Time-Regularized Blind Deconvolution Approach for Radio Interferometry

Pierre-Antoine Thouvenin, Audrey Repetti, Arwa Dabbech, and Yves Wiaux
Institute of Sensors, Signals and Systems, Heriot-Watt University,
Edinburgh EH14 4AS, United Kingdom
Email: {P.Thouvenin, A.Repetti, A.Dabbech, Y.Wiaux}@hw.ac.uk

***Abstract*—Radio-interferometric imaging aims to estimate a sky intensity image from degraded undersampled Fourier measurements. At the dynamic range of interest to modern radio telescopes, the image reconstruction quality will be limited by the unknown time-dependent calibration kernels. Hence the need of performing joint image reconstruction and calibration, and consequently of solving a non-convex blind deconvolution problem. Extending our recent work where the calibration kernels are assumed to be smooth in space, we further assume in this work that the calibration kernels are smooth in time. In addition, an average sparsity prior is used for the estimation of the image of interest. The resulting high dimensional non-convex non-smooth minimization problem is then solved by leveraging an alternating forward-backward algorithm which benefits from well-established convergence guarantees. Our results show that time-regularization is effective in enhancing imaging quality.**

***Index Terms*—Radio-interferometric imaging, calibration, alternating forward-backward approach**

I. Introduction

Radio interferometry is a technique allowing for probing the radio sky at high sensitivity and angular resolution via a collection of antennas. Each antenna pair gives access to a radio-interferometric (RI) measurement. Assuming a co-planar and perfectly calibrated instrument, RI data reduce to noisy Fourier components of the sky intensity [1]. However, each RI measurement corresponds in practice to a noisy Fourier component of the sky intensity modulated with unknown time-variable complex antenna gains. These encompass time-variations of the antenna beam, ionospheric phase shifts, antenna pointing errors. When it comes to the traditional instruments, the effects of these modulations are buried within the thermal noise, thus no calibration is required. In this context, recovering the unknown image of the radio sky is an inverse problem which has been extensively studied in the literature. State-of-the-art approaches leverage compressive sensing and optimization theories [2]–[12]. This is not the case of the next-generation instruments, e.g. the upgraded VLA [13] and the upcoming SKA [14], which are characterized by an extreme sensitivity and sub-arcsec angular resolution. Hence, calibration of the antenna gains is necessary to obtain maps of the radio sky meeting these instruments' capabilities.

Antenna gains are typically classified into two categories. (i) Direction dependent effects (DDEs); these are spatially-variable complex-valued modulations in the image domain, equivalently convolutional kernels in the Fourier space. (ii) Direction independent effects (DIEs); these are constant complex modulations in the image domain, equivalently complex scalar multipliers in the Fourier domain. Traditionally, the calibration process is limited to the estimation of DIEs. The general approach consists in alternating between a DIEs calibration step and an imaging step [15], [16]. However, DIEs modeling errors and lack of DDEs models result in sub-optimal image recovery in terms of dynamic range. Novel calibration approaches aiming to estimate DDEs have been recently devised. [17] proposes a parametric model for the ionospheric phase errors, whereas [18] proposes the use of non-linear Kalman filters. More recently, a faceting-based framework has been adopted in [19]–[21], that is based on the assumption of piece-wise constant DDEs across the field of view. All these methods consist of alternating calibration and imaging severally. In our recent works, we have proposed a new framework for RI, where we estimate jointly the unknown DDEs and the image of interest [22]–[25]. The approach leverages compressive sensing and non-convex optimization theories and is shipped with convergence guarantees. Data fidelity is imposed by minimizing a regularized non-linear least squares (LS) criterion with respect to both the image and the DDEs. Furthermore, a spatial smoothness prior is adopted for the DDEs, on the one hand. Positivity and sparsity of the unknown image in a given dictionary are promoted, on the other hand. In this work, we extend our approach to incorporate a time prior on the unknown DDEs.

In the remainder of the article, we present the RI imaging problem in Section II. In Section III, we first describe our proposed minimization task for joint calibration and imaging. Then, we explain the alternating forward-backward algorithmic structure to solve it. A description of the realistic simulations setting and a discussion of the results are provided in Section IV. Conclusions are finally reported in Section V.

II. Problem description

The imaging problem for RI can be formulated as an inverse problem, where the objective consists in finding an estimate $\boldsymbol{x}^\star \in \mathbb{R}^N$ of an unknown radio-sky image $\overline{\boldsymbol{x}} \in \mathbb{R}^N$ from complex observations, namely the visibilities $\boldsymbol{y} \in \mathbb{C}^M$. The visibilities are determined by the relative position between each antenna pair indexed by $(\alpha, \beta) \in \{1, \dots, n_a\}^2$, with $\alpha < \beta$, for different time instants $t \in \{1, \dots, T\}$. In total, an interferometer

This work was supported by EPSRC, grants EP/M011089/1, EP/M008843/1 and EP/M019306/1.

