A multilevel framework for accelerating uSARA in radio-interferometric imaging
Guillaume Lauga, Audrey Repetti, Elisa Riccietti, Nelly Pustelnik, Paulo Gonçalves, Yves Wiaux

To cite this version:
Guillaume Lauga, Audrey Repetti, Elisa Riccietti, Nelly Pustelnik, Paulo Gonçalves, et al.. A multilevel framework for accelerating uSARA in radio-interferometric imaging. 2024. hal-04509285

HAL Id: hal-04509285
https://inria.hal.science/hal-04509285
Preprint submitted on 19 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A multilevel framework for accelerating uSARA in radio-interferometric imaging

Guillaume Lauga†, Audrey Repetti‡**, Elisa Riccietti†, Nelly Pustelnik*, Paulo Gonçalves†, Yves Wiaux**

†Univ Lyon, Inria, EnsL, UCBL, CNRS, LIP, UMR 5668, F-69342
*Univ Lyon, CNRS, Laboratoire de Physique, F-69342
‡Department of Actuarial Mathematics & Statistics, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom
** Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

{guillaume.lauga,elisa.riccietti,nelly.pustelnik,paulo.goncalves}@ens-lyon.fr ; {a.repetti,y.wiaux}@hw.ac.uk

Abstract—This paper presents a multilevel algorithm specifically designed for radio-interferometric imaging in astronomy. The proposed algorithm is used to solve the uSARA (unconstrained Sparsity Averaging Reweighting Analysis) formulation of this image restoration problem. Multilevel algorithms rely on a hierarchy of approximations of the objective function to accelerate its optimization. In contrast to the usual multilevel approaches where this hierarchy is derived in the parameter space, here we construct the hierarchy of approximations in the observation space. The proposed approach is compared to a reweighted forward-backward procedure, which is the backbone iteration scheme for solving the uSARA problem.

Index Terms—Multilevel, proximal methods, radio-interferometry, dimensionality reduction, astronomy

I. INTRODUCTION

Radio-interferometric (RI) imaging aims to reconstruct a sky brightness distribution from noisy Fourier observations (named visibilities). The Fourier frequency sample distribution is dictated by the position of the antennas used to probe the sky. The number of antennas within an interferometer being finite, this leads to undersampling the Fourier measurements (see Figure 1). Obtaining the target image requires to deconvolve and denoise the Fourier measurements.

The goal is thus to reconstruct the intensity image \( \mathbf{\tau} \in \mathbb{R}^n \) from a set of measured complex visibilities \( \mathbf{y} \in \mathbb{C}^m \) acquired in the Fourier space. The corresponding discretized forward model can be formulated as follows [2]:

\[
\mathbf{y} = \mathbf{G} \mathbf{F} \mathbf{Z} \mathbf{\tau} + \mathbf{\epsilon} := \mathbf{\Phi} \mathbf{\tau} + \mathbf{\epsilon}
\]

(1)

where \( \mathbf{G} \in \mathbb{C}^{m \times d} \) is a sparse interpolation operator that maps the Fourier coefficients on a regular grid to the non-uniformly located visibilities, \( \mathbf{F} \in \mathbb{C}^{d \times d} \) is the 2D Discrete Fourier Transform, and \( \mathbf{Z} \in \mathbb{R}^{d \times n} \) is a zero-padding operator to properly map \( \mathbf{\tau} \) for the convolution performed through the operator \( \mathbf{G} \). The term \( \mathbf{\epsilon} \) stands for a centered white Gaussian noise. Typically \( m \gg n \) and is of the order of 10 millions of visibilities (see [7] for more details).

This problem presents two main challenges. First, it is severely ill-posed, hence requiring advanced imaging techniques. Second, the size of the data streams coming from radio-telescopes is expected to be ever increasing, thus raising the challenge of designing highly scalable methods.

