COMPUTATIONALLY EFFICIENT RADIO ASTRONOMICAL IMAGE FORMATION USING CONSTRAINED LEAST SQUARES AND AND THE MVDR BEAMFORMER

A. M Sardarabadi[1] * *and A. Leshem*[2†] *and A-J van der Veen*[1]

[1]Department of electrical engineering, Delft University of Technology, Delft, The Netherlands
[2]Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel

ABSTRACT

Linear image deconvolution for radio–astronomy isan ill–posed problem. For this reason a-priori knowledge is crucial for improving the performance of the deconvolution. In this paper we show that combining non–negativity constraints with an upper bound on the magnitude of each pixel in the image can significantly improve the image formation algorithm. We also show that the minimum variance distortionless response (MVDR) dirty image provides the tightest upper bound among all beamformers. We then show how LS-MVI image formation a;gorithm can be reformulated as a preconditioned weighted least squares algorithm. The resulting algorithm can be efficiently solved using the active–set method. The performance of the algorithm is demonstrated in simulation and compared with constrained least squares based on the classical dirty image.

Index Terms— Radio astronomy, array signal processing, constrained optimization, Krylov subspace, LSQR, MVDR, image deconvolution

1. INTRODUCTION

Since the early days of radio astronomy many deconvolution techniques have been developed to solve the image formation problem. The basic idea behind a deconvolution algorithm is to exploit a-priori knowledge about the image in order to solve the ill-posed imaging problem. The first algorithm and the most popular of these techniques is the CLEAN method proposed by Högbom [1]. Subsequently the maximum entropy algorithm (MEM) with various entropy functions was proposed in [2], [3], [4] and [5] and the current implementation by Cornwell and Evans [6] is the most widely used. Beyond these two techniques there are several extensions in various directions: extensions of the CLEAN algorithm to support multi-resolution and wavelets as well as non co-planar arrays and multiple wavelengths such as the W-projection [7] and the A-projection algorithm (see the overview paper [8]). MEM techniques have been also extended to take into account source structure through the use of multiresolution and wavelet based techniques [9], global non-negative least squares was proposed by Briggs [10], matrix based parametric imaging such as the Least Squares Minimum Variance Imaging (LS–MVI), maximum likelihood based techniques in [11], [12], [13] and sparse L_1 reconstruction in [14] and [15]. Source modeling is an important issue and various techniques to improve modeling over simple point source models by using shapelets, wavelets and Gaussians [16] have been implemented. A more extensive overview of classical techniques and implementation issues is given in [17] or [18].

Better performance analysis of imaging as well as the development of computationally efficient techniques are some of the major challenges for the radio astronomical signal and image processing community. This is likely to become a more critical problem for the future generation of radio interferometers that will be built in the next two decades such as the square kilometer array (SKA) and its prototypes, the Low Frequency Array (LOFAR), the Allen Telescope Array (ATA) and the Long Wavelength Array (LWA). These radio-telescopes are composed of many stations (each station made up of multiple antennas that are combined using adaptive beamforming). They will have significantly increased sensitivity and bandwidth compared to traditional radio interferometers, and some of them will operate at much lower frequencies. Improved sensitivity will therefore require a much better calibration, the capability to perform imaging with much higher dynamic range in order to reduce the effect of the residuals of powerful foreground sources inside and outside the field of view and better handling of non-coplanar arrays.