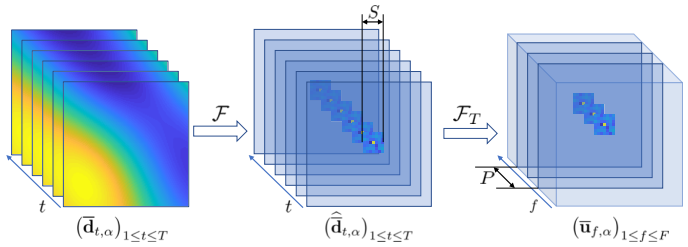

Fig. 1: Illustration of the proposed smoothness prior: the DDEs are represented by compact-support kernels in both the spatial and the temporal Fourier domains ($\mathcal{F}$ denotes a 2D Fourier transform along the spatial dimensions, whereas $\mathcal{F}_T$ represents a 1D Fourier transform along the temporal dimension).

acquires $M = Tn_a(n_a-1)/2$ different measurements in the Fourier domain of the image of interest. Formally, each visibility $y_{t,\alpha,\beta} \in \mathbb{C}$ measured by the antenna pair (α,β) at instant t at the spatial frequency $k_{t,\alpha,\beta}$ can be written as

$$y_{t,\alpha,\beta} = \sum_{n=-N/2}^{N/2-1} \overline{d}_{t,\alpha}(n)\overline{d}_{t,\beta}(n)^*\overline{x}(n)e^{-2\mathrm{i}\pi k_{t,\alpha,\beta}\frac{n}{N}}, \quad (1)$$

where $\overline{\boldsymbol{d}}_{t,\alpha} = (\overline{d}_{t,\alpha}(n))_{-N/2\leq n\leq N/2-1} \in \mathbb{C}^N$ represents the DDEs related to antenna α at instant t. Note that, using these notations, the DIEs can be seen as a special case of the DDEs where $\overline{\boldsymbol{d}}_{t,\alpha} = \delta_{t,\alpha}\mathbf{1}_N$, with $\delta_{t,\alpha} \in \mathbb{C}$ and $\mathbf{1}_N$ being the unitary vector of dimension N. The objective of this work is to estimate simultaneously the unknown image of the sky and the DDE images. To this aim, we assume that calibration transfer has been performed, hence that the zero spatial frequency coefficients of the DDEs (i.e. the DIEs) in (1) are normalized to 1. As described in [24], this prior information can be used to obtain a first estimate of $\overline{\boldsymbol{x}}$, denoted by $\boldsymbol{x}_0^\star$, using any RI imaging method. In this context, the original unknown image $\overline{\boldsymbol{x}}$ can be expressed as the sum of two images, i.e. $\overline{\boldsymbol{x}} = \boldsymbol{x}_0 + \overline{\boldsymbol{\epsilon}}$, where $\boldsymbol{x}_0$ contains the brightest coefficients of $\boldsymbol{x}_0^\star$ and $\overline{\boldsymbol{\epsilon}}$ represents the remaining low-amplitude sources that must be estimated simultaneously with the DDEs.

Problem (1) can be reformulated as follows

$$\boldsymbol{y} = \check{\Phi}\Big((\overline{\boldsymbol{d}}_{t,\alpha})_{1\leq t\leq T,\,1\leq\alpha\leq n_a}, \overline{\boldsymbol{\epsilon}}\Big) + \boldsymbol{w}, \quad (2)$$

where $\boldsymbol{w} \in \mathbb{C}^M$ is a realization of an additive i.i.d. random Gaussian noise, and $\Phi\colon (\mathbb{C}^N)^{n_aT}\times\mathbb{R}^N \to \mathbb{C}^M$ is the measurement operator described in (1). Note that this measurement operator is linear in $\overline{\boldsymbol{\epsilon}}$, but is not linear with respect to the DDEs. In [25], we have proposed to incorporate an auxiliary variable to obtain a tri-linear inverse problem. In this context, (2) can be rewritten as follows

$$\boldsymbol{y} = \widetilde{\Phi}\Big((\overline{\boldsymbol{d}}_{1,t,\alpha}, \overline{\boldsymbol{d}}_{2,t,\alpha})_{1\leq t\leq T,\,1\leq\alpha\leq n_a}, \overline{\boldsymbol{\epsilon}}\Big) + \boldsymbol{w}, \quad (3)$$

where, for every $t \in \{1,\dots,T\}$ and $\alpha \in \{1,\dots,n_a\}$, $\overline{\boldsymbol{d}}_{1,t,\alpha} = \overline{\boldsymbol{d}}_{2,t,\alpha} = \overline{\boldsymbol{d}}_{t,\alpha}$.

