Radioastronomical image reconstruction with regularized least squares

Naghibzadeh, Shahrzad; Mouri Sardarabadi, Ahmad; van der Veen, Alle-Jan

DOI
10.1109/icassp.2016.7472291

Publication date
2016

Document Version
Accepted author manuscript

Published in
2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

RADIOASTRONOMICAL IMAGE RECONSTRUCTION WITH REGULARIZED LEAST SQUARES

Shahrzad Naghibzadeh, Ahmad Mouri Sardarabadi, Alle-Jan van der Veen

Faculty of EEMCS, Delft University of Technology, The Netherlands

ABSTRACT

Image formation using the data from an array of sensors is a familiar problem in many fields such as radio astronomy, biomedical and geodetic imaging. The problem can be formulated as a least squares (LS) estimation problem and becomes ill-posed at high resolutions, i.e. large number of image pixels. In this paper we propose two regularization methods, one based on weighted truncation of the eigenvalue decomposition of the image deconvolution matrix and the other based on the prior knowledge of the "dirty image" using the available data. The methods are evaluated by simulations as well as actual data from a phased-array radio telescope in the Netherlands, the Low Frequency Array Radio Telescope (LOFAR).

Index Terms— Array signal processing, image formation, interferometry, regularization, radio astronomy

1. INTRODUCTION

In radio astronomy, the image reconstruction problem is to estimate the sky spatial intensity distribution over the field of view of the radio telescope. The radio telescope is synthesized by an array of antenna stations and a technique called "interferometry" is used to reconstruct the images [1]. The measurement equation relates the measured covariance data to the source intensities. Assuming a discrete point source model [2], estimating the source intensities becomes a linear least squares (LS) estimation problem. Traditionally, the LS problem is solved by first constructing the "dirty image" from the measured noisy covariance data, or " visibilities", followed by a deconvolution step to remove the effect of the antenna sampling pattern from the image [3]. It was shown [4] that the deconvolution step becomes ill-conditioned for a large number of image pixels.

In this paper, we introduce two regularization methods for the deconvolution process based on the physical properties of the data. The first method is based on windowing the spectral representation of the deconvolution matrix. We show that this method is related to baseline weighting [3] in the traditional radio astronomical interferometry. The second method is accomplished by introducing a conditioning of the deconvolution matrix based on the acquired dirty image from the covariance data.

This work was supported in part by the NWO DRIFT project.

The effectiveness of the proposed regularization methods is demonstrated through simulations and on actual data from the Low Frequency Array Radio Telescope (LOFAR).

2. DATA MODEL

To formulate the problem, we follow the notations as proposed in [2] and [5]. In these notations, $(.)^T$,$(.)^H$, $(.)^*$, $\circ$ and $\otimes$ denote transpose, the Hermitian transpose, complex conjugate, Khatri-Rao product and the Kronecker product, respectively. For our data model, P distinct receiving elements and a discrete point source model [2] with Q sources, or image pixels, is assumed. The signals received on each antenna are first time-sampled and divided into narrow frequency bands over each time snapshot. The sampled received signals on all the receiving elements for the kth frequency band and the nth time sample are denoted as

$$\mathbf{x}_k[n] = \mathbf{A}_k\mathbf{s}[n] + \mathbf{n}_k[n]. \tag{1}$$

In this notation, $\mathbf{s}[n]$ is the $Q \times 1$ vector of the sampled source signals, $\mathbf{n}_k[n]$ is a $P \times 1$ vector denoting the sampled receiver noise on all the receivers and $\mathbf{A}_k$ is a $P \times Q$ matrix in which the columns represent the array response vectors.

The astronomical signals and the receiver noise can be considered Gaussian. Therefore, the covariance matrices are sufficient statistics to represent the received signals [6]. The autocovariance of the received signals is defined as

$$\mathbf{R}_k = E\{\mathbf{x}_k[n]\mathbf{x}_k^H[n]\}. \tag{2}$$

Assuming the signals and the receiver noise are uncorrelated and the sky sources are stationary, the covariance data model [2] is stated as

$$\mathbf{R}_k = \mathbf{A}_k\mathbf{\Sigma}_\mathbf{s}\mathbf{A}_k^H + \mathbf{\Sigma}_{\mathbf{n},k}, \tag{3}$$

where $\mathbf{\Sigma}_\mathbf{s} = \text{diag}\{\boldsymbol{\sigma}\}$ and $\mathbf{\Sigma}_{\mathbf{n},k} = \text{diag}\{\boldsymbol{\sigma}_{\mathbf{n},k}\}$ represent the source and the noise covariance matrices respectively. The data covariance matrix is estimated using the available received data samples. Assuming N time samples, the sample covariance matrix is computed as

$$\hat{\mathbf{R}}_k = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}_k[n]\mathbf{x}_k^H[n], \tag{4}$$