In contrast to non-linear deconvolution techniques, least squares based techniques offer a computationally efficient closed form approach [19]. Unfortunately, the deconvolution problem becomes ill–posed as the resolution increases [19]. To overcome this problem, a non-negativity constraint has been proposed using the non–negative least squares algorithm (NNLS) [20]. In order to benefit from the vast literature related to solving least square problems on one hand and to gain from the non-linear processing offered by standard deconvolution techniques on the other hand, we propose to reformulate the imaging problem using an active set approach with two additional constraints as well as a low-dimensional fine fitting of the parameters. We generalize several of the sequential parametric techniques into an active–set weighted least squares algorithm [21] with weighting derived to allow the introduction of high resolution techniques such as the LS-MVI ([12]). This reformulation will allow us to obtain computationally efficient imaging algorithms that are closely related to existing non-linear sequential source estimation techniques with the benefit of accelerated convergence due to tighter upper bounds on the power distribution over the complete image. The Karush Kuhn Tucker (KKT) total power constraint is enforced over the complete image not only at the location of the source. This in turn eliminates the inclusion of negative flux sources and other anomalies that appear in some existing sequential techniques. Specifically, we show that the pixel values are bounded also from above by the minimum variance distortionless response (MVDR) dirty image [11] and then we extend the idea behind NNLS and formulate the multichannel imaging as an optimization problem with both lower and upper bounds.

The structure of the paper is as follows: In section 2 we describe

*Corresponding author, email: a.mourisardarabadi@tudelft.nl. This research was supported by NWO-TOP 2010, 614.00.005

†The research of A. Leshem was also supported by the Israeli Science foundation, grant 1240-2009

the basic data model. Section 3 we describe the imaging algorithm. We end up with a simulated experiments comparing the algorithm to previously proposed constrained least square algorithm [22]. We end up with conclusions and extensions. Due to space limitations, the implementation details of the active set technique as well as comparison of the algorithm to other algorithms on real data will appear in the full version of this paper [23].

2. DATA MODEL

Assume that a radio telescope with p antennas is observing the sky. In this paper we use the data model proposed by Leshem et al. [11], [24]. For simplicity we assume a single frequency measurement and the generalization to multi-frequency synthesis follows the model of [11]. To distinguish between the true (and unknown) position of the sources and the position of each pixel in the image we use tilde to represent parameters that depend on the true position of the sources. Following [11] we are given estimates $\hat{\mathbf{R}}_k$ of the antenna measurements covariance (also known as the visibility) at snapshots $k = 1, ..., K$. Since the received signals and noise are Gaussian, these covariance matrices form sufficient statistics for the imaging problem [11]. We also define

$$\hat{\mathbf{r}}_k = \text{vect}(\hat{\mathbf{R}}_k) \tag{1}$$

Stacking these measurements in a vector we can form the total measurement vector for the system which becomes

$$\hat{\mathbf{r}} = \begin{bmatrix} \hat{\mathbf{r}}_1^T & , \ldots , & \hat{\mathbf{r}}_K^T \end{bmatrix}^T. \tag{2}$$

.

The measured covariance matrices are noisy versions of the true covariance matrices. These are composed of the sky contribution and the receivers noise covariance matrices. Following [24] the model covariance matrices are given by:

$$\mathbf{R}_k = \mathcal{E}\{\mathbf{y}_k \mathbf{y}_k^H\} = \tilde{\mathbf{A}}_k \tilde{\mathbf{\Sigma}} \tilde{\mathbf{A}}_k^H + \mathbf{R}_{\mathbf{n},k}, \tag{3}$$

where $\tilde{\mathbf{\Sigma}} = \mathcal{E}\{\tilde{\mathbf{s}}\tilde{\mathbf{s}}^H\}$ and $\mathbf{R}_{\mathbf{n},k}$ are sky brightness and the noise covariance matrix respectively. By assuming that sky sources are independent and stationary, we can model $\tilde{\mathbf{\Sigma}} = \text{diag}(\tilde{\boldsymbol{\sigma}})$ where

$$\tilde{\boldsymbol{\sigma}} = \begin{bmatrix} \tilde{\sigma}_1 & , \ldots , & \tilde{\sigma}_q \end{bmatrix}^T \tag{4}$$

represents the the brightness of the sources. We will use the terms brightness and source power interchangeably for the rest of this paper.

