

RADIOASTRONOMICAL LEAST SQUARES IMAGE RECONSTRUCTION WITH ITERATION REGULARIZED KRYLOV SUBSPACES AND BEAMFORMING-BASED PRIOR CONDITIONING

Shahrzad Naghibzadeh, Alle-Jan van der Veen

Delft University of Technology

ABSTRACT

A fast iterative method based on projection onto Krylov subspaces has been proposed for Radio Astronomical (RA) image reconstruction from telescope array measurements. The image formation problem is formulated as a linear least squares (LS) estimation problem by discretizing the Field of View (FoV) of the telescope array into a number of pixels. The ill-posed imaging problem is regularized by the Krylov iterations and the system matrix is prior conditioned by the weights attained from the matched filter beamformed data. The performance of the proposed method is shown based on simulated data from a single station of the the Low Frequency Array Radio Telescope (LOFAR) antenna configuration on a test radio astronomical image. It has been shown that the prior conditioning of the system matrix results in a more accurate image estimate by reducing the artifacts introduced in the empty parts of the image. Furthermore, it was shown that Krylov-based methods fit very well in the context of large scale RA image reconstruction due to their speed and computational benefits.

Index Terms— Image formation, interferometry, regularization, radio astronomy, Krylov-based imaging

1. INTRODUCTION

In many array processing and image reconstruction applications, it is required to estimate the location and the intensity of the sources or image pixels given an incomplete and noisy set of measurements from an array of antennas. One such application area is RA imaging in which the goal is to obtain an sky map over the FoV of the radio telescope. In a technique called interferometry [1, 2], the signals from two or more radio telescopes are combined through correlators to synthesize a larger radio telescope. In this context, the term baseline is used to refer to the vector pointing between each pair of the antennas. The resolution of the telescope array is dictated by the Point Spread Function (PSF) of the array and is inversely proportional to the maximum baseline length.

In a point source modeling of the RA imaging problem, the FoV of the array is decomposed in a number of pixels over which the intensity is estimated [3, 4]. In the current and future radio telescopes, such as the Low Frequency Array (LOFAR) and the Square Kilometer Array (SKA), fine resolution images with a high number of pixels are required. We

define the RA image reconstruction as a linear LS regression problem with the noisy covariance measurements as the available data and the pixel powers as the unknowns. The system matrix performs the transform from the correlation domain to the image domain. High resolution imaging requirements leads to a large system matrix and an underdetermined system of equations. Furthermore, noisy and incomplete covariance measurements result in an ill-posed inverse problem. To ensure a unique solution and avoid noise amplification, regularization schemes are necessary.

A direct solution method for the LS RA imaging problem based on the spectral decompositions of the system matrix have been proposed in [5]. Moreover, two methods for the regularization of the direct solution have been proposed in [6]; (i) based on spectral weighting and (ii) based on conditioning of the system matrix based on an initial estimate of the image obtained from the matched filter beamformed data.

Direct methods require explicit formation and storage of the system matrix and computation of a Singular Value Decomposition (SVD) on the system matrix. High resolution RA imaging leads to large system matrices where the spectral decomposition methods are infeasible due to the computation and storage constraints. Krylov subspace-based methods [7] appear to be a good candidate as an alternative iterative solution method for LS RA imaging problems.

Krylov subspace-based methods were first investigated in the context of RA imaging by Mouri Sardarabadi et al. [8]. The algorithm, dubbed Nonnegative Least Squares (NNLS), consists of (i) an outer loop based on the greedy method of active set [9] to iteratively find the sparse support of the image and (ii) an inner loop in which a dimension reduced version of the LS problem is solved using the Krylov subspace-based method of LSQR [10]. It has been shown that the algorithm has similar working principle and improved reconstruction quality compared to the most widely used image reconstruction method in radio astronomy, the method CLEAN [4]. NNLS algorithm was further extended to account for the extended emissions by modeling the emissions in terms of point sources and Gaussian beams in [11].

We present an efficient iterative imaging algorithm based on the Krylov subspace-based method of LSQR and introduce an initial estimate of the support of the image by prior conditioning the system matrix by an estimate of the image obtained from the matched filter beamformed data. Furthermore, we exploit the preconvergence characteristics of the krylov subspace-based methods to regularize the imaging

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

problem by the iteration count. The performance of the proposed method is investigated based on simulated data from a single station LOFAR antenna configuration on a radio astronomical image.

