

A Faceted Prior for Scalable Wideband Imaging: Application to Radio Astronomy

Pierre-Antoine Thouvenin*, Abdullah Abdulaziz*, Ming Jiang†, Audrey Repetti‡*, Arwa Dabbech* and Yves Wiaux*

*Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

†Signal Processing Laboratory 5 (LTS5), École Polytechnique Fédérale de Lausanne, CH-1015, Lausanne, Switzerland

‡Department of Actuarial Mathematics & Statistics, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

Abstract—Wideband radio-interferometric (RI) imaging consists in estimating images of the sky across a whole frequency band from incomplete Fourier data. Powerful prior information is needed to regularize the inverse imaging problem. At the extreme resolution and dynamic range of interest to modern telescopes, image cubes will far exceed Terabyte sizes, with data volumes orders of magnitude larger, making image estimation a very challenging task. The computational cost and memory requirements of corresponding iterative image recovery algorithms are extreme and call for high parallelism. A *data-splitting* strategy was recently introduced to parallelize computations over data blocks within an advanced primal-dual convex optimization algorithm. Building on the same algorithm, we propose an *image faceting* approach that consists in splitting the image cube into 3D overlapping facets with their own prior, reducing the computational bottleneck from full image to facet size. Simulation results suggest our prior provides similar if not superior reconstruction quality to the corresponding state-of-the-art non-faceted approach, with facet parallelization offering acceleration and therefore increased potential of scalability to large data and image sizes.

Index Terms—Wideband radio-interferometric imaging, facet-based prior, preconditioned primal-dual algorithm.

I. INTRODUCTION

The advent of modern radio interferometers [1]–[3] will enable wide-field astronomical surveys with unprecedented resolution and sensitivity over thousands of spectral channels ($\approx 10^{14}$ pixels for the Square Kilometer Array (SKA) [3]). Exploiting the sheer amount of data produced by these instruments while leveraging sophisticated processing techniques is a technical challenge. Wideband radio-interferometric (RI) imaging consists in recovering an image of radio sources in many spectral channels to study the distinctive spectral signature of physical sources in the sky, *e.g.*, radio galaxies [4], fast radio bursts [5] and radio pulsars [6]. In this context, several algorithms have been proposed in the literature to jointly exploit the spatial and spectral information.

CLEAN-based algorithms [7]–[9] progressively reconstruct the wideband model cube by locally removing the contribution of the so-called dirty beam from the back-projected data. Albeit efficient, these algorithms require many parameters to be manually tuned while providing limited imaging quality [10], [11]. Another class of methods is based on Bayesian statistical inference techniques [12], which assume that the sky brightness follows log-normal statistics. In practice, the

This work was supported by EPSRC, grants EP/M011089/1, EP/M008843/1, EP/M019306/1, EP/R026173/1, by the Swiss-South Africa Joint Research Program (IzLSZ2_170863/1), and used the Cirrus UK National Tier-2 HPC Service at EPCC (<http://www.cirrus.ac.uk>) funded by the University of Edinburgh and EPSRC (EP/P020267/1).

high computational cost of these methods prevents them from scaling to large data and image sizes. Over the past few years, scalable optimization-driven imaging approaches have been proposed in the compressed sensing framework [13]–[16]. The single-channel SARA approach [11] has shown the efficiency of using an averaged sparsity prior for RI imaging [11]. It can be applied to each channel separately in a wideband setting, but does not capture the spectral correlations inherent to the data. To address this shortcoming, HyperSARA [14] includes a low-rankness and a joint average sparsity prior, to respectively promote higher resolution and sensitivity images when compared to state-of-the-art imaging techniques [14]. However, this approach entails memory and computing requirements scaling with the size of the full image cube, precluding its use to reconstruct large images. Although the main computational bottleneck induced by the number of measurements has been tackled in [17], another one arises when reconstructing large wideband images. To address this issue, a wideband, wide-field calibration and faceted imaging approach based on CLEAN, namely DDFacet, has been proposed in [18]. This method has been primarily developed for calibration purposes, introducing a piece-wise constant calibration model based on image facets. The facets considered in this framework reflect the underlying flux distribution, and are thus not tailored to reconstruct emissions extending across multiple facets. In addition, this method does not benefit from any convergence guarantee.

