Fast and Accurate Radio Interferometric Imaging Using Krylov Subspaces

Shahrzad Naghibzadeh* and Alle-Jan van der Veen*

*Facultuy of EEMCS, Delft University of Technology, Delft, The Netherlands

Abstract—We propose a fast iterative method for image formation in Radio Astronomy (RA). We formulate the image formation problem as a maximum likelihood estimation problem to estimate the image pixel powers via array covariance measurements. We use an iterative solution method based on projections onto Krylov subspaces and exploit the sample covariance error estimate via discrepancy principle as the stopping criterion. We propose to regularize the ill-posed imaging problem based on a Bayesian framework using MVDR beamformed data applied as a right preconditioner to the system matrix. We compare the proposed method with the state-of-the-art sparse sensing methods and show that the proposed method obtains comparably accurate solutions with a significant reduction in computation.

I. INTRODUCTION

The goal in RA imaging is to obtain an estimate of the spatial distribution of the intensity of the celestial sources over the Field of View (FoV) of the radio telescope array. Radio interferometry [1], [2] is a technique in which an image of the radio sky is obtained from the correlation data of the observing telescope array. This technique finds applications in many other array processing and image reconstruction problems where the estimation of the location and intensity of the sources present at the FoV of the measuring array via noisy and incomplete array measurements is required.

Next generation radio telescopes, such as the LOw Frequency ARray (LOFAR) [3] and the Square Kilometer Array (SKA) [4] are designed to generate very high resolution images using large volumes of correlation data. Therefore, the imaging problem in RA is highly large scale. The quest for imaging algorithms that can handle the scale of the RA imaging problem efficiently is one of the main focuses of the radio interferometric research. Another main problem in RA imaging is the inherent ill-posedness of the problem. Unless additional information about the sky map is exploited, it is not possible to obtain a unique and physically meaningful image from the finite set of correlation measurements.

Due to the large scale of the RA imaging problem and computation and storage limitations, iterative methods are exploited to obtain the image. Sparsity constraint is one of the main assumptions of the state-of-the-art RA imaging methods especially in high resolution point source recovery. Classical greedy methods and methods based on convex optimization and compressive sensing are used to find sparse solutions [5], [6], [7], [8]. The sheer amount of correlation data cannot be stored and has to be processed quasi-realtime. High processing power required by the current imaging algorithms and quasi-realtime requirements motivate the design of fast and efficient new algorithms. In the current research we focus on the science cases where a considerable amount of complex diffuse emissions are present such as in the studies of Galactic magnetism and the epoch of reionization. In these cases, the classical greedy methods are prohibitively slow.

In [9] the authors showed the applicability of methods based on Krylov subspaces in RA imaging. In this paper we modify the results via stochastic problem formulation based on maximum likelihood estimation. Furthermore, we suggest a regularization scheme based on Bayesian estimation and show that we can obtain an estimate of the source power covariance using the available correlation data through Minimum Variance Distortionless Response (MVDR) beamforming. We show via simulations that the method outperforms the methods based on compressed sensing in computations while maintaining similar sparse reconstruction quality.

II. SIGNAL PROCESSING DATA MODEL

We employ the array signal processing framework and data model for RA imaging as suggested in [10], [11], [12]. The notations $(.)^T, (.)^H, (.)^*, \circ, \otimes, \odot$ and $\mathbf{I}$ respectively denote transpose, the Hermitian transpose, complex conjugate, Khatri-Rao product, the Kronecker product, element-wise product and the identity matrix of proper size. Furthermore, the operator diag(.) converts the diagonal of a matrix to a vector and converts a vector to a diagonal matrix, vec(.) forms a vector from a matrix by stacking the columns of the matrix and trace(.) computes the sum of the diagonal elements of a matrix.

Assuming a telescope array of P distinct receiving elements, the baseband output signal of the array elements can be stacked in a $P \times 1$ vector $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_P(t)]^T$. where $x_p(t)$ denotes the output signal of the pth array element. Similarly, assuming Q mutually independent source signals $s_q(t)$ impinging on the array, we stack all the $Q \times 1$ source signals into a vector $\mathbf{s}(t)$. The output of the telescope array is a linear combination of the phase-delay compensated source signals and receiver noise. We model the receiver noise as mutually independent Gaussian signals stacked in a $P \times 1$ vector $\mathbf{n}(t)$. Assuming that narrow-band condition holds, the source signals are received with different phase delays on different antenna elements. The array response vector $\mathbf{a}_q$ compensates for the phase delays of the qth source on different array elements and is known based on the known source and array positions. We can write the output of the telescope array as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_Q]$ is the array response vector.

