Abstract— Blind source separation is one of the major analysis tools to extract relevant information from multichannel data. While being central, joint deconvolution and blind source separation (DBSS) methods are scarce. To that purpose, a DBSS algorithm coined SDecGMCA is proposed. It is designed to process data sampled on the sphere, allowing large-field data analysis in radio-astronomy.

Context

With the forthcoming large-scale radio-telescope, such as the Square Kilometer Array (SKA) standard blind source separation algorithms are faced with a key bottleneck: accounting for instrumental response calls for jointly tackling a separation and deconvolution problem. Additionally, dedicated methods must be designed to specifically address the spherical data that large arrays of radio-telescopes will produce. In this context, multichannel data are considered where the number of sources is unknown. In this work, we extend the algorithm DecGMCA [7], which is an iterative algorithm coined SDecGMCA is proposed. It is designed to process data sampled on the sphere, allowing large-field data analysis in radio-astronomy.

Methodology

In the scope of the joint DBSS problem, the objective is to estimate \( \mathbf{A} \) and \( \mathbf{S} \) from \( \mathbf{Y} \), knowing \( \mathbf{H} \) and the level of the noise \( \mathbf{N} \). The problem amounts to minimizing an objective function:

\[
\arg \min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \left( \sum_{l,m} \left\| \mathbf{Y}^{l,m} - \operatorname{diag} \left( \mathbf{H} \right) \mathbf{A} \mathbf{S}^{l,m} \right\|_2^2 + \left\| \mathbf{A} \otimes \left( \mathbf{S}^\mathbb{H} \right)^T \right\|_1 \right)
\]

The sources are assumed to be sparse in a representation \( \Phi \), hence the \( \ell_1 \)-penalization of sparsity parameters \( \mathbf{A} \). To mitigate the scale indeterminacy of the product \( \mathbf{A} \mathbf{S} \), the columns of \( \mathbf{A} \) are enforced to be on the \( \ell_2 \)-hypersphere or oblique ensemble \( \mathcal{O} \).

The problem in Eq. (1) is not convex but multiconvex, which calls for an alternate minimization according to each variable \( \mathbf{A} \) and \( \mathbf{S} \). However, traditional proximal algorithms such as the BCD [11] or the PALM [2] generally exhibit a clear lack of robustness with respect to the often spurious local critical points of the above cost function. Projected alternate least-squares (pALS) [8] has long been advocated as allowing for more robust minimization schemes [5][8]. Furthermore, pALS allows for simple and robust heuristics to fix the sparse regularization parameter \( \Lambda \) [8]. Hence, and following the architecture of DecGMCA [7], the proposed algorithm will build upon a sparsity-enforcing pALS, which iterates the following:

- **Estimation of \( \mathbf{A} \) with \( \mathbf{S} \) fixed:** Solving the least-square problem
  \[
  \mathbf{A}_c = \left( \sum_{l,m} \mathbf{Y}^{l,m} \mathbf{H} \mathbf{S}^{l,m} \right)^{-1} \left( \sum_{l,m} \mathbf{H} \mathbf{S}^{l,m} \mathbf{Y}^{l,m} \right)
  \]
  Since the number of frequencies is much greater than the number of sources \( N_s \), the matrix \( \left( \sum_{l,m} \mathbf{H} \mathbf{S}^{l,m} \mathbf{Y}^{l,m} \right) \) is well conditioned and safe to invert. The solution is then projected on the multidimensional \( \ell_2 \)-hypersphere \( \mathcal{O} \).

- **Estimation of \( \mathbf{S} \) with \( \mathbf{A} \) fixed:** The quadratic term of the joint deconvolution/separation problem is likely ill-conditioned, if not ill-posed. To alleviate this problem, an ad hoc regularization is required, which turns the following additional regularization term to
  \[
  \frac{1}{2} \sum_{l,m} c_{n,l} \left\| \mathbf{S}^{l,m} - \mathbf{c} \right\|_2^2
  \]
  are the regularization coefficients, which depend on the frequency \( l \) and on the source \( n \). In [7], these parameters were fixed to an ad hoc small value \( \epsilon \). However, these parameters largely impact the quality of the separation. In the sequel, we investigate different strategies allowing more efficient and adaptive way of tuning these key parameters.