In this work, we design a highly scalable wideband imaging approach building on HyperSARA, relying on a flux-agnostic facet-based prior promoting local spectral correlations. We solve the resulting convex optimization problem with a scalable iterative preconditioned primal-dual (PD) algorithm. This algorithm comes with convergence guarantees and allows all the functions to be handled in parallel without resorting to sub-iterations or costly operator inversions [19]. We show through simulations on realistic synthetic data that the proposed method performs well on extended sources, independently of the choice of the facet size. The article is organized as follows. Section II introduces the problem, while the proposed approach is described in Section III. Simulation results are reported in Section IV. Section V concludes this work and outlines future research perspectives.

II. PROBLEM STATEMENT

A radio interferometer acquires *visibilities* resulting from the mutual coherence of signals received by antenna pairs. The visibilities are 2D Fourier coefficients of the sky image,

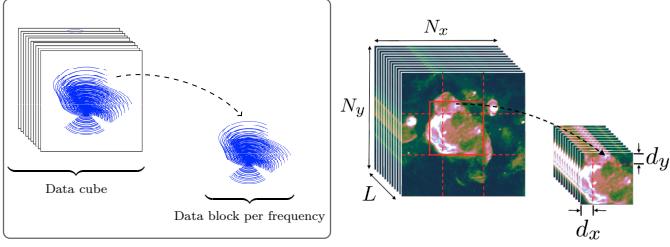


Fig. 1: Data blocking strategy (left panel) and 3D facets (right panel) considered in this work. A facet (red lines) is defined as an extended version of a tile, resulting from a decomposition of the image into non-overlapping regions (dashed red lines). Each facets overlaps with its left and top neighbours by d_x and d_y pixels respectively.

probed at non-uniformly spaced spatial frequencies depending on the observation wavelength and the configuration of the instrument. Due to the limited number of antennas, the sampling pattern, referred to as the *uv*-coverage, does not cover the entire Fourier space. Estimating the wideband sky image $\bar{\mathbf{X}} = (\bar{\mathbf{x}}_l)_{1 \leq l \leq L} \in \mathbb{R}^{N \times L}$ from incomplete, degraded measurements is thus an ill-posed inverse problem. The acquisition process in the frequency channel l can be modeled as

$$\mathbf{y}_l = \Phi_l \bar{\mathbf{x}}_l + \mathbf{n}_l, \text{ with } \Phi_l = \mathbf{G}_l \mathbf{F} \mathbf{Z} \quad (1)$$

where $\mathbf{y}_l \in \mathbb{C}^M$ are the visibilities acquired in the channel $l \in \{1, \dots, L\}$, $\bar{\mathbf{x}}_l \in \mathbb{R}^N$ is the corresponding image, and $\mathbf{n}_l \in \mathbb{C}^M$ is a realization of a complex Gaussian random vector accounting for both instrumental noise and errors due to imperfection of the calibration process. The measurement operator Φ_l is composed of a zero-padding and scaling operator $\mathbf{Z} \in \mathbb{R}^{K \times N}$, the Fourier transform $\mathbf{F} \in \mathbb{C}^{K \times K}$ and an interpolation matrix $\mathbf{G}_l \in \mathbb{C}^{M \times N}$. Each row of \mathbf{G}_l contains compact support kernels representing both the interpolation kernels involved in the non-uniform fast Fourier transform [20] and calibration effects [21]. The main computational bottleneck induced by the large number of measurements is addressed by decomposing the visibility vectors \mathbf{y}_l into B blocks of data $\{\mathbf{y}_{b,l}\}_{b=1}^B$, leading to per-block measurement operators $\Phi_{b,l}$ [17]. Different blocking strategies can be adopted, *e.g.*, based on a spatial tessellation of the *uv*-space to obtain well-balanced sets of visibilities [17], or on a per channel decomposition of the data into groups of snapshots [22] (see Fig. 1). In this context, all blocks can be handled in parallel by advanced algorithmic structures, such as the PD algorithm. The dimension of the wideband image to be formed from the data induces a second computational bottleneck, specifically addressed by the faceted prior introduced in the next section.

III. PROPOSED APPROACH

A. Minimization problem

Wideband RI imaging can be formulated as the problem

$$\underset{\mathbf{X} = (\mathbf{x}_l)_{l=1}^L \in \mathbb{R}_+^{N \times L}}{\text{minimize}} \sum_{l=1}^L \sum_{b=1}^B \iota_{\mathcal{B}(\mathbf{y}_{b,l}, \varepsilon_{b,l})}(\Phi_{b,l} \mathbf{x}_l) + r(\mathbf{X}) \quad (2)$$

where the indices $(b, l) \in \{1, \dots, B\} \times \{1, \dots, L\}$ refer to a data block b in the channel l , and $\mathcal{B}(\mathbf{y}_{b,l}, \varepsilon_{b,l})$ is the ℓ_2 -norm ball centred in $\mathbf{y}_{b,l}$, of radius $\varepsilon_{b,l} > 0$. The indicator functions

