SF2A 2016
C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse and I. Vauglin (eds)

ALGORITHMIC ASPECTS FOR THE RECONSTRUCTION OF SPATIO-SPECTRAL DATA CUBES IN THE PERSPECTIVE OF THE SKA

D. Mary[1], A. Ferrari[1], C. Ferrari[1], J. Deguignet[1] and M. Vannier[1]

Abstract. With millions of receivers leading to TerraByte data cubes, the story of the giant SKA telescope is also that of collaborative efforts from radioastronomy, signal processing, optimization and computer sciences. Reconstructing SKA cubes poses two challenges. First, the majority of existing algorithms work in 2D and cannot be directly translated into 3D. Second, the reconstruction implies solving an inverse problem and it is not clear what ultimate limit we can expect on the error of this solution. This study addresses (of course partially) both challenges. We consider an extremely simple data acquisition model, and we focus on strategies making it possible to implement 3D reconstruction algorithms that use state-of-the-art image/spectral regularization. The proposed approach has two main features: (i) reduced memory storage with respect to a previous approach; (ii) efficient parallelization and ventilation of the computational load over the spectral bands. This work will allow to implement and compare various 3D reconstruction approaches in a large scale framework.

Keywords: SKA, radio, inverse problems, parallelization, spatio-spectral cubes.

1 Introduction

The SKA* is an ambitious international project aimed at building the world's largest radio telescope. The full array will be built over two sites in Australian and African deserts. The frequency range that the SKA is expected to cover is unprecedented, from approximately 50 MHz to 15 GHz (and possibly up to $\sim$ 30 GHz in its final phase).

In the Murchison desert of Western Australia will reside the lowest frequency part of the instrument ($\sim$50 to 350 MHz). The telescope will consist there of hundreds of thousands of log–periodic dual–polarised antenna elements, which will be arranged in hundreds of stations of a few metres in diameters. The signals of all elements within one station will be combined numerically and all the stations will be combined together, forming a so-called "aperture array". The inter-station distances will range from a few tens of meters in the central core area to several tens (even hundreds) of kilometres in the outer distribution, which will probably include spiral arms.

In the Karoo desert (800 km North of Cape Town) will be built the part of SKA operating at higher frequencies (above 350 MHz). The array will consist of hundreds of 15 m diameter dishes, with a distribution eventually spreading in different states of Africa. The maximum baselines will be progressively increased from several tens to at least several hundreds of kilometres. Further technical developments are planned, aiming to cover the intermediate frequency part of the SKA also with dense aperture arrays.

Extending significantly the performances of other contemporary antenna arrays in radio, SKA will allow for increased survey speed, sensitivity, angular resolution and for observing simultaneously in a large number of frequency bands. In turn, imaging algorithms will have to be able to reconstruct high fidelity and high dynamic range sky cubes. The final quality of the astrophysical data of SKA rely heavily on joint efforts from radioastronomy, signal processing, optimization and computer sciences. Owing to the gigantic complexity of SKA telescope, this is perhaps more critical than for any other ground-based astronomical instrument.

The purpose of this communication is first to outline the organization, development and challenges posed by the SKA. We focus in these points in Sec. 2.

[1] Lab. J.-L. Lagrange, Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, Parc Valrose, F-06108 Nice cedex 02, France
*For an extended version of information presented below on SKA see Ferrari (2016).

We then turn to the reconstruction of SKA cubes, which poses two challenges. First, the majority of existing algorithms work in 2D. Even if the spectral dimension can be tackled with similar methods as for the spatial dimensions, similar algorithms cannot be translated into 3D. The reason is that when the number of data and optimization variables is too large, classical algorithms suffer from one of two opposite, but equally lethal issues: an immediate stop (related to memory overflows) or an endless iterative process (related to computational load). Second, the reconstruction implies solving an inverse problem and it is not clear what ultimate limit we can expect on the error of this solution.

The study below addresses (partially of course) both challenges. We consider an extremely simplistic data acquisition model (Sec. 3), ignoring for instance calibration errors, direction depending effects or gridding issues. The focus here is on strategies allowing to make possible the implementation of large-scale 3D reconstruction algorithms benefitting from state-of-the-art image/spectral reconstruction methods (Sec. 4). We present some results and perspectives in Sec. 5.

The present paper is a review of some contributions from (Ferrari et al. 2015; Ferrari 2016; Deguignet et al. 2016).

2 SKA organization, developments and challenges

2.1 Project organization

The current status of this ambitious international project is the result of a quite long development history. Planning future infrastructures and data processing strategies is a major management challenge, involving engineers, astronomers, and project managers (see Ferrari (2016) and references therein), and requiring governance.

The SKA governance, called SKA Organisation (SKAO) includes today ten formal members (Australia, Canada, China, India, Italy, New Zealand, South Africa, Sweden, the Netherlands and the United Kingdom), with in addition several countries that have expressed their potential interest in joining the Organisation.

