

Scalable Algorithm for Polarization Constrained Sparse Interferometric Stokes Imaging

Jasleen Birdi, Audrey Repetti, and Yves Wiaux

Institute of Sensors, Signals and Systems

Heriot-Watt University, Edinburgh EH14 4AS, UK

Email: jb36@hw.ac.uk, a.repetti@hw.ac.uk, and y.wiaux@hw.ac.uk

Abstract—Polarized emissions from various astrophysical sources provide an additional means to infer the properties of the source. This information can be extracted by imaging the Stokes parameters characterizing these emissions, from the incomplete Fourier sampling measurements provided by the radio interferometer. To solve the corresponding ill-posed inverse problem, we recently proposed the Polarized SARA method which exploits the polarization constraint along with imposing an average sparsity prior for each of the Stokes parameters, encompassed in a reweighting approach. Motivated by the large amounts of data produced by the next-generation radio interferometers, in the current work, we propose a scalable version of this method, combined with an accelerating strategy. We showcase the performance of this improved method on a real dataset acquired by the Very Large Array (VLA).

I. INTRODUCTION

Polarimetric imaging is an important analysis tool in many applications. In particular for radio astronomy, it gives insights into the magnetic field distributions around the polarized source of interest as well as about various properties of the medium of propagation of these emissions along the line of sight [1]. This all calls for the need of polarimetric imaging. Moreover, next-generation radio telescopes will produce huge amounts of data, and thus the scalability of the underlying imaging algorithms is of paramount importance.

In order to characterize these polarized emissions, the Stokes parameters- I , Q , U and V , all real-valued, are used. While I corresponds to the total intensity distribution of the target sky area, Q and U denote the linear polarization, and V refers to the circular polarization. From these parameters, the linear polarization image $P = Q + \mathbf{i}U$ can be defined, whose magnitude gives the linear polarization intensity and the phase gives the electric vector polarization angle. Interestingly, all these parameters are linked via a physical non-linear constraint, namely the polarization constraint, according to which the polarized intensity provides a lower bound on the total intensity, i.e. $\sqrt{Q^2 + U^2} + V^2 \leq I$.

In the context of radio interferometry (RI), the measurements correspond to an incomplete Fourier sampling of the brightness distribution of the sky area of interest, posing a highly under-determined image reconstruction problem. Many algorithms have been proposed in the past to tackle this problem, mainly focusing on Stokes I imaging though. Conventionally speaking, extension of these algorithms to polarimetric imaging is straightforward. It consists in employing the same

approach, as developed for Stokes I imaging, for all the other Stokes parameters as well. The standard RI imaging algorithm is CLEAN, which works on iterative non-linear deconvolution approach, assuming implicitly the sparsity of the underlying image [2]. Many improvements of CLEAN have been suggested in the past years, such as its multi-scale version MS-CLEAN [3]. Recently, compressive sensing (CS) based techniques have also been applied for radio interferometric imaging [4]–[7]. These methods rely on reconstructing the sought image by imposing its sparsity in some sparsifying domain. The first application of these CS techniques for polarimetric imaging in RI has been done in [8], where the authors promote the sparsity of the images of interest using the ℓ_1 norm and total variation (TV) regularization.

It is worth mentioning that the above mentioned approaches estimate each of the Stokes parameters totally independently, which may produce unphysical reconstructions not satisfying the polarization constraint. It motivates the enforcement of this physical constraint in the reconstruction process. In this direction, Maximum Entropy Methods (MEM) offer an entropy function suitably designed to impose this constraint [9]–[11]. Note that, traditionally, these methods use gradient based optimization techniques, hence requiring the objective function to be differentiable. Another possibility to impose this constraint is by making use of change of variables, wherein the polarized images are instead represented in terms of their fractional polarization and position angle, thereby solving directly for these quantities [12]. However, solving for these variables leads to a non-convex optimization problem, and thus convergence to global minimum may not be guaranteed.

The work herein builds on our recently proposed approach for sparse imaging of the Stokes parameters under the polarization constraint, namely Polarized SARA [13]. In particular, it imposes the polarization constraint explicitly in the reconstruction process, using epigraphical projection techniques [14], such that the resultant minimization problem is convex. At the same time, it can deal with both smooth and non-smooth terms, and leverages the CS framework to promote the sparsity of each of the Stokes parameter using the Sparsity Averaging Reweighting Analysis (SARA) approach [5], [15]. In the current work, we propose an improved version of the Polarized SARA method, employing three complementary techniques foundational to handle real datasets. Firstly, we introduce a parallel data-block strategy [7], in order to be able to deal

with large data volumes. Secondly, a preconditioning scheme is used [16] to accelerate the algorithm. Lastly, when dealing with real datasets, the model may not be perfectly known, and must account for calibration errors. In this context, we leverage the adaptive approach developed in [17] to estimate the algorithm parameters (i.e. the bounds on the noise). We apply the resulting proposed method on a VLA dataset and compare it with MS-CLEAN.