Solving the newly formed quadratic term yields
\[
\mathbf{S}^{l,m} = \left( \mathbf{M}[l] + \left( \mathbf{c}_{\mathbf{S}_{\mathbf{n}}}[l] \right)^{-1} \mathbf{A}^T \mathbf{H} \mathbf{Y}^{l,m} \right)^{-1} \mathbf{M}[l] = \mathbf{A}^T \left( \mathbf{H} \right)^{-1} \mathbf{A}
\]
Four strategies that reduce the choice of the parameters to a single one, called the regularization hyperparameter and denoted \( c \), are considered:

- **Strategy #1** (naive strategy): the regularization parameters are chosen independently of the frequency \( l \) and the source \( n \); \( c_{n,l} = c \).
- **Strategy #2** (strategy used in DecGMCA [7]): \( c_{n,l} = c \lambda_{\max} \left( \mathbf{M}[l] \right) \), where \( \lambda_{\max} \left( \cdot \right) \) returns the greatest eigenvalue.
- **Strategy #3:** \( c_{n,l} = \max \left( 0, c - \lambda_{\min} \left( \mathbf{M}[l] \right) \right) \), where \( \lambda_{\min} \left( \cdot \right) \) returns the smallest eigenvalue. This strategy allows to limit the noise amplification to \( c \lambda_{\min} \left( \mathbf{A}^T \mathbf{A} \right) \).
- **Strategy #4** (SNR strategy): \( c_{n,l} = c / \text{SNR}_n[l] = c \sigma_{\mathbf{S}_{\mathbf{n}}}[l] / \left( \lambda_{\min} \left( \mathbf{M}[l] \right) \sigma_{\mathbf{S}_{\mathbf{n}}}[l] \right) \), where \( \sigma_{\mathbf{S}_{\mathbf{n}}}[l] \) and \( \sigma_{\mathbf{S}_{\mathbf{n}}}[\cdot] \) are the angular power spectra of the sources and the noise, respectively. This strategy, which supposes to know the angular power spectra of the sources, is reminiscent of a Wiener deconvolution filter.

The solution is then soft-thresholded in the transformed domain \( \mathbf{A} \). A \( \lambda_{\ell_1} \)-rewighting strategy is implemented [3] to adapt the thresholds to the pixel values, thus reducing the bias introduced by the soft-thresholding and improving the separation performances. The choice of the corresponding sparse regularization parameter \( \Lambda \) follows [8].

SDecGMCA is initialized using Principal Component Analysis. During the first iterations, regularization strategy #3 is used; when the estimated sources converge, the strategy is switched to #4.
The first stage allows to have a first estimation of the sources, whose angular power spectra are close enough to the ground-truth ones (warm-up). The second stage allows to refine the results, by using a more precise regularization strategy (refinement); more specifically, the regularization parameters are calculated with the angular power spectra of the sources estimated at last iteration.

SDecGMCA needs to be provided the regularization hyperparameters at warm-up $c_{wu}$ and refinement $c_{ref}$. As proposed in DecGMCA, a decrease of the warm-up regularization hyperparameter is implemented to improve the robustness of the algorithm.

Algorithm 1 SDecGMCA

Regularization strategy selection: #3 (warm-up)

For $i = 1, \ldots, i_{\text{max}}$ do
    (1) Estimate $\hat{S}$ with $A$ fixed
    - Tikhonov-penalized least-square update of $\hat{S}$
    - Soft-thresholding of $S\Phi$
    (2) Estimate $A$ with $\hat{S}$ fixed
    - Least-square update of $A$
    - Projection of $A$ on $O$
    if $S$ has converged then
        Regularization strategy selection: #4 (refinement)