Recovery techniques in RI – CLEAN [9], [10] is the most used algorithm in RI imaging. It is similar to a matching pursuit method, but shows limitations when probing extended complex emissions, or for large numbers of point sources [11]. Penalized variational procedures have hence been proposed to improve the quality of the reconstructed images in a more general framework [7], [6]. The state-of-the-art variational formulation in RI is SARA [2], promoting average sparsity of the solution in a concatenation of bases through a reweighted-\( \ell_1 \) procedure. Specifically, it aims at minimizing a log-sum prior by solving a sequence of weighted \( \ell_1 \) problems [3], [12], [13]. Two SARA formulations have been proposed for imaging [2]–[4]: a constrained and an unconstrained one aka uSARA, on which we will focus in this work.

Scaling to high-dimensional data – Many algorithms have been proposed to reduce the computational load induced by the number of visibilities. Most often, these algorithms consider only a subset of visibilities at each iteration. A first idea is to split the visibilities into blocks and to parallelize the action of \( \mathbf{\Phi} \) blockwise [6]. In [14], this approach was also extended to multi-spectral data with a modification of SARA to take into account spectral correlations. Other approaches solve approximations of the original problem in smaller dimensions, by selecting relevant visibilities in a sketching fashion (suitable random projections for instance) [15] or updating only with a fraction of the visibilities at each optimization step in an online manner [16]. Finally acceleration schemes such as preconditioning strategies can be considered to better take into account the specific RI Fourier distribution [17].

Contributions – We leverage a recent work [18] of the authors to design an accelerated algorithm that combines a multilevel (ML) procedure with FISTA iterations [19], [20] for solving the uSARA problem. ML algorithms have been shown to significantly accelerate the resolution of problems whose structure endow a hierarchy of functions defined on nested subspaces of the parameters space, able to approximate the objective function [18], [21]–[24]. In this work,
we propose an original ML framework adapted to RI, where lower dimensional nested subspaces are constructed in the visibility domain, rather than in the image domain as usually done for ML approaches. We show through simulations that the resulting ML scheme yields significant acceleration when solving the uSARA problem.

II. THE MULTILEVEL FRAMEWORK

Inverse problems of the form of (1) can be solved by defining iterations to
\[
\min_{x \in \mathbb{R}^n} F(x) := \frac{1}{2} \|\Phi x - y\|_2^2 + R(x),
\]
where \( R : \mathbb{R}^n \to [-\infty, +\infty] \) is a regularization function incorporating prior information on the target solution.

Without loss of generality, we present the proposed ML strategy on a two-level case. In this setting, we index the functions at the coarse level with subscript \( H \), i.e. \( F_H, L_H \) and \( R_H \) for \( F, L \), and \( R \), respectively.

Spirit of the method – Given an objective function at the fine level \( F \), the goal of ML approaches is to build a coarse approximation \( F_H \) of \( F \), which is cheaper to optimize, to accelerate the minimization of \( F \). Then, a ML algorithm consists of alternating iterations at the coarse level on \( F_H \) (ML steps) and at the fine level on \( F \). Within the ML framework developed in [18], the minimization of (2) at the fine level can be computed using either forward-backward iterations, the sequence (4)
\[
\frac{1}{2}\|\Phi x - y\|_2^2 + R(x) \tag{1}
\]

is standard in the sketching literature, where \( S \) is typically a Gaussian random matrix so that minimizing (4) guarantees signal recovery [23]. Such a strategy has however many drawbacks for RI imaging that were investigated in [15]. Notably operator \( S \Phi \) is dense and thus computationally intensive to apply in iterative optimization. In general, choosing \( S \) to reduce computation complexity without sacrificing reconstruction accuracy is challenging, and some choices may lead to sub-optimal reconstruction [15]. However in our framework, \( L_H \) is only used to propel the minimization of the fine level objective function. We propose to sub-sample the \( u-v \) coverage, that will enable preserving the reconstruction quality while reducing the computation complexity of the overall minimization method. This choice will be further discussed in Section II-C.