2. DATA MODEL

Radio astronomical source recovery was first considered in the context of array signal processing by van der Veen et al. [3, 12]. We employ a similar array processing framework and data model as suggested in [3, 11]. The notations $(\cdot)^T, (\cdot)^H, (\cdot)^*$, $\circ$, $\otimes$ and $(\cdot)^\dagger$ respectively denote transpose, the Hermitian transpose, complex conjugate, Khatri-Rao product, the Kronecker product and Moore-Pensore pseudo inverse. The telescope array under consideration is composed of P distinct receiving elements and the FoV of the array is decomposed into Q pixels.

The celestial sources are assumed to be stationary. Due to the earth rotation, the observation time of the celestial sources is divided into a number of time snapshots over which the observed positions of the celestial sources by the earth-bound telescope array are considered stationary. The signals received on each antenna element and over each time snapshot k are first time-sampled into N samples and divided into narrow frequency bands. The sampled received signals on all the array elements for one frequency band is represented as

$$\mathbf{y}_k[n] = \mathbf{A}_k \mathbf{s}[n] + \mathbf{n}_k[n], \quad n = 1, \dots, N. \quad (1)$$

In this representation, $\mathbf{y}_k[n]$, $\mathbf{s}[n]$ and $\mathbf{n}_k[n]$ respectively denote the $P \times 1$ vector of the received signal sample over all the antennas, the $Q \times 1$ vector of the sampled source signals and the $P \times 1$ vector indicating the sampled noise signal on all the receivers. $\mathbf{A}_k$ is the $P \times Q$ array response matrix. Without loss of generality, in the rest of the paper we consider a single time snapshot and frequency band and therefore drop the index k . Each element of the array response matrix is computed as

$$a_{pq} = \frac{1}{\sqrt{P}} e^{-j \frac{2\pi}{\lambda} \mathbf{v}_p^T \mathbf{z}_q} \quad (2)$$

where λ is the wavelength of the received signal, $\mathbf{v}_p$ is a 3×1 vector of the Cartesian location of the p th array element with respect to a chosen origin in the field of array and $\mathbf{z}_q$ contains the direction cosines of the q th pixel in the image plane. Assuming the signals and the receiver noise are uncorrelated, the autocovariance of the received signals is computed as

$$\mathbf{R} = E\{\mathbf{y}[n]\mathbf{y}^H[n]\} = \mathbf{A}\Sigma_s\mathbf{A}^H + \Sigma_n, \quad (3)$$

where $\Sigma_s = \text{diag}\{\sigma\}$ and $\Sigma_n = \text{diag}\{\sigma_n\}$ represent the covariance matrices associated with the source signals and the received noise respectively. An estimate of the data covariance matrix is obtained using the available received data samples. The sample covariance matrix is calculated as

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}[n]\mathbf{y}^H[n]. \quad (4)$$

The linear measurement equation is obtained by vectorizing the covariance data model and the covariance measurement data

$$\hat{\mathbf{r}} = \mathbf{r} + \mathbf{e}, \quad (5)$$

where $\hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}})$, $\mathbf{r} = \mathbf{M}\sigma + \mathbf{r}_n$, in which $\mathbf{M} = \mathbf{A}^* \circ \mathbf{A}$ and $\mathbf{r}_n = \text{vec}(\Sigma_n) = (\mathbf{I} \circ \mathbf{I})\sigma_n$, and $\mathbf{e}$ represents the error due to the finite sample modeling of the covariance data. The system matrix $\mathbf{M}$ has dimensions $P^2 \times Q$. One element of $\mathbf{M}$ corresponding to the baseline between the i th and j th antenna and the q th pixel is computed as:

$$\mathbf{M}_{ij,q} = a_{iq}^* a_{jq} = e^{-j \frac{2\pi}{\lambda} (\mathbf{v}_i - \mathbf{v}_j)^T \mathbf{z}_q}. \quad (6)$$