Numerical experiments

We set the Healpix parameters $n_{\text{side}} = 128$ and $l_{\text{max}} = 384$. We generate toy example sources, which are sparse in the spherical starlet domain and band-limited to $l_{\text{max}}/6 = 64$. We take $N_s = 4$, $N_c = 8$ and $\text{cond}(A) = 2$. The convolution kernels are Gaussian, with resolutions evenly spread between the minimum resolution $l_{\text{max}}/8 = 48$ and $l_{\text{max}}$. The overall SNR is 10 dB. The channels are unmixed and deconvolved at the resolution of the best-resolved channel. This amounts to replacing $\hat{H}_i$ by $\tilde{H}_i / \hat{H}_{i0}$, where $\tilde{v}_b$ is the best-resolved channel number. The performance metrics employed to assess the results are: (i) the normalized mean square error $\text{NMSE} = ||H_{i0} * S^* - S||_F^2 / ||H_{i0} * S||_F^2$, with $S^*$ the ground truth sources and $S$ the estimated sources, (ii) the mixing matrix criterion $C_A = \text{mean}(A^T A^* - I)$, with $A^*$ the ground-truth mixing matrix and $A$ the estimated mixing matrix.

In order to compare the impact of the 4 regularization strategies, non-blind estimations of $S$ are performed on a wide range of SNR, $N_c$, cond($A$) and channel resolutions; the results are reported in Table 1. Unsurprisingly, strategy #4 clearly provides the best reconstruction qualities. Among the other strategies, that do not assume the sources to be known, strategy #3 achieves better results. It is mostly thanks to the non-linear max operator, which allows to keep the lower frequencies unbiased, where most of the sources energy is located. Strategy #2 gives poor results; indeed, it biases more significantly the lower frequencies than the higher ones.

Table 2 shows the mean NMSE of SDecGMCA as a function of the regularization hyperparameters. The choice of $c_{wu}$ has little impact on the NMSE. On the contrary, the selection of $c_{ref}$ is more critical. However, in a range of one order of magnitude around the optimal hyperparameter, the NMSE loss is contained. It is noted that the reconstruction errors are dominated by the deconvolution artifacts (see example Figure 1). The mean $C_A$ varies between 22.07 and 25.60 dB for the same ranges considered in Table 2. Therefore, both $c_{wu}$ and $c_{ref}$ have little impact on the quality of the estimation of $\hat{S}$.

SDecGMCA is finally compared to an optimized version of DecGMCA (SDecGMCA with strategy #2). The results are reported in Table 3. SDecGMCA performs a significant gain in NMSE and a moderate increase in $C_A$. SDecGMCA is also compared to two non-deconvolving BSS algorithms. For the latter, the data are deteriorated to a common resolution (the worse one) beforehand. They achieve poor results; indeed, crucial information is lost when the data are deteriorated.

Conclusion

We proposed an enhanced version of DecGMCA, coined SDecGMCA, extended for spherical data. We investigated in particular the regularization and proposed better suited regularization strategies. The results showed that SDecGMCA clearly outperformed DecGMCA. During the workshop, results on realistic simulation data will be presented.
References

[1] J. Bobin, J. Rapin, J.-L. Starck, and A. Larue. Sparsity and adaptivity for the blind separation of partially correlated sources. *IEEE Transactions on Signal Processing*, 63(5):1199–1213, 2015.

[2] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, pages 1–36, 2013.

[3] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted L1 minimization. *Journal of Fourier Analysis and Applications*, 14(5), 2008. 877–905.

[4] M. Chan. Hyperspectral image unmixing program code. Available on https://github.com/MichelsonChan/hyperspec_unmix, last access on 2020-02-13.

[5] N. Gillis and F. Glineur. Accelerated Multiplicative Updates and Hierarchical ALS Algorithms for Nonnegative Matrix Factorization. *Neural Computation*, 24(4):1085–1105, 2012.

[6] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *Astrophysical Journal*, 622, April 2005. 759–771.

[7] M. Jiang, J. Bobin, and Starck, J.-L. Joint multichannel deconvolution and blind source separation. *SIAM Journal on Imaging Sciences*, 10(4):1997–2021, Jan 2017.

[8] C. Kervazo, J. Bobin, C. Chenot, and F. Sureau. Use of palm for $\ell_1$ sparse matrix factorization: Difficulty and rationalization of an heuristic approach. *Digital Signal Processing*, 97, February 2020.

[9] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994.

[10] J.-L. Starck, Y. Moudden, Abrial P., and M. Nguyen. Wavelets, ridgelets and curvelets on the sphere. *Astronomy and Astrophysics*, 446(3):1191–1204, 2006.

[11] Y. Xu and W. Yin. A Block Coordinate Descent Method for Regularized Multiconvex Optimization