Vectorizing both sides of (3) we get

$$\mathbf{r}_k = \text{vect}(\mathbf{R}_k) = (\tilde{\mathbf{A}}_k^* \circ \tilde{\mathbf{A}}_k)\tilde{\boldsymbol{\sigma}} + \mathbf{r}_{\mathbf{n},k} \tag{5}$$

where $\mathbf{r}_{\mathbf{n},k} = \text{vect}(\mathbf{R}_{\mathbf{n},k})$. Stacking the vectorized covariances for all of snapshots together we obtain

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1^T, ..., \mathbf{r}_K^T \end{bmatrix}^T = \tilde{\mathbf{\Psi}}\tilde{\boldsymbol{\sigma}} + \mathbf{r}_\mathbf{n} \tag{6}$$

where $\tilde{\mathbf{\Psi}} = \left[\left(\tilde{\mathbf{A}}_1^* \circ \tilde{\mathbf{A}}_1 \right)^T, ..., \left(\tilde{\mathbf{A}}_K^* \circ \tilde{\mathbf{A}}_K \right)^T \right]^T$ and $\mathbf{r}_\mathbf{n} = \begin{bmatrix} \mathbf{r}_{\mathbf{n},1}^T, ..., \mathbf{r}_{\mathbf{n},K}^T \end{bmatrix}^T$. The imaging problem amounts to estimating the locations and brightness of all the sky sources from the sample covariance estimates $\hat{\mathbf{R}}_1, ..., \hat{\mathbf{R}}_K$. We also define the block diagonal matrices $\mathbf{R} = \text{diag}\{\mathbf{R}_1, ..., \mathbf{R}_K\}$ and $\hat{\mathbf{R}} = \text{diag}\{\hat{\mathbf{R}}_1, ..., \hat{\mathbf{R}}_K\}$. These will be used for computing the asymptotically optimal weighting matrix for the weighted least squares algorithm.

3. THE IMAGE FORMATION ALGORIHTM AS A REGULARIZED CONSTRAINED LEAST SQUARES ALGORITHM

The imaging problem aims to find the spatial power distribution over the sky, given a set of covariance matrices at snapshots, $k = 1, ..., K$. We define a grid for the image and try to estimate the power on the grid. Assuming that the grid is sufficiently fine this will provide a good estimate of the image. Threfore we replace $\tilde{\mathbf{A}}$ with the array response matrix towards the grid pixels $\mathbf{A}$ and $\tilde{\boldsymbol{\sigma}}$ with $\boldsymbol{\sigma}$ in (3) and obtain

$$\mathbf{R}_k = \mathbf{A}_k \text{diag}(\boldsymbol{\sigma}) \mathbf{A}_k^H + \mathbf{R}_{\mathbf{n},k}. \tag{7}$$

In this new model $\mathbf{A}$ is a $p \times M$ matrix and $\boldsymbol{\sigma}$ is a $M \times 1$ vector, where M is the number of pixels. Similarly $\tilde{\mathbf{\Psi}}$ changes to

$$\mathbf{\Psi} = \left[\left(\mathbf{A}_1^* \circ \mathbf{A}_1 \right)^T, ..., \left(\mathbf{A}_K^* \circ \mathbf{A}_K \right)^T \right]^T. \tag{8}$$

Hence for a given $\hat{\mathbf{r}}$ in (2), we want to estimate the power of each pixel. Note that $\mathbf{A}$, $\mathbf{\Psi}$ and $\boldsymbol{\sigma}$ depend on the position of the pixels on the gridded image. The change of notation is to indicate the difference between the pixel locations and the true (and unknown) location of the sources.

