma^2 \quad \text{subject to} \quad \mathbf{R} - \sigma^2 \mathbf{a} \mathbf{a}^H \geq 0 \quad (6)$$

$$\mathbf{a} = \bar{\mathbf{a}} + \mathbf{E} \mathbf{u}, \|\mathbf{u}\| \leq 1 \quad (7)$$

where \mathbf{a} is not parametrized, $\bar{\mathbf{a}} = \mathbf{a}(\ell, m)$ is the (modeled) direction vector for the current look direction, and \mathbf{E} describes an ellipsoidal uncertainty set. Since there is no information on the calibration errors, the uncertainty set is taken to be a sphere, i.e. $\mathbf{E} = \epsilon \mathbf{I}$. An algorithm for computing the solution is given in [7].

It is shown in [7] that, at the optimum \mathbf{a} , the corresponding $\hat{\sigma}$ is given by $\hat{\sigma}^2 = (\bar{\mathbf{a}}^H \mathbf{R}^{-1} \bar{\mathbf{a}})^{-1}$. Therefore, the solution to (6) is not unique: a factor can be exchanged between \mathbf{a} and σ^2 . It is indicated in [7] that an accurate estimate of the power is obtained by scaling \mathbf{a} such that it has the same norm as $\bar{\mathbf{a}}$. As a result, the image is formed as

$$\hat{I}_{\text{RCB}}(\ell, m) = \hat{\sigma}^2 \frac{\|\mathbf{a}\|^2}{\|\bar{\mathbf{a}}\|^2}.$$

The images corresponding to our LOFAR example are shown in figure 1(c) and 2(c), where we have taken $\epsilon = 0.1$. Compared to panel (b) the power is estimated more accurately while the side-lobe levels are still low.

The method proposed by Li to solve (6) requires $O(p^3)$ flops for computing the eigenvalue decomposition of \mathbf{R} , where p is the number of antennas. However, solving for multiple different $\bar{\mathbf{a}}$ and a constant \mathbf{R} as is done in imaging requires an initial step of $O(p^3)$

flops, but for each new $\bar{\mathbf{a}}$ only $O(p^2)$ flops are needed. Imaging costs $O(p^3) + O(Np^2)$ flops, where N is the number of pixels.

3.4. RCB with additional constraints

The RCB algorithm searches for a maximum value of the Capon beamformer within an uncertainty set. Up to now, only simple spherical sets specified by $\mathbf{E} = \epsilon \mathbf{I}$ have been used. A problem with this is that the sets around neighboring points in the image plane will overlap, such that a (local) maximum is included in the uncertainty sets of all points close to that maximum. Hence, the RCB algorithm will return the same value for all points in the neighborhood of the maximum. The result is that every peak shows up as a disc shaped plateau where the value of the image is constant.

To correct this problem, we may pose additional restrictions on the uncertainty sets so that they do not overlap. This can be achieved by making the uncertainty set orthogonal to the surface described by the function $\mathbf{a}(\ell, m)$, since these are directions that will be scanned by neighboring points. A local linearization of this surface is given by the plane

$$\mathbf{a}(\ell + \delta\ell, m + \delta m) = \mathbf{a}(\ell, m) + \mathbf{d}_\ell(\ell, m) \delta\ell + \mathbf{d}_m(\ell, m) \delta m$$

$$\mathbf{d}_\ell(\ell, m) = \frac{\partial \mathbf{a}(\ell, m)}{\partial \ell}$$

$$\mathbf{d}_m(\ell, m) = \frac{\partial \mathbf{a}(\ell, m)}{\partial m}$$

The uncertainty set orthogonal to this plane is given by $\mathbf{a} = \bar{\mathbf{a}} + \mathbf{E} \mathbf{u}$, ($\|\mathbf{u}\| \leq 1$), and

$$\mathbf{E} = \epsilon \mathbf{D}^\perp$$

where the columns of \mathbf{D}^\perp are an orthonormal basis of the space orthogonal to the column span of $\mathbf{D} = [\mathbf{d}_\ell \ \mathbf{d}_m]$.