The rest of the paper is organized as follows. In Section II, we describe the RI imaging model and pose the inverse problem for polarimetric imaging. Section III details the proposed approach. Simulations and results are presented in Section IV, with the conclusion provided in Section V.

II. OBSERVATION MODEL

A radio interferometer consists of an array of antennas, designed to image the radio sources in a given sky area, such that each antenna pair provides a set of measurements. Let the Stokes matrix $\bar{\mathbf{S}} \in \mathbb{R}^{N \times 4}$ be the concatenation of the discretization of the Stokes images I, Q, U and V . Each row of the measurement matrix, denoted by $\mathbf{Y} \in \mathbb{C}^{M \times 4}$, corresponds to the full polarization measurements acquired by a given antenna pair, and is given by

$$\mathbf{Y} = \Phi(\bar{\mathbf{S}}) + \mathbf{E}, \quad (1)$$

where each column of the matrix $\mathbf{E} \in \mathbb{C}^{M \times 4}$ is a realization of an additive i.i.d. Gaussian noise. The measurement operator $\Phi : \mathbb{C}^{N \times 4} \rightarrow \mathbb{C}^{M \times 4}$ is given by $\Phi(\bar{\mathbf{S}}) = \mathbf{\Theta} \mathbf{G} \mathbf{F} \mathbf{Z} \bar{\mathbf{S}} \mathbf{L}$. More precisely, \mathbf{L} denotes the linear operator acting on the Stokes matrix $\bar{\mathbf{S}}$ to give the brightness matrix [18]. In addition, \mathbf{F} is the Fourier matrix acting on the oversampled brightness matrix, $\bar{\mathbf{S}} \mathbf{L}$. This oversampling is done by the matrix \mathbf{Z} , which also takes into account scale factors pre-compensating for the interpolation errors introduced by the matrix \mathbf{G} [7]. In particular, each row of matrix \mathbf{G} consists of convolution kernels to compute continuous Fourier samples from the discrete Fourier coefficients. Furthermore, the noise statistics are incorporated by the matrix $\mathbf{\Theta}$, whose elements are the square root of the natural weights, used to whiten the noise [16].

Following the block data splitting technique described in [7], the measurement (1) can be decomposed into d blocks, such that $\mathbf{Y} = (\mathbf{Y}_j)_{1 \leq j \leq d}$, where, for every $j \in \{1, \dots, d\}$, $\mathbf{Y}_j \in \mathbb{C}^{M_j \times 4}$ is given by

$$\mathbf{Y}_j = \Phi_j(\bar{\mathbf{S}}) + \mathbf{E}_j. \quad (2)$$

In (2), \mathbf{E}_j corresponds to the j -th block of the additive noise \mathbf{E} , and the measurement operator Φ_j is given by $\Phi_j(\bar{\mathbf{S}}) = \mathbf{\Theta}_j \mathbf{G}_j \mathbf{F} \mathbf{Z} \bar{\mathbf{S}} \mathbf{L}$, $\mathbf{\Theta}_j$ and \mathbf{G}_j being the j -th block matrices of $\mathbf{\Theta}$ and \mathbf{G} , respectively.

III. DISTRIBUTED POLARIZED SARA

A. Minimization problem

In order to recover the Stokes images from the degraded measurements (1), as proposed in [13], we aim to solve the following minimization problem

$$\underset{\mathbf{S} \in \mathbb{R}^{N \times 4}}{\text{minimize}} f(\Phi(\mathbf{S})) + \iota_{\mathbb{U}}(\mathbf{S}) + \gamma g(\Psi^\dagger \mathbf{S}) + \iota_{\mathbb{P}}(\mathbf{S}), \quad (3)$$

where the first term is the data fidelity term, whereas the other terms are the regularization terms incorporating *a priori* information on the sought images, with $\gamma > 0$ being a free parameter. In (3), for every closed convex non-empty set C , the function $\iota_C(\mathbf{S})$ denotes the indicator function of C , evaluated in \mathbf{S} , which is equal to 0 if $\mathbf{S} \in C$, and $+\infty$ otherwise.