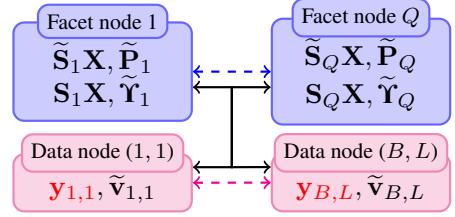


Fig. 2: Data nodes (red group) and facet nodes (in blue) in the implementation of Algorithm 1, given with the main variables they are responsible of. Black (respectively coloured) arrows represent communications between (respectively within) the two groups.

$\iota_{\mathcal{B}(\mathbf{y}_{b,l}, \varepsilon_{b,l})}$ ensure the consistency of the modeled data with the measurements¹, while the non-negativity constraint enforces the physical consistency of the wideband image. The function r encodes additional *a priori* knowledge on the structure of the image to be estimated, depending on the approach considered.

SARA prior [11]: This single-channel approach considers $r(\mathbf{X}) = \mu \|\Psi^\dagger \mathbf{X}\|_{1,1,\omega}$, where $\mu > 0$, $\Psi^\dagger \in \mathbb{R}^{R \times N}$ represents the SARA dictionary (i.e. concatenation of the first 8 Daubechies wavelets and the Dirac basis) [11], and $\|\cdot\|_{1,1,\omega}$ is a weighted $\ell_{1,1}$ norm with weights $\omega \in (\mathbb{R}_+^*)^R$. This prior is extremely parallelizable with respect to the number of spectral channels, but gives a sub-optimal reconstruction quality due to the absence of spectral correlations [14].

HyperSARA prior [14]: This prior is defined as $r(\mathbf{X}) = \|\mathbf{X}\|_{*,\overline{\omega}} + \mu \|\Psi^\dagger \mathbf{X}\|_{2,1,\omega}$, where $\|\cdot\|_{*,\overline{\omega}}$ denotes a weighted nuclear norm with weights $\overline{\omega} \in (\mathbb{R}_+^*)^J$ ($J = \min(N, L)$), and $\|\cdot\|_{2,1,\omega}$ is a weighted $\ell_{2,1}$ norm with weights $\omega \in (\mathbb{R}_+^*)^R$. This approach was shown to give a significantly better reconstruction quality than both SARA and CLEAN-based methods [14]. In practice, the above prior induces memory and computing requirements scaling directly with the size of the full wideband image, which represents a significant bottleneck for very large image cubes. Faceting thus stands out as an efficient technique to tackle this issue, while leading to additional degrees of parallelism.

Proposed faceted prior: We propose to decompose the image into $Q \in \mathbb{N}^*$ overlapping facets to reduce the computing and storage requirements induced by the dimension of the image cube, and thus consider a prior adapted to this decomposition. On the one hand, we observe that the full-size wavelet transforms contained in Ψ^\dagger can be decomposed into a collection of facet-based operators Ψ_q^\dagger to implement an exact but faceted version of Ψ^\dagger [23, Chapter 4]. On the other hand, we introduce a tessellated nuclear norm to obtain a more tractable alternative to the HyperSARA prior, leading to

$$r(\mathbf{X}) = \sum_{q=1}^Q \|\mathbf{W}_q \tilde{\mathbf{S}}_q \mathbf{X}\|_{*,\omega_q} + \mu_q \|\Psi_q^\dagger \mathbf{S}_q \mathbf{X}\|_{2,1,\overline{\omega}_q} \quad (3)$$

where, for every $q \in \{1, \dots, Q\}$, $\mu_q > 0$, and both $\tilde{\mathbf{S}}_q$ and \mathbf{S}_q yield rectangular overlapping facets. These operators differ in the size of the overlap, taken as an adjustable parameter for $\tilde{\mathbf{S}}_q$, and prescribed by [23] for \mathbf{S}_q (see Fig. 1). Diagonal weighting

¹Let C be a closed convex subset of \mathbb{C}^N . The function ι_C denotes the indicator function of C , i.e., $\iota_C(\mathbf{z}) = +\infty$ if $\mathbf{z} \in C$, 0 otherwise.

Algorithm 1: PD algorithm to solve (3).