III. Proposed approach

A. Minimization problem

In this work, we assume that the DDEs are smooth functions of both sky and time. In this context, for every $\alpha \in \{1,\dots,n_a\}$ and $i \in \{1,2\}$, we propose to represent the DDEs $(\overline{\boldsymbol{d}}_{i,t,\alpha})_{1\leq t\leq T}$ as compact-support kernels $\overline{\boldsymbol{u}}_{i,\alpha} \in \mathbb{C}^{S^2\times P}$ in both the spatial and temporal Fourier domains, where S is the size of the square kernels in the spatial frequency domain, and P represents their size in the temporal frequency domain. An illustration is provided in Fig. 1. Let $\mathcal{Z}\colon \mathbb{C}^{S^2\times P} \to \mathbb{C}^{N\times T}$ be the linear operator defined such that $\mathcal{Z}(\overline{\boldsymbol{u}}_{i,\alpha}) = (\overline{\boldsymbol{d}}_{i,t,\alpha})_{1\leq t\leq T}$, and let $\Phi = \widetilde{\Phi}\big(\mathcal{Z}(.),\mathcal{Z}(.),.\big)$. Then we have

$$\boldsymbol{y} = \Phi(\boldsymbol{U}_1, \boldsymbol{U}_2, \overline{\boldsymbol{\epsilon}}) + \boldsymbol{w}, \quad (4)$$

with, for every $i \in \{1,2\}$, $\boldsymbol{U}_i = (\boldsymbol{u}_{i,\alpha})_{1\leq\alpha\leq n_a} \in \mathbb{C}^{S^2\times P\times n_a}$. Using this formulation, we propose to define the estimate of the sky image and the DDEs as a solution to

$$\underset{\boldsymbol{\epsilon},\boldsymbol{U}_1,\boldsymbol{U}_2}{\text{minimize}}\ h(\boldsymbol{\epsilon},\boldsymbol{U}_1,\boldsymbol{U}_2) + r(\boldsymbol{\epsilon}) + p(\boldsymbol{U}_1,\boldsymbol{U}_2), \quad (5)$$

where h is the data fidelity term associated with model (4), whereas r and p are regularization terms incorporating prior information on the target objects. Assuming the noise is white and Gaussian, we choose to use a least squares data criterion for the data fidelity term, given by

$$h(\boldsymbol{\epsilon},\boldsymbol{U}_1,\boldsymbol{U}_2) = \frac{1}{2}\|\Phi(\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{\epsilon}) - \boldsymbol{y}\|_2^2. \quad (6)$$

Concerning the regularization term for the image, we propose to use a hybrid function, expressed as follows:

$$r(\boldsymbol{\epsilon}) = \lambda\|\boldsymbol{\Psi}^\dagger(\boldsymbol{x}_0+\boldsymbol{\epsilon})\|_1 + \iota_{\mathbb{E}}(\boldsymbol{\epsilon}) \quad (7)$$

where $\lambda > 0$ is a regularization parameter, $\|.\|_1$ denotes the ℓ_1 norm, $\boldsymbol{\Psi} \in \mathbb{R}^{N\times L}$ is the sparsity average basis introduced in [6], and $\iota_{\mathbb{E}}$ is the indicator function of the set $\mathbb{E}$ defined as

$$\mathbb{E} = \big\{\boldsymbol{\epsilon}\in\mathbb{R}^N \,|(\forall n\in\mathbb{S}_0)\ -\vartheta x_0(n) \leq \epsilon(n) \leq \vartheta x_0(n), \\ (\forall n\in\mathbb{S}_0^c)\ 0 \leq \epsilon(n)\big\}, \quad (8)$$

$\mathbb{S}_0$ being the support of $\boldsymbol{x}_0$ and $\vartheta \in [0,1]$ representing the percentage of error assumed on $\boldsymbol{x}_0$. For every $\boldsymbol{\epsilon} \in \mathbb{R}^N$, the indicator function $\iota_{\mathbb{R}}(\boldsymbol{\epsilon})$ is equal to 0 if $\boldsymbol{\epsilon} \in \mathbb{E}$, and to $+\infty$ otherwise. For the DDE regularization, we define the function p as

$$p(\boldsymbol{U}_1,\boldsymbol{U}_2) = \frac{\eta}{2}\|\boldsymbol{U}_1 - \boldsymbol{U}_2\|_2^2 + \iota_{\mathbb{D}}(\boldsymbol{U}_1) + \iota_{\mathbb{D}}(\boldsymbol{U}_2), \quad (9)$$

where $\eta > 0$ is a regularization parameter, the ℓ_2-norm term controls the distance between $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$, and the set $\mathbb{D}$ defines a subset of $\mathbb{C}^{S^2\times P\times n_a}$. More precisely, $\mathbb{D}$ is defined such that the DIEs belong to an ℓ_∞ complex ball centered in 1 with radius $\nu > 0$, whereas the other DDE coefficients are contained in an ℓ_∞ complex ball centered in 0 with radius ν.