In the current work, we propose to enforce the data fidelity term in a distributed manner. Based on the approach developed in [7], we propose to constrain the residual blocks $\mathbf{Y}_j - \Phi_j(\mathbf{S})$ to belong to the ℓ_2 ball centered in $\mathbf{0}$ with radius $\epsilon_j > 0$, by choosing

$$f(\Phi(\mathbf{S})) = \sum_{j=1}^d \iota_{\mathcal{B}_j(\mathbf{Y}_j, \epsilon_j)}(\Phi_j(\mathbf{S})), \quad (4)$$

where, for every $j \in \{1, \dots, d\}$, $\mathcal{B}_j(\mathbf{Y}_j, \epsilon_j) = \{\mathbf{B} \in \mathbb{C}^{M_j \times 4} : \|\mathbf{B} - \mathbf{Y}_j\|_2 \leq \epsilon_j\}$ is the ℓ_2 ball centered in \mathbf{Y}_j with radius $\epsilon_j > 0$, chosen according to the theoretical noise level.

Concerning the regularization terms, firstly the term $\iota_{\mathbb{U}}(\mathbf{S})$ enforces real-valuedness of the Stokes images by projecting \mathbf{S} onto the set $\mathbb{U} = \mathbb{R}^{N \times 4}$. Secondly, we leverage the CS framework and promote the sparsity of the Stokes images in a sparsifying dictionary Ψ using the SARA regularization [15], [19]. It considers Ψ as the concatenation of the first eight Daubechies wavelet bases and the Dirac basis, i.e. $\Psi = [\Psi_1, \dots, \Psi_9]$, coupled with a reweighting scheme. The latter consists in iteratively solving the problems with weighted ℓ_1 norm, and thus approaching towards the solution in an ℓ_0 sense [20]. We propose to split the corresponding weighted ℓ_1 term as a sum over the different bases involved as well as over the different Stokes images¹, i.e.,

$$g(\Psi^\dagger \mathbf{S}) = \sum_{i=1}^4 \sum_{l=1}^9 \|[\mathbf{W}_l]_{:,i} \Psi_l^\dagger \mathbf{S}_{:,i}\|_1, \quad (5)$$

where for each $l \in \{1, \dots, 9\}$, $\Psi_l^\dagger \in \mathbb{R}^{N \times N}$ and $\mathbf{W}_l \in \mathbb{R}_+^{N \times 4}$ is the weighting matrix. For all l and i , at each reweighting iteration, the elements of $[\mathbf{W}_l]_{:,i}$ are computed from the inverse of the estimated images in the previous iteration [13].

Furthermore, we exploit the physical link between the Stokes images provided by the polarization constraint. The associated polarization constraint set is defined as

$$\mathbb{P} = \{\mathbf{S} \in \mathbb{R}^{N \times 4} \mid (\forall n \in \{1, \dots, N\}) -\mathbf{S}_{n,1} + \|\mathbf{S}_{n,2:4}\|_2 \leq 0\}. \quad (6)$$

In (3), the indicator function of \mathbb{P} is used to impose this constraint, which also implicitly enforces the positivity of the total intensity image (Stokes I). However, the associated projection onto the set \mathbb{P} does not have a closed form and to overcome this problem, we split this set into simpler constraint sets leveraging the epigraphical projection technique [14]. This technique consists in introducing an auxiliary variable

¹For a matrix \mathbf{S} , i and j in the notation $\mathbf{S}_{i,j}$ respectively refer to the row and column indices. If $i = :$, it selects the whole j -th column and vice-a-versa.

$\mathbf{Z} \in \mathbb{R}^{N \times 2}$, and reformulating the constraint $\mathbf{S} \in \mathbb{P}$ as follows, for every $n \in \{1, \dots, N\}$,

$$\begin{cases} h_1(\mathbf{S}_{n,1}) = -\mathbf{S}_{n,1} \leq \mathbf{Z}_{n,1}, \\ h_2(\mathbf{S}_{n,2:4}) = \|\mathbf{S}_{n,2:4}\|_2 \leq \mathbf{Z}_{n,2}, \\ \mathbf{Z}_{n,1} + \mathbf{Z}_{n,2} \leq 0. \end{cases} \quad (7)$$

$$\begin{cases} h_2(\mathbf{S}_{n,2:4}) = \|\mathbf{S}_{n,2:4}\|_2 \leq \mathbf{Z}_{n,2}, \\ \mathbf{Z}_{n,1} + \mathbf{Z}_{n,2} \leq 0. \end{cases} \quad (8)$$