In the case of France, the SKA project is completing the pre-construction phase without France as a formal partner, but France is invited as an observer to International SKA Organisation meetings. During 2014, the French astronomy community developed a roadmap for future research at the CNRS - National Institute for Science of the Universe (INSU). Recognizing the undeniable scientific and technological interest of the SKA, the CNRS-INSU, together with national Observatories and Universities already involved in the SKA project, created in 2016 the SKA-France organization[†], which coordinates the activities of scientific, technological and industrial participants from France. SKA-France is intended to complement the work done by the CNRS Action Spécifique SKA-LOFAR, which has a mandate to encourage the scientific involvement of the French community in current and future radio astronomical facilities.

The general director of SKAO chairs a board, which includes representatives of member countries and which is called to take all relevant decisions for the development of the project. A scientific director is responsible for the parallel and complementary advancement of the SKA scientific preparation.

In agreement with SKA technical, methodological and scientific ambitions, the SKAO has identified several key work packages[‡] (WP). These WP will allow for assembling the different telescope components, building the necessary infrastructures, dealing with data transport, energy issues related to this "big machine", managing the telescope, and, last but not least, processing the data, store them and making them usable by the astronomical community.

For technical WP, the SKA largely builds on the construction and operation of precursors and pathfinders instruments: SKA precursors are the three instruments that are located on the future SKA sites (MWA, ASKAP and MeerKAT), while SKA pathfinders are radio telescopes engaged in SKA related technology and science studies (a census of these instruments is available at the SKA web page[§]). Through the huge progress in antenna design and wide bandwidth feeds, as well as in the ability to transport and process massive amounts of data, a new generation of radio telescopes is being built with these instruments, that pave the way to SKA.

At low frequencies, instruments such as the Low-Frequency Array (LOFAR, in Europe, with its major extension NenuFAR, in France) and the Murchison Widefield Array (MWA, in Australia) employ several thousands of inexpensive dipole antennas arranged within stations without moving parts; the signals from these stations are digitized, transported to a central processor and combined to emulate a conventional dish antenna. These

[†]see `https://ska-france.oca.eu/index.php/en/home-ska-en`

[‡]see `https://www.skatelescope.org/wp/` for a list of the SKA Work Packages

[§]`https://www.skatelescope.org/precursors-pathfinders-design-studies/`

telescopes process the signal on sufficiently short time scales to correct for the ionospheric changes, severely affecting radio observations below a few hundreds MHz.

At higher frequencies (from hundreds of MHz to a few GHz), two major arrays are being built, which are formed by tens of big (∼12-15m diameter) dishes: ASKAP in Australia and MeerKAT in South Africa. In addition, several major radio facilities have been recently upgraded to improve their performance, including, for instance the Jansky Very Large Array (JVLA, in New Mexico), the Giant Meterwave Radio Telescope (GMRT, in India) and the Westerbork Synthesis Radio Telescope (WSRT, in the Netherlands).

2.2 SKA development

In 2012 the SKA board proceeded to the selection of the SKA construction sites. Both Australia and South African deserts, sites of the three precursors, were considered as excellent locations for building the arrays covering the low- and mid-frequency part of the electromagnetic spectrum. At that phase, in particular, both precursors going to GHz frequencies (ASKAP and MeerKAT) were planned to be integrated to the future SKA antennas. In 2015, however, a redefinition of the design needed to be developed (known as "re-baselining") due to cost issues. This led to the definition of the first phase of SKA (generally referred to as "SKA1"), that will consist of two arrays:

- **SKA1-LOW** in Australia, including ∼131,000 dipoles covering the frequency range from ∼50 MHz to ∼350 MHz. The array will be in the same region of ASKAP and will have a maximum baseline of approximately 80 km;
- **SKA1-MID** in South Africa, including ∼130 15m diameter dishes and observing frequencies from 350 MHz to 13.8 GHz, with a $\Delta\nu \approx 3$ GHz hole around 1.5 GHz. The instrument will integrate the MeerKAT antennas, for a total number of ∼200 dishes, separated by a maximum distance of ∼150 km.

Based on these decisions, the construction of SKA1, which corresponds to ≈10% of the full final instrument, is planned to start in 2018, with the first science operations beginning in early 2020's. SKA1 will be a single observatory built over three sites, among which one will host the head-quarter (at Jodrell Bank, in UK, site decision taken in 2015) and the other the two telescopes, SKA1-LOW in Australia and SKA1-MID in South Africa.

After 2023, the instrument is planned to be further developed towards the full square kilometre total collecting area. This will of course require the necessary preparation from the technological and scientific point of view, for which SKA1 will play a crucial role, but also a likewise important budget and managing strategy. In order to provide a long-term government commitment and funding stability, the SKAO is thus currently evolving towards an intergovernmental organisation (IGO), similar to other big research infrastructures (such as ESO, ESA and CERN).