3. PROBLEM FORMULATION

The RA image formation problem can be formulated as a linear LS regression problem [3]. In this problem formulation, the aim is to fit the available noisy and incomplete covariance data to the covariance model. Assuming the knowledge of the receiver noise powers, we denote $\hat{\mathbf{r}} - \mathbf{r}_n$ by $\tilde{\mathbf{r}}$. Therefore, the RA imaging problem reduces to

$$\hat{\sigma} = \underset{\sigma}{\text{argmin}} \|\tilde{\mathbf{r}} - \mathbf{M}\sigma\|_2^2. \quad (7)$$

4. SOLUTION METHOD BASED ON KRYLOV SUBSPACES

Instead of performing a direct inversion on the system matrix, the Krylov subspace-based methods solve Equation 7 by iteratively and implicitly forming a polynomial approximation of the pseudo inverse. This is done implicitly by iteratively projecting the system matrix over the Krylov subspaces

$$\mathcal{K}_t = \text{span}\{\mathbf{M}^H \tilde{\mathbf{r}}, (\mathbf{M}^H \mathbf{M}) \mathbf{M}^H \tilde{\mathbf{r}}, \dots, (\mathbf{M}^H \mathbf{M})^{t-1} \mathbf{M}^H \tilde{\mathbf{r}}\}, \quad (8)$$

where $\mathcal{K}$ denotes the Krylov subspace and t shows the iteration count. As denoted in the algorithm description in [7], the subspace is formed iteratively by matrix-vector multiplications of the form $\mathbf{M}\mathbf{u}$ and $\mathbf{M}^H \mathbf{v}$ for arbitrary vectors $\mathbf{u}$ and $\mathbf{v}$. Due to the storage constraints in RA imaging, using the available knowledge on the generation function of $\mathbf{M}$ as indicated in Equation 6, we can implement a subroutine that directly computes the solution of the aforementioned matrix vector products without the need to store $\mathbf{M}$ in the memory. Furthermore, since the basis vectors that span the Krylov subspace are formed from the system matrix and the available data they adapt to the particular problem. This is in contrast with the SVD based basis vectors that are only based on the system matrix.

As mentioned in [10], the Krylov subspace-based method of LSQR is the preferred method when dealing with ill-conditioned system matrices. The reason is that LSQR avoids squaring of the condition number of the system matrix by not forming the normal equations in contrast with the standard method of conjugate gradients. Therefore, we employ the LSQR method for the solution of our image formation problem.

4.1. Iterative regularization

As mentioned before, the inverse imaging problem introduced in Equation 7 is ill-posed. Instead of explicitly introducing additional regularizing constraints, we show that the Krylov subspaces perform regularization by spectral filtering. The image at iteration t is a linear combination of the Krylov basis vectors. Therefore, there exist coefficients $b_1, b_2, \dots, b_t$ such that the obtained image at iteration t , $\hat{\sigma}^{(t)}$, can be represented as

$$\hat{\sigma}^{(t)} = b_1 \mathbf{M}^H \tilde{\mathbf{r}} + b_2 (\mathbf{M}^H \mathbf{M}) \mathbf{M}^H \tilde{\mathbf{r}} + \dots + b_t (\mathbf{M}^H \mathbf{M})^{t-1} \mathbf{M}^H \tilde{\mathbf{r}}. \quad (9)$$

To show the regularizing effect of the Krylov subspace-based methods, following the explanation in [13], we show that $\hat{\sigma}^{(t)}$ has a filtered SVD expansion. The SVD expansion of the system matrix can be represented as $\mathbf{M} = \mathbf{U} \mathbf{S} \mathbf{V}^H$ where $\mathbf{U}$ and $\mathbf{V}$ contain the left and the right singular vectors, respectively and $\mathbf{S}$ is a diagonal matrix that contains the singular values in decreasing order. Inserting the SVD expansion, we can rewrite Equation 9 as

$$\begin{aligned} \hat{\sigma}^{(t)} &= (b_1 + b_2 \mathbf{V} \mathbf{S}^2 \mathbf{V}^H + \dots + b_t \mathbf{V} \mathbf{S}^{2(t-1)} \mathbf{V}^H) \mathbf{V} \mathbf{S} \mathbf{U}^H \tilde{\mathbf{r}} \\ &= \mathbf{V} (b_1 \mathbf{S}^2 + b_2 \mathbf{S}^4 + \dots + b_t \mathbf{S}^{2t}) \mathbf{S}^{-1} \mathbf{U}^H \tilde{\mathbf{r}} \\ &= \mathbf{V} \Phi \mathbf{S}^{-1} \mathbf{U}^H \tilde{\mathbf{r}} \end{aligned} \quad (10)$$

where Φ is the diagonal weighting matrix. As can be seen, the filtered SVD weights obtained from the Krylov subspace expansion depend on the iteration count. Choosing the right iteration count amounts to filtering out the insignificant and noisy singular values and hence provides regularization. We employ this semiconvergence property of LSQR to regularize our imaging problem.