In practice, the position of the celestial sources are unknown. We decompose the FoV of the telescope array into a fine grid where each grid point denotes an image pixel, with a total of Q image pixels. Moreover, 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

$$\mathbf{x}_k[n] = \mathbf{A}_k\mathbf{s}[n] + \mathbf{n}_k[n], \ n = 1, \cdots, N. \qquad (2)$$

In this representation, $\mathbf{x}_k[n]$, $\mathbf{s}[n]$ 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 and $\mathbf{A}_k$ 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

$$a_{pq} = \frac{1}{\sqrt{P}} e^{-j\frac{2\pi}{\lambda}\mathbf{v}_p^T\mathbf{z}_q} \qquad (3)$$

where λ is the wavelength of the received signal , $\mathbf{v}_p$ is a 3×1 vector of the Cartesian location of the pth array element with respect to a chosen origin in the field of array and $\mathbf{z}_q$ contains the direction cosines of the qth pixel in the image plane. With the knowledge of the antenna and pixel positions, the array response matrix is known. Assuming the signals and the receiver noise are uncorrelated and the noise on different antennas are mutually uncorrelated, the autocovariance of the received signals is computed as

$$\mathbf{R} = E\{\mathbf{x}[n]\mathbf{x}^H[n]\} = \mathbf{A}\boldsymbol{\Sigma}_\mathbf{s}\mathbf{A}^H + \boldsymbol{\Sigma}_\mathbf{n}, \qquad (4)$$

where $\boldsymbol{\Sigma}_\mathbf{s} = \text{diag}\{\boldsymbol{\sigma}\}$ and $\boldsymbol{\Sigma}_\mathbf{n} = \text{diag}\{\boldsymbol{\sigma}_\mathbf{n}\}$ represent the covariance matrices associated with the source signals and the received noise respectively where $\boldsymbol{\sigma} = [\sigma_{s1}^2, \sigma_{s2}^2, \cdots, \sigma_{sQ}^2]^T$ and $\boldsymbol{\sigma}_\mathbf{n} = [\sigma_{n1}^2, \sigma_{n2}^2, \cdots, \sigma_{nP2}^2]^T$. An estimate of the data covariance matrix is obtained using the available received data samples. The sample covariance matrix is calculated as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}[n]\mathbf{x}^H[n], \qquad (5)$$

and is used as an estimate of the true covariance matrix $\mathbf{R}$.

III. PROBLEM FORMULATION

A. Measurement model

The radio astronomical imaging process amounts to estimating the image pixel intensities $\boldsymbol{\sigma}$ based on the covariance data measured by a telescope array $\hat{\mathbf{R}}$ over the FoV of the array. To obtain a linear measurement model, we vectorize the covariance data model (4) as

$$\mathbf{r} = (\mathbf{A}^* \circ \mathbf{A})^H\boldsymbol{\sigma} + \mathbf{r}_\mathbf{n}, \qquad (6)$$

where $\mathbf{r} = \text{vec}(\mathbf{R})$ and $\mathbf{r}_\mathbf{n} = \text{vec}(\mathbf{R}_\mathbf{n}) = (\mathbf{I} \circ \mathbf{I})\boldsymbol{\sigma}_\mathbf{n}$. Subsequently, we vectorize the covariance measurement matrix as

$$\hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}}). \qquad (7)$$

This brings us to the linear measurement equation for estimating $\boldsymbol{\sigma}$ based on the measured $\hat{\mathbf{r}}$

$$\hat{\mathbf{r}} = \mathbf{r} + \mathbf{e} = \mathbf{M}\boldsymbol{\sigma} + \mathbf{r}_\mathbf{n} + \mathbf{e}, \qquad (8)$$

where $\mathbf{e}$ represents the error due to the finite sample estimation of the covariance. For large number of samples N we can assume that $\mathbf{e}$ is distributed according to a zero-mean complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \mathbf{C}_\mathbf{e})$ where $\mathbf{C}_\mathbf{e} = (\mathbf{R}^T \otimes \mathbf{R})/N$ [13], [14]. $\mathbf{M} = \mathbf{A}^* \circ \mathbf{A}$ is the system matrix of the linear measurement model of size $P^2 \times Q$. One element of $\mathbf{M}$ corresponding to the baseline between the ith and jth

antenna and the qth pixel is computed as:

$$\mathbf{M}_{ij,q} = a_{iq}^*a_{jq} = e^{j\frac{2\pi}{\lambda}(\mathbf{v}_i - \mathbf{v}_j)^T\mathbf{z}_q}. \qquad (9)$$

