A distributed algorithm for wide-band radio-interferometry

A distributed algorithm for wide-band radio-interferometry

Abdullah Abdulaziz, Alexandru Onose, Arwa Dabbech 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 ν_i, and is correlated at each antenna pair, producing radio measurements $\boldsymbol{y}_i \in \mathbb{C}^M$ for each band ν_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 $\mathbf{Y} = \boldsymbol{\Phi}(\mathbf{X}) + \mathbf{N}$, where $\mathbf{Y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_b) \in \mathbb{C}^{M\times b}$ denotes the wide-band measured data at b bands, corrupted by additive white Gaussian noise $\mathbf{N} = (\boldsymbol{n}_1, \dots, \boldsymbol{n}_b) \in \mathbb{C}^{M\times b}$ and $\mathbf{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_b) \in \mathbb{R}_+^{N\times b}$ is the unknown hyper-spectral image. The linear operator $\boldsymbol{\Phi}(\mathbf{X}) = \left([\boldsymbol{\Phi}_i \boldsymbol{x}_i]_{i=1:b}\right)$ 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, join-sparsity and positivity of the image cube $\mathbf{X}$ [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

$$\min_{\mathbf{X}} f(\mathbf{X}) + \mu g_1(\boldsymbol{\Psi}^\dagger \mathbf{X}) + g_2(\mathbf{X}) + \sum_{i=1}^{b} h_i\left(\overline{\boldsymbol{\Phi}}_i(\mathbf{X})\right), \qquad (1)$$

with the functions involved: $f = \iota_{\mathcal{D}}, \mathcal{D} = \mathbb{R}_+^{N\times b}$ accounting for the positivity constraint; $g_1(\mathbf{Z}) = \|\mathbf{Z}\|_{\ell_{2,1}}$ imposing joint-sparsity in a concatenation of wavelet basis $\boldsymbol{\Psi}$; $g_2(\mathbf{Z}) = \|\mathbf{Z}\|_*$ imposing low-rankness onto the desired solution; $h_i = \iota_{\mathcal{B}_i}, \mathcal{B}_i = \{\mathbf{Z} \in \mathbb{C}^{M\times b} : \|\mathbf{Z} - \mathbf{Y}_i\|_F \leq \epsilon_i\}$ enforcing data fidelity by constraining the solution to belong to the ϵ_i-balls defined by the known noise statistics. We denote with $\mathbf{Y}_i = (\alpha_1 \boldsymbol{y}_1, \dots, \alpha_b \boldsymbol{y}_b) \in \mathbb{C}^{M\times b}$ the measurement matrix active only at the band ν_i such that $\alpha_j = 0, \forall j \neq i$. The associated linear operator is $\overline{\boldsymbol{\Phi}}_i(\mathbf{X}) = ([\alpha_j \boldsymbol{\Phi}_i \boldsymbol{x}_i]_{i=1:b})$ with $\alpha_j = 0, \forall j \neq i$.

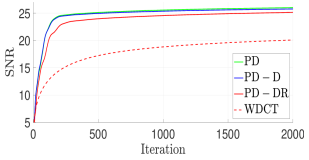
To solve (1), we use a randomised primal-dual algorithm [2] that relies on forward-backward (FB) iterations to manage the non-smooth functions. The algorithmic structure has been employed for distributed, single-band imaging [3] and for non-distributed wide-band imaging [1]. The operations are detailed 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 $(\boldsymbol{\mathcal{S}}_\alpha^{\ell_{2,1}}(\mathbf{Z}))_{k,:} = \frac{\bar{\boldsymbol{z}}(\|\bar{\boldsymbol{z}}\|_{\ell_2} - \alpha)}{\|\bar{\boldsymbol{z}}\|_{\ell_2}}$ if $\|\bar{\boldsymbol{z}}\|_{\ell_2} > \alpha$ and $(\boldsymbol{\mathcal{S}}_\alpha^{\ell_{2,1}}(\mathbf{Z}))_{k,:} = 0$ otherwise. The nuclear norm produces the soft-thresholding of the eigenvalues of $\mathbf{Z}$, $\boldsymbol{\mathcal{S}}_\alpha^*(\mathbf{Z}) = \mathbf{H}_1 \boldsymbol{\mathcal{S}}_\alpha^{\ell_1}(\boldsymbol{\Sigma})\mathbf{H}_2^\dagger$. Data fidelity is enforced by the projections $\boldsymbol{\mathcal{P}}_{\mathcal{B}_i}$ onto the ϵ_i sized ℓ_2 balls, for each band and positivity is imposed via the projection $\boldsymbol{\mathcal{P}}_{\mathcal{D}}$ onto the positive orthant $\mathcal{D}$.