$$\mathbf{Z}_{n,1} + \mathbf{Z}_{n,2} \leq 0. \quad (9)$$

Then, problem (3) can be rewritten as follows

$$\begin{aligned} \underset{\mathbf{S} \in \mathbb{R}^{N \times 4}, \mathbf{Z} \in \mathbb{R}^{N \times 2}}{\text{minimize}} \quad & f(\Phi(\mathbf{S})) + \iota_{\mathbb{U}}(\mathbf{S}) + \gamma g(\Psi^\dagger \mathbf{S}) + \iota_{\mathbb{V}}(\mathbf{Z}) \\ & + \iota_{\mathbb{E}_1}(\mathbf{S}_{:,1}, \mathbf{Z}_{:,1}) + \iota_{\mathbb{E}_2}(\mathbf{S}_{:,2:4}, \mathbf{Z}_{:,2}), \end{aligned} \quad (10)$$

where $\mathbb{V} = \{\mathbf{Z} \in \mathbb{R}^{N \times 2} \mid (\forall n \in \{1, \dots, N\}) \mathbf{Z}_{n,1} + \mathbf{Z}_{n,2} \leq 0\}$ is defined to enforce the constraint (9), and \mathbb{E}_1 and \mathbb{E}_2 are the epigraphs² of functions h_1 and h_2 , respectively, used to impose the constraints (7) and (8) (for details, see [13]).

B. Proposed algorithm

We propose to solve problem (10) using the primal-dual forward-backward algorithm developed in [21]–[23]. We refer the reader to [24], [25] for further details on the primal-dual framework. The proposed algorithm is described in Algorithm 1, where following four main steps can be identified.

1) *Data constraint*: The data constraint is enforced in steps 5-16. In these steps, for each data block $j \in \{1, \dots, d\}$, a preconditioning diagonal operator $\mathbf{U}_j \in \mathbb{R}_+^{M_j \times M_j}$ is incorporated [23]. This operator is used to accelerate the convergence of the algorithm by utilizing additional prior information about the data (i.e. the Fourier sampling density), and is built as described in [16]. The ℓ_2 data constraints are handled in step 7, by performing projections onto the ellipsoids $\mathcal{E}_j(\mathbf{Y}_j, \epsilon_j^{(k)}) = \{\mathbf{A} \in \mathbb{C}^{M_j \times 4} : \|\mathbf{U}_j^{-1/2} \mathbf{A} - \mathbf{Y}_j\|_2 \leq \epsilon_j\}$. The projection of $\tilde{\mathbf{A}}$ onto a closed convex non-empty set \mathcal{C} is defined by $\mathcal{P}_{\mathcal{C}}(\tilde{\mathbf{A}}) = \arg\min_{\mathbf{A} \in \mathcal{C}} \|\mathbf{A} - \tilde{\mathbf{A}}\|_2^2$.

Note that the projection onto the ellipsoids $\mathcal{E}_j(\mathbf{Y}_j, \epsilon_j^{(k)})$ necessitates to know the bounds ϵ_j for each data block. However, in practice, these bounds not only incorporate the thermal noise, but also the calibration errors which are unknown. In this context, we showcase an adaptive scheme to estimate these bounds within Algorithm 1, described in steps 9-15. For every $j \in \{1, \dots, d\}$, ϵ_j is updated as a weighted mean of the current value of this bound and the ℓ_2 norm of the corresponding residual data block, provided technical conditions are met [17].

2) *Sparsity regularization*: The sparsity is promoted in steps 17-19, via soft-thresholding operation [26], [27]. Let $\mathbf{w} \in \mathbb{R}^N$ be a threshold level. The soft-thresholding operator $\mathcal{T}_{\mathbf{w}}$, applied to $\mathbf{b} \in \mathbb{R}^N$, is defined by $\mathcal{T}_{\mathbf{w}}(\mathbf{b}) = (p_n)_{1 \leq n \leq N}$ where, for every $n \in \{1, \dots, N\}$,

$$p_n = \begin{cases} -b_n + w_n & \text{if } b_n < -w_n, \\ 0 & \text{if } -w_n \leq b_n \leq w_n, \\ b_n - w_n & \text{otherwise.} \end{cases} \quad (11)$$

²The epigraph of a proper, lower semi-continuous function $\psi: \mathbb{R}^N \rightarrow]-\infty, +\infty]$ corresponds to $\text{epi } \psi = \{(z, \gamma) \in \mathbb{R}^N \times \mathbb{R} \mid \psi(z) \leq \gamma\}$.