Vectorizing the covariance data model and the covariance measurement data, the measurement equation can be stated as

$$\hat{\mathbf{r}}_k = \mathbf{r}_k + \mathbf{w}_k, \tag{5}$$

where $\hat{\mathbf{r}}_k = \text{vec}(\hat{\mathbf{R}}_k)$, $\mathbf{r}_k = \mathbf{M}_k\boldsymbol{\sigma} + \mathbf{r}_{\mathbf{n},k}$, where $\mathbf{M}_k = \mathbf{A}_k^* \circ \mathbf{A}_k$ and $\mathbf{r}_{\mathbf{n},k} = \text{vec}(\boldsymbol{\Sigma}_{n,k}) = (\mathbf{I} \circ \mathbf{I})\boldsymbol{\sigma}_{\mathbf{n},k}$, and $\mathbf{w}$ is the zero-mean additive noise with covariance [7]

$$\mathbf{C}_w = E\{(\hat{\mathbf{r}}_k - \mathbf{r}_k)(\hat{\mathbf{r}}_k - \mathbf{r}_k)^H\} = \frac{1}{N}(\mathbf{R}_k^* \otimes \mathbf{R}_k). \tag{6}$$

For K frequency bands and time snapshots, the measurement equation is formed by stacking the vectorized covariances for all the frequency bands, that is, $\mathbf{r} = \left[\mathbf{r}_1, \cdots, \mathbf{r}_K\right]^T$, $\mathbf{M} = \left[\mathbf{M}_1, \cdots, \mathbf{M}_K\right]^T$, $\mathbf{r}_\mathbf{n} = \left[\mathbf{r}_{\mathbf{n},1}, \cdots, \mathbf{r}_{\mathbf{n},K}\right]^T$, $\hat{\mathbf{r}} = \left[\hat{\mathbf{r}}_1, \cdots, \hat{\mathbf{r}}_K\right]^T$.

3. PROBLEM STATEMENT

The imaging problem is to estimate the signal powers $\boldsymbol{\sigma}$ from the received covariance data $\hat{\mathbf{r}}$. A discrete point source model [2] is considered in which each pixel of the image corresponds to a point source. Therefore, knowing the position of the antenna stations, $\mathbf{A}_k$ is known a priori. Assuming for the moment that there is no noise signal, the imaging problem can be stated as a least squares estimation problem

$$\hat{\boldsymbol{\sigma}} = \underset{\boldsymbol{\sigma}}{\text{argmin}} \parallel \hat{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma} \parallel_2^2 . \tag{7}$$

The solution to this problem can be obtained by solving the linear system of equations

$$\mathbf{H}\hat{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}_d, \tag{8}$$

where $\hat{\boldsymbol{\sigma}}_d = \mathbf{M}^H\hat{\mathbf{r}}$ is the so-called "dirty image" and $\mathbf{H} = \mathbf{M}^H\mathbf{M}$ is called the deconvolution matrix. Finding the source power estimates $\hat{\boldsymbol{\sigma}}$ amounts to inverting the deconvolution matrix $\mathbf{H}$ to remove the effect of the spatial pattern of the sampling pattern of the array. This process is called deconvolution in the conventional radio astronomy [3]. The condition number of $\mathbf{H}$ is dependent on the number of chosen image pixels Q. By increasing the resolution of the image, matrix $\mathbf{M}$ becomes "wider" and the deconvolution matrix becomes ill-conditioned. Thus, to obtain the source power estimates, regularization is required.

