Why CLEAN when you can PURIFY? A new approach for next-generation radio-interferometric imaging

Rafael E. Carrillo*, Jason D. McEwen†, and Yves Wiaux‡

* Institute of Electrical Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.
† Mullard Space Science Laboratory, University College London (UCL), Holmbury St Mary, Surrey RH5 6NT, UK.
‡ Institute of Sensors, Signals, and Systems, Heriot-Watt University, Edinburgh EH14 4AS, UK.

Abstract—In recent works, sparse models and convex optimization techniques have been applied to radio-interferometric (RI) imaging showing the potential to outperform state-of-the-art imaging algorithms in the field. In this talk, I will review our latest contributions in RI imaging, which leverage the versatility of convex optimization to both handle realistic continuous visibilities and offer a highly parallelizable structure paving the way to high-dimensional data scalability. Firstly, I will review our recently proposed average sparsity approach, SARA, which relies on the observation that natural images exhibit strong average sparsity over multiple coherent bases. Secondly, I will discuss efficient implementations of SARA, and sparse regularization problems in general, for large-scale imaging problems in a new toolbox dubbed PURIFY.

The advent of next-generation radio telescopes, such as the new LOw Frequency ARray (LOFAR), the recently upgraded Karl G. Jansky Very Large Array (VLA) and the future Square Kilometer Array (SKA), has posed several challenges for image reconstruction and the design of data processing systems [1]. The new telescopes will achieve much higher dynamic range than current instruments at a higher angular resolution. Also, these telescopes will acquire a massive amount of data, thus posing large-scale inverse problems in the perspective of image reconstruction. These challenges have triggered an intense research in the community to reformulate imaging and calibration techniques for radio interferometry (RI).

The RI measurement equation can be discretized as $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{y} \in \mathbb{C}^M$ denotes the vector of measured visibilities, $\boldsymbol{\Phi} \in \mathbb{C}^{M \times N}$ is a discretization of the measurement operator and $\boldsymbol{n} \in \mathbb{C}^M$ represents the observation noise. In [2] we propose an imaging algorithm dubbed sparsity averaging reweighted analysis (SARA) based on average sparsity over multiple bases, showing superior reconstruction qualities relative to state-of-the-art imaging methods in the field. A sparsity dictionary composed of a concatenation of q coherent bases, $\boldsymbol{\Psi} = [\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_q]$, is used and average sparsity is promoted through the minimization of an analysis ℓ_0 prior, $\|\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{x}}\|_0$, where $\boldsymbol{\Psi}^\dagger$ denotes the adjoint operator of $\boldsymbol{\Psi}$ [3].

SARA adopts a reweighted ℓ_1 minimization scheme to promote average sparsity through the prior $\|\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{x}}\|_0$. The algorithm replaces the ℓ_0 norm by a weighted ℓ_1 norm and solves a sequence of weighted ℓ_1 problems where the weights are essentially the inverse of the values of the solution of the previous problem [2], [3]. The weighted ℓ_1 problem is defined as:

$$\min_{\bar{\boldsymbol{x}} \in \mathbb{R}_+^N} \|\mathsf{W}\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{x}}\|_1 \text{ subject to } \|\boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}}\|_2 \leq \epsilon, \quad (1)$$

where $\mathsf{W} \in \mathbb{R}^{D \times D}$ denotes the diagonal matrix with positive weights, $\mathbb{R}_+^N$ denotes the positive orthant in $\mathbb{R}^N$ and ϵ is an upper bound on the ℓ_2 norm of the noise, which can be accurately estimated. Hence, we focus our attention on solving problem (1) efficiently, especially for large-scale data problems, i.e. when the number of visibilities is very large ($M \gg N$). In this case, we propose to split the data vector $\boldsymbol{y}$

and the measurement operator into R blocks in the following manner:

$$\boldsymbol{y} = [\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_R^T]^T \text{ and } \boldsymbol{\Phi} = [\boldsymbol{\Phi}_1^T, \ldots, \boldsymbol{\Phi}_R^T]^T, \quad (2)$$

where each $\boldsymbol{y}_i$ is modelled as $\boldsymbol{y}_i = \boldsymbol{\Phi}_i\boldsymbol{x} + \boldsymbol{n}_i$ and $\boldsymbol{n}_i$ denotes the noise vector. With this partition the optimization problem in (1) can be reformulated as

$$\min_{\bar{\boldsymbol{x}} \in \mathbb{R}_+^N} \|\mathsf{W}\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{x}}\|_1 \text{ subject to } \|\boldsymbol{y}_i - \boldsymbol{\Phi}_i\bar{\boldsymbol{x}}\|_2 \leq \epsilon_i, \ i = 1, \ldots, R, \quad (3)$$

where each ϵ_i is an appropriate bound for the ℓ_2 norm of $\boldsymbol{n}_i$.

In [4] we propose a general algorithmic framework based on the simultaneous-direction method of multipliers (SDMM) to solve (3). The proposed framework offers a parallel implementation structure that decomposes the original problem into several small simple problems, hence allowing implementation in multicore architectures or in computer clusters, or on graphics processing units. These implementations provide both flexibility in memory requirements and a significant gain in terms of speed, thus enabling scalability to large-scale problems. A beta version of an SDMM-based imaging software written in C and dubbed PURIFY was released that handles various sparsity priors, including SARA, thus providing a new powerful framework for RI imaging (toolbox available at http://basp-group.github.io/purify/). Even though this beta version of PURIFY is not parallelized yet, we discuss in detail the extraordinary parallel and distributed optimization potential of SDMM to be exploited in future versions. We also discuss other possible research avenues for big data scalability. One possibility is to incorporate ideas from stochastic gradient methods into proximal splitting and augmented Lagrangian methods. The key idea is to use only one data block $\boldsymbol{y}_i$, or a subset of blocks, at each iteration of the reconstruction. By doing so, the computational complexity per iteration will be reduced. Thus, the total processing time of the algorithm will also be reduced if the convergence rate of the original problem is preserved. See for example [5] and references therein for first theoretical results.

REFERENCES

[1] S. Wijnholds, A.-J. van der Veen, F. De Stefani, E. La Rosa, and A. Farina, "Signal processing challenges for radio astronomical arrays," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 5382–5386.

[2] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (SARA): a novel algorithm for radio-interferometric imaging," *MNRAS*, vol. 426, no. 2, pp. 1223–1234, 2012.

[3] R. E. Carrillo, J. D. McEwen, D. V. D. Ville, J.-P. Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging," *IEEE Signal Process. Letters*, vol. 20, no. 6, pp. 591–594, 2013.

[4] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "PURIFY: a new approach to radio-interferometric imaging," *MNRAS*, vol. 439, no. 4, pp. 3591–3604, 2014.

[5] S. Azadi and S. Sra, "Towards an optimal stochastic alternating direction method of multipliers," in *Proceedings of the 31st International Conference on Machine Learning*, 2014, pp. 620–628.