Now that we have defined a grid the imaging equation (6) becomes

$$\mathbf{r} = \mathbf{\Psi}\boldsymbol{\sigma} + \mathbf{r}_n. \tag{9}$$

For a sufficiently fine grid this approximates the solution of the discrete source model. However, working entirely in the image domain leads to a gridding related noise floor. This is solved by fine adaptation of the location of the sources and estimating the true locations in the visibility domain. We can now reformulate (9 as a Least Squares (LS) estimate of $\boldsymbol{\sigma}$ as is done in [19]. Solving the imaging problem with LS reduces to the following minimization problem

$$\min_{\boldsymbol{\sigma}} \frac{1}{2K} \|\hat{\mathbf{r}} - \mathbf{\Psi}\boldsymbol{\sigma}\|^2. \tag{10}$$

Unfortunately when the number of pixels is large the problem is ill-posed and (9) has infinitely many solutions [19]. Therefore, solving this problem requires some kind of regularization. Typically image formation algorithms exploit external information regarding the image in order to regularize the ill-posed problem. For example maximum entropy techniques [2, 3] impose a smoothness condition on the image while the CLEAN algorithm [1] exploits the fact that most of the image is empty. To regularize the problem we follow [20] who required non-negativity of every pixel in the image. This leads to a lower bound $\boldsymbol{\sigma} \geq \mathbf{0}$. The NNLS problem can thus be given as:

$$\begin{array}{rl} & \min_{\boldsymbol{\sigma}} \frac{1}{2K} \|\hat{\mathbf{r}} - \mathbf{\Psi}\boldsymbol{\sigma}\|^2 \\ \text{subject to:} & \mathbf{0} \leq \boldsymbol{\sigma} \end{array} \tag{11}$$

We now improve the NNLS regularization by exploiting our knowledge regarding the dirty image and bounding the pixel powers also from above. such that $\boldsymbol{\sigma} \leq \boldsymbol{\gamma}$ for some upper bound $\boldsymbol{\gamma}$. We will motivate and discuss different choices for $\boldsymbol{\gamma}$.

By closer inspection of the ith pixel on the dirty image, we see that for ith pixel we have

$$\sigma_{\mathbf{w},i} = \sum_k \mathbf{w}_{i,k}^H \mathbf{R}_k \mathbf{w}_{i,k} = \sum_k \mathbf{w}_{i,k}^H \mathbf{a}_{i,k} \sigma_i \mathbf{a}_i^H \mathbf{w}_{i,k} + \sum_k \mathbf{w}_{i,k}^H \mathbf{R}_{r,k} \mathbf{w}_{i,k} \tag{12}$$

If we require that

$$\mathbf{w}_i^H \mathbf{a}_i = \sqrt{K} \tag{13}$$

where

$$\mathbf{w}_i = \begin{bmatrix}\mathbf{w}_{i,1}^T & \dots & \mathbf{w}_{i,K}^T\end{bmatrix}^T \tag{14}$$

we obtain that

$$\sigma_{\mathbf{w},i} = \sigma_i + \mathbf{w}_i^H \mathbf{R}_r \mathbf{w}_i. \tag{15}$$

Since $\mathbf{R}_r$ is positive definite, we obtain that we always have:

$$\sigma_i \leq \sigma_{\mathbf{w},i}. \tag{16}$$

Following [25]the optimal beamforming vectors $\mathbf{w}_i$ which satisfy (13) and minimize the second term on the right hand side of (15). The beamforming vectors are given by:

$$\mathbf{w}_i = \frac{\sqrt{K}}{\mathbf{a}_i^H \mathbf{R}^{-1} \mathbf{a}_i} \mathbf{R}^{-1} \mathbf{a}_i. \tag{17}$$

The dirty image formed by substituting (17) into (15) is called the MVDR dirty image [12]. Each pixel of the MVDR dirty image is given by

$$\begin{aligned}\sigma_{\text{MVDR},i} &= \frac{K \sum_k \mathbf{a}_{i,k}^H \mathbf{R}_k^{-1} \mathbf{a}_{i,k}}{\left(\sum_k \mathbf{a}_{i,k}^H \mathbf{R}_k^{-1} \mathbf{a}_{i,k}\right)^2} \\ &= \frac{1}{\frac{1}{K}\sum_k \mathbf{a}_{i,k}^H \mathbf{R}_k^{-1} \mathbf{a}_{i,k}}.\end{aligned} \tag{18}$$