A detail which we have overlooked so far is that \mathbf{a} is a complex vector, while ℓ and m are real. We only want to remove the

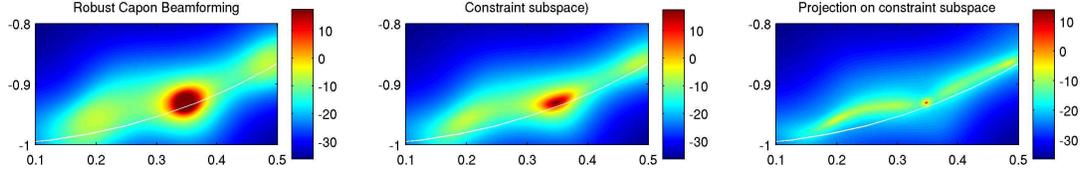


Figure 3. Skymap obtained by (a) Robust Capon beamformer; (b) Robust Capon beamformer, reduced uncertainty set; (c) Robust Capon beamformer, projection afterwards; Observation frequency is 18.77MHz, bandwidth is 9.77kHz, $\epsilon = 0.1$.

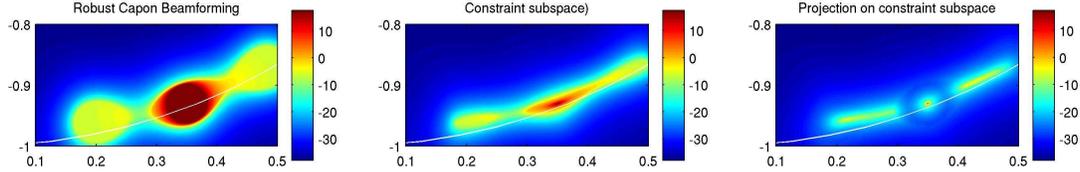


Figure 4. Similar to figure 3, but the result of a simulation of a point source model

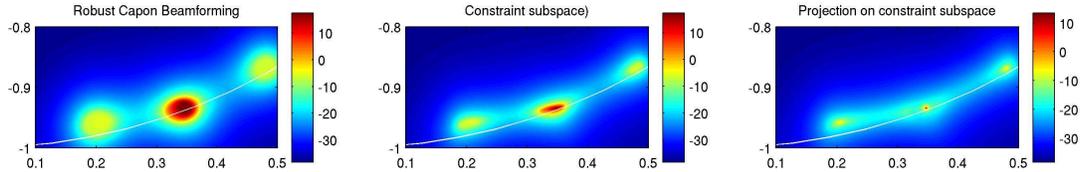


Figure 5. Similar to figure 3, but the result of a simulation of a point source model with calibration errors

“real” span of \mathbf{d}_ℓ and \mathbf{d}_m from the uncertainty set. To do this, the problem needs to be transformed to the real domain. We will use the notation $\tilde{\cdot}$ for “real versions” of complex vectors and matrices. In particular, the real version of the hermitian matrix \mathbf{R} is

$$\tilde{\mathbf{R}} = \begin{bmatrix} \text{Re}\mathbf{R} & -\text{Im}\mathbf{R} \\ \text{Im}\mathbf{R} & \text{Re}\mathbf{R} \end{bmatrix}$$

whereas the real version of vectors \mathbf{a} , \mathbf{d}_m , \mathbf{d}_ℓ are given by

$$\tilde{\mathbf{a}} = \begin{bmatrix} \text{Re}\mathbf{a} \\ \text{Im}\mathbf{a} \end{bmatrix}, \quad \tilde{\mathbf{d}}_\ell = \begin{bmatrix} \text{Re}\mathbf{d}_\ell \\ \text{Im}\mathbf{d}_\ell \end{bmatrix}, \quad \tilde{\mathbf{d}}_m = \begin{bmatrix} \text{Re}\mathbf{d}_m \\ \text{Im}\mathbf{d}_m \end{bmatrix}.$$

After transforming \mathbf{D} vector-wise into $\tilde{\mathbf{D}} = [\tilde{\mathbf{d}}_\ell \ \tilde{\mathbf{d}}_m]$, we can apply the algorithm without further modification in the real domain.