Algorithm 1 Randomised PD for distributed WB RI.

given $\mathbf{X}^{(0)}, \tilde{\mathbf{X}}^{(0)}, \mathbf{V}_1^{(0)}, \mathbf{V}_2^{(0)}, \mathbf{U}_1^{(0)}, \dots, \mathbf{U}_b^{(0)}, \mu, \tau, \sigma_1, \sigma_2, \sigma_3$
repeat for $t = 1, \dots$
  generate active set $\mathcal{A} \subset \{1, \dots, b\}$
  do in parallel
    $\mathbf{V}_1^{(t)} = \boldsymbol{V}_1^{(t-1)} + \boldsymbol{\Psi}^\dagger \tilde{\mathbf{X}}^{(t-1)} - \boldsymbol{\mathcal{S}}_{\mu/\sigma_1}^{\ell_{2,1}}\left(\mathbf{V}_1^{(t-1)} + \boldsymbol{\Psi}^\dagger \tilde{\mathbf{X}}^{(t-1)}\right)$
    $\mathbf{V}_2^{(t)} = \mathbf{V}_2^{(t-1)} + \tilde{\mathbf{X}}^{(t-1)} - \boldsymbol{\mathcal{S}}_{1/\sigma_2}^{*}\left(\mathbf{V}_2^{(t-1)} + \tilde{\mathbf{X}}^{(t-1)}\right)$
    $\forall i \in \mathcal{A}$ **do in parallel**
      $\mathbf{U}_i^{(t)} = \mathbf{U}_i^{(t-1)} + \overline{\boldsymbol{\Phi}}_i\left(\tilde{\mathbf{X}}^{(t-1)}\right) - \boldsymbol{\mathcal{P}}_{\mathcal{B}_i}\left(\mathbf{U}_i^{(t-1)} + \overline{\boldsymbol{\Phi}}_i\left(\tilde{\mathbf{X}}^{(t-1)}\right)\right)$
    end
  end
  $\mathbf{X}^{(t)} = \boldsymbol{\mathcal{P}}_{\mathcal{D}}\left(\mathbf{X}^{(t-1)} - \tau\left(\sigma_1 \boldsymbol{\Psi}\mathbf{V}_1^{(t)} + \sigma_2 \mathbf{V}_2^{(t)} + \sigma_3 \sum_{i=1}^{b} \overline{\boldsymbol{\Phi}}_i^\dagger\left(\mathbf{U}_i^{(t)}\right)\right)\right)$
  $\tilde{\mathbf{X}}^{(t)} = 2\mathbf{X}^{(t)} - \mathbf{X}^{(t-1)}$
until convergence

III. Simulations and results

We simulate a wide-band image cube following the spectral curvature model $\mathbf{x}_i = \mathbf{x}_0 \left(\nu_i/\nu_0\right)^{-\boldsymbol{\gamma} + \boldsymbol{\beta}\log(\nu_i/\nu_0)}$, where $\mathbf{x}_0$ is a 256×256 sized image of a radio region in the M31 galaxy; $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ 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 uv-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 $\mathcal{A}$ from the full data $\mathbf{Y}$ 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

[1] A. Abdulaziz, A. Dabbech, A. Onose, and Y. Wiaux, in *Proceedings of the 24th EUSIPCO*, 2016, pp. 388–392.
[2] J.-C. Pesquet and A. Repetti, *J. Nonlinear Convex Anal.*, vol. 16, no. 12, pp. 2453–2490, 2015.
[3] A. Onose, R. E. Carrillo, A. Repetti, J. D. McEwen, J.-P. Thiran, J.-C. Pesquet, and Y. Wiaux, *MNRAS*, vol. 462, no. 4, pp. 4314–4335, 2016.
[4] A. Ferrari, J. Deguignet, C. Ferrari, D. Mary, A. Schutz, and O. Smirnov, *arXiv preprint arXiv:1504.06847*, 2015.