Algorithm 1 Distributed algorithm for joint Stokes imaging

1: **Let** $\mathbf{S}^{(0)} \in \mathbb{R}^{N \times 4}$, $\mathbf{Z}^{(0)} \in \mathbb{R}^{N \times 2}$, $(\forall j \in \{1, \dots, d\}) \mathbf{A}_j^{(0)} \in \mathbb{R}^{M_j \times 4}$, $(\forall l \in \{1, \dots, 9\}) \mathbf{B}_l^{(0)} \in \mathbb{R}^{N \times 4}$, $\mathbf{C}^{(0)} \in \mathbb{R}^{N \times 4}$, $\mathbf{D}^{(0)} \in \mathbb{R}^{N \times 2}$, $(\rho_1, \rho_2, \rho_3, \tau) \in \mathbb{R}_+^4$ such that $\rho_1 \|\Phi\|_S^2 + \rho_2 \|\Psi\|_S^2 + \rho_3 < \tau^{-1}$, $\lambda \in]0, 1[$, $T \in \mathbb{N}$, and $(\epsilon_1, \epsilon_2) \in \mathbb{R}_+^2$.

2: **For** $k = 0, 1, \dots$
 Stokes matrix update:

3: $\mathbf{S}^{(k+1)} = \mathcal{P}_{\mathbb{U}}\left(\mathbf{S}^{(k)} - \tau(\rho_1 \sum_{j=1}^d \Phi_j^\dagger(\mathbf{A}_j^{(k)}) + \rho_2 \sum_{l=1}^9 \mathbf{B}_l^{(k)} + \rho_3 \mathbf{C}^{(k)})\right)$

4: $\tilde{\mathbf{S}}^{(k+1)} = 2\mathbf{S}^{(k+1)} - \mathbf{S}^{(k)}$

 Enforcing data constraint:

5: **for** $j \in \{1, \dots, d\}$ **do in parallel**

6: $\tilde{\mathbf{A}}_j^{(k)} = \mathbf{A}_j^{(k)} + \mathbf{U}_j \Phi_j(\tilde{\mathbf{S}}_{:,i}^{(k+1)})$

7: $\mathbf{A}_j^{(k+1)} = \mathbf{U}_j^{1/2} \left(I - \mathcal{P}_{\mathcal{E}_j(\mathbf{Y}_j, \epsilon_j^{(k)})} \right) \left(\mathbf{U}_j^{-1/2} \tilde{\mathbf{A}}_j^{(k)} \right)$

ℓ_2 bound update

8: $\mu_j^{(k+1)} = \|\mathbf{Y}_j - \Phi_j(\mathbf{S}_j^{(k+1)})\|$

9: **if** $\left(\frac{\|\mathbf{S}^{(k+1)} - \mathbf{S}^{(k)}\|_2}{\|\mathbf{S}^{(k+1)}\|_2} < \epsilon_1 \right) \& \left(k - t_j^{(k)} \geq T \right) \& \left(\frac{\|\mu_j^{(k+1)} - \epsilon_j^{(k)}\|}{\epsilon_j^{(k)}} > \epsilon_2 \right)$

10: $\epsilon_j^{(k+1)} = \lambda \mu_j^{(k+1)} + (1 - \lambda) \epsilon_j^{(k)}$

11: $t_j^{(k+1)} = k$

12: **else**

13: $\epsilon_j^{(k+1)} = \epsilon_j^{(k)}$

14: $t_j^{(k+1)} = t_j^{(k)}$

15: **end if**

16: **end for**

 Promoting sparsity:

17: **for** $l \in \{1, \dots, 9\}$ and $\forall i \in \{1, 2, 3, 4\}$ **do in parallel**

18: $[\mathbf{B}_l^{(k+1)}]_{:,i} = \Psi_l \left(I - \mathcal{T}_{\gamma[\mathbf{w}_l]_{:,i}} \right) \left([\mathbf{B}_l^{(k)}]_{:,i} + \Psi_l^\dagger \tilde{\mathbf{S}}_{:,i}^{(k+1)} \right)$

19: **end for**

 Enforcing polarization constraint:

20: $\mathbf{Z}^{(k+1)} = \mathcal{P}_{\mathbb{V}}\left(\mathbf{Z}^{(k)} - \tau \mathbf{D}^{(k)}\right)$

21: $\tilde{\mathbf{Z}}^{(k+1)} = 2\mathbf{Z}^{(k+1)} - \mathbf{Z}^{(k)}$

22: $\begin{bmatrix} \mathbf{C}_{:,1}^{(k+1)} \\ \mathbf{D}_{:,1}^{(k+1)} \end{bmatrix} = \left(I - \mathcal{P}_{\mathbb{E}_1} \right) \left(\begin{bmatrix} \mathbf{C}_{:,1}^{(k)} + \tilde{\mathbf{S}}_{:,1}^{(k+1)} \\ \mathbf{D}_{:,1}^{(k)} + \tilde{\mathbf{Z}}_{:,1}^{(k+1)} \end{bmatrix} \right)$