5. BEAMFORMING-BASED PRIOR CONDITIONING

RA images contain substantial black background of radio quiet zones. The least squares problem defined in 7 does not contain any prior information on the potential position of the empty sky and thus would lead to a more evenly-filled estimated image. We propose to include the prior information on the sky brightness distribution obtained from the matched filter beamformer in the problem formulation to attain a more realistic sky brightness distribution.

The matched filtered dirty image, denoted as σ_d , is computed as $\sigma_d = \mathbf{M}^H \tilde{\mathbf{r}}$. denotes a weighted sum of the columns of $\mathbf{M}$. Each pixel of the dirty image contains the sum of the received signals by all the antennas pointed towards that pixel position in the FoV. Thus, for example, if a pixel is empty, i.e. does not contain any radio source, the corresponding pixel in the dirty image will only contain noise and sidelobes from the other sources. As a result, the dirty image provides a crude estimate of the support of σ .

We introduce a weighting matrix $\mathbf{W} = \text{diag}(\sigma_d)$ and weight the columns of $\mathbf{M}$ based on the corresponding pixel

in the dirty image as

$$\tilde{\mathbf{M}} = \mathbf{M} \mathbf{W}. \quad (11)$$

$\tilde{\mathbf{M}}$ denotes the prior conditioned system matrix with columns $\tilde{\mathbf{M}} = [\sigma_{d,1} \mathbf{m}_1, \sigma_{d,2} \mathbf{m}_2, \dots, \sigma_{d,Q} \mathbf{m}_Q]$, where $\sigma_{d,q}$ and $\mathbf{m}_q$ respectively denote the q th pixel of the image and the q th column of $\mathbf{M}$. Therefore, we recast the imaging problem as

$$\tilde{\sigma} = \underset{\sigma}{\text{argmin}} \|\tilde{\mathbf{r}} - \tilde{\mathbf{M}} \sigma\|_2^2 \quad (12)$$

and apply LSQR to find the solution. To show the equivalence of the solution of Equations 7 and 12 we proceed as follows. At the optimal point of problem 12 the normal equations

$$\tilde{\mathbf{M}}^H \tilde{\mathbf{M}} \tilde{\sigma} = \tilde{\mathbf{M}}^H \tilde{\mathbf{r}}. \quad (13)$$

must hold. Substituting (11) in (13) we have

$$\begin{aligned} \mathbf{W}^H \mathbf{M}^H \mathbf{M} \mathbf{W} \tilde{\sigma} &= \mathbf{W}^H \mathbf{M}^H \tilde{\mathbf{r}} \\ \mathbf{M}^H \mathbf{M} \mathbf{W} \tilde{\sigma} &= \mathbf{M}^H \tilde{\mathbf{r}} \end{aligned} \quad (14)$$

Comparing with the normal equations of the original problem 7, we conclude that

$$\hat{\sigma} = \mathbf{W} \tilde{\sigma}. \quad (15)$$

6. SIMULATION RESULTS

The proposed method has been tested on noisy simulated data using the configuration of antennas from a single station of the LOFAR telescope array. The test image was chosen as a the HII region in the Large Magellanic Cloud of Tarantula Nebula radio image shown in 1(a) (Available at <https://casaguides.nrao.edu/index.php>). The station contains $P = 48$ antennas with maximum baseline length of about 63 m as shown in Figure 1(b). The operating frequency is chosen to be 60 MHz and a single time snapshot is considered. Figure 1(c) illustrates the PSF of the array showing the limited resolution of the array and the existence of sidelobes. To construct the sampled covariance matrix, Gaussian receiver noise with variance $\sigma_n = 0.5$ is added to the covariance $\mathbf{R}$ and $N = 10^5$ data samples are used to construct $\hat{\mathbf{R}}$. The image is discretized into $Q = 64009$ pixels. The dirty image obtained from matched filtered beamformer is shown in Figure 1(d). The simulations were performed on a laptop with Intel i5-2430 CPU 2.4 GHz under 64-bit Windows 7.