B. Alternating forward-backward algorithm

We propose to solve problem (5) using an alternating forward-backward algorithm [26]–[31] described in Algorithm 1, where $\cdot$ denotes the Hadamard product and, for every $k \in \mathbb{N}$, $\tau^{(k)}$ is a positive constant, and $\Gamma_1^{(k)}$ and $\Gamma_2^{(k)}$ are positive matrices. At each iteration $k \in \mathbb{N}$ of the algorithm, the user decides to update either the DDEs $(\overline{\boldsymbol{U}}_1, \overline{\boldsymbol{U}}_2)$ (steps 5 to 16) or the image $\boldsymbol{\epsilon}$ (steps 17 to 24). In both cases, the updates are performed by computing a gradient step followed by a proximal step. The partial gradient of h with respect to $\boldsymbol{U}_1$ (resp. $\boldsymbol{U}_2$ and $\boldsymbol{\epsilon}$) is denoted by $\nabla_{\boldsymbol{U}_1} h$ (resp. $\nabla_{\boldsymbol{U}_2} h$ and $\nabla_{\boldsymbol{\epsilon}} h$). The proximal operator of a convex, proper, l.s.c function f at a point $\boldsymbol{x}$ is denoted by $\mathrm{prox}_f(\boldsymbol{x})$, and is defined as the unique minimizer of $f + \frac{1}{2}\|\boldsymbol{x} - .\|_2^2$. When $f = \iota_{\mathcal{C}}$, where $\mathcal{C}$ is a closed, convex non-empty set, then $\mathrm{prox}_{\iota_{\mathcal{C}}}$ reduces to the usual Euclidean projection onto $\mathcal{C}$, denoted by $\Pi_{\mathcal{C}}$.

Let $(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}^{(k)}, \boldsymbol{U}^{(k)})_{k\in\mathbb{N}}$ be a sequence generated by Algorithm 1. We assume that the blocks $(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}^{(k)}, \boldsymbol{U}^{(k)})_{k\in\mathbb{N}}$ are updated following an essentially cyclic rule, i.e. that each variable is updated at least once within a given finite number of iterations. In addition, for every $k \in \mathbb{N}$,

$$0 < \tau^{(k)} < 1/\|\Phi(., \boldsymbol{U}_1^{(k+1)}, \boldsymbol{U}_2^{(k+1)})\|_S^2, \tag{10}$$

where $\|.\|_S$ denotes the spectral norm, and

$$(\forall i \in \{1,2\}) \quad \Gamma_i^{(k)} = \left[\gamma_{i,1,1}^{(k)} \mathbf{1}_S | \ldots | \gamma_{i,T,n_a}^{(k)} \mathbf{1}_S\right]^\top, \tag{11}$$

where, for every $\alpha \in \{1, \ldots, n_a\}$, $0 < \gamma_{i,t,\alpha}^{(k)} < 1/(\eta + \zeta_{i,t,\alpha}^{(k)})$, with $\zeta_{1,t,\alpha}^{(k)} = \|\Phi(\boldsymbol{\epsilon}^{(k)}, ., \boldsymbol{U}_2^{(k)})\|_S^2$ and $\zeta_{1,t,\alpha}^{(k)} = \|\Phi(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}_1^{(k+1)}, .)\|_S^2$.

Under these assumptions, the sequence of iterates $\left(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}_1^{(k)}, \boldsymbol{U}_2^{(k)}\right)_{k\in\mathbb{N}}$ generated by Algorithm 1 converges to a critical point $\left(\boldsymbol{\epsilon}^\star, \boldsymbol{U}_1^\star, \boldsymbol{U}_2^\star\right)$ of the objective function minimized in problem (5). Moreover, the objective function value is decreasing along the iterations.

IV. Experiments on synthetic data

To assess the performance of the proposed approach, we adopt two 256×256 images of M31 and W28. Their associated data are simulated using the VLA antenna's configuration, where the total number of antennas is $n_a = 27$. We consider $T = 200$ snapshots. In the non-uniform Fourier transform, embedded in the measurement operator $\tilde{\Phi}$ (see (2)), we consider the Kaiser-Bessel interpolation kernels [32] of size 5×5. The DDE kernels have been randomly generated in the Fourier domain, with a spatial Fourier support of size $S = 5$ and a temporal Fourier support of size $P = 3$. Note that both S and P are assumed to be known in these experiments.