It is useful to write (18) in a vector form:

$$\boldsymbol{\sigma}_{\text{MVDR}} = \mathbf{D}_{\text{MVDR}}^{-1} \boldsymbol{\Psi}^H (\mathbf{R}^{-T} \otimes \mathbf{R}^{-1}) \mathbf{r}, \tag{19}$$

where

$$\mathbf{D}_{\text{MVDR}} = \frac{1}{K} \text{diag}^2 \left(\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A}\right). \tag{20}$$

By substituting this vector into (16) the upper bound becomes

$$\boldsymbol{\sigma} \leq \boldsymbol{\sigma}_{\text{MVDR}}. \tag{21}$$

Given the fact that $\mathbf{w}_i$ minimizes the positive error in output of the beamformer shows that the MVDR dirty image forms the tightest upper bound on the power of each pixel among all beamforming vectors. Using this as an additional constraint can improve the speed of convergence and also the quality of the solution. This improves the results of [22] which was based on the classical dirty image. The constrained LS (CLS) imaging problem can now be formulated as:

$$\begin{aligned}&\min_{\boldsymbol{\sigma}} \tfrac{1}{2K} \|\hat{\mathbf{r}} - \boldsymbol{\Psi}\boldsymbol{\sigma}\|^2 \\ \text{subject to:}\quad &\mathbf{0} \leq \boldsymbol{\sigma} \leq \boldsymbol{\gamma}\end{aligned} \tag{22}$$

where $\boldsymbol{\gamma}$ can be chosen either as $\boldsymbol{\gamma} = \boldsymbol{\sigma}_{\text{MF}}$ for the matched filter dirty image or $\boldsymbol{\gamma} = \boldsymbol{\sigma}_{\text{MVDR}}$ for the MVDR dirty image.

The upper bound (21) assumes that we know the true covariance matrices $\mathbf{R}_k$. However in practice we only measure $\hat{\mathbf{R}}_k$ which is subject to statistical fluctuations. Choosing a confidence level of 6 times the standard deviation of the dirty images ensures that the upper bound will hold with probability 99.9% for all the pixels in an image of 1000x1000 pixels.

$$\boldsymbol{\sigma} \leq \alpha\, \hat{\boldsymbol{\sigma}}_{\text{MVDR}} \tag{23}$$

where

$$\hat{\sigma}_{\text{MVDR},i} = \frac{C}{\frac{1}{K}\sum_k \mathbf{a}_{i,k} \hat{\mathbf{R}}_k^{-1} \mathbf{a}_{i,k}} \tag{24}$$

is an unbiased estimate of the MVDR dirty image, and $C = N/(N-p)$ is a bias correction constant. The unbiased estimate can also be written in vector form as

$$\hat{\boldsymbol{\sigma}}_{\text{MVDR}} = \mathbf{D}^{-1} \boldsymbol{\Psi}^H (\hat{\mathbf{R}}^{-T} \otimes \hat{\mathbf{R}}^{-1}) \hat{\mathbf{r}}, \tag{25}$$

where

$$\mathbf{D} = \frac{1}{KC} \text{diag}^2 \left(\mathbf{A}^H \hat{\mathbf{R}}^{-1} \mathbf{A}\right). \tag{26}$$

It is well known that the statistical properties of a LS solution can be improved in the Gaussian case by applying a proper weighting which leads to a weighted least squares (WLS) problem [26]. The proper weighting uses $\mathbf{R}^{-1}$. However, if $\hat{\mathbf{R}}$ is a consistent estimate of $\mathbf{R}$ using $\hat{\mathbf{R}}$ provides an asymptotically optimal weighting. This leads to the cost function:

$$f_{\text{WLS}}(\boldsymbol{\sigma}) = \frac{1}{2} \| \left(\hat{\mathbf{R}}^{-T/2} \otimes \hat{\mathbf{R}}^{-1/2}\right) (\hat{\mathbf{r}} - \boldsymbol{\Psi}\boldsymbol{\sigma}) \|^2 \tag{27}$$