The results of applying the CLEAN algorithm, NNLS algorithm, iteration regularized LSQR (shortly represented as LSQR) and beamforming-based prior conditioned LSQR (shortly represented as P-LSQR) are shown in Figure 2(a), (b), (c) and (d), respectively.

CLEAN and NNLS algorithms start from an empty image and build up the support of the image iteratively by nonlinear greedy methods, assuming the image can be represented as a set of point sources. 500 iterations were chosen for the termination of these methods. The computation time of CLEAN

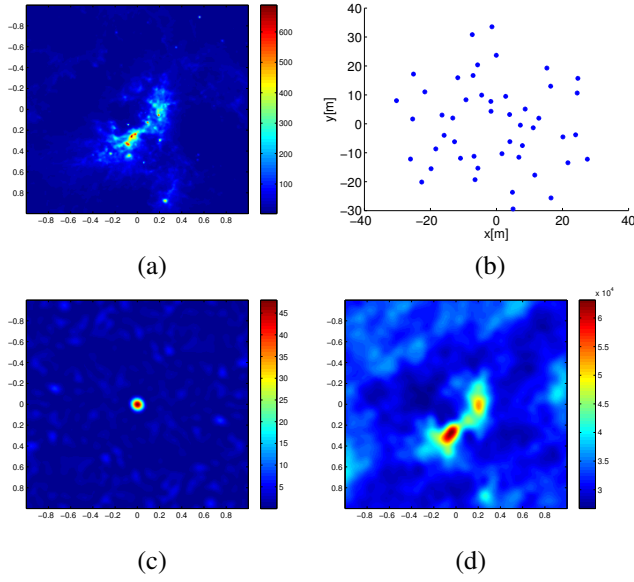


Fig. 1: (a) True sky image, (b) antenna placement, (c) PSF and (d) dirty image

and NNLS were 2.8 and 4.6 minutes, respectively. Furthermore, the results were post-processed by a clean Gaussian beam to restore the resolution of the main beam of the array. As can be seen, these algorithms obtain a sufficiently good estimate of the sky image at the resolution of the telescope array. However, the convergence speed for the greedy methods is prohibitively slow for radio images with a sufficient amount of extended emission.

On the other hand, the LSQR-based methods consider the complete image rather than individual point sources in each iteration and proceed by projections onto Krylov subspaces. In these simulations, the LSQR and P-LSQR methods were terminated after only 10 iterations. The required time for the generation of the results were 2.9 and 3.5 seconds for LSQR and P-LSQR methods, respectively. Therefore, the LSQR-based methods exhibit about a 60 times computational saving compared to the widely-used greedy methods. The disadvantage of the first LSQR method is that it fills up the empty parts of the sky. Prior conditioning the system matrix with matched-filtered dirty image weights results in a better estimate of the sky brightness distribution while maintaining the empty parts of the sky.

Furthermore, the performance of the LSQR-based algorithms have been investigated by comparing the L_2 and L_1 norm errors per iteration as shown in logarithmic scale for 100 iterations in Figures 3(a) and (b), respectively. The blue and the black curves correspond to the LSQR and P-LSQR method, respectively. As can be seen, the L_2 and L_1 estimation errors are less for the P-LSQR method. The L_1 norm error is an indicator of the support of the estimated image and illustrates a big gap between the two methods indicating that the P-LSQR method is indeed better capable of estimating the support of the image. Furthermore, as the iteration number increases, the noisy smaller eigenvalues start to corrupt the image and the estimation error increases. Therefore,