23: $\begin{bmatrix} \mathbf{C}_{:,2:4}^{(k+1)} \\ \mathbf{D}_{:,2}^{(k+1)} \end{bmatrix} = \left(I - \mathcal{P}_{\mathbb{E}_2} \right) \left(\begin{bmatrix} \mathbf{C}_{:,2:4}^{(k)} + \tilde{\mathbf{S}}_{:,2:4}^{(k+1)} \\ \mathbf{D}_{:,2}^{(k)} + \tilde{\mathbf{Z}}_{:,2}^{(k+1)} \end{bmatrix} \right)$

24: **end For**

3) *Polarization constraint*: Steps 20-23 are used to enforce the polarization constraint. This constraint is handled by performing three projections, in step 20 for the set \mathbb{V} , in step 22 for the set \mathbb{E}_1 , and in step 23 for the set \mathbb{E}_2 .

4) *Stokes matrix update*: Steps 3-4 are dedicated to update the Stokes parameters. Step 3 consists of a consensus step, aggregating the information obtained from the main steps described in Sections III-B1 to III-B3, followed by a projection onto the set \mathbb{U} .

IV. SIMULATIONS AND RESULTS

We apply our method on the VLA observations of the radio galaxy Cygnus A, centered at the frequency 8.422 GHz (X band). The dataset corresponds to the B configuration of the VLA, with size $M = 3 \times 10^5$. We implement Algorithm 1 in

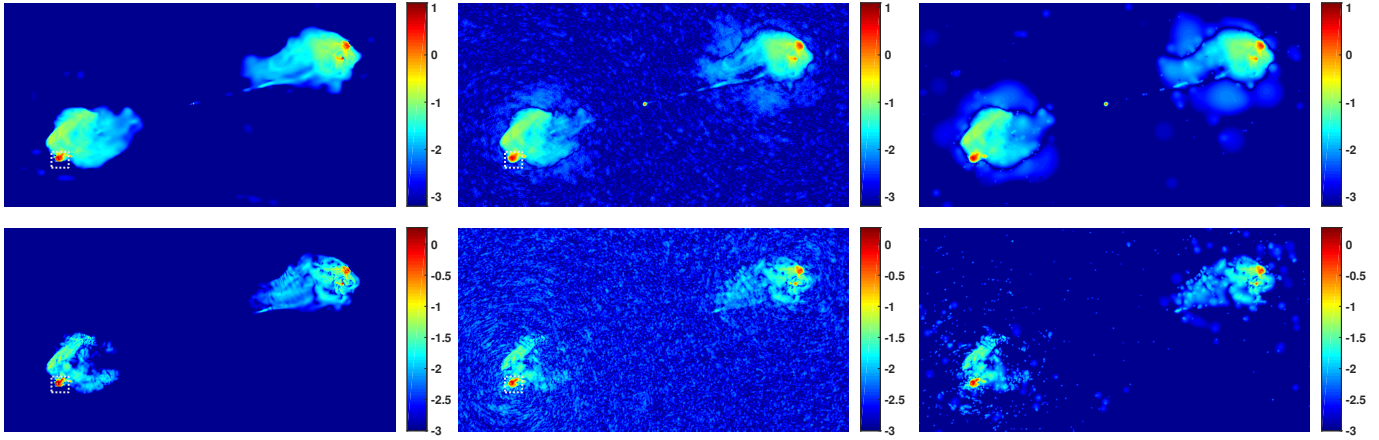


Fig. 1. Results obtained for Stokes I (first row) and linear polarization intensity ($|P|$) images, displayed in log scale. Column-wise, from left to right: Results from our method; MS-CLEAN restored images using Briggs weighting; and CLEAN beam convolved model images of MS-CLEAN using Briggs weighting.

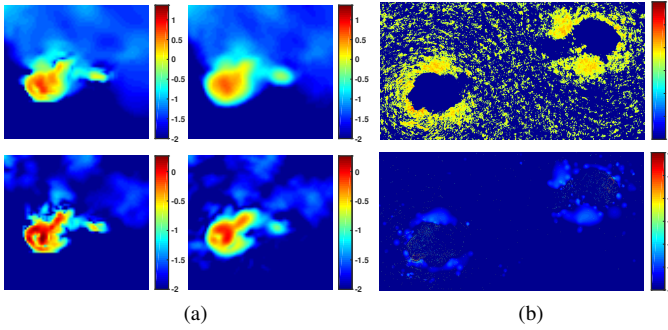


Fig. 2. (a) Zoomed west jet hotspot in Stokes I (first row) and linear polarization intensity (second row) images from our method's reconstructions (first column) and MS-CLEAN restored images (second column). (b) Polarization error images showing the pixels in the image ($|P| - I$) violating the polarization constraint, corresponding to MS-CLEAN restored (first row) and model (second row) images. All the images are shown in log scale.