III. ML APPROACH FOR USARA ACCELERATION

A. uSARA approach in a nutshell

The uSARA problem can be written as in (2), where \( R \) corresponds to a log-sum penalization to promote sparsity in the concatenation of the first eight Daubechies wavelet bases and the Dirac basis. Such a regularization is then handled using a reweighted \( l_1 \) approach, that aims at solving a sequence of weighted \( l_1 \) problems [3], [4], [12]. The resulting reweighting procedure can then be written as
\[
\begin{align*}
& \text{for } i = 0, \ldots, I \\
& \quad \bar{x}_{i+1} \in \text{Argmin}_{x \in \mathbb{R}^n} F_i(x) := L(x) + R(x, W_i), \\
& \quad W_{i+1} = \text{Diag} \left( \lambda(x) + |\Psi^* \bar{x}_{i+1}|^{-1} \right),
\end{align*}
\]

where \( I > 0 \) is the maximum number of reweighting steps, \( R(x, W_i) = \| \Psi^* x \|_1 + \tau R (x) \), with \( \Psi \) being the SARA dictionary, and \( \tau_{\mathbb{R}^+} \) being the indicator function associated with the positive orthant, \( \lambda > 0 \) is a regularization parameter balancing the contribution of the regularization and the data fidelity terms, and \( \rho > 0 \) ensures stability of the method. As \( \rho \) tends to 0 the solution of the weighted \( l_1 \)-norm problem approaches that of the \( l_0 \) pseudo-norm problem.

It has been shown in [13] that when solving approximately the minimization problem in (3) with a fixed number of forward-backward iterations, the sequence \((\bar{x}_i)_{i \in \mathbb{N}}\) converges to a critical point of \( F \) in (2).

B. Proposed IML-FISTA for uSARA

In the context of RI imaging we adapt the inexact ML FISTA (IML-FISTA) proposed in [18] for minimizing \( F_i \) at each reweighting step \( i \in \{0, \ldots, I\} \). Then, at the \( i \)-th reweighting step, the proposed IML-FISTA iterations for uSARA read
\[
\begin{align*}
& \text{for } k = 0, 1, \ldots \\
& \quad \tau_k = \text{ML}(z_k) \\
& \quad x_{k+1} \approx \text{prox}_{\tau R(\cdot, W_i)} \left( z_k - \tau \nabla L(z_k) \right), \\
& \quad z_{k+1} = x_{k+1} + \alpha_k (x_{k+1} - x_k),
\end{align*}
\]

where \( S \colon C^m \to C^{m_H} \) is an operator reducing the data dimensionality \( m \) to a lower dimension \( m_H < m \). Note that \( L_H : \mathbb{R}^n \to \mathbb{R} \) is defined on the same space \( \mathbb{R}^n \) as \( L \), thus information transfer operators between levels, commonly used in standard multilevel algorithms [18], [21]–[24], are not required in the proposed setting.
where \( \tau > 0 \) and \((\alpha_k)_{k \in \mathbb{N}}\) are chosen according to [26], and the approximation errors on the proximal operator are assumed to be summable [18, 26]. The multilevel (ML) step consists of updating the variable \( z_k \) at certain iterations with a correction from coarse models to obtain a better update \( z_k \)

The detailed version of this step and the variables involved are presented in Algorithm 1.

**Algorithm 1 ML step in (6)**

```plaintext
if Coarse correction at iteration \( k \) then
  Set \( \tau_H > 0 \) and \( \alpha_H > 0 \) according to [26]
  \[ z_+ = (\text{Id} - \tau_H \nabla F_H) \circ \ldots \circ (\text{Id} - \tau_H \nabla F_H)(z_k) \]
  \[ z_k = z_k + \alpha_H (z_+ - z_k) \]
else
  \[ z_k = z_k \]
end if
```

At iteration \( k \) of algorithm (6) the coarse objective function \( F_H \) is given by

\[ F_H = L_H + R_{H,\gamma}(\cdot, W_i) + \langle v_H, \cdot \rangle, \]  