Weighted least squares estimation with covariance-matched weightings asymptotically leads to Maximum Likelihood (ML) estimates [7] and provides a statistically efficient source power estimate. The weighted least square estimation problem is stated as

$$\hat{\boldsymbol{\sigma}} = \underset{\boldsymbol{\sigma}}{\text{argmin}} \parallel \mathbf{C}_w^{-1/2}(\hat{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma}) \parallel_2^2 \tag{9}$$

which leads to similar closed form solutions for the dirty image and the deconvolution matrix.

$$\mathbf{H}_{\text{WLS}} = \mathbf{M}^H\mathbf{C}_w^{-1}\mathbf{M},\ \hat{\boldsymbol{\sigma}}_d = \mathbf{M}^H\mathbf{C}_w^{-1}\hat{\mathbf{r}} \tag{10}$$

In practice, we use the modified weighting method described in [8, 5] to remove the effect of the noise source from the data.

4. PROPOSED METHODS

Traditionally, iterative methods such as CLEAN [9] or sequential source removal techniques are used to perform the deconvolution. The authors in [5] have proposed a direct data-driven model-based least squares method based on Karhunen-Loeve transform (KLT) to obtain the signal power estimates $\hat{\boldsymbol{\sigma}}$ from the noisy covariance data. As the results in [5] suggest, the KLT-based deconvolution process introduces a ripple effect in the reconstructed image due to the truncation of the eigenvalues with a rectangular window. We propose two regularization methods to improve the reconstructed image quality.

4.1. Weighted truncated eigenvalue decomposition

The first method is based on introducing a weighting matrix $\boldsymbol{\Phi}$ in the eigenvalue decomposition of the deconvolution matrix $\mathbf{H}$ to perform smoother spectral windowing [10] and to reduce the ripple effect caused in the KLT-based method proposed in [5]. The eigenvalue decomposition of the deconvolution matrix is stated as

$$\mathbf{H} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H, \tag{11}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{H}$ sorted in decreasing order and columns of the matrix $\mathbf{V}$ contain the corresponding eigenvectors. According to [5], by truncating the eigenvalues of $\mathbf{H}$, the deconvolution matrix can be represented by the set of dominant eigenvalues and eigenvectors as

$$\mathbf{H} \approx \hat{\mathbf{V}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{V}}^H, \tag{12}$$

where $\hat{\boldsymbol{\Lambda}}$ is a diagonal matrix composed of the significant eigenvalues of $\mathbf{H}$ and the columns of $\hat{\mathbf{V}}$ contains the corresponding significant eigenvectors. This method is known as numerical filtering [11]. To reduce the effect of spectral truncation with a sharp rectangular window, we propose to introduce a diagonal weighting matrix $\boldsymbol{\Phi}$ to correct for the different dominance of the remaining spatial frequencies. Defining $\mathbf{H}^\dagger$ as the weighted inverse of the truncated eigenvalue decomposition of $\mathbf{H}$, the proposed method computes the image estimate as

$$\hat{\boldsymbol{\sigma}} = \mathbf{H}^\dagger\hat{\boldsymbol{\sigma}}_d = \hat{\mathbf{V}}\boldsymbol{\Phi}\hat{\boldsymbol{\Lambda}}^{-1}\hat{\mathbf{V}}^H\hat{\boldsymbol{\sigma}}_d. \tag{13}$$

Truncating the eigenvalues acts in a similar way as low-pass spatial filtering. Smaller eigenvalues correspond to higher spatial frequencies and therefore longer baselines. Since the longer baselines are sparser they contain less information and the corresponding eigenvalues have smaller magnitude which affects the condition number of the decon-

volution matrix $\mathbf{H}$ adversely. Using different filter shapes such as a triangular window instead of the rectangular window, that is normally used in numerical filtering, reduces the sidelobes at the cost of decreased resolution [12].

Note that in conventional radio astronomy, a weighting function is used to weight the set of baseline vectors contributing to the image [3]. The weighting is used to control the resulting beamshape and the sidelobes that are caused due to the finite extent of the array and the missing baseline spacings. The weighting method proposed in this paper has a similar effect as the conventional baseline weighting.