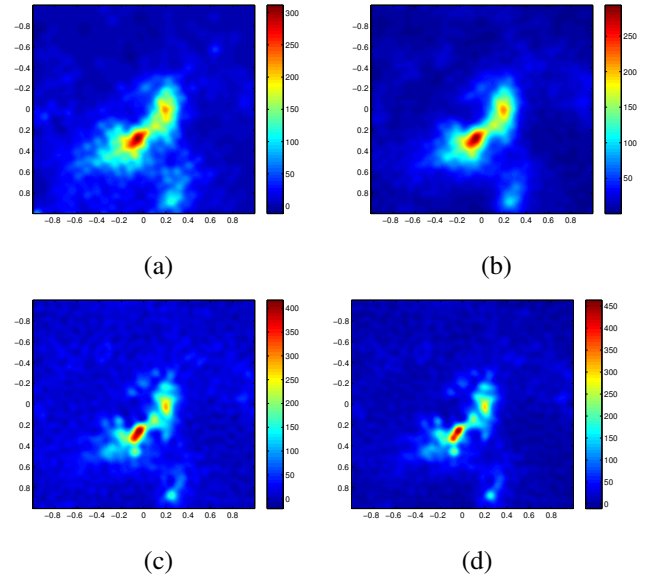


Fig. 2: Image estimates based on (a) CLEAN, (b) NNLS, (c) LSQR and (d) P-LSQR

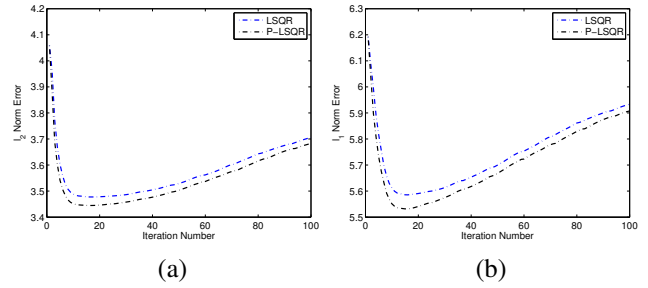


Fig. 3: (a) L_2 norm error, (b) L_1 norm error

exploiting the preconvergence behavior of the Krylov-based methods helps in regularizing the solution.

7. CONCLUSIONS

A fast iterative approach to the LS radio astronomical imaging problem has been investigated. The proposed method is based on the Krylov subspace-based method of LSQR. The regularization of the ill-posed inverse imaging problem is performed by the Krylov iterations and a prior conditioning on the system matrix is obtained based on the crude estimate of the image from the matched filter beamformer. The performance of the algorithm has been tested on a radio astronomical image with simulated data from a single LOFAR station configuration. It has been shown that LSQR-based methods exhibit favorable computational characteristics for large scale RA imaging problems. Furthermore, prior conditioning the system matrix provides an effective way to enforce the support of the image into the Krylov subspaces. For the future work it will be interesting to compare this method with the existing L_1 -norm regularized compressed sensing imaging algorithms in terms of reconstruction quality and computations.

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] R. A. Perley, F. R. Schwab, and A. H. Bridle, “Synthesis imaging in radio astronomy,” 1989.
- [3] 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.
- [4] J. Högbom, “Aperture synthesis with a non-regular distribution of interferometer baselines,” *Astronomy and Astrophysics Supplement Series*, vol. 15, p. 417, 1974.
- [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] S. Naghibzadeh, A. M. Sardarabadi, and A.-J. van der Veen, “Radioastronomical image reconstruction with regularized least squares,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3316–3320, IEEE, 2016.
- [7] Y. Saad, “Krylov subspace methods for solving large unsymmetric linear systems,” *Mathematics of computation*, vol. 37, no. 155, pp. 105–126, 1981.
- [8] A. M. Sardarabadi, A. Leshem, and A.-J. van der Veen, “Radio astronomical image formation using constrained least squares and krylov subspaces,” *Astronomy & Astrophysics*, vol. 588, p. A95, 2016.
- [9] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, vol. 161. SIAM, 1974.
- [10] C. C. Paige and M. A. Saunders, “LSQR: An algorithm for sparse linear equations and sparse least squares,” *ACM transactions on mathematical software*, vol. 8, no. 1, pp. 43–71, 1982.
- [11] S. Naghibzadeh, A. Mouri Sardarabadi, and A.-J. van der Veen, “Point and beam-sparse radio astronomical source recovery using non-negative least squares,” in *The Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop*, p. to appear, IEEE, 2016.
- [12] A.-J. van der Veen, A. Leshem, and A.-J. Boonstra, “Array signal processing for radio astronomy,” in *The Square Kilometre Array: An Engineering Perspective*, pp. 231–249, Springer, 2005.
- [13] P. C. Hansen, *Discrete inverse problems: insight and algorithms*, vol. 7. Siam, 2010.