Solving this problem also requires $O(p^3)$ flops for computing the eigenvalue decomposition of $\mathbf{E}^H \tilde{\mathbf{R}}^{-1} \mathbf{E}$, but because \mathbf{E} is different for each pixel, imaging costs $O(Np^3)$. A faster (but not equivalent) method to get solutions orthogonal to the derivative vectors is applying the projection to the solution instead of to the uncertainty set, i.e., $\hat{\mathbf{a}} = \mathbf{P}\hat{\mathbf{a}}_0$ where $\hat{\mathbf{a}}_0$ is the solution of (6) for $\mathbf{E} = \epsilon\mathbf{I}$. See figure 3 to compare the algorithms.

4. SIMULATIONS

We have done simulations to compare the results in figure 3 with images based on a point source model. From local maxima above -30dB in the map of figure 2(a) we have created a point source model \mathbf{R}_1 consisting of 11 sources. A second model \mathbf{R}_2 also takes the calibration errors into account. The models are given by

$$\mathbf{R}_1 = \sum_{n=1}^{11} \sigma_n^2 \mathbf{a}(\ell_n, m_n) \mathbf{a}(\ell_n, m_n)^H,$$

$$\mathbf{R}_2 = \sum_{n=1}^{11} \sigma_n^2 (\mathbf{a}(\ell_n, m_n) + \mathbf{e}_n) (\mathbf{a}(\ell_n, m_n) + \mathbf{e}_n)^H,$$

where σ_n^2 is the power of the n -th point source and (ℓ_n, m_n) its location. Vector \mathbf{e}_n is a random error vector. The vector $\mathbf{a}_n + \mathbf{e}_n$ will be

within the uncertainty set defined by (7) and $\mathbf{E} = \epsilon\mathbf{I}$, if $\|\mathbf{e}_n\|^2 < \epsilon$. In our simulation we have chosen $\|\mathbf{e}_n\| = 0.5\sqrt{\epsilon}$. The results in figure 4 and 5 show that closely separated point sources tend to get connected especially when ϵ is much bigger than the actual error. Overall the Robust Capon Beamformer and its derivatives perform much better than classical or standard Capon beamforming.

References

- [1] A.J. van der Veen and A. Leshem, “Signal processing for radio astronomical arrays,” in *IEEE workshop on Sensor Array and Multichannel*, Sitges (Spain), July 2004.
- [2] “LOFAR web site,” <http://www.lofar.org>.
- [3] J.D. Bregman, “Design concepts for sky noise limited low frequency array,” in *Perspectives on Radio Astronomy - Technologies for Large Antenna Arrays*, A.B. Smolders and M.P. van Haarlem, Eds. ASTRON, 1999.
- [4] A. Leshem and A.J. van der Veen, “Radio astronomical imaging in the presence of strong radio interference,” *IEEE Trans. Information Th.*, vol. 46, no. 5, pp. 1730–1747, Aug. 2000.
- [5] S.A. Vorobyov, A.B. Gershman, and Zhi-Quan Luo, “Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem,” *IEEE Tr. on Sign. Proc.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
- [6] Robert G. Lorenz and Stephen P. Boyd, “Robust minimum variance beamforming,” in *Asilomar Conf. on Signals, Systems and Computers*, 2003, vol. 2, pp. 1345–1352.
- [7] Jian Li, P. Stoica, and Zhisong Wang, “On robust capon beamforming and diagonal loading,” *IEEE Trans. on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, 2003.
- [8] R.A. Perley, F. Schwab, and A.H. Bridle, Eds., *Synthesis imaging in radio astronomy*, Astronomical society of the pacific, 1989.
- [9] J.A. Hogbom, “Aperture synthesis with non-regular distribution of interferometer baselines,” *Astronomy and Astrophysics Supp.*, vol. 15, pp. 417–426, 1974.