2.3 Challenges

The power of the SKA comes from the large number of receivers, the large fields of view made available and the high numbers of spectral bands. The large scale nature of this instrument poses automatically several challenges, which make mandatory dedicated research in various fields working in close connexion:

- ***Computer sciences***. An immediate challenge is that of the processing and storage of the data produced simultaneously by myriads of receivers. In its first phase (SKA1), when the total collecting area will be approximately one tenth of the final expected array, the total raw data output for the low- and mid-frequency parts of the telescope will be of the order of 150 and 2 Tbytes per second, respectively. This results in several Zbytes (10^{21} bytes) of raw data per year, which exceeds the total global internet traffic rate at present day. Even after data reduction, the archived data rate for astronomical exploitations is expected to be of the order of 50 to 300 Pbytes per year. Data processing and storage will require by early 2020's super–computers characterised by about 10 times better performances than the fastest machines available today.
- ***Signal Processing***. An instrument like SKA requires the outputs of many different elements to be combined. Beyond issues related to the data rate, all array elements must be modeled and calibrated for this massive combination to be reliable. The calibration part of the array must estimate two main quantities. First, the gain response and noise power of each antenna (or group of dipoles) (Wijnholds &

van der Veen 2009; van der Veen & Wijnholds 2013; van der Tol et al. 2007). Second, the atmospheric perturbation on the propagation of radio waves, especially the phase delays caused by the ionosphere, which are time and wavelength dependent (van der Tol 2009; Thompson et al. 2001; Intema et al. 2009). Many calibration models and algorithms have been devised in the signal processing litterature. Another important branch where Signal Processing is very active is the imaging part, where sparsity based models in particular have proved very useful in radiointerferometric imaging (see below).

- ***Optimization***. Optimization techniques express problems in mathematical form by means of objectives and constraints. These techniques allow to address some properties of the solution (*e.g.* existence, unicity,...), to derive systematic methods to solve the problem, and to study the convergence rate towards the solution (see *e.g.* Canu et al. (2016) for an introduction). Optimization consequently plays a central role for designing calibration, imaging or more generally processing algorithms that are both computationally efficient and able to cope with several disturbance sources, see *e.g.* Brossard et al. (2016) and references therein.

The success of the interplay between the expertises above in the context of large scale data is recognized as a crucial point from the SKAO, which has proposed several grants within the program *AstroCompute in the Cloud* in conjunction with Amazon Web Services (AWS). This program encourages development and code optimization for massive data for image analysis, calibration and data mining, visualization, management and sharing.

The study below reflects some works going in the direction of allying signal processing, optimization and parallelization tools. Our focus here is not on the data model, which is taken very simple and ignores for instance calibration, gridding issues or direction dependent effects. Our focus is on devising reconstruction methods able to cope with large (multiband) data while using efficient regularization techniques. In fact, building algorithms that are both able to reconstruct large data cubes with affordable computing power and high fidelity is a challenge *per se* and this work is a step in this direction.

3 Data model and reconstruction approach

The considered data acquisition model is sketched in Fig. 1. The unknown reference sky is discretized and represented by means of parameters (pixels or voxels values) corresponding to fluxes integrated over some sky patch and spectral band. The interferometer samples information on the Fourier spectrum of $\boldsymbol{x}^\star$ (bottom left panel) at locations specified by the antennas positions and the observing wavelength (bottom right panel). The samples are called complex visibilities. Because the antennas are sparsely disseminated on the ground, the sampling in the Fourier space is sparse as well. When observing in a set of L frequency bands, the interferometer provides L radially aligned Fourier samples for each antenna/station baseline (in fact twice this number by symmetry). As Earth rotates, the position of the antennas with respect to the sky changes and the number of collected samples increases. A grey image of the sky $\boldsymbol{i}^d$ can be formed by computing the inverse Fourier Transform (top right panel). We shall assume that this image is related to the grey image by $\boldsymbol{i}^d = \boldsymbol{H}\boldsymbol{x}^\star$, and that $\boldsymbol{H}$ represents a convolution.

Of course, it is also possible to produce L narrow band sky images $\{\boldsymbol{i}_\ell^d\}, \ell = 1, \ldots, L$, by considering only the samples in a given spectral band (Fig. 2). In this case, the dirty image in each channel is $\boldsymbol{i}_\ell^d = \boldsymbol{H}_\ell \boldsymbol{x}_\ell^\star$, where $\boldsymbol{H}_\ell$ and $\boldsymbol{x}_\ell^\star$ represent respectively the convolution matrix and the reference sky image in band ℓ. For convenience, the L reconstructed images (N pixels each) will be stacked in an $N \times L$ matrix $\boldsymbol{X} := [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell]$, whose ℓ^{th} column represents the optimization variables (voxel values) for channel ℓ. Inverting the operators $\boldsymbol{H}_\ell$ to retrieve $\mathbf{x}_\ell^\star$ from $\boldsymbol{i}_\ell^d$ is not possible owing to the 'holes' in the Fourier space: the system is blind to some frequency contents and simply provides no information on it. (The problem is even worse in practice, because samples are never on regular grids and the operator from image to complex visibilities often departs from a simple Fourier Transform, which increases the causes of non invertibility.)