To comparatively appreciate the performance of the proposed approach, we consider the following methods.

(i) **Imaging with the ground truth DDEs**: the case where the DDEs are perfectly known represents the best configuration possible to estimate the image of interest, and can thus be taken as a reference. We consequently define the image $\widetilde{\boldsymbol{x}}$ as the solution to the following imaging problem

Algorithm 1 Alternating forward-backward algorithm

1: **Initialization:** Let $\boldsymbol{\epsilon}^{(0)} \in \mathrm{dom}\, r$ and $(\boldsymbol{U}_1^{(0)}, \boldsymbol{U}_2^{(0)}) \in \mathbb{D}^2$. Let, for every $k \in \mathbb{N}$, $(L^{(k)}, J^{(k)}) \in \mathbb{N}^2$.
2: **Iterations:**
3: **For** $k = 0, 1, \ldots$
4: Choose to update either the DDEs or the image.
5: **If the DDEs are updated:**
6: $(\boldsymbol{U}_1^{(k,0)}, \boldsymbol{U}_2^{(k,0)}) = (\boldsymbol{U}_1^{(k)}, \boldsymbol{U}_2^{(k)})$
7: **For** $\ell = 0, \ldots, L^{(k)} - 1$
8: $\boldsymbol{U}_1^{(k,\ell+1)} = \Pi_{\mathbb{D}}\Big(\boldsymbol{U}_1^{(k,\ell)} - \Gamma_1^{(k)} \cdot \nabla_{\boldsymbol{U}_1} h(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}_1^{(k,\ell)}, \boldsymbol{U}_2^{(k)}) - \Gamma_1^{(k)} \cdot \eta(\boldsymbol{U}_1^{(k,\ell)} - \boldsymbol{U}_2^{(k)})\Big)$
9: **end for**
10: $\boldsymbol{U}_1^{(k+1)} = \boldsymbol{U}_1^{(k,L^{(k)})}$
11: **For** $\ell = 0, \ldots, L^{(k)} - 1$
12: $\boldsymbol{U}_2^{(k,\ell+1)} = \Pi_{\mathbb{D}}\Big(\boldsymbol{U}_2^{(k,\ell)} - \Gamma_2^{(k)} \cdot \nabla_{\boldsymbol{U}_2} h(\boldsymbol{\epsilon}^{(k)}, \boldsymbol{U}_1^{(k+1)}, \boldsymbol{U}_2^{(k,\ell)}) - \Gamma_2^{(k)} \cdot \eta(\boldsymbol{U}_2^{(k,\ell)} - \boldsymbol{U}_1^{(k+1)})\Big)$
13: **end for**
14: $\boldsymbol{U}_2^{(k+1)} = \boldsymbol{U}_2^{(k,L^{(k)})}$
15: $\boldsymbol{\epsilon}^{(k+1)} = \boldsymbol{\epsilon}^{(k)}$
16: **end if**
17: **If the image is updated:**
18: $\boldsymbol{\epsilon}^{(k,0)} = \boldsymbol{\epsilon}^{(k)}$
19: **For** $j = 0, \ldots, J^{(k)} - 1$
20: $\boldsymbol{\epsilon}^{(k,j+1)} = \mathrm{prox}_{\tau^{(k)} r}\big(\boldsymbol{\epsilon}^{(k,j)} - \tau^{(k)} \nabla_{\boldsymbol{\epsilon}} h(\boldsymbol{\epsilon}^{(k,j)}, \boldsymbol{U}_1^{(k+1)}, \boldsymbol{U}_2^{(k+1)})\big)$
21: **end for**
22: $\boldsymbol{\epsilon}^{(i+1)} = \boldsymbol{\epsilon}^{(k,J^{(k)})}$
23: $(\boldsymbol{U}_1^{(k+1)}, \boldsymbol{U}_2^{(k+1)}) = (\boldsymbol{U}_1^{(k)}, \boldsymbol{U}_2^{(k)})$
24: **end if**
25: **end for**

$$\underset{\boldsymbol{x}\in\mathbb{R}^N}{\text{minimize}} \ \|\Phi\big((\overline{\boldsymbol{d}}_{1,t,\alpha}, \overline{\boldsymbol{d}}_{2,t,\alpha})_{t,\alpha}, \boldsymbol{x}\big) - \boldsymbol{y}\|_2^2 + \tilde{r}(\boldsymbol{x}) \tag{12}$$

with $\tilde{r}(\boldsymbol{x}) = \lambda\|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1 + \iota_{[0,+\infty[^N}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^N$.