(7)

where

\[ v_H = \nabla (L + R_{\gamma}(\cdot, W_i) - L_H - R_{H,\gamma}(\cdot, W_i))(z_k). \]  

(8)

In [8], \( R_{\gamma}(\cdot, W_i) \) corresponds to a smooth approximation of \( R(\cdot, W_i) \) with parameter \( \gamma > 0 \) [27, Definition 2.1], obtained using a Moreau-Yosida smoothing technique (see [18, 27] for details). The advantage of this technique is that the gradient has a closed form expression. Similarly, \( R_{H,\gamma} \) is a smooth approximation of the coarse approximation \( R_H \), built using the same technique.

In [7], \( v_H \) imposes first-order coherence between the smoothed versions of the functions at fine and coarse levels. According to [18, Thm 2.15 and Thm 2.16], we then have the following theoretical guarantees:

**Theorem 1:** Let \( i \in \{1, \ldots, I\} \). Let \((x_k)_{k \in \mathbb{N}}\) and \((z_k)_{k \in \mathbb{N}}\) be sequences generated by algorithm (6). Assume that, for every \( k \in \mathbb{N} \), the coarse model defined in Algorithm 1 decreases, i.e. \( F_H(z_+) \leq F_H(z_k) \). Then, the following assertions hold:

1. \((F(x_k) - F^*)_{k \in \mathbb{N}}\) is decreasing at a rate of \( 1/k^2 \)
2. \((x_k)_{k \in \mathbb{N}}\) converges to a minimizer of \( F \) when \( k \to \infty \).

**C. Algorithmic settings and implementation**

**Proximity operator computation** – In [6], the proximity operator of \( R(\cdot, W_i) \) is defined, for every \( x \in \mathbb{R}^n \), as

\[ \text{prox}_{\tau R(\cdot, W_i)}(x) = \arg \min_{u \in \mathbb{R}^n} \frac{1}{2\tau} \| u - x \|^2 + \| \psi u \|_1. \]

Since this proximity operator does not have a closed form expression, it can be computed with sub-iterations. In particular, the dual forward-backward algorithm proposed in [28] produces a sequence of feasible iterates converging to \( \text{prox}_{\tau R(\cdot, W_i)}(x) \) [28, Thm 3.7].

\[ \text{This is ensured as soon as } \tau_H < \beta_H^{-1}, \text{ where } \beta_H > 0 \text{ is the Lipschitz constant of } \nabla F_H. \]

Fig. 1. \( u - v \) coverage for the fine (in blue) and coarse level (in red) when using the MeerKAT telescope [29].

**Construction of \( S \)** – To demonstrate the potential of the proposed IML-FISTA for RI imaging, we choose \( S \) in [4] to select low-frequency coefficients in the Fourier \( u - v \) coverage, and preserving the ellipses arcs (i.e. corresponding to antenna pairs selecting low-frequency components). An example with a \( u - v \) coverage simulated from a subset of 64 antennas of the MeerKAT telescope [29] is displayed in Figure 1 where \( S \) selects \( m/2 \) coefficients (in red) out of the \( m = 10,080,000 \) total observations. In other words we keep the visibilities produced by pair of antennas with the smallest distance to each others in the physical world. This choice is also based on the fact that most of the signal energy is usually concentrated around low frequencies [2] to accelerate the reconstruction of the image, a common idea in RI imaging [17].

Formally \( S \) is a sub-sampling operator that selects a subset \( J \subseteq \{1, \ldots, m\} \) of the available visibilities. We then construct \( \Phi_H = S \Phi \) to map the DFT of the image \( x \) to \( S y \).

**Choice of coarse model** – We choose \( R_{H,\gamma} = 0 \), i.e. the coarse level is not regularized explicitly. This choice is due to the fact that through \( S \) we are only working with low-frequencies, and we observed that adding a coarse regularization in this case was not making a quantitative difference in practice. Thus, the coarse objective function \( F_H \) boils down to

\[ (\forall x \in \mathbb{R}^n) \quad F_H(x) := \frac{1}{2} \| S \Phi x - S y \|^2 + \langle v_H, x \rangle. \]  

(9)

Hence, the coarse model is still guided by the fine level regularization through \( v_{H,k} \).