In such situations a classical approach is to formulate the reconstruction problem as an inverse, ill-posed problem and to a introduce a cost function balancing a data fidelity term and a regularization term. This function will be minimized with dedicated optimization techniques. In the current description, we consider a cost function leading to a minimization problem of the type

$$\min_{\boldsymbol{X}} \sum_{\ell=1}^{L} \frac{1}{2\sigma_\ell^2} \|\mathbf{i}_\ell^d - \boldsymbol{H}_\ell \boldsymbol{x}_\ell\|^2 + f_{\text{reg}}(\boldsymbol{X}) \tag{3.1}$$

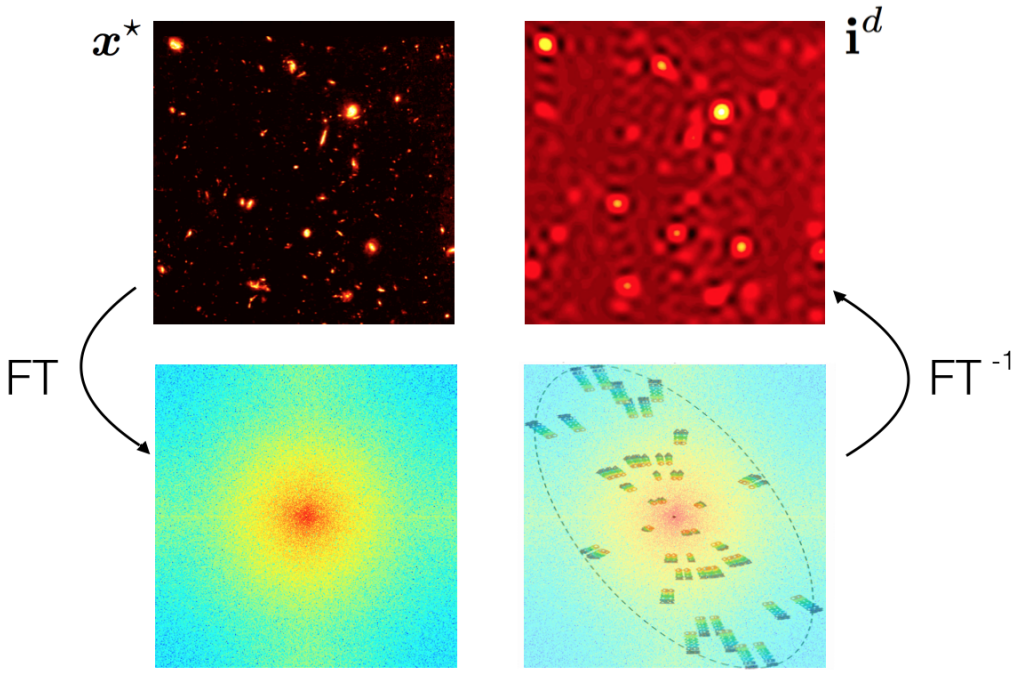

Fig. 1. Top left: the reference sky $\boldsymbol{x}^\star$ (sum over several bands). **Bottom left:** modulus of the Fourier Transform of $\boldsymbol{x}^\star$. The origin is at the center of the image. **Bottom right**: Sketch of an interferometric sampling pattern. The Fourier samples collected by the radio array are located by color points (from red to blue). Points of the same color correspond to samples acquired in the same spectral band. **Top right**: Sky estimate by inverse Fourier transform considering all samples available (*i.e.*, the L subsets if there are L bands) and setting missing samples to 0: 'grey' dirty image, $\boldsymbol{i}^d$.