Data: $\{\mathbf{y}_{b,l}\}_{b,l}$
Input: $\mathbf{X}^{(0)}$, $\{\mathbf{P}_q^{(0)}\}_q$, $\{\boldsymbol{\Upsilon}_q^{(0)}\}_q$, $\{\mathbf{v}_{b,l}^{(0)}\}_{b,l}$
Parameters: $\{\mathbf{U}_{b,l}\}_{b,l}$, $\{\varepsilon_{b,l}\}_{b,l}$, $\{\mu_q\}_q$, τ , ζ , η , ν

- 1 $k \leftarrow 0$, $\xi = +\infty$;
- 2 $\tilde{\mathbf{X}}^{(0)} = \mathbf{X}^{(0)}$;
- 3 **while** $\xi > 10^{-5}$ **do**
 - // Broadcast auxiliary variables
 - 4 **for** $q = 1$ to Q **do**
 - 5 $\tilde{\mathbf{X}}_q^{(k)} = \tilde{\mathbf{S}}_q \tilde{\mathbf{X}}^{(k)}$, $\tilde{\mathbf{X}}_q^{(k)} = \mathbf{S}_q \tilde{\mathbf{X}}^{(k)}$;
 - 6 **for** $l = 1$ to L **do**
 - 7 $\mathbf{z}_l^{(k)} = \mathbf{F} \tilde{\mathbf{X}}_l^{(k)}$; // Fourier transforms
 - 8 **for** $b = l$ to B **do**
 - 9 $\mathbf{z}_{b,l}^{(k)} = \mathbf{M}_{b,l} \mathbf{z}_l^{(k)}$; // broadcast to data nodes
 - 10 // Update low-rankness variables [facet nodes]
 - 11 **for** $q = 1$ to Q **do**
 - 12 $\mathbf{P}_q^{(k+1)} = (\mathbf{I} - \text{prox}_{\|\cdot\|_*, \omega_q / \zeta})(\mathbf{P}_q^{(k)} + \mathbf{W}_q \tilde{\mathbf{X}}_q^{(k)})$;
 - 13 $\tilde{\mathbf{P}}_q^{(k+1)} = \mathbf{W}_q^\dagger \mathbf{P}_q^{(k+1)}$;
 - 14 // Update joint-sparsity variables [facet nodes]
 - 15 **for** $q = 1$ to Q **do**
 - 16 $\boldsymbol{\Upsilon}_q^{(k+1)} = (\mathbf{I} - \text{prox}_{\mu_q \|\cdot\|_{2,1}, \bar{\omega}_q / \nu})(\boldsymbol{\Upsilon}_q^{(k)} + \boldsymbol{\Psi}_q^\dagger \tilde{\mathbf{X}}_q^{(k)})$;
 - 17 $\tilde{\boldsymbol{\Upsilon}}_q^{(k+1)} = \boldsymbol{\Psi}_q \boldsymbol{\Upsilon}_q^{(k+1)}$;
 - 18 // Update data fidelity variables [data nodes]
 - 19 **for** $(b, l) = (1, 1)$ to (B, L) **do**
 - 20 $\mathbf{v}_{b,l}^{(k+1)} = \mathbf{U}_{b,l} (\mathbf{I} - \text{prox}_{\mathcal{B}(\mathbf{y}_{b,l}, \varepsilon_{b,l})}^{\mathbf{U}_{b,l}})(\mathbf{U}_{b,l}^{-1} \mathbf{v}_{b,l}^{(k)} + \mathbf{G}_{b,l} \mathbf{z}_{b,l}^{(k)})$;
 - 21 $\tilde{\mathbf{v}}_{b,l}^{(k+1)} = \mathbf{G}_{b,l}^\dagger \mathbf{v}_{b,l}^{(k+1)}$;
 - 22 // Inter node communications
 - 23 $\mathbf{X}^{(k)} = \sum_{q=1}^Q (\zeta \tilde{\mathbf{S}}_q^\dagger \tilde{\mathbf{P}}_q^{(k+1)} + \nu \mathbf{S}_q^\dagger \tilde{\boldsymbol{\Upsilon}}_q^{(k+1)}) + \eta \mathbf{Z}^\dagger \mathbf{F}^\dagger \sum_{b,l} \tilde{\mathbf{v}}_{b,l}^{(k+1)} \mathbf{e}_l^\dagger$;
 - 24 // Update image [facet nodes]
 - 25 $\mathbf{X}^{(k+1)} = \text{prox}_{\mathbb{R}_{+}^{N \times L}}(\mathbf{X}^{(k)} - \tau \mathbf{I}^{(k)})$;
 - 26 $\tilde{\mathbf{X}}^{(k)} = 2\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}$; // communicate facet borders
 - 27 $\xi = \|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F / \|\mathbf{X}^{(k)}\|_F$;
 - 28 $k \leftarrow k + 1$;