The WLS problem is now given by

$$\begin{aligned}&\min_{\boldsymbol{\sigma}} f_{\text{WLS}}(\boldsymbol{\sigma}) \\ \text{subject to:}\quad &\mathbf{0} \leq \boldsymbol{\sigma} \leq \boldsymbol{\gamma}.\end{aligned} \tag{28}$$

As we have shown in (10), the LS solution is related to the classical (matched filter) dirty image. Similarly we can show that by defining a diagonal preconditioner and applying it to the WLS we can relate this new preconditioned WLS (PWLS) problem to the MVDR dirty image. This will also allow us to understand the relation between the WLS solution and the LS-MVI algorithm [12]. The diagonal preconditioner with this property given by (26). Now we can rewrite (27) as

$$f_{\text{WLS}}(\boldsymbol{\sigma}) = \frac{1}{2} \| \left(\hat{\mathbf{R}}^{-T/2} \otimes \hat{\mathbf{R}}^{-1/2}\right) (\hat{\mathbf{r}} - \boldsymbol{\Psi}\mathbf{D}^{-1}\mathbf{D}\boldsymbol{\sigma}) \|^2 \tag{29}$$

Substituting $\check{\boldsymbol{\sigma}} = \mathbf{D}\boldsymbol{\sigma}$ we obtain that (28) is equivalent to solving

$$\begin{aligned}\check{\boldsymbol{\sigma}} = \quad &\arg\min_{\check{\boldsymbol{\sigma}}} \tfrac{1}{2} \| \left(\hat{\mathbf{R}}^{-T/2} \otimes \hat{\mathbf{R}}^{-1/2}\right) (\hat{\mathbf{r}} - \boldsymbol{\Psi}\mathbf{D}^{-1}\check{\boldsymbol{\sigma}}) \|^2 \\ \text{subject to:}\quad &\mathbf{0} \leq \check{\boldsymbol{\sigma}} \leq \mathbf{D}\boldsymbol{\gamma}\end{aligned} \tag{30}$$

and setting $\boldsymbol{\sigma} = \mathbf{D}^{-1}\check{\boldsymbol{\sigma}}$. This is correct since $\mathbf{D}$ is a positive diagonal matrix. The relation between this function and MVDR leads to better detection and hence better estimation results when we solve the constrained problem.

The problem (29) can be solved efficiently using an active set algorithm together with a Krylov space based updates. These issues will be discussed in the full version of this paper [23].

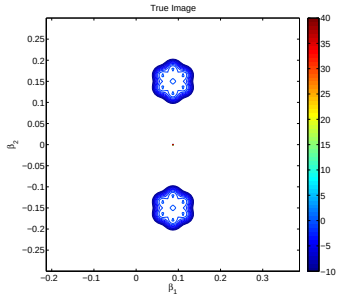

Fig. 1: True source

4. SIMULATIONS

In this section we will evaluate the performance of the proposed method on a simulated image. An array of 100 dipoles ($p = 100$) with random distribution is used. 3 frequency channels each with a bandwidth of 195kHz and two snapshots (K=2) have been used. The

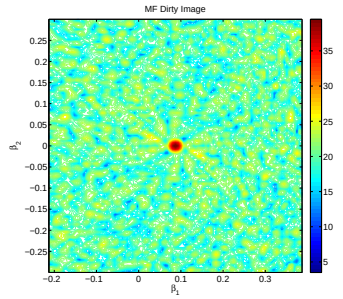

Fig. 2: MF Dirty Image

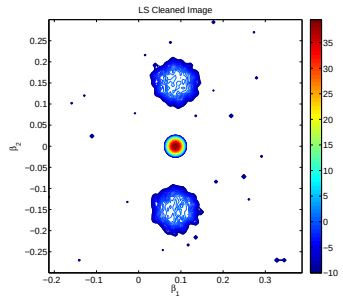

Fig. 3: Solution of the constrained LS image after convolution with a Gaussian beam.