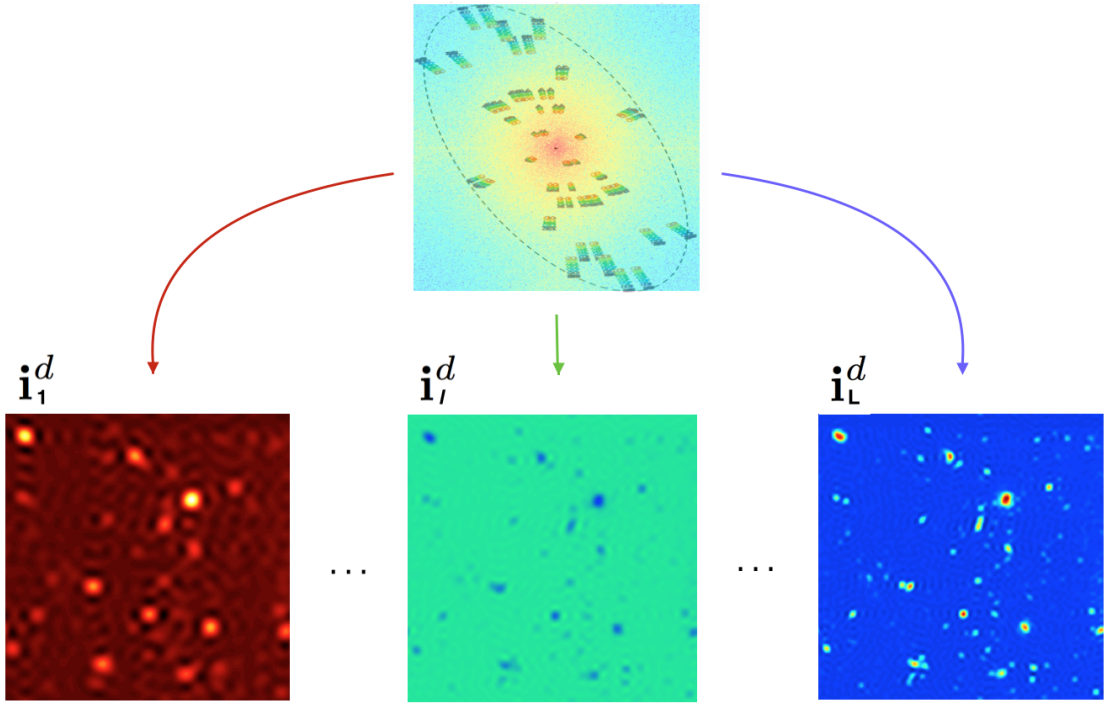

Fig. 2. L narrow band dirty images produced from the L subsets of complex visibilities.

where σ_ℓ^2 is the noise variance in dirty image $\boldsymbol{y}_\ell$, and where a positivity constraint can be further imposed on $\boldsymbol{X}$. The first term of (3.1) guarantees that the proposed solution is in agreement (in the specified ℓ_2 sense) with the data. This term alone is not sufficient, because infinitely many sky models may fulfill any fixed level of data agreement (for a given candidate solution $\boldsymbol{x}_\ell$, adding anything in the kernel of $\boldsymbol{H}_\ell$ does not change anything to the first term). The role of the second term is to help choosing among all solutions one that is compatible with respect to some regularity criterion. The last decade has shown that regularization based on sparse representations in appropriate transform domains can be very effective. Such regularization terms can be formulated in an analysis or in a synthesis framework. These two formalisms are discussed and compared e.g. in Elad et al. (2007). For both approaches, redundant dictionaries improve over non redundant (orthogonal) ones. For narrow-band radio-interferometric imaging, state-of-the-art results appear so far to be obtained with union of bases (Carrillo et al. 2012, 2014; Onose et al. 2016) and IUWT (Garsden et al. 2015; Dabbech et al. 2015). We consider a regularization of the form:

$$f_{\text{reg}}(\boldsymbol{X}) := \mathbf{1}_{\mathbb{R}^+}(\boldsymbol{X}) + \mu_S \sum_{\ell=1}^{L} \|\mathbf{W}_S \boldsymbol{x}_\ell\|_1 + \mu_\nu \sum_{n=1}^{N} \|\mathbf{W}_\nu \boldsymbol{x}^n\|_1. \tag{3.2}$$

In (3.2), $\boldsymbol{x}_\ell$ (the ℓ^{th} column of $\boldsymbol{X}$) corresponds to the image at wavelength ℓ and $\boldsymbol{x}^n$ (the n^{th} row of $\boldsymbol{X}$) is the spectrum associated to pixel n. $\mathbf{W}_S$ and $\mathbf{W}_\lambda$ are the operators corresponding to the spatial and spectral decomposition respectively. IUWT will here be considered for the spatial regularization and a cosine decomposition (DCT) for the spectral model.

Note that the first sum in (3.2) decouples into L terms (one per image) and this is also true for the sum in (3.1). In contrast, the second sum in (3.2) decouples into N terms, each related to one spectrum. This structure is important as it allows to parallelize some computations spectrally and other spatially. Note, however, that each voxel is tied spatially to all voxels at the same wavelength (through the first two sums above) and also tied spectrally to all voxels in the same spectrum (through the second sum). Hence, the problem cannot be globally decoupled and it is truly spatio-spectral.

It is interesting to compare the few existing multi-frequency reconstruction algorithms with approach (3.2). Most of the existing approaches rely on a physical model for the spectral dependence of astrophysical radio sources. Rau & Cornwell (2011) use a Taylor expansion of a model power-law. Other (parametric) models for this spectral dependence have been proposed by Junklewitz et al. (2014), who address the estimation problem using a Bayesian framework and by Bajkova & Pushkarev (2011), who propose a constrained maximum entropy estimation algorithm.

These methods rely on spectral models and thus clearly offer advantages and estimation accuracy when the model is indeed appropriate. However, across the broad frequency coverage of current radio facilities, radio sources exhibiting complex spectral shapes (not simple power laws) are expected. For instance, the works of Kellermann (1974) evidence that some sources may exhibit one or more relative minima, breaks, and turnovers. For the new generation of low frequency telescopes such as LOFAR, recent studies have also shown that second order broadband spectral models are often insufficient (Scaife & Heald 2012). Attempts to relax the spectral power-law model are thus necessary. One such attempt, by Wenger & Magnor (2014), formulates the problem as an inverse problem with a smooth spectral regularization allowing for local deviations. The present approach is another attempt of this kind, as it does not either rely on a spectral power-law model. Examples of situations where this approach can be advantageous are shown in Deguignet et al. (2016).