Result: $\mathbf{X}^{(k)}$, $\{\mathbf{P}_q^{(k)}\}_q$, $\{\boldsymbol{\Upsilon}_q^{(k)}\}_q$, $\{\mathbf{v}_{b,l}^{(k)}\}_{b,l}$

matrices $\{\mathbf{W}_q\}_{q=1}^Q$ correct for the fact that pixels within overlapping regions appear multiple times in the problem (2)-(3). In practice, \mathbf{W}_q ensures that each pixel in the facet q is divided by the total number of facets it is involved in.

B. A primal-dual imaging algorithm

An iterative PD algorithm has been advocated for wideband imaging in [14] due to its efficiency in addressing each non-smooth function in (2) in parallel through its proximity operator², avoiding sub-iterations or operator inversions [24], [25]. We thus propose to solve problem (2), for r defined in (3), with the preconditioned PD algorithm described in Algorithm 1. This method is a scalable extension of HyperSARA where the variables associated with each data fidelity term (Algorithm 1 line 17) and facet (lines 11 and 14) can be updated independently.

In practice, the computing nodes involved in the reconstruction algorithm are divided into two groups (see Fig. 2). The first group of nodes (data nodes) handles variables $\tilde{\mathbf{v}}_{b,l}$ of full image size in a single (or a few) channel(s) (see Algorithm 1

²Let \mathbf{U} be a positive definite matrix. The proximity operator of a function f is defined as [19]: $\text{prox}_{f}^{\mathbf{U}}(\mathbf{z}) = \underset{\mathbf{x}}{\operatorname{argmin}}\{f(\mathbf{x}) + \langle \mathbf{x}, \mathbf{U}\mathbf{z} \rangle / 2\}$.

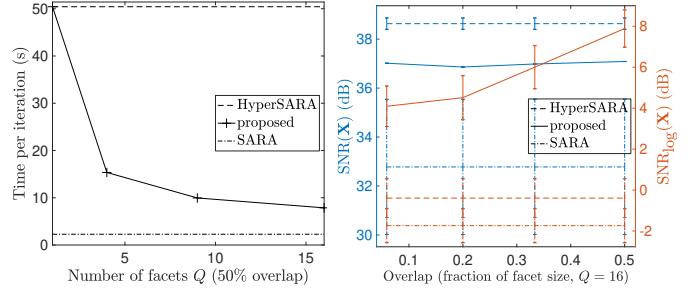


Fig. 3: Per iteration computing time as a function of the number of facets (overlap of 50% of the size of a facet), and average SNR (SNR, SNR_{\log} in dB) as a function of the overlap for $Q = 16$ (error bars correspond to the standard deviation over the spectral channels).

line 18). The second group (facet nodes) handles variables $\tilde{\mathbf{P}}_q$ and $\tilde{\boldsymbol{\Upsilon}}_q$ of facet size over the full spectrum (Algorithm 1 lines 12 and 15). Note that the storage of the 3D image is distributed between several facet nodes, which are responsible for updating the portion of the image they own (line 20). In addition, the Fourier transform is computed for each channel l only on the data nodes $(1, l)$ (line 7). Parts of the Fourier plane, selected by the masks $\mathbf{M}_{b,l}$, are then transferred to the remaining data nodes $(b, l)_{2 \leq b \leq B}$ (line 9). The vector $\mathbf{e}_l \in \mathbb{R}^L$ in line 19 contains a 1 in position l , 0 elsewhere.

To further accelerate the convergence of the algorithm, we leverage a preconditioning strategy based on the inverse of the sampling density [26] to update the data fidelity terms (line 17). Finally, the weights $\bar{\omega}_q$ and ω_q (ω for SARA, and $\bar{\omega}$, ω for HyperSARA) are adjusted following a standard reweighting scheme [27], leading to a sharper regularization of the problem [14]. Under technical assumptions [19], Algorithm 1 is ensured to converge to a solution to problem (2)-(3).

IV. EXPERIMENTS ON SYNTHETIC DATA

Following [14], we generate a wideband image of the W28 supernova remnant composed of $N = 1024 \times 1024$ pixels in $L = 20$ spectral channels ($\approx 10^7$ pixels in total). The measurement operator relies on a realistic spatial Fourier sampling, yielding $M \approx N/2$ measurements per channel. The resulting visibilities are affected by an additive zero-mean white Gaussian noise, leading to an input signal-to-noise ratio $i\text{SNR} = 10 \log_{10} \left(\frac{\sum_i \|\Phi_i \bar{\mathbf{x}}_i\|_2^2}{LM\sigma^2} \right)$ of 60 dB, with σ^2 the noise variance. We investigate the performance of the proposed faceted prior for a varying number of facets Q , and a varying size of overlap between contiguous facets. Reconstruction performance is evaluated in terms of computing time and quality using two criteria: the average signal-to-noise ratio (SNR), defined as $\text{SNR}(\mathbf{X}) = \frac{20}{L} \sum_l \log_{10} \left(\frac{\|\mathbf{x}_l\|_2}{\|\mathbf{x}_l - \bar{\mathbf{x}}_l\|_2} \right)$, and the average SNR computed over the logarithm of the images, denoted by SNR_{\log} . The SNR_{\log} is chosen to evaluate the reconstruction of faint emissions (to which the standard SNR is not sensitive). Note that in RI, an accurate reconstruction of faint emissions is as important as a precise estimation of bright sources, given the high dynamic range of sky images.

Reconstructions are performed using three methods:

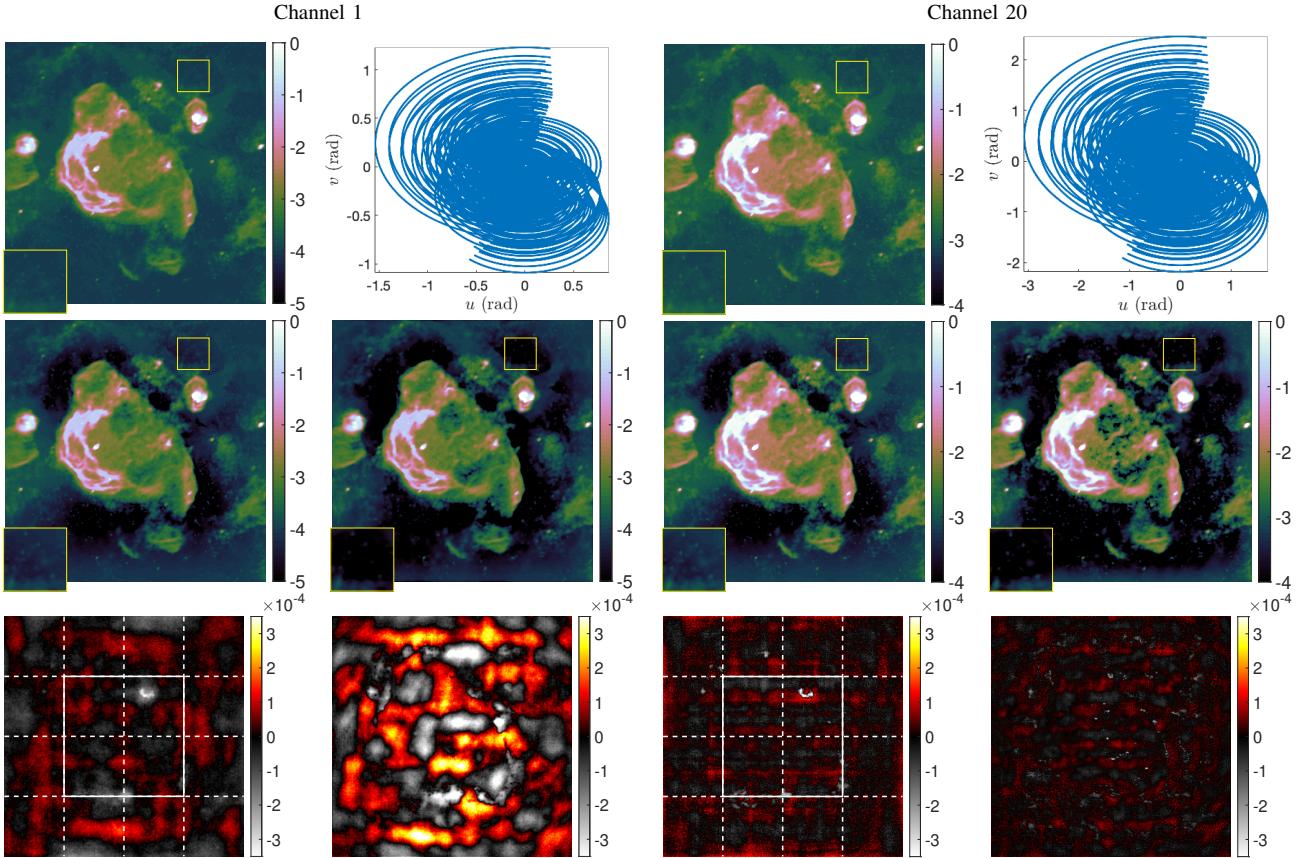


Fig. 4: Synthetic data results reported for the channel 1 (first two columns) and 20 (last two columns) for the faceted (column 1 and 3) and standard HyperSARA (column 2 and 4). From top to bottom are reported the ground truth with the associated normalized uv -coverage. A number of 16 facets has been considered, with an overlap of 50% of the size of a facet. The non-overlapping tiles underlying the facets are delineated on the relevant residual images in white dotted lines, with the central facet in continuous lines. Our approach gives an imaging quality similar to HyperSARA for high intensity emissions, and outperforms it for faint emissions as can be seen on the reconstructed images (zoomed region). The faceted approach yields better or comparable residual images of very small amplitude when compared to HyperSARA (last row). Note that the square-shaped patterns observed on the residual images are not likely to be related to the faceting as (i) they are not aligned with the facet borders, (ii) they also appear for the standard HyperSARA.

- (i) **SARA** [11] is taken as reference in terms of computing time due to its lower computational cost, with $\mu = 10^{-3}$;
- (ii) **HyperSARA** [14] is taken as reference in terms of reconstruction quality, with $\mu = 10^{-3}$.
- (iii) **Proposed faceted approach**, with, for every $q \in \{1, \dots, Q\}$, $\mu_q = 10^{-5}$. The computing time of Algorithm 1 is evaluated for a varying number of facets $Q \in \{4, 9, 16\}$, and a fixed overlap corresponding to 50% of the size of a facet. Reconstruction performance is evaluated for $Q = 16$ and a varying amount of overlap (6%, 20%, 33% and 50% of the size of a facet, corresponding to 16, 64, 128 and 256 pixels respectively).

The results reported in Fig. 3 show that the proposed approach gives a good reconstruction of high intensity pixels (reflected by an SNR close to HyperSARA) for a computing time significantly closer to the SARA approach as the number of facets increases. Note that the reconstruction performance of SARA is limited in terms of resolution and sensitivity, given the nature of the Fourier sampling in radio-astronomy and the absence of prior specifically exploiting spectral correlations. Even if the performance of the proposed approach does not

vary significantly in terms of SNR as the overlap increases, the SNR taken over the log of the image improves significantly. This criterion reflects the ability of the faceted prior to enhance the estimation of faint emissions, beyond HyperSARA, by promoting local spectral correlations. This observation is further confirmed by the reconstructed images reported in Fig. 4 for the channels 1 and 20. The associated residual images (last row of Fig. 4) are comparable to or better than HyperSARA.

V. CONCLUSION AND FUTURE WORK

We have proposed a highly parallelizable faceted low-rankness and joint average sparsity prior for wideband RI imaging. In comparison with HyperSARA, experiments conducted on realistic synthetic data show that the proposed approach offers acceleration through facet parallelization, and therefore an increased potential of scalability to large data and image sizes, for a similar imaging quality. The results suggest the proposed prior promotes local spatial correlations, in that it consistently enhances the reconstruction of faint emissions for a varying number of facets and overlap size. Future work will be focused on validating the proposed prior on real data,

and integrating the associated imaging algorithm in the C++ library Puri-Psi (<https://basp-group.github.io/Puri-Psi/>).

REFERENCES

- [1] M. Van Haarlem *et al.*, “Lofar: The low-frequency array,” *Astronomy & Astrophysics*, vol. 556, p. A2, 2013.
- [2] R. Perley *et al.*, “The expanded very large array: A new telescope for new science,” *The Astrophysical Journal Letters*, vol. 739, no. 1, p. L1, 2011.
- [3] P. Dewdney *et al.*, “SKA1 system baseline design,” *Document number SKA-TEL-SKO-DD-001 Revision*, vol. 1, no. 1, 2013.
- [4] A. D. Kapińska *et al.*, “Spectral Energy Distribution and Radio Halo of NGC 253 at Low Radio Frequencies,” *The Astrophysical Journal*, vol. 838, no. 1, p. 68, Mar. 2017.
- [5] S. Chatterjee *et al.*, “A direct localization of a fast radio burst and its host,” *Nature*, vol. 541, no. 7635, pp. 58–61, Jan. 2017. [Online]. Available: <https://www.nature.com/articles/nature20797>
- [6] T. Murphy *et al.*, “Low-Frequency Spectral Energy Distributions of Radio Pulsars Detected with the Murchison Widefield Array,” *Publications of the Astronomical Society of Australia*, vol. 34, 2017.
- [7] R. Sault and M. Wieringa, “Multi-frequency synthesis techniques in radio interferometric imaging,” *Astronomy and Astrophysics Supplement Series*, vol. 108, 1994.
- [8] U. Rau and T. J. Cornwell, “A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry,” *Astronomy & Astrophysics*, vol. 532, p. A71, 2011.
- [9] A. Offringa and O. Smirnov, “An optimized algorithm for multiscale wideband deconvolution of radio astronomical images,” *Monthly Notices of the Royal Astronomical Society*, vol. 471, no. 1, pp. 301–316, 2017.
- [10] Li, F., Cornwell, T. J., and de Hoog, F., “The application of compressive sampling to radio astronomy - i. deconvolution,” *Astronomy & Astrophysics*, vol. 528, p. A31, 2011.
- [11] R. E. Carrillo, J. D. 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.
- [12] H. Junkleitz, M. R. Bell, and T. Enßlin, “A new approach to multi-frequency synthesis in radio interferometry,” *Astronomy & Astrophysics*, vol. 581, p. A59, 2015.
- [13] M. Jiang, J. Bobin, and J.-L. Starck, “Joint multichannel deconvolution and blind source separation,” *SIAM J. Imaging Sci.*, vol. 10, no. 4, pp. 1997–2021, 2017.
- [14] A. Abdulaziz, A. Dabbech, and Y. Wiaux, “Wideband super-resolution imaging in radio interferometry via low rankness and joint average sparsity models (HyperSARA),” 2018, submitted. [Online]. Available: <https://arxiv.org/abs/1806.04596>
- [15] R. Ammanouil, A. Ferrari, and C. Richard, “Ada-pt: An adaptive parameter tuning strategy based on the weighted stein unbiased risk estimator,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018, pp. 4449–4453.
- [16] C. Meillier *et al.*, “Distribution strategies for very large 3d image deconvolution algorithms,” *Signal Processing: Image Communication*, vol. 67, pp. 149 – 160, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0923596518306416>
- [17] A. Onose *et al.*, “Scalable splitting algorithms for big-data interferometric imaging in the SKA era,” *Monthly Notices of the Royal Astronomical Society*, vol. 462, no. 4, pp. 4314–4335, 2016. [Online]. Available: <http://dx.doi.org/10.1093/mnras/stw1859>
- [18] C. Tasse *et al.*, “Faceting for direction-dependent spectral deconvolution,” *Astronomy & Astrophysics*, vol. 611, p. A87, Mar. 2018.
- [19] J.-C. Pesquet and A. Repetti, “A class of randomized primal-dual algorithms for distributed optimization,” *Journal of nonlinear and convex analysis*, vol. 16, no. 12, pp. 2453–2490, Nov. 2015.
- [20] J. A. Fessler and B. P. Sutton, “Nonuniform fast Fourier transforms using min-max interpolation,” *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Jan. 2003.
- [21] A. Repetti *et al.*, “Non-convex optimization for self-calibration of direction-dependent effects in radio interferometric imaging,” *Monthly Notices of the Royal Astronomical Society*, vol. 470, no. 4, pp. 3981–4006, Oct. 2017.
- [22] A. Dabbech *et al.*, “Cygnus A super-resolved via convex optimization from VLA data,” *Monthly Notices of the Royal Astronomical Society*, vol. 476, no. 3, pp. 2853–2866, Apr. 2018.
- [23] Z. Průša, “Segmentwise discrete wavelet transform,” Ph.D. dissertation, Brno university of technology, 2012.
- [24] L. Condat, “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms,” *J. Optim. Theory Appl.*, vol. 158, pp. 460–479, 2013.
- [25] B. C. Vũ, “A splitting algorithm for dual monotone inclusions involving cocoercive operators,” *Advances in Computational Mathematics*, vol. 38, no. 3, pp. 667–681, Apr. 2013.
- [26] A. Onose, A. Dabbech, and Y. Wiaux, “An accelerated splitting algorithm for radio-interferometric imaging: when natural and uniform weighting meet,” *Monthly Notices of the Royal Astronomical Society*, vol. 469, no. 1, pp. 938–949, 2017.
- [27] E. J. Candès, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *J. Fourier Anal. Appl.*, vol. 4, no. 5–6, pp. 877–905, 2008.