Assuming the knowledge of the receiver noise powers, we denote $\hat{\mathbf{r}} - \mathbf{r}_\mathbf{n}$ by $\tilde{\mathbf{r}}$. RA imaging problem reduces to estimating $\boldsymbol{\sigma}$ from $\tilde{\mathbf{r}}$ assuming the knowledge of $\mathbf{M}$. In high resolution imaging, matrix $\mathbf{M}$ is wide and singular and (8) denotes an underdetermined set of equations.

B. Beamforming-based estimation

An initial estimate of the image can be obtained via beamforming. We consider the two main beamforming approaches; Matched Filtering (MF) and MVDR [15]. The "dirty" image obtained by the MF beamformer is computed as

$$\boldsymbol{\sigma}_{\text{MF},i} = \mathbf{a}_i^H\mathbf{R}\mathbf{a}_i, \ i = 1, 2, \cdots, Q \qquad (10)$$

where $\boldsymbol{\sigma}_{\text{MF},i}$ denotes the ith pixel of the MF image. Similarly, the MVDR dirty image is computed as [11]

$$\boldsymbol{\sigma}_{\text{MVDR},i} = \frac{1}{\mathbf{a}_i^H\mathbf{R}^{-1}\mathbf{a}_i}, \ i = 1, 2, \cdots, Q \qquad (11)$$

where $\boldsymbol{\sigma}_{\text{MVDR},i}$ is the ith pixel of the MVDR image. Since we don't have access to the true covariance matrix, we use the sample covariance matrix $\hat{\mathbf{R}}$ instead to obtain the beamformed images.

C. Maximum likelihood estimation

Equation (8) shows the linear measurement model for RA imaging. It has been shown that for linear measurement models, Maximum Likelihood Estimation (MLE) results in an efficient estimator that is also a Minimum Variance Unbiased (MVU) estimator [16]. MLE with the complex Gaussian prior on the error leads to weighted least squares (WLS) formulation of the imaging problem as shown in [12], [11]

$$\{\hat{\boldsymbol{\sigma}}\} = \underset{\boldsymbol{\sigma}}{\text{argmin}} \parallel \boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma}) \parallel_2^2, \qquad (12)$$

where $\{\hat{\boldsymbol{\sigma}}\}$ represents the set of possible solutions to the underdetermined imaging problem and $\mathbf{C}_\mathbf{e}^{-1} = \boldsymbol{\Gamma}^H\boldsymbol{\Gamma}$. To obtain a unique and regularized solution, additional information is necessary. We can write the general regularized WLS RA imaging problem as

$$\hat{\boldsymbol{\sigma}} = \underset{\boldsymbol{\sigma}}{\text{argmin}} \parallel \boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma}) \parallel_2^2 + \tau\mathcal{R}(\boldsymbol{\sigma}) \qquad (13)$$

where τ is the regularization parameter and $\mathcal{R}(.)$ denotes the regularization operator, typically an ℓ_2 or ℓ_1 constraint on the solution. Equivalently we can rewrite the problem formulation (13) as

$$\underset{\boldsymbol{\sigma}}{\text{minimize}} \ \mathcal{R}(\boldsymbol{\sigma}) \ \text{subject to} \ \parallel \boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma}) \parallel_2^2 \leqslant \epsilon. \qquad (14)$$

There is a data-dependent one-to-one correspondence between τ in (13) and ϵ in (14) that makes the two optimization problems have the same solution. Therefore, they can be used interchangeably where ϵ and τ are related to the level of noise in the data.

