A distributed algorithm for wide-band radio-interferometry
Abdulaziz, Abdullah; Onose, Alex; Dabbech, Arwa; Wiaux, Yves

Publication date:
2017

Document Version
Peer reviewed version

Citation for published version (APA):
A distributed algorithm for wide-band radio-interferometry

Abdullah Abdulaziz, Alexandru O nose, Arwa Dabbache and Yves Wiaux
Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh EH14 4AS, UK

Abstract—We propose a scalable, randomised algorithm to solve the inverse imaging problem in wide-band radio-interferometry. In the big-data context of the next-generation radio-telescopes, the scalability is paramount due to the large-scale of the problem to be solved. The proposed method distributes the data measured at each frequency and processes it in parallel. We showcase the algorithm capabilities through realistic simulations.

I. INTRODUCTION

In wide-band radio-interferometry (RI), the electromagnetic signal coming from the sky is probed by an array of antennas, at multiple frequencies \( \nu_i \), and is correlated at each antenna pair, producing radio measurements \( y_{ij} \in \mathbb{C}^M \) for each band \( \nu_i \). To recover a hyper-spectral image of the sky, an ill-posed problem has to be solved, which under simplifying assumptions, can be modelled as \( Y = \Phi(X) + N \), where \( Y = (y_{ij}, \ldots, y_{ij}) \in \mathbb{C}^{M \times b} \) denotes the wide-band measured data at \( b \) bands, corrupted by additive white Gaussian noise \( N = (n_{1i}, \ldots, n_{1i}) \in \mathbb{C}^{M \times b} \) and \( X = (x_{1i}, \ldots, x_{1i}) \in \mathbb{R}^{N \times b} \) is the unknown hyper-spectral image. The linear operator \( \Phi(X) = (\Phi(x_{1i}), \ldots, \Phi(x_{1i})) \) models the acquisition process, that is an incomplete Fourier sampling.

II. CONVEX MINIMISATION PROBLEM

We assume a linear mixture model for the image cube and solve a convex minimisation problem imposing low-rankness, joint-sparsity and positivity of the image cube \( X \) [1]. We introduce multiple data fidelity terms defined for each frequency band, to achieve a high degree of parallelism. The minimisation problem can be defined as

\[
\min_X f(X) + \mu g_1(\Psi X) + g_2(X) + \sum_{i=1}^{b} h_i(\Phi_i(X)),
\]

with the functions involved: \( f = tr, D = \mathbb{R}^{N \times b} \) accounting for the positivity constraint; \( g_1(Z) = \|Z\|_2^2 \) imposing joint-sparsity in a concatenation of wavelet basis \( \Psi \); \( g_2(Z) = \|Z\|_1 \) imposing low-rankness on the desired solution; \( h_i = \|g_i, B_i = \{Z \in \mathbb{C}^{M \times b} : \|Z - Y_i\|_2 \leq \epsilon_i \} \) enforcing data fidelity by constraining the solution to belong to the \( \epsilon_i \)-balls defined by the known noise statistics. We denote with \( Y_i = (\alpha_1 x_{1i}, \ldots, \alpha_b y_{1i}) \in \mathbb{C}^{M \times b} \) the measurement matrix active only at the band \( \nu_i \) such that \( \alpha_j = 0, \forall j \neq i \). The associated linear operator is \( \Phi_i(X) = (\alpha_j \Phi(x_{1i}), \ldots, \alpha_j \Phi(x_{1i})) \) with \( \alpha_j = 0, \forall j \neq i \).

To solve (1), we use a randomised primal-dual algorithm [2] that relies on forward-backward (FB) iterations to manage the non-smooth functions. The algorithmic structure has been employed for distributed, single-band imaging [3] and for non-distributed wide-band imaging [1]. The operations are detailed in Algorithm 1. All the proximal FB steps have closed-form solutions. The proximity operator for the joint-sparsity prior is a row-wise soft-thresholding operation, for row \( k \) defined as \( (S_{\alpha}^{\epsilon_k} \Psi^\dagger(Z))_{kj} = \frac{\Psi^\dagger_{kj} - \alpha}{\|\Psi^\dagger_{kj}\|} \) if \( \|\Psi^\dagger_{kj}\| > \alpha \) and \( (S_{\alpha}^{\epsilon_k} \Psi^\dagger(Z))_{kj} = 0 \) otherwise. The nuclear norm produces the soft-thresholding of the eigenvalues of \( Z, S_k^\dagger(Z) = \mathbf{H}_k S_k^\dagger(\Sigma) \mathbf{H}_k^\dagger \). Data fidelity is enforced by the projections \( P_{\Phi_i} \) onto the \( \epsilon_i \) sized \( \ell_2 \) balls, for each band and positivity is imposed via the projection \( P_{\Phi_i} \) on the positive orthant \( D \).

Algorithm 1 Randomised PD for distributed WB RI

\[
\begin{align*}
given X^{(0)}, X^{(0)}, V_1^{(0)}, V_2^{(0)}, U_1^{(0)}, \ldots, U_b^{(0)}, \mu, \sigma_1, \sigma_2, \sigma_3 & \text{  } \text{repeat for } i = 1, \ldots \text{ } \text{generate active set } A \subset \{1, \ldots, b\} \text{  } \text{do in parallel } \\
V_1^{(i)} = V_1^{(i-1)} + \Psi^\dagger(X^{(i-1)} - S_{\sigma_1}^1(X^{(i-1)})) & \text{  } \forall i \in A \text{ do in parallel } \\
V_2^{(i)} = V_2^{(i-1)} + X^{(i-1)} - S_{\sigma_2}^2(X^{(i-1)} + X^{(i-1)}) & U_1^{(i)} = U_1^{(i-1)} + P_{\Phi_1}(U_1^{(i-1)} + P_{\Phi_2}(X^{(i-1)})) \\
& \text{end } \\
X^{(i)} = P_{\Phi_i}(X^{(i-1)} - \mu(\Psi^\dagger V_1^{(i)} + \sigma_2 V_2^{(i)} + \sigma_3 \sum_{i=1}^{b} P_{\Phi_i}(U_i^{(i)}))) & \text{until convergence }
\end{align*}
\]

III. SIMULATIONS AND RESULTS

We simulate a wide-band image cube following the spectral curvature model \( x_i = x_0(\nu_i/\nu_0)^{-\gamma} + \beta \log(\nu_i/\nu_0) \), where \( x_0 \) is a 256 \times 256 sized image of a radio region in the M31 galaxy; \( \gamma \) and \( \beta \) are the spectral index maps of size \( N \) and modelled as correlated Gaussian random fields. The wide-band cube is generated for \( b = 16 \) bands in the range \([1.4, 2.8] \) GHz. The wide-band data are simulated using realistic u/v-coverages from the VLA array-configuration with \( M = 33120 \) measurements at each band and are corrupted with zero-mean Gaussian noise with an input signal-to-noise ratio (SNR) of 30 dB.

The figure reveals the SNR evolution for the different algorithms. We can see that the non-distributed primal-dual algorithm denoted by PD [1] and the distributed version PD-D exhibit comparable behaviour, reaching a SNR = 26 dB. For the proposed distributed randomised algorithm PD-DR, we fix the probability of selecting an active subset \( A \) from the full data \( Y \) to 0.5. This has the advantage of lower infrastructure and memory requirements, at the expense of an increased number of iterations to achieve convergence. Also, when compared to the approach proposed in [4] and denoted by WDCT, our proposed algorithm presents superior performance: PD-DR reaches a SNR = 25 dB that is 5 dB higher than WDCT.

REFERENCES