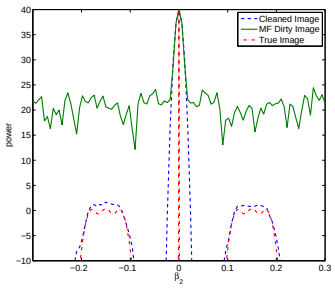

Fig. 4: Image cross–section.

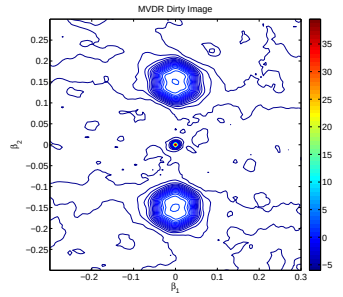

Fig. 5: MVDR dirty image.

simulated source is a combination of a strong point source and two extended structures. The extended sources are composed from seven Gaussian shaped sources, one in the middle and 6 on a hexagon

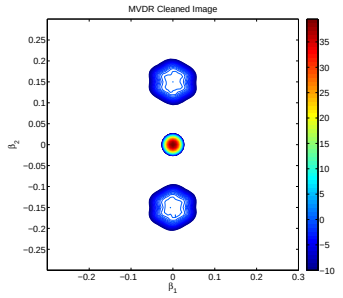

Fig. 6: Preconditioned WLS Image after convolution with a Gaussian beam.

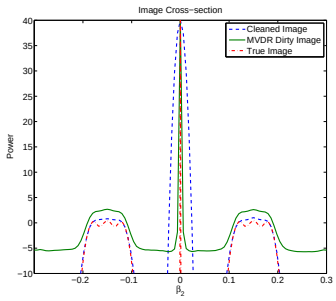

Fig. 7: MVDR dirty image and preconditioned WLS cross–section.

around it. Figure 1 shows the simulated image in dB scale. The background noise level that is added is 10 dB below the the extended sources. Figures 2 and 5 show the matched filter and MVDR dirty images respectively. Figures 3 and 6 show the reconstructed images, after deconvolution and smoothing with a Gaussian clean beam, for the LS and MVDR deconvolution with MF and MVDR dirty images as upper bounds respectively. A cross section of the images has been illustrated in Figures 4 and 7. Remarks:

- As expected MVDR dirty image has a much better dynamic range and lower side–lobes;
- due to better initial dirty image and upper bound the preconditioned WLS deconvolution gives a better cleaned image. However a trade off is made between the resolution of the point source and correct shape of the extended sources when we use the Gaussian beam to smoothen the image.
- the cross sections show the accuracy of the magnitudes. This shows that not only the shape but also the magnitude of the sources are better estimated using preconditioned WLS.

5. CONCLUSION

In this paper we provided an improved constrained least squares (CLS) image formation algorithm. We have demonstrated its superiority over previously proposed CLS algorithms in a simulated experiments. The full version of this paper [23] will provide examples for the simulated 3C catalog as well as measured data. It will also provide the full implementation of the algorithm which is computationally simple but requires careful use of Krylov spaces to prevent the need to store very large matrices. Future work will show how to combine the proposed approach with the robust adaptive selective sidelobe canceler [27].

6. REFERENCES

[1] J. A. Högbom, "Aperture synthesis with nonregular distribution of interefero meter baselines," *Astron. Astrophys. Suppl*, vol. 15, pp. 417–426, 1974.

[2] B. Frieden, "Restoring with maximum likelihood and maximum entropy," *Journal of the Optical Society of America*, vol. 62, pp. 511–518, 1972.

[3] S. Gull and G. Daniell, "Image reconstruction from incomplete and noisy data," *Nature*, vol. 272, pp. 686–690, 1978.