4 Optimization

One drawback of the works by Ferrari et al. (2015) was the necessity to solve a large linear system at each iteration during the minimization. This was kept computationally tractable by using for $\boldsymbol{W}_S$ a concatenation of orthogonal wavelet bases in (Ferrari et al. 2015; Carrillo et al. 2014). Another drawback is the amount of memory. In order to reduce the required memory, we have proposed in (Deguignet et al. 2016) to use the primal-dual optimization algorithm by Condat (2014) and Vũ (2011). Application of Condat (2014) and Vũ (2011) to (3.1, 3.2) leads to Algorithm 1, where:

$$\text{sat}(u) := \begin{cases} -1 & \text{if} \quad u < -1 \\ 1 & \text{if} \quad u > 1 \\ u & \text{if} \quad |u| \leq 1 \end{cases} \tag{4.1}$$

and $(\cdot)_+$ is the projection on the positive orthant. Parameters ρ, σ and τ are fixed according to Condat (2014) in order to guarantee the convergence of the algorithm.

Algorithm 1: MUFFIN algorithm.

Initialize: $\boldsymbol{x}$, $\mathbf{U}$ and $\mathbf{V}$

repeat

$\boldsymbol{\nabla} = \left(\boldsymbol{H}_1^\dagger(\boldsymbol{H}_1\boldsymbol{x}_1 - \mathbf{i}_1^d) \mid \cdots \mid \boldsymbol{H}_\ell^\dagger(\boldsymbol{H}_\ell\boldsymbol{x}_\ell - \mathbf{i}_\ell^d)\right)$ $\tilde{\boldsymbol{X}} = \left(\boldsymbol{X} - \tau(\boldsymbol{\nabla} + \mu_S \boldsymbol{W}_S^\dagger \boldsymbol{U} + \mu_\nu \boldsymbol{V}\boldsymbol{W}_\nu^\dagger)\right)_+$

$\tilde{\boldsymbol{U}} = \mathrm{sat}\left(\boldsymbol{U} + \sigma\mu_S \boldsymbol{W}_S(2\tilde{\boldsymbol{X}} - \boldsymbol{X})\right)$ $\tilde{\boldsymbol{V}} = \mathrm{sat}\left(\boldsymbol{V} + \sigma\mu_\nu (2\tilde{\boldsymbol{X}} - \boldsymbol{X})\boldsymbol{W}_\nu\right)$

$(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{V}) = \rho(\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) + (1-\rho)(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{V})$

until *stopping criterion is satisfied*;

Return : $\boldsymbol{X}$

Algorithm 1 requires 6 variables $(\tilde{\boldsymbol{X}}, \boldsymbol{X}, \tilde{\boldsymbol{U}}, \boldsymbol{U}, \tilde{\boldsymbol{V}}, \boldsymbol{V})$ in addition to the gradient and 5 if $\rho = 1$, while 10 variables were necessary in the previous approach (Ferrari et al. 2015). Moreover, in contrast to Ferrari et al. (2015) and Carrillo et al. (2014), Algorithm 1 does not require to solve at each iteration large linear systems. This allows the use of highly redundant, translation invariant wavelet transforms like, for instance, IUWT (Starck et al. 2007).

A major advantage of Algorithm 1 is that the most demanding steps are separable with respect to the wavelengths, leading to the following parallel implementation. MUFFIN is distributed on a cluster where the master node centralises the reconstructed data cube and each wavelength is associated to a compute node ℓ. The algorithm iterates as follows:

1. The master node computes $\boldsymbol{T} = \mu_\lambda \boldsymbol{V}\boldsymbol{W}_\lambda^\dagger$ and sends the column ℓ of $\boldsymbol{T}$, denoted as $\boldsymbol{t}_\ell$, to node ℓ.
2. Each node $\ell = 1 \dots L$ computes sequentially:

$$\boldsymbol{\nabla}_\ell = \boldsymbol{H}_\ell^\dagger(\boldsymbol{H}_\ell\boldsymbol{x}_\ell - \boldsymbol{y}_\ell) \tag{4.2}$$

$$\boldsymbol{s}_\ell = \mu_S \boldsymbol{W}_S^\dagger \boldsymbol{u}_\ell \tag{4.3}$$

$$\tilde{\boldsymbol{x}}_\ell = (\boldsymbol{x}_\ell - \tau(\boldsymbol{\nabla}_\ell + \boldsymbol{s}_\ell + \boldsymbol{t}_\ell))_+ \tag{4.4}$$

$$\tilde{\boldsymbol{u}}_\ell = \mathrm{sat}\left(\boldsymbol{u}_\ell + \sigma\mu_S \boldsymbol{W}_S(2\tilde{\boldsymbol{x}}_\ell - \boldsymbol{x}_\ell)\right) \tag{4.5}$$

$$(\boldsymbol{x}_\ell, \boldsymbol{u}_\ell) = \rho(\tilde{\boldsymbol{x}}_\ell, \tilde{\boldsymbol{u}}_\ell) + (1-\rho)(\boldsymbol{x}_\ell, \boldsymbol{u}_\ell) \tag{4.6}$$