(ii) **Imaging with normalized DIEs**: in practice, we can assume that calibration transfer has been performed on the data prior to the imaging step, thus leading to normalized DIEs. In this case, the image can be estimated by setting the DIEs equal to 1 and solving an imaging problem similar to (12), leading to an estimate $\check{\boldsymbol{x}}$.

(iii) **Joint DIE calibration and imaging**: this approach consists in applying the joint calibration and imaging method introduced in [24], [25] with $S = 1$. The reference image $\boldsymbol{x}_0$ is obtained by thresholding the image $\check{\boldsymbol{x}}$ obtained in (ii) with a threshold set to $0.05 \max_n\big(\check{\boldsymbol{x}}(n)\big)$, and $\boldsymbol{\epsilon}$ is initialized with the value $\mathbf{0}$.

(iv) **Joint DDE calibration and imaging**: initialized as the method (iii), the joint calibration and imaging approach introduced in [24] is selected to precisely evaluate the interest of the temporal prior introduced in Section III.

(v) **Proposed approach**: the proposed procedure is initialized exactly as the approach described in (iii).

The aforementioned algorithms have been run with the same hyper-parameter values whenever appropriate, with $(\lambda, \eta, \vartheta) = (50, 10, 0.5)$ for M31, and $(\lambda, \eta, \vartheta) = (1, 10, 0.5)$ for W28.

The best reconstruction performance, obtained when the true DDEs are known, is reported in Fig 2 in terms of the signal-to-