4.2. Dirty image based conditioning

The second proposed regularization method uses the prior knowledge that the expected value of the dirty image is an upper bound on the desired source power and that the source powers are positive [13], that is,

$$0 \leqslant \boldsymbol{\sigma} \leqslant \boldsymbol{\sigma}_d, \tag{14}$$

where $\boldsymbol{\sigma}_d = \mathbf{M}^H \mathbf{r}$ represents the expected value of the dirty image. By defining a conditioning vector $\mathbf{d}$ based on the obtained dirty image from the covariance data, the inequality condition (14) can be translated into conditioning weights d_i satisfying $0 \leqslant d_i \leqslant 1$, that is,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_d \odot \mathbf{d}, \tag{15}$$

where $\odot$ represents the element-wise Hadamard product. The conditioning weights are then computed as

$$\boldsymbol{\sigma}_d = \mathbf{H}\boldsymbol{\sigma} = \mathbf{H}\text{diag}\{\boldsymbol{\sigma}_d\}\mathbf{d} \tag{16}$$

$$\mathbf{d} = (\mathbf{H}\text{diag}\{\boldsymbol{\sigma}_d\})^{\dagger}\boldsymbol{\sigma}_d. \tag{17}$$

In practice, the dirty image obtained from the noisy data, i.e. $\hat{\boldsymbol{\sigma}}_d$, is used to obtain an estimate of $\mathbf{d}$ denoted as $\hat{\mathbf{d}}$ and the pseudo-inverse mentioned in (17) by $(.)^{\dagger}$ is computed by a truncated singular value decomposition of the deconvolution matrix $\mathbf{H}$ conditioned by weights obtained from the dirty image. That is,

$$(\mathbf{H}\text{diag}\{\hat{\boldsymbol{\sigma}}_d\})^{\dagger} = \hat{\mathbf{Y}}\hat{\boldsymbol{\Psi}}^{-1}\hat{\mathbf{U}}^H,$$

where $\boldsymbol{\Psi}$ is a diagonal matrix containing the significant singular values of $(\mathbf{H}\text{diag}\{\hat{\boldsymbol{\sigma}}_d\})$ and the columns of $\hat{\mathbf{Y}}$ and $\hat{\mathbf{U}}$ contain the corresponding right and left singular vectors. Therefore, the source power estimates are then obtained by

$$\hat{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}_d \odot \hat{\mathbf{d}}.$$

The idea behind conditioning the deconvolution matrix with the dirty image is that the discrete source model assumes a point source per image pixel. However, if the resolution is chosen high, source distributions are spread over multiple pixels which causes a linear dependence in the columns of the deconvolution matrix. Applying the conditioning weights to the deconvolution matrix $\mathbf{H}$ based on the dirty image promotes the columns with more available information, i.e. source power while demoting the columns related to the empty parts of the sky, that contribute to less power in the dirty image, before applying the inversion.

5. SIMULATION RESULTS

In this section, the performances of the proposed methods are evaluated using a set of simulations. A uniform linear array (ULA) configuration with $P = 15$ antennas, placed at half-wavelength spacing, with scanning angles in the range $[-90^\circ, 90^\circ]$ is considered. The resolution of the image is chosen to be 0.1 of the main beam which results in $Q = 235$ image pixels. 5 point sources with intensities $[5, 5, 0.8, 5, 5]$ are placed at angles $[-23^\circ, -17^\circ, 7.5^\circ, 23^\circ, 32.5^\circ]$. $N = 10^5$ time samples are used to construct the sample covariance matrix and the eigenvalues are truncated to the length $T = 2P - 1$ due to the fact that in the special case of a ULA matrix $\mathbf{V}$ represents a spatial Fourier transform with the unique weights defined by the number of unique baselines, i.e. $2P - 1$. The obtained dirty image is shown in figure 1(a). Following section 4.1, figure 1(b) shows the source power estimates obtained by a truncated eigenvalue decomposition. As can be seen, the truncation causes sinc-like ripple effects, the two nearby sources are not distinguishable and the smallest source is completely buried in the sidelobes of the other sources. Figure 1(c) illustrates the effect of applying the triangular weighting function on the truncated eigenvalues. The result shows significantly reduced sidelobes at the expense of reduced resolution, the close sources are still not distinguishable but the small source is recovered. The triangular weighting function with weights ϕ_i is computed as [10]

$$\phi_i = \begin{cases} 1 - \frac{i-1}{T} & i \leqslant T \\ 0 & i \geqslant T \end{cases}. \tag{18}$$

and is scaled by window power per length, i.e. $\frac{1}{T}\sum_{i=1}^{T} |\phi_i|^2$.