3. Each node sends $\boldsymbol{x}_\ell$ and $\tilde{\boldsymbol{x}}_\ell$ to the master and the master computes sequentially:

$$\tilde{\boldsymbol{V}} = \mathrm{sat}\left(\boldsymbol{V} + \sigma\mu_\lambda(2\tilde{\boldsymbol{X}} - \boldsymbol{X})\boldsymbol{W}_\lambda\right) \tag{4.7}$$

$$\boldsymbol{V} = \rho\tilde{\boldsymbol{V}} + (1-\rho)\boldsymbol{V} \tag{4.8}$$

Note that the particularly time consuming steps associated to (4.2,4.3,4.5) are computed in parallel at each wavelength. This is particularly important for (4.3) when the transform is not orthogonal. In such cases, the adjoint operator differs from the perfect reconstruction synthesis operator and its implementation may not benefit from the same fast algorithm.

A distributed memory implementation of MUFFIN should be available soon[¶]. The algorithm has been implemented in *Julia* (Bezanson et al. 2014).

5 Some results and future works

We show here one example illustrating the efficiency of using a joint spatio-spectral approach with respect to a channel per channel reconstruction. Simulations use PSFs obtained with the HI-inator package[‖] based on MeqTrees software (Noordam & Smirnov 2010) with MeerKAT arrays configuration. The Fourier coverage

[¶] https://github.com/andferrari/muffin.jl

[‖] https://github.com/SpheMakh/HI-Inator

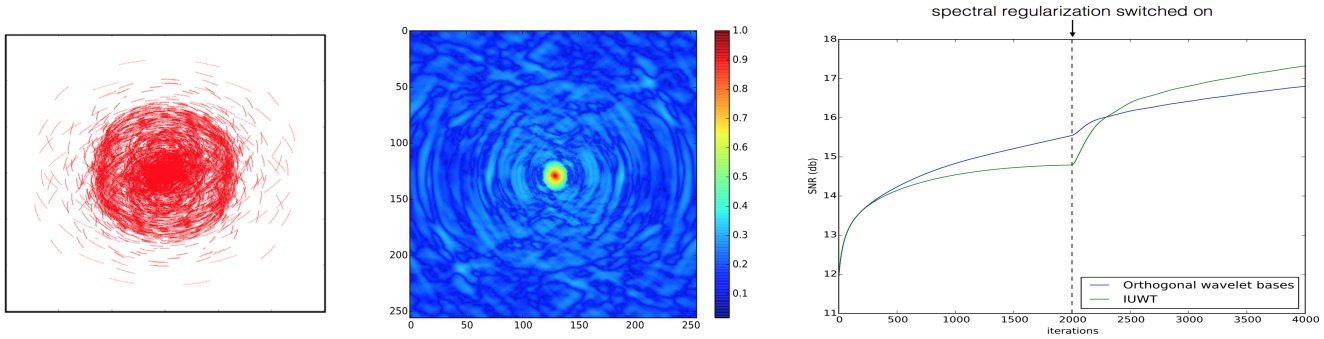

Fig. 3. Left: considered (u, v) coverage. **Center:** PSF in one band. **Right:** Comparison of the SNR for union of orthogonal bases and IUWT. Spectral regularization is turned on at iteration 2000.

correspond to a total observation time of 8 hours. For the purpose of making Monte Carlo simulations, we simulated small cubes of 15 frequency bands with images of 256×256 pixels.

In Algorithm 1, we compare the case where $\boldsymbol{W}_S$ in (4.5) corresponds to "2^{nd} generation" IUWT (Starck et al. 2007) (and $\boldsymbol{W}_S^\dagger$ in (4.3) is the exact corresponding adjoint operator), or to the union of orthogonal bases used by Ferrari et al. (2015). The sky corresponds to the radio emission of an HII region in the M31 galaxy. A sky cube is computed from this real sky image by applying a first order power-law spectrum model. The 256×256 map of spectral indices was constructed following the procedure detailed in Junklewitz et al. (2014): for each pixel, the spectral index is a linear combination of an homogeneous Gaussian field and the reference sky image. A Gaussian noise corresponding to 10 dB was finally added to the dirty images to simulate instrumental and model errors. The parameters of the optimization algorithm are set to $\rho = 1$, $\sigma = 1$ and $\tau = 10^{-5}$.

A critical problem for the deconvolution of large data cubes is the calibration of the regularization parameters μ_S and μ_λ. We decouple the calibration in two steps:

1. μ_λ is first set to 0: the problem is separable with respect to the wavelengths and each node independently iterates Eqs. (4.2-4.6) with $\boldsymbol{t}_\ell = \boldsymbol{0}$. This setting avoids data transfers with the master node. It is relatively fast and allows for multiple runs to calibrate μ_S, *e.g.* by cross-validation.

2. The second step keeps μ_S and the $\boldsymbol{X}$ estimated in step 1) and calibrates μ_λ using the full algorithm with $\boldsymbol{X}$ as an initial condition.