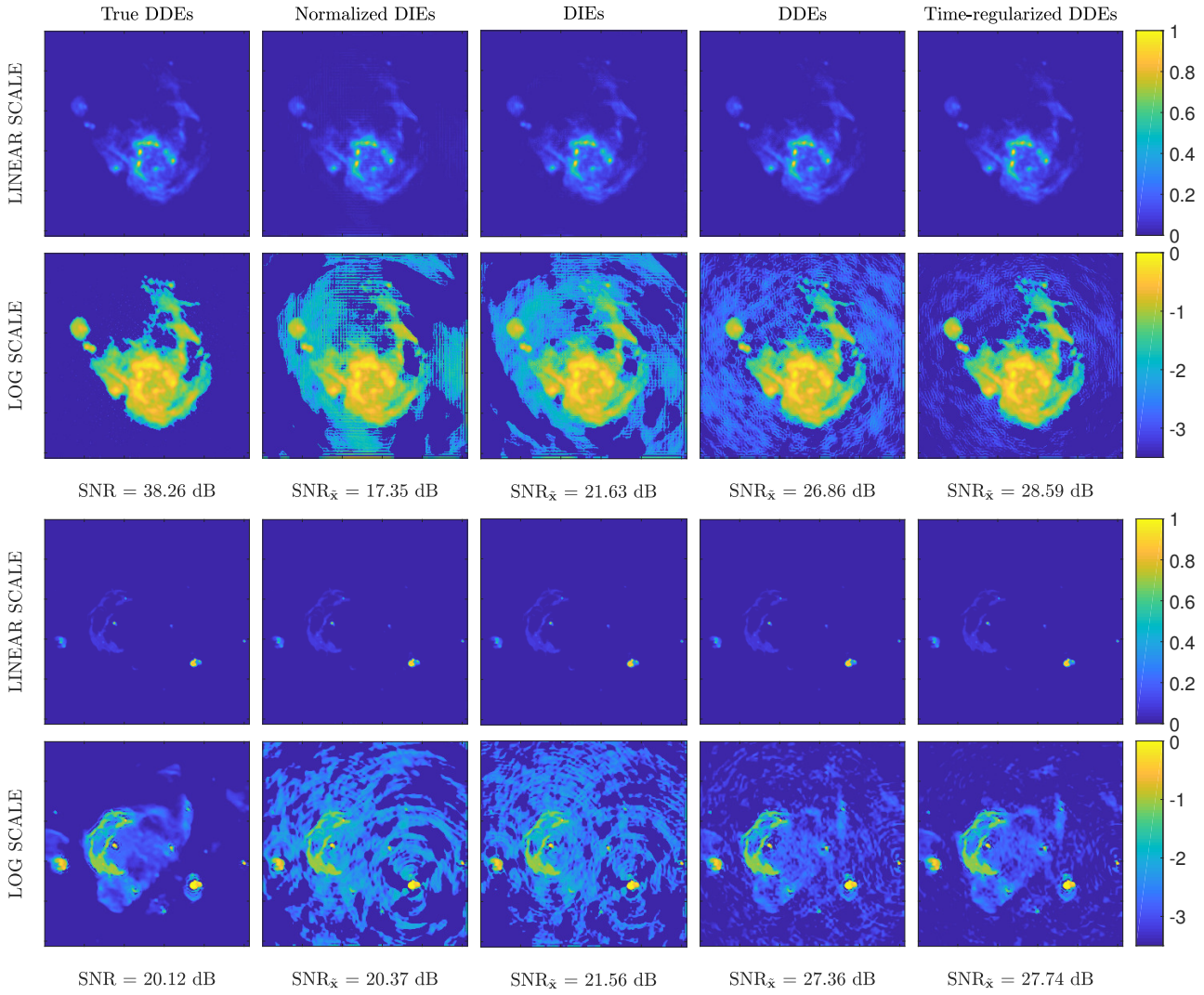

Fig. 2: Results obtained with the image of M31 (first two rows) and W28 (last two rows), below which the corresponding SNR and SNR$_{\widetilde{\boldsymbol{x}}}$ values are reported for each method (given in column). For each scene, the estimated images are displayed in a linear and logarithmic scale along the rows. The images displayed in the first two columns correspond to the solution to the imaging problem (12) with the true DDEs ("True DDEs") and normalized DIEs ("Normalized DIEs"), respectively. The third column corresponds to the image estimated by the joint DIE calibration and imaging approach ("DIEs"). The last two columns correspond to the images estimated with the joint calibration and imaging approach introduced in [24] ("DDEs") and the proposed method ("Time-regularized DDEs"), respectively.

noise ratio $\mathrm{SNR} = 20\log_{10}\big(\frac{\|\overline{\boldsymbol{x}}\|_2}{\|\boldsymbol{x}^*-\overline{\boldsymbol{x}}\|_2}\big)$. The performance of the other approaches is assessed relatively to the reference image $\tilde{\boldsymbol{x}}$. To alleviate the scaling ambiguity inherent to the problem under study, we consider the criterion $\mathrm{SNR}_{\widetilde{\boldsymbol{x}}} = 20\log_{10}\big(\frac{\|\overline{\boldsymbol{x}}\|_2}{\|\nu^*\boldsymbol{x}^*-\widetilde{\boldsymbol{x}}\|_2}\big)$ with $\nu^* = \underset{\nu>0}{\operatorname{argmin}} \ \|\nu\boldsymbol{x}^* - \widetilde{\boldsymbol{x}}\|_2^2$.

The results reported for both datasets in Fig. 2 clearly illustrate the relevance of the proposed approach to improve the reconstruction performance. The temporal prior apparently leads to a moderate improvement in terms of the modified SNR when compared to the approach introduced in [24] (i.e., $+1.7$ and $+0.4$ dB for M31 and W28 respectively). This results from the fact that the SNR mostly accounts for reconstruction improvements in the estimation of high amplitude pixels. Note that a visual inspection of the images recovered with the proposed method shows a significant improvement in the estimation of the low-amplitude background coefficients (last column of Fig. 2), thus illustrating the effectiveness of the proposed prior to mitigate the artifacts resulting from calibration errors.

V. Conclusions

We have extended the joint calibration and imaging approach for radio interferometry proposed in [25] to account for both smooth time-variations of the DDEs and non-uniform Fourier transform embedded in the measurement operator. The parameters involved in the resulting non-convex calibration and imaging problem have been estimated with an alternating forward-backward algorithm, which benefits from well-established convergence guarantees [30], [31]. Simulations show that time-regularization is effective in enhancing the image quality.

REFERENCES

[1] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy*. New York: Wiley-Interscience, 2001.

[2] Y. Wiaux, L. Jacques, A. M. M. Puy, G.and Scaife *et al.*, "Compressed sensing imaging techniques for radio interferometry," *Mon. Not. R. Astron. Soc.*, vol. 395, pp. 1733–1742, 2009.

[3] Y. Wiaux, G. Puy, Y. Boursier *et al.*, "Spread spectrum for imaging techniques in radio interferometry," *Mon. Not. R. Astron. Soc.*, vol. 400, no. 2, pp. 1029–1038, 2009.

[4] S. Wenger, M. Magnor, Y. Pihlström *et al.*, "SparseRI: A compressed sensing framework for aperture synthesis imaging in radio astronomy," *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 897, pp. 1367–1374, 2010.

[5] F. Li, T. J. Cornwell, and F. de Hoog, "The application of compressive sampling to radio astronomy," *A & A*, vol. 528, no. A31, p. 10, 2011.

[6] R. Carrillo, J. McEwen, and Y. Wiaux, "Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging," *Mon. Not. R. Astron. Soc.*, vol. 426, pp. 1223–1234, 2012.