Furthermore, the effect of the proposed conditioning method on the point source estimates, as discussed in section 4.2, is demonstrated in Figure 1(d). As can be seen, by using the prior knowledge from the dirty image in the deconvolution process, the dynamic range is significantly increased, the sidelobes are reduced and a sharper estimate of the source positions with less bias is obtained. Moreover, the nearby sources are clearly distinguishable.

6. EXPERIMENTAL RESULTS

Actual data from the low frequency array (LOFAR) telescope [14] was used to investigate the effect of the proposed methods on the reconstructed image quality. For comparison reasons, the same data set as introduced in [5] was used. The data is captured by 48 semi-randomly spaced planar stations and consists of 25 time snapshots of 10 seconds each in 25 distinct frequency channels of bandwidth 156 kHz between 45.156 and 67.188 MHz. The image is sampled regularly on

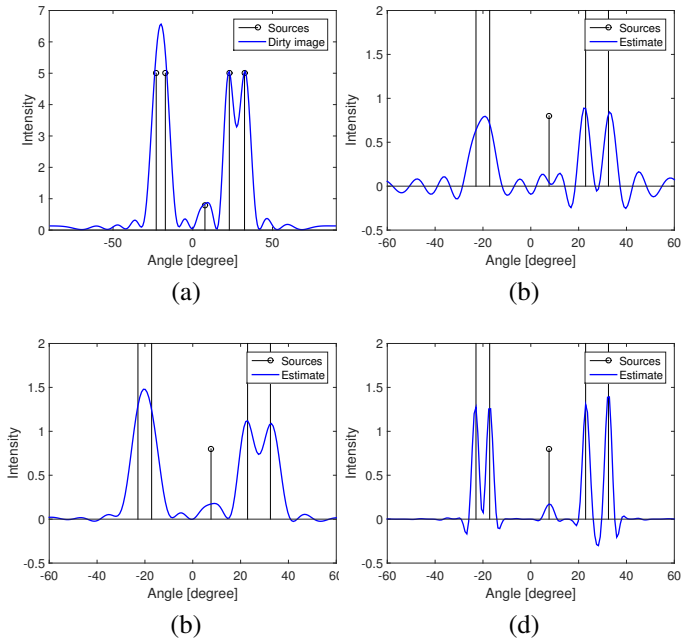

Fig. 1. Simulation results for (a) dirty image, and point source estimates based on (b) truncated eigenvalue decomposition, (c) triangular-weighted eigenvalue decomposition, (d) dirty image based conditioning.

the projected plane and the same number of image pixels, $Q = 8937$ and eigenvalue truncation threshold of $\lambda_{max}/200$ as in [5] was chosen, where λ_{max} denotes the maximum eigenvalue of the deconvolution matrix $\mathbf{H}$.

Using a weighted LS formulation, the resulting KLT-based image as proposed in [5] is shown in Figure 2(a). Next, the triangular spectral weighting method was used to reduce the ripple effect in the resulting image. The resulting reconstructed image is shown in Figure 2(b). Moreover, the reconstructed image using the dirty image based conditioning method is shown in Figure 2(c). As the figure suggests, conditioning based on the computed dirty image results in a sharper image with less sidelobes and higher dynamic range as compared with the KLT-based method.

7. CONCLUSIONS

Image deconvolution for radio astronomy becomes ill-posed for a large number of image pixels. In this paper, two regularization methods based on the physical properties of the covariance data were proposed. The first method uses a weighting function to control the dominance of the available spatial frequencies in the data and the second uses the prior information on the dirty image to condition the spatial distribution of the celestial sources. The proposed regularization methods are validated by actual data from the LOFAR radio telescope and through simulations.