Regarding the smoothing of \( R(\cdot, W_i) \) in (8), we choose to only smooth the SARA weighted-\( \ell_1 \) regularization without enforcing the first order coherence with respect to \( i_{R_1^T} \cdot \). In practice we have not observed unfeasible coarse iterates.

**Multilevel setup** – For all of the performed experiments, our ML algorithm will consist of 3 levels. Each coarse level has half of the available measurements of its corresponding fine level. At each level we perform \( p = 5 \) iterations of gradient descent in algorithm [1] with \( F_H \) given in (9).

IV. NUMERICAL EXPERIMENTS

**A. Simulated data**

We use a subset of 64 antennas from the MeerKAT array [29]. Each antenna pair acquires 5,000 visibilities, leading to
In this section we will compare three optimization methods for solving Equation (2): FB, FISTA, and IML-FISTA. Each algorithm is given a budget of CPU time to reach the best reconstruction (λ is chosen via grid search). Our main goal is to demonstrate that IML-FISTA is faster than FISTA to solve this problem. First and foremost we are interested in the quality of the reconstruction so we will plot two criteria.

We now focus on solving the complete uSARA problem. As we solve a sequence of optimization problems (5) that will be different for each optimization method, the easiest way to evaluate each method is to only compare FB, FISTA and IML-FISTA on the SNR evolution with respect to the CPU time computation. The results are shown in Figure 2-right. One can see that at each "reweighting" step a small jump in the SNR of the iterates occurs (for both FISTA and IML-FISTA) due to the reset of the inertia parameters (for convergence reasons [26]). With the given coverage we can only slightly improve the SNR of the reconstruction, but nevertheless IML-FISTA reaches an upper bound faster than FISTA.

**V. CONCLUSION AND PERSPECTIVES**

**Conclusion** – We proposed a ML approach for solving the uSARA problem in RI imaging, where the coarse level enables working with low-dimensional data, while the fine model ensures consistency with the full data and promotes averaging sparsity. We have also integrated the resulting IML-FISTA iterations, reminiscent from the ML approach proposed in [18].
within a reweighting framework, further enhancing sparsity. We have shown through simulations on RI imaging that the proposed IML-FISTA leads to impressive acceleration with respect to FB to solve uSARA by exploiting approximations in the observation space of the problem. Our method shows promising results on simulations when integrated within the reweighting procedure.

**Perspectives** – The proposed ML acceleration for uSARA being very promising, we have identified future research directions to better assess its potential for RI imaging.

On the one hand, there exist approaches to efficiently handle high-dimensional data based on a parallel implementation of the measurement operator $\Phi$. Our method could be coupled with such a parallelisation strategy to benefit at the same time from the dimensionality reduction at the coarse levels and from an efficient parallel implementation of $\Phi$ at the fine levels. Also, preconditioning strategies enabling natural weighting (leveraging the local density of the Fourier sampling) could be considered for comparison and/or further acceleration of the proposed method [17].

Moreover, sophisticated coarse models for RI could be investigated to potentially improve the results (for instance sketching approaches [15]).

Furthermore, connections of ML approaches with CLEAN algorithm [9] and its learned version R2D2 [30] could be studied. Both methods are built on major-minor cycles reminiscent of matching pursuit. During the minor cycles, an approximate data term is used, ultimately enabling a much smaller number of major cycles (requiring passing through the full data). This is akin to the proposed ML method.

On the other hand, a few theoretical research directions could be pursued. Leveraging approximation theory as in [13], the global convergence of the ML strategy within a reweighting framework could be studied. The ML framework could also be extended to primal-dual algorithms, to enable solving the constrained formulation of SARA, which in this RI context yields better reconstruction quality.

**References**