MATLAB to image the Stokes parameters I , Q and U (Stokes V being of negligible intensity), each of size 1024×1024 pixels with the corresponding pixel size $\delta l = 0.16''$. It corresponds to recovering the images at 1.5 times the nominal resolution, i.e. the band-limit of the observations. We split the dataset into three blocks and perform 15 reweighting iterations in total, where each iteration is stopped when the relative variation between the consecutively estimated Stokes images is less than 10^{-5} and the percentage of pixels (in the estimated images) not satisfying the polarization constraint is less than 0.1.

We compare our method with the RI imaging algorithm MS-CLEAN with Briggs weighting, implemented using the WSCLEAN program [28]. For visual comparison, Fig. 1 shows the reconstructed images for Stokes I (first row) and the linear polarization intensity image $|P|$ (second row), generated from the reconstructed Stokes Q and U images. We scale recovered images from our method with the flux of the CLEAN beam, and since the MS-CLEAN restored images are dominated by the noise, we also show the beam convolved model images. Note that the shown images are cropped to the size 512×1024 pixels. Furthermore, in Fig. 2(a), we display the zoomed west

jet hotspot (marked with a white box in the associated full images in Fig. 1). The results indicate the super-resolution capability of our method, capturing the finer details in the source structure, in contrast with MS-CLEAN which provides images at the nominal resolution by convolving with the CLEAN beam. Moreover, the model images from MS-CLEAN have around 15% of the pixels with non-physical values, violating the polarization constraint. This holds true for the restored images as well, primarily due to the added residual, having around 28% of such pixels, as illustrated in Fig. 2(b). It is in contrast to our method which explicitly enforces the polarization constraint and thus, is bound to produce images with physical values.

Additionally, comparing our method's reconstruction with the MS-CLEAN model image, both smoothed at the nominal resolution, gives a similarity of 29.4 dB for Stokes I and 19.3 dB for linear polarization image, implying that the reconstructions from the two methods, especially for Stokes I , are in agreement in terms of their low spatial frequency content.

V. CONCLUSION

We have presented a scalable version of the Polarized SARA method for Stokes imaging which promotes sparsity of the sought images while imposing the polarization constraint. The scalability of the method is achieved by employing distributed and parallel processing coupled with a preconditioning strategy. Application on a real-dataset shows the superiority of our method over MS-CLEAN, both in terms of encapsulating the source structure details going beyond the nominal resolution, and in producing physical images satisfying the polarization constraint. While in the current work, the adaptive scheme incorporating unmodelled calibration terms with the noise within the ℓ_2 bounds is employed, we plan to take into account the calibration terms in our model for future work [29], [30].

ACKNOWLEDGMENT

This work was supported by EPSRC, grants EP/M008843/1 and EP/M019306/1. The authors would like to thank R. Perley from NRAO for providing calibrated VLA data.