IV. PROPOSED SOLUTION METHOD

A. Problem reformulation

We focus on the general Tikhonov regularized WLS problem formulation

$$\hat{\boldsymbol{\sigma}} = \underset{\boldsymbol{\sigma}}{\text{argmin}} \|\boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\boldsymbol{\sigma})\|_2^2 + \tau\|\mathbf{L}\boldsymbol{\sigma}\|_2^2, \qquad (15)$$

where $\mathbf{L}$ is a linear regularization operator. Assuming $\mathbf{L}$ is invertible, with a change of variables $\boldsymbol{\alpha} = \mathbf{L}\boldsymbol{\sigma}$ we can rewrite the objective function in terms of $\boldsymbol{\alpha}$ as

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \|\boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\mathbf{L}^{-1}\boldsymbol{\alpha})\|_2^2 + \tau\|\boldsymbol{\alpha}\|_2^2, \qquad (16)$$

The image can be obtained by the linear transform $\hat{\boldsymbol{\sigma}} = \mathbf{L}^{-1}\hat{\boldsymbol{\alpha}}$. Equation (16) is equivalent to the constrained optimization problem

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \ \|\boldsymbol{\alpha}\|_2^2 \ \text{subject to} \ \|\boldsymbol{\Gamma}(\tilde{\mathbf{r}} - \mathbf{M}\mathbf{L}^{-1}\boldsymbol{\alpha})\|_2^2 \leqslant \epsilon, \quad (17)$$

for some equivalent data-dependent value of τ and ϵ. In this context, $\mathbf{L}^{-1}$ can be viewed as a right preconditioner directly applied to the system matrix.

B. Regularization by beamforming

Starting from a Bayesian framework with Gaussian prior distribution on the image, the Maximum A Posteriori (MAP) estimator results in a similar formulation as (15) with $\tau = 1$ and $\mathbf{C}_{\boldsymbol{\sigma}}^{-1} = \mathbf{L}^T\mathbf{L}$ where $\mathbf{C}_{\boldsymbol{\sigma}}$ is the covariance of the image.

To obtain an estimate of $\mathbf{C}_{\boldsymbol{\sigma}}$, we consider for the moment that $\mathbf{M}$ is full rank. Thus, the MLE estimate in (12) has a closed form solution

$$\hat{\boldsymbol{\sigma}}_{\text{MLE}} = (\mathbf{M}^H\mathbf{C_e}^{-1}\mathbf{M})^{-1}\mathbf{M}^H\mathbf{C_e}^{-1}\tilde{\mathbf{r}}. \qquad (18)$$

It is shown that this estimator is an efficient MVU estimator [16] with covariance

$$\mathbf{C}_{\hat{\boldsymbol{\sigma}}} = (\mathbf{M}^H\mathbf{C_e}^{-1}\mathbf{M})^{-1}. \qquad (19)$$

We can write

$$\begin{aligned}\mathbf{M}^H\mathbf{C_e}^{-1}\mathbf{M} &= N(\mathbf{A}^*\circ\mathbf{A})^H(\mathbf{R}^{-T}\otimes\mathbf{R}^{-1})(\mathbf{A}^*\circ\mathbf{A})\\ &= N(\mathbf{A}^T\mathbf{R}^{-T}\mathbf{A}^*)\odot(\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}).\end{aligned} \qquad (20)$$

Assuming the estimated pixel powers are uncorrelated we can write $\mathbf{C}_{\hat{\boldsymbol{\sigma}}} \approx \text{diag}(\text{var}(\hat{\boldsymbol{\sigma}}))$ where

$$\text{var}_i(\hat{\boldsymbol{\sigma}}) = \frac{1}{N(\mathbf{a}_i^H\mathbf{R}^{-1}\mathbf{a}_i)^2}, \ i = 1, 2, \ldots, Q, \qquad (21)$$

with $\text{var}_i(\hat{\boldsymbol{\sigma}})$ denoting the variance at the ith pixel. Comparing (11) and (21) we conclude that

$$\mathbf{C}_{\hat{\boldsymbol{\sigma}}} \approx \text{diag}(\text{var}(\hat{\boldsymbol{\sigma}})) = \frac{\text{diag}(\boldsymbol{\sigma}_{\text{MVDR}})^2}{N}. \qquad (22)$$

Therefore, by using the sample covariance data we obtain the sample covariance MVDR image $\hat{\boldsymbol{\sigma}}_{\text{MVDR}}$ and set $\mathbf{L}^{-1} = \text{diag}(\hat{\boldsymbol{\sigma}}_{\text{MVDR}})/\sqrt{N}$. We have previously suggested in [9] to use the MF dirty image for regularization purposes. We show via simulations that using the MVDR image reduces the estimation error.