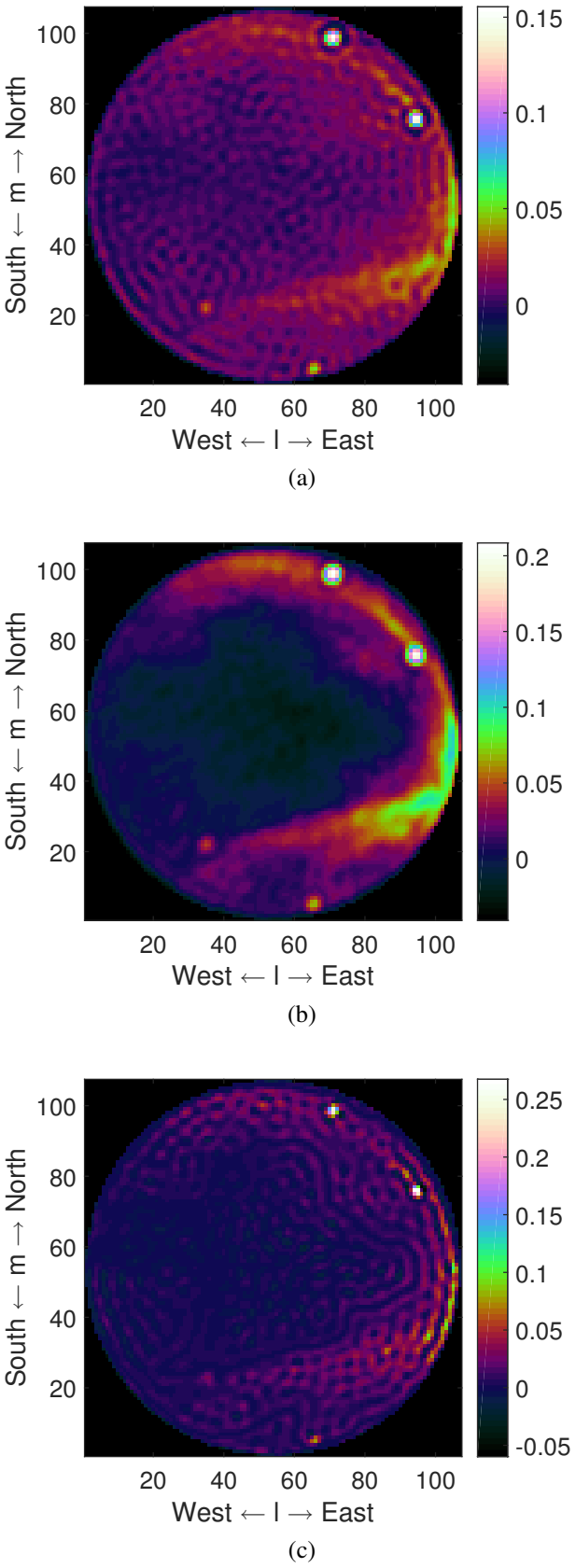

Fig. 2. Image estimates based on (a) truncated eigenvalue decomposition, (b) triangular-weighted eigenvalue decomposition, (c) dirty image based conditioning.

8. REFERENCES

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr, *Interferometry and synthesis in radio astronomy*. John Wiley & Sons, 2008.

[2] A.-J. van der Veen and S. J. Wijnholds, "Signal processing tools for radio astronomy," in *Handbook of Signal Processing Systems*, pp. 421–463, Springer, 2013.

[3] R. A. Perley, F. R. Schwab, and A. H. Bridle, "Synthesis imaging in radio astronomy," 1989.

[4] S. J. Wijnholds and A.-J. van der Veen, "Fundamental imaging limits of radio telescope arrays," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 5, pp. 613–623, 2008.

[5] S. J. Wijnholds and A.-J. Van der Veen, "Data driven model based least squares image reconstruction for radio astronomy," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 2704–2707, IEEE, 2011.

[6] A. Leshem and A.-J. Van der Veen, "Radio-astronomical imaging in the presence of strong radio interference," *Information Theory, IEEE Transactions on*, vol. 46, no. 5, pp. 1730–1747, 2000.

[7] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.

[8] S. J. Wijnholds, *Fish-eye observing with phased array radio telescopes*. TU Delft, Delft University of Technology, 2010.

[9] J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," *Astronomy and Astrophysics Supplement Series*, vol. 15, p. 417, 1974.

[10] M. Bertero, C. De Mol, and E. R. Pike, "Linear inverse problems with discrete data: Ii. stability and regularisation," *Inverse problems*, vol. 4, no. 3, p. 573, 1988.

[11] S. Twomey, "The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements," *Journal of the Franklin Institute*, vol. 279, no. 2, pp. 95–109, 1965.

[12] M. H. Hayes, *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.

[13] A. M. Sardarabadi, A. Leshem, and A.-J. van der Veen, "Radio astronomical image formation using constrained least squares and Krylov subspaces," *Astronomy and Astrophysics*, 2016.

[14] M. De Vos, A. W. Gunst, and R. Nijboer, "The LOFAR telescope: System architecture and signal processing," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1431–1437, 2009.