REFERENCES

- [1] M. Brentjens and A. De Bruyn, “Faraday rotation measure synthesis,” *A&A*, vol. 441, pp. 1217–1228, 2005.
- [2] J. A. Högbom, “Aperture synthesis with a non-regular distribution of interferometer baselines,” *A&A*, vol. 15, pp. 417–426, 1974.
- [3] T. J. Cornwell, “Multiscale CLEAN deconvolution of radio synthesis images,” *IEEE J. Sel. Top. Sig. Process.*, vol. 2, no. 5, pp. 793–801, 2008.
- [4] Y. Wiaux, L. Jacques, G. Puy, A. Scaife, and P. Vanderghenst, “Compressed sensing imaging techniques for radio interferometry,” *MNRAS*, vol. 395, pp. 1733–1742, 2009.
- [5] R. Carrillo, J. McEwen, and Y. Wiaux, “PURIFY: a new approach to radio-interferometric imaging,” *MNRAS*, vol. 439, no. 4, pp. 3591–3604, 2014.
- [6] H. Garsden, J. Girard, J. Starck, S. Corbel, C. Tasse, A. Woiselle, J. McKean, A. Van Amesfoort, J. Anderson, I. Avruch, and R. Beck, “Lofar sparse image reconstruction,” *A&A*, vol. 575, p. A90, 2015.
- [7] A. Onose, R. Carrillo, A. Repetti, J. McEwen, J. Thiran, J.-C. Pesquet, and Y. Wiaux, “Scalable splitting algorithms for big-data interferometric imaging in the SKA era,” *MNRAS*, vol. 462, pp. 4314–4335, 2016.
- [8] K. Akiyama, S. Ikeda, M. Pleau, V. Fish, F. Tazaki, K. Kuramochi, A. Broderick, J. Dexter, M. Mościbrodzka, M. Gowanlock, M. Honma, and S. S. Doeleman, “Superresolution Full-polarimetric Imaging for Radio Interferometry with Sparse Modeling.”
- [9] R. Narayan and R. Nityananda, “Maximum entropy image restoration in astronomy,” *Ann. Rev. Astron. Astrophys.*, vol. 24, pp. 127–170, 1986.
- [10] M. Holdaway and J. Wardle, “Maximum entropy imaging of polarization in very long baseline interferometry,” *Digital Image Synthesis and Inverse Optics*, vol. 1351, pp. 714–724, 1990.
- [11] C. Coughlan and D. Gabuzda, “High resolution VLBI polarization imaging of AGN with the maximum entropy method,” *MNRAS*, vol. 463, pp. 1980–2001, 2016.
- [12] A. Chael, M. Johnson, R. Narayan, S. Doeleman, J. Wardle, and K. Bouman, “High-resolution Linear Polarimetric Imaging for the Event Horizon Telescope,” *AJ*, vol. 829, no. 1, p. 11, 2016.
- [13] J. Birdi, A. Repetti, and Y. Wiaux, “Sparse interferometric Stokes imaging under polarization constraint (Polarized SARA).” *ArXiv e-prints*, 2018. [Online]. Available: <https://arxiv.org/abs/1801.02417>
- [14] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, “Epigraphical projection and proximal tools for solving constrained convex optimization problems,” *Sig. Image Vid. Process.*, vol. 9, no. 8, pp. 1737–1749, 2015.
- [15] R. Carrillo, J. McEwen, D. Van De Ville, J.-P. Thiran, and Y. Wiaux, “Sparsity averaging for compressive imaging,” *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 591–594, 2013.
- [16] A. Onose, A. Dabbech, and Y. Wiaux, “An accelerated splitting algorithm for radio-interferometric imaging: when natural and uniform weighting meet,” *MNRAS*, vol. 469, no. 1, p. 938, 2017.
- [17] A. Dabbech, A. Onose, A. Abdulaziz, R. A. Perley, O. M. Smirnov, and Y. Wiaux, “Cygnus A super-resolved via convex optimisation from VLA data,” *MNRAS*, 2018. [Online]. Available: <http://dx.doi.org/10.1093/mnras/sty372>
- [18] O. Smirnov, “Revisiting the radio interferometer measurement equation-I. A full-sky jones formalism,” *A&A*, vol. 527, p. A106, 2011.
- [19] R. Carrillo, J. McEwen, and Y. Wiaux, “Sparsity Averaging Reweighted Analysis (SARA): A novel algorithm for radio-interferometric imaging,” *MNRAS*, vol. 426, no. 2, pp. 1223–1234, 2012.
- [20] E. Candès, M. Wakin, and S. Boyd, “Enhancing Sparsity by Reweighted ℓ_1 Minimization,” *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, 2008.
- [21] L. Condat, “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms,” *J. Opt. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.
- [22] B. Vũ, “A splitting algorithm for dual monotone inclusions involving cocoercive operators,” *Adv. Comp. Math.*, vol. 38, no. 3, pp. 667–681, 2013.
- [23] J.-C. Pesquet and A. Repetti, “A class of randomized primal-dual algorithms for distributed optimization,” *J. Nonlinear Convex Anal.*, vol. 16, no. 12, pp. 2453–2490, 2015.
- [24] N. Komodakis and J.-C. Pesquet, “Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems,” *IEEE Sig. Proc. Mag.*, vol. 32, no. 6, pp. 31–54, 2015.
- [25] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer-Verlag New York, 2011.
- [26] P. L. Combettes and J.-C. Pesquet, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, New York, 2011.
- [27] C. Chau, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, “A variational formulation for frame-based inverse problems,” *Inverse Problems*, vol. 23, no. 4, pp. 1495–1518, 2007.
- [28] A. Offringa, B. McKinley, N. Hurley-Walker, F. Briggs, R. Wayth, D. Kaplan, M. Bell, L. Feng, A. Neben, J. Hughes, and J. Rhee, “WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy,” *MNRAS*, vol. 444, pp. 606–619, 2014.
- [29] A. Repetti, J. Birdi, A. Dabbech, and Y. Wiaux, “Non-convex optimization for self-calibration of direction-dependent effects in radio interferometric imaging,” *MNRAS*, vol. 470, no. 4, pp. 3981–4006, 2017.
- [30] A. Repetti and Y. Wiaux, “A non-convex perspective on calibration and imaging in radio interferometry,” in *Wavelets and Sparsity XVII*, vol. 10394, San Diego, United States, 6-9 August 2017, p. 103941W.