C. Krylov subspace method

To solve problem (17) we use the Krylov subspace based method of LSQR [17] as discussed in [9]. Denoting $\tilde{\mathbf{M}} =$ $\boldsymbol{\Gamma}\mathbf{M}\mathbf{L}^{-1}$, the solution $\hat{\boldsymbol{\alpha}}^{(t)}$ after t iterations is obtained by solving the optimization problem

$$\begin{aligned}\hat{\boldsymbol{\alpha}}^{(t)} = &\underset{\boldsymbol{\alpha}}{\text{argmin}} \ \| \boldsymbol{\Gamma}\tilde{\mathbf{r}} - \tilde{\mathbf{M}}\boldsymbol{\alpha} \|_2^2\\ &\text{subject to } \boldsymbol{\alpha} \in \mathcal{K}_t(\tilde{\mathbf{M}}^H\tilde{\mathbf{M}}, \tilde{\mathbf{M}}^H\boldsymbol{\Gamma}\tilde{\mathbf{r}}).\end{aligned} \qquad (23)$$

where $\mathcal{K}$ denotes the Krylov subspace and t shows the iteration count. $\mathcal{K}_t$ is constructed as [18],

$$\begin{aligned}&\mathcal{K}_t(\tilde{\mathbf{M}}^H\tilde{\mathbf{M}}, \tilde{\mathbf{M}}^H\boldsymbol{\Gamma}\tilde{\mathbf{r}})\\ &= \text{span}\{\tilde{\mathbf{M}}^H\boldsymbol{\Gamma}\tilde{\mathbf{r}}, (\tilde{\mathbf{M}}^H\tilde{\mathbf{M}})\tilde{\mathbf{M}}^H\boldsymbol{\Gamma}\tilde{\mathbf{r}}, \cdots, (\tilde{\mathbf{M}}^H\tilde{\mathbf{M}})^{t-1}\tilde{\mathbf{M}}^H\boldsymbol{\Gamma}\tilde{\mathbf{r}}\}.\end{aligned} \qquad (24)$$

We employ the discrepancy principle [19] as the stopping rule in solving equation (17). The threshold on the residual norm can be set using the error on the data based on the expression for the sample noise covariance matrix $\mathbf{C_e}$ as

$$\|\boldsymbol{\Gamma}\tilde{\mathbf{r}} - \tilde{\mathbf{M}}\boldsymbol{\alpha}\|_2^2 = \|\boldsymbol{\Gamma}\mathbf{e}\|_2^2 = \text{trace}(\text{Cov}(\boldsymbol{\Gamma}\mathbf{e})) \qquad (25)$$

where $\boldsymbol{\Gamma}\mathbf{e}$ is the whitened error on the data,

$$\text{Cov}(\boldsymbol{\Gamma}\mathbf{e}) = E\{\boldsymbol{\Gamma}\mathbf{e}\mathbf{e}^H\boldsymbol{\Gamma}^H\} = \mathbf{I}, \qquad (26)$$

where $\mathbf{I}$ denotes a $P^2 \times P^2$ identity matrix, $\text{Cov}(.)$ denotes the covariance and $E\{.\}$ denotes the expected value. The Krylov iterations are stopped as soon as

$$\|\boldsymbol{\Gamma}\tilde{\mathbf{r}} - \tilde{\mathbf{M}}\boldsymbol{\alpha}_t\|_2^2 \leqslant \epsilon \qquad (27)$$

is satisfied for the first time at iteration t. Based on the above, we set $\epsilon = P^2$.

V. 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 M31 galaxy shown in Figure 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. To construct the sampled covariance matrix, Gaussian receiver noise with variance $\sigma_n^2 = 0.5$ 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 and the MVDR image are shown in Figure 1(c) and (d), respectively . The simulations were performed in MATLAB R2014b on a computer with Intel i5-4670 CPU 3.40 GHz under 64-bit Windows 7.