Fig. 3 compares the reconstruction Signal to Noise Ratio (SNR) for the union of bases (blue) and IUWT (green) as a function of the iterations. SNR is here defined as:

$$\mathrm{SNR}(\boldsymbol{X}, \boldsymbol{X}^\star) := 10 \log_{10} \left(\frac{\|\boldsymbol{X}^\star\|_2^2}{\|\boldsymbol{X} - \boldsymbol{X}^\star\|_2^2} \right) \tag{5.1}$$

where $\boldsymbol{X}$ is the estimated solution and $\boldsymbol{X}^\star$ the "sky truth". The first 2000 iterations correspond to step 1) with $\mu_\lambda = 0$ and $\mu_S = 0.25$, and the subsequent iterations to step 2) with $\mu_S = 0.25$ and $\mu_\lambda = 3.0$. The value of $\mu_S = 0.25$ in 1) and $\mu_\lambda = 3$ in 2) were set, for both types of wavelets, after trials and errors in the range $[10^{-4}\ 10^{1}]$ and best performances were retained.

The evolution of the SNRs after iteration 2000, i.e. when $\mu_\lambda > 0$ clearly evidences the gain obtained through a joint spatio-spectral reconstruction for both approaches. We see that while performances of both approaches are comparable, they relative behavior depend on the regularization and on the number of iterations (which is an important point in a large scale framework). Indeed, such questions deserve further studies. Those are outside the scope of the present paper but are made possible with the parallel implementation proposed in this contribution.

6 Acnkowledgements

This work was supported by ANR project MAGELLAN (ANR-14-CE23-0004-01).

References

Bajkova, A. & Pushkarev, A. 2011, Monthly Notices of the Royal Astronomical Society, 417, 434

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2014, CoRR, abs/1411.1607

Brossard, M., El Korso, M. N., Pesavento, M., et al. 2016, ArXiv e-print: 1609.02448

Canu, S., Flamary, R., & Mary, D. 2016, in Introduction to Optimization with Applications in Astronomy and Astrophysics, EAS Publications series, ed. D. Mary et al., 127–162

Carrillo, R., McEwen, J., & Wiaux, Y. 2012, Monthly Notices of the Royal Astronomical Society, 426, 1223

Carrillo, R., McEwen, J., & Wiaux, Y. 2014, Monthly Notices of the Royal Astronomical Society, 439, 3591

Condat, L. 2014, Signal Processing Letters, IEEE, 21, 985

Dabbech, A., Ferrari, C., Mary, D., et al. 2015, Astronomy and Astrophysics, 576, A7

Deguignet, J., Ferrari, A., Mary, D., & Ferrari, C. 2016, in Proceedings EUSIPCO 2016

Elad, M., Milanfar, P., & Rubinstein, R. 2007, Inverse Problems, 23, 947

Ferrari, A., Deguignet, J., Ferrari, C., et al. 2015, in SKA Pathfinders Radio Continuum Surveys (SPARCS)

Ferrari, C. 2016, in The Renaissance of Radio Astronomy: towards the SKA, EAS Publications Series, ed. D. Mary et al., Vol. 78-79, 21–44

Garsden, H., Girard, J. N., Starck, J.-L., et al. 2015, Astronomy and Astrophysics, 575, A90

Intema, H. T., van der Tol, S., Cotton, W. D., et al. 2009, Astronomy & Astrophysics, 501, 1185

Junklewitz, H., Bell, M., & Enßlin, T. 2014, arXiv.org, 4711

Kellermann, K. I. 1974, Radio Galaxies and Quasars, ed. G. L. Verschuur, K. I. Kellermann, & V. van Brunt, 127

Noordam, J. & Smirnov, O. 2010, Astronomy and Astrophysics, 524, A61

Onose, A., Carrillo, R. E., Repetti, A., et al. 2016, MNRAS, 462, 4314

Rau, U. & Cornwell, T. J. 2011, Astronomy and Astrophysics, 532, 71

Scaife, A. & Heald, G. 2012, Monthly Notices of the Royal Astronomical Society: Letters, 423, L30

Starck, J.-L., Fadili, J., & Murtagh, F. 2007, Image Processing, IEEE Transactions on, 16, 297

Thompson, A. R., Moran, J. M., & Swenson, G. W. 2001, Interferometry and Synthesis in Radio Astronomy; 2nd ed. (Wiley-VCH)

van der Tol, S. 2009, PhD thesis, Delft University of Technology

van der Tol, S., Jeffs, B. D., & van der Veen, A. J. 2007, IEEE Transactions on Signal Processing, 55, 4497

van der Veen, A.-J. & Wijnholds, S. J. 2013, in Handbook of Signal Processing Systems (Springer), 421–463

Vũ, B. C. 2011, Advances in Computational Mathematics, 38, 667

Wenger, S. & Magnor, M. 2014, A Sparse Reconstruction Algorithm for Multi-Frequency Radio Images, Tech. rep., Computer Graphics Lab, TU Braunschweig

Wijnholds, S. J. & van der Veen, A.-J. 2009, IEEE Transactions on Signal Processing, 57, 3512