[4] J. Ables, "Maximum entropy spectreal analysis," *AAS*, vol. 15, pp. 383–393, 1974.

[5] S. Wernecke, "Two dimensional maximum entropy reconstruction of radio brightness," *Radio Science*, vol. 12, pp. 831–844, 1977.

[6] T. Cornwell and K. Evans, "A simple maximum entropy deconvolution algorithm," *Astronomy and Astrophysics*, vol. 143, pp. 77–83, 1985.

[7] T. Cornwell, K. Golap, and S. Bhatnagar, "The non-coplanar baselines effect in radio interferometry: The w-projetion algorithm," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 647–657, October 2008.

[8] U. Rau, S. Bhatnagar, M. Voronkov, and T. Cornwell, "Advances in calibration and imaging techniques in radio interferometry," *Proceeding of the IEEE*, vol. 97, pp. 1472–1481, Aug 2009.

[9] E. Pantin and J.-L. Starck, "Deconvolution of astronomical images using the multiscale maximum entropy method.," *Astronomy and Astrophysics Supplements*, vol. 118, pp. 575–585, Sept. 1996.

[10] D. S. Briggs, *High fidelity deconvolution of moderately resolved sources*. PhD thesis, The new Mexico Institute of Mining and Technology, Socorro, New Mexico, 1995.

[11] A. Leshem and A. van der Veen, "Radio-astronomical imaging in the presence of strong radio interference," *IEEE Trans. on Information Theory, Special issue on information theoretic imaging*, pp. 1730–1747, August 2000.

[12] C. Ben-David and A. Leshem, "Parametric high resolution techniques for radio astronomical imaging," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, pp. 670–684, Oct 2008.

[13] R. Levanda and A. Leshem, "Synthetic aperture radio telescopes," *Signal Processing Magazine, IEEE*, vol. 27, no. 1, pp. 14–29, 2010.

[14] R. Levanda and A. Leshem, "Radio astronomical image formation using sparse reconstruction techniques," *Electrical and Electronics Engineers in Israel, 2008. IEEEI 2008. IEEE 25th Convention of*, pp. 716–720, Dec. 2008.

[15] Y. Wiaux, L. Jacques, G. Puy, A. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of The Royal Astonomical Society*, Submitted 2009.

[16] R. Reid, "Smear fitting: a new image-deconvolution method for interferometric data," *Monthly Notices of the Royal Astronomical Society*, vol. 367, no. 4, pp. 1766–1780, 2006.

[17] A. Thompson, J. Moran, and G. Swenson, eds., *Interferometry and Synthesis in Radio astronomy*. John Wiley and Sons, 1986.

[18] G. Taylor, C. Carilli, and R. Perley, *Synthesis Imaging in Radio-Astronomy*. Astronomical Society of the Pacific, 1999.

[19] S. 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.

[20] D. Briggs, *High Fidelity Deconvolution of Moderately Resolved Sources*. PhD thesis, The New Mexico Institute of Mining and Technology, 1995.

[21] C. C. Paige and M. A. Saunders, "LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares," *ACM Trans. Math. Softw.*, vol. 8, pp. 43–71, Mar. 1982.

[22] A. Sardarabadi and A.-J. van der Veen, "Constrained imaging for radio astronomy," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on*, pp. 344–347, Dec 2013.

[23] A. Sardarabadi, A. Leshem, and A.-J. van der Veen, "Radio astronomical image formation using constrained imaging and Krylov spaces." 2014.

[24] A. Leshem, A. van der Veen, and A. J. Boonstra, "Multichannel interference mitigation techniques in radio-astronomy," *The Astrophysical Journal Supplements*, pp. 355–373, November 2000.

[25] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, pp. 1408–1418, 1969.

[26] 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.

[27] R. Levanda and A. Leshem, "Adaptive selective sidelobe canceller beamformer with applications to interference mitigation in radio astronomy," *Signal Processing, IEEE Transactions on*, vol. 61, pp. 5063–5074, Oct 2013.