The reconstructed images for different imaging algorithms have been compared in Figure 2. Figures 2(a) shows the reconstruction based on the classical greedy method CLEAN [5] after applying the post-processing with the main beam. Figure 2(b) shows the Alternating Direction Method of Multipliers (ADMM) [20] reconstruction with ℓ_1 sparsity constraints. Figure 2(c) is the reconstructed image based on Sparsity Averaging Reweighted Analysis (SARA) formalism [8] implemented with ADMM. Figures 2(d),(e) and (f) share

The convergence in terms of relative residual and the reconstruction quality in terms of ℓ_1 norm error and ℓ_2 norm error per iteration have been compared in Figure 3. ADMM, LSQR, PLSQR and MLSQR are shown in blue, black, red, magenta and green graphs, respectively. For better visualization, LSQR-based methods are run beyond the stopping iterations. As can
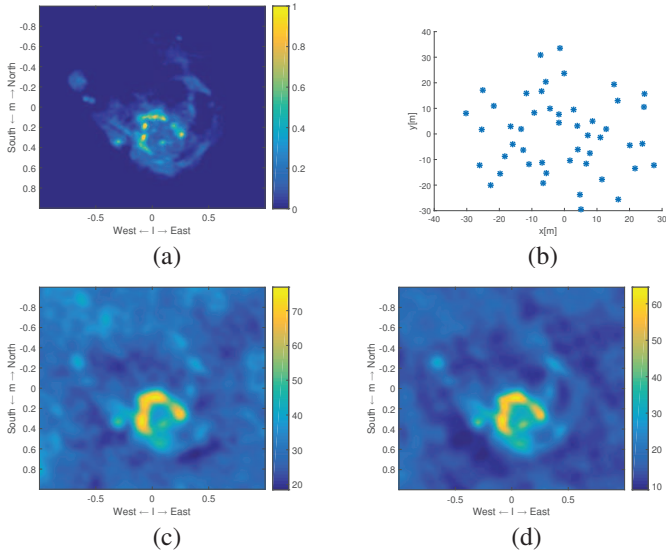
Fig. 1: (a) Test image, (b) antenna placement, (c) dirty image, (d) MVDR dirty image
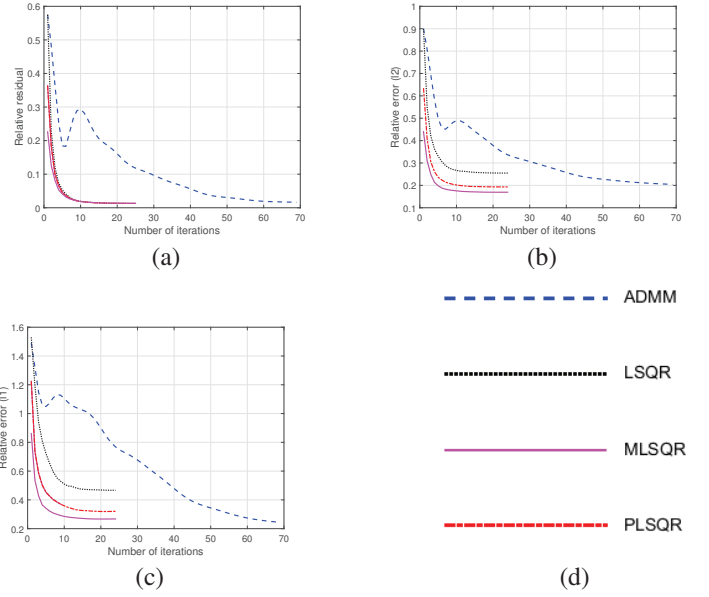

Fig. 2: (a) CLEAN, (b) ADMM, (c) SARA, (d) LSQR, (e) PLSQR, (f) MLSQR

Fig. 3: (a) Relative residual per iteration, (b) ℓ_2 norm error, (c) ℓ_1 norm error

	Number of iterations	Reconstruction time (seconds)	$\|\hat{\sigma} - \sigma\|_2$	$\|\hat{\sigma} - \sigma\|_1$
CLEAN	500	89.4	3.62	10.77
ADMM	69	10.8	1.94	5.94
SARA	140	42.7	1.65	4.97
LSQR	9	1.27	2.45	8.95
PLSQR	9	1.74	1.69	5.82
MLSQR	8	1.75	1.36	4.52

TABLE I: Performance analysis

be concluded form Figure 3, methods based on LSQR exhibit substantially faster convergence than the steepest descent based ADMM while maintaining comparable reconstruction quality.

The performance of the imaging algorithms are compared in Table I in terms of number of iterations, reconstruction time and error norm. Comparing the reconstruction times and quality, MLSQR is the most efficient method with the least reconstruction error. Furthermore, we compare the Mean Square estimation Error (MSE) of PLSQR and MLSQR. The MSE is computed as $\text{MSE} = \frac{1}{M} \sum_{m=1}^{M} \|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}\|_2^2$, where M is the number of Monte-Carlo experiments. After $M = 1000$ runs, the MSE for PLSQR and MLSQR are respectively 3.06 and 2.05. This shows the reduction in estimation error when applying regularization via the MVDR image in comparison with the dirty image.