[7] A. Dabbech, C. Ferrari, D. Mary *et al.*, "MORESANE: MOdel RE-construction by Synthesis-ANalysis Estimators. A sparse deconvolution algorithm for radio interferometric imaging," *A & A*, vol. 576, no. A7, p. 16, 2015.

[8] H. Garsden, J. N. Girard, J. L. Starck *et al.*, "LOFAR sparse image reconstruction," *A&A*, vol. 575, p. A90, Mar. 2015.

[9] A. Onose, R. E. Carrillo, A. Repetti *et al.*, "Scalable splitting algorithms for big-data interferometric imaging in the SKA era," *Mon. Not. R. Astron. Soc.*, vol. 462, pp. 4314–4335, 2016.

[10] A. Onose, A. Dabbech, and Y. Wiaux, "An accelerated splitting algorithm for radio-interferometric imaging: when natural and uniform weighting meet," Tech. Rep., 2017, https://arxiv.org/abs/1701.01748.

[11] L. Pratley, J. D. McEwen, M. d'Avezac *et al.*, "Robust sparse image reconstruction of radio interferometric observations with PURIFY," *Monthly Notices of the Royal Astronomical Society*, vol. 473, no. 1, pp. 1038–1058, 2018. [Online]. Available: + http://dx.doi.org/10.1093/mnras/stx2237

[12] A. Dabbech, A. Onose, A. Abdulaziz *et al.*, "Cygnus A super-resolved via convex optimisation from VLA data," *Monthly Notices of the Royal Astronomical Society*, p. sty372, 2018. [Online]. Available: + http://dx.doi.org/10.1093/mnras/sty372

[13] "The Karl J. Jansky Very Large Array (VLA)." [Online]. Available: http://www.vla.nrao.edu

[14] "The Square Kilometre Array (SKA)." [Online]. Available: https://www.skatelescope.org/

[15] D. A. Mitchell, L. J. Greenhill, R. B. Wayth *et al.*, "Real-time calibration of the Murchison widefield array," *IEEE J. Sel. Topics Signal Process*, vol. 2, no. 5, pp. 707–717, Oct. 2008.

[16] S. Salvini and S. J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *A&A*, vol. 571, p. A97, 2014.

[17] H. T. Intema, S. Van der Tol, W. D. Cotton *et al.*, "Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results," *A&A*, vol. 501, no. 3, pp. 1185–1205, 2009.

[18] C. Tasse, "Nonlinear Kalman filters for calibration in radio interferometry," *A & A*, vol. 556, no. A127, p. 11, 2014.

[19] ——, "Applying Wirtinger derivatives to the radio interferometry calibration problem," Tech. Rep., Oct. 2014, https://arxiv.org/abs/1410.8706.

[20] O. Smirnov and C. Tasse, "Radio interferometric gain calibration as a complex optimization problem," *Mon. Not. R. Astron. Soc.*, vol. 449, pp. 2668–2684, 2015.

[21] R. J. van Weeren, W. L. Williams, M. J. Hardcastle *et al.*, "LOFAR facet calibration," *The Astrophysical Journal Supplement Series*, vol. 223, p. 17, 2016.

[22] A. Repetti, J. Birdi, and Y. Wiaux, "Joint imaging and DDEs calibration for radio interferometry." Villars-sur-Ollon, Switzerland: Biomedical and Astronomical Signal Processing (BASP) Frontiers workshop, 29 Jan.–3 Feb. 2017, p. 25.

[23] ——, "Non-convex blind deconvolution approach for sparse radio-interferometric imaging." Lisbon, Portugal: SPARS, 5-8 June 2017, p. 130.

[24] 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.

[25] A. Repetti, J. Birdi, A. Dabbech *et al.*, "Non-convex optimization for self-calibration of direction-dependent effects in radio interferometric imaging," *Mon. Not. R. Astron. Soc.*, vol. 470, no. 4, pp. 3981–4006, Oct. 2017.

[26] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, no. 1, pp. 7–35, 1992.

[27] J. Bolte, P. L. Combettes, and J.-C. Pesquet, "Alternating proximal algorithm for blind image recovery," in *Int. Conf. Image Process. (ICIP 2010)*, vol. 4, Hong-Kong, 26-29 Sep. 2010, pp. 1673–1676.

[28] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," vol. 6, no. 3, pp. 1758–1789, 2013.

[29] P. Frankel, G. Garrigos, and J. Peypouquet, "Splitting methods with variable metric for Kurdyka-Łojasiewicz functions and general convergence rates," *J. Optim. Theory Appl.*, vol. 165, no. 3, pp. 874–900, June 2015.

[30] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Program.*, vol. 146, no. 1, pp. 459–494, 2014.

[31] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," *J. Global Optim.*, vol. 66, no. 3, pp. 457–485, Nov. 2016.

[32] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Image Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.