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

Citation for published version:

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Publisher Rights Statement:
© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
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

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 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,
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 \( \mathbf{X} = (\mathbf{x}_{l})_{1 \leq l \leq L} \in \mathbb{R}^{N \times L} \) from incomplete, degraded measurements is thus an ill-posed inverse problem. The acquisition process in the frequency channel \( l \in \{1, \ldots, L\} \) can be modeled as

\[
y_{l} = \Phi_{l}\mathbf{x}_{l} + \mathbf{n}_{l}, \quad \text{with} \quad \Phi_{l} = G_{l}\mathbf{FZ}
\]  

where \( y_{l} \in \mathbb{C}^{M} \) are the visibilities acquired in the channel \( l \in \{1, \ldots, L\} \), \( \mathbf{x}_{l} \in \mathbb{R}^{N} \) 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 \( \Phi_{l} \) 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 \( G_{l} \in \mathbb{C}^{M \times N} \). Each row of \( G_{l} \) 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 \( y_{l} \) into \( B \) blocks of data \( \{\mathbf{x}_{b,l}\}_{b=1}^{B} \), 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 estimated, depending on the approach considered.

### III. PROPOSED APPROACH

#### A. Minimization problem

Wideband RI imaging can be formulated as the problem

\[
\text{minimize} \quad \mathbf{x} = (\mathbf{x}_{l})_{l=1}^{L} \in \mathbb{R}_{+}^{N \times L} \quad \sum_{l=1}^{L} \sum_{b=1}^{B} \varepsilon(y_{b,l}, \varepsilon_{b,l}) (\Phi_{b,l}\mathbf{x}_{l}) + r(\mathbf{X})
\]  

where the indices \( (b, l) \in \{1, \ldots, B\} \times \{1, \ldots, L\} \) refer to a data block \( b \) in the channel \( l \), and \( \varepsilon(y_{b,l}, \varepsilon_{b,l}) \) is the \( \ell_{2} \)-norm of the overlap ball in \( y_{b,l} \), of radius \( \varepsilon_{b,l} > 0 \). The indicator functions \( \varepsilon(y_{b,l}, \varepsilon_{b,l}) \) ensure the consistency of the modeled data with the measurements, 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

\[
r(\mathbf{X}) = \mu \|\Psi^{\mathbf{F}}\mathbf{X}\|_{1, \omega}, \quad \text{where} \quad \mu > 0, \quad \Psi^{\mathbf{F}} \in \mathbb{R}^{R \times N}
\]

represents the SARA dictionary (i.e. concatenation of the first 8 Daubechies wavelets and the Dirac basis) [11], and \( \|\cdot\|_{1, \omega} \) is a weighted \( \ell_{1} \)-norm with weights \( \omega \in (\mathbb{R}_{+}^{*})^{J} \). 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

\[
r(\mathbf{X}) = \|\mathbf{X}\|_{\star, \omega} + \mu \|\Psi_{\mathbf{F}}\mathbf{X}\|_{2, \omega}, \quad \text{where} \quad \|\cdot\|_{\star, \omega} \text{ denotes a weighted nuclear norm with weights } \omega \in (\mathbb{R}_{+}^{*})^{J} (J = \min(N, L)), \text{ and } \|\cdot\|_{2, \omega} \text{ is a weighted } \ell_{2,1} \text{ norm with weights } \omega \in (\mathbb{R}_{+}^{*})^{B}.
\]

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 computational 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 \( \Psi^{\mathbf{F}} \) can be decomposed into a collection of facet-based operators \( \Psi_{Q}^{\mathbf{F}} \) to implement an exact but faceted version of \( \Psi^{\mathbf{F}} \) [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

\[
r(\mathbf{X}) = \sum_{q=1}^{Q} \|W_{q}S_{q}\mathbf{X}\|_{\star, \omega_{q}} + \mu_{q} \|\Psi_{q}\mathbf{S}_{q}\mathbf{X}\|_{2, \omega_{q}}
\]  

where, for every \( q \in \{1, \ldots, Q\} \), \( \mu_{q} > 0 \), and both \( S_{q} \) and \( \mathbf{S}_{q} \) yield rectangular overlapping facets. These operators differ in the size of the overlap, taken as an adjustable parameter for \( S_{q} \), 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 \( \iota_{C}(\mathbf{x}) \) denotes the indicator function of \( C \), i.e., \( \iota_{C}(\mathbf{x}) = +\infty \) if \( \mathbf{x} \in C \), 0 otherwise.
matrices $\{W_q\}_{q=1}^Q$ correct for the fact that pixels within overlapping regions appear multiple times in the problem (2)-(3). In practice, $W_q$ 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 operator, 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 Hyper-SARA 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 groups (as Fig. 2). The first group of nodes (data nodes) handles variables $\tilde{V}_{h,l}$ of full image size in a single (or a few) channel(s) (see Algorithm 1 line 18). The second group (facet nodes) handles variables $\tilde{V}_q$ of facet size over the full spectrum (Algorithm 1 line 17). Parts of the Fourier plane, selected by the masks $M_{b,l}$, are then transferred to the remaining data nodes ($b,l$) for $q$ (line 9). The vector $\omega_q \in \mathbb{R}^L$ 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 $\omega_q$ and $\omega_q$ (for SARA, $\omega_q$ 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 $iSNR = 10 \log_{10} \left( \frac{\|\Psi_x\|^2}{L\sigma_x^2} \right)$ of 60 dB, with $\sigma^2$ 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 $SNR(X) = \frac{1}{L} \sum_{l=1}^L \log_{10} \left( \frac{\|\Psi_x\|^2}{\|\Psi_x\|^2} \right)$, 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:

(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 $q \in \{1, \ldots, Q\}$, $\mu_q = 10^{-5}$. 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/).