VI. CONCLUSION

We consider RA imaging problem as a maximum likelihood estimation problem which results in a WLS formulation through covariance preconditioning. We show that right preconditioning provides a way to efficiently add regularizing assumptions and additional constrains into the methods based on Krylov subspaces without the need to increase the iteration level. We consider the image formation problem in the context of Bayesian estimation and show that the preconditioner in the Bayesian framework can be considered as the MVDR beamformered data. The proposed method is compared with the state-of-the art sparse imaging methods and is shown to provide comparably accurate solutions with significantly reduced computations. The results of the current research can be applied in different imaging fields.
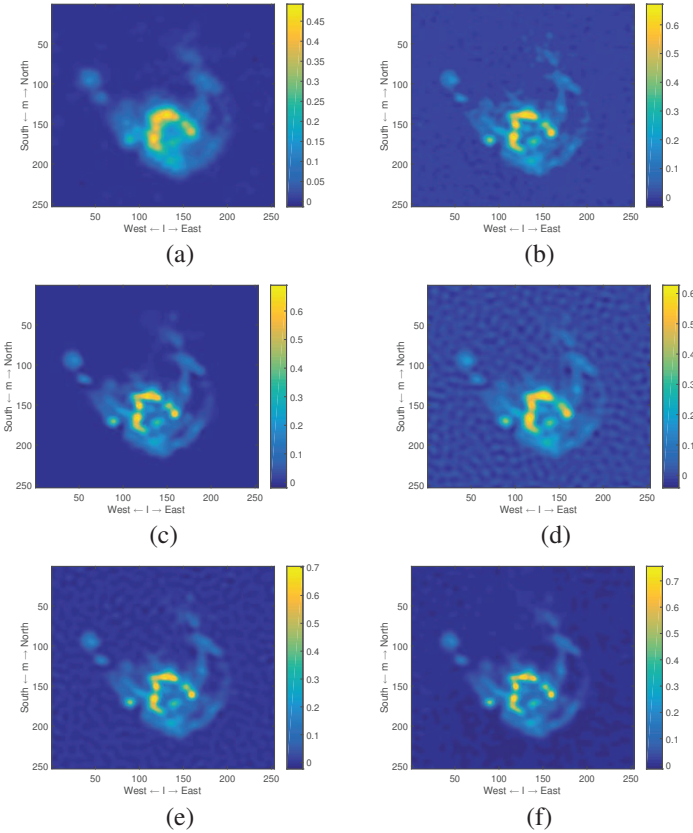
REFERENCES

[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] M. Van Haarlem, M. Wise, A. Gunst, G. Heald, J. McKean, J. Hessels, A. De Bruyn, R. Nijboer, J. Swinbank, R. Fallows *et al.*, "Lofar: The low-frequency array," *Astronomy & Astrophysics*, vol. 556, p. A2, 2013.

[4] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. Lazio, "The square kilometre array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.

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

[6] Y. Wiaux, L. Jacques, G. Puy, A. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1733–1742, 2009.

[7] J. McEwen and Y. Wiaux, "Compressed sensing for wide-field radio interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 413, no. 2, pp. 1318–1332, 2011.

[8] R. Carrillo, J. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (SARA): a novel algorithm for radio-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 426, no. 2, pp. 1223–1234, 2012.

[9] S. Naghibzadeh and A.-J. van der Veen, "Radioastronomical least squares image reconstruction with iteration regularized krylov subspaces and beamforming-based prior conditioning," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 3385–3389.

[10] 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*. Springer, 2005, pp. 231–249.

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

[12] 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.

[13] 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.

[14] 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.

[15] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[16] S. M. Kay, "Fundamentals of statistical signal processing, volume i: estimation theory," 1993.

[17] 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.

[18] Y. Saad, "Krylov subspace methods for solving large unsymmetric linear systems," *Mathematics of computation*, vol. 37, no. 155, pp. 105–126, 1981.

[19] V. A. Morozov, "On the discrepancy principle in solving operator equations by the regularization method," *Zh. vychisl. mat. mat. fiz*, vol. 8, pp. 295–309, 1968.

[20] S. Boyd, "Alternating direction method of multipliers," in *Talk at NIPS Workshop on Optimization and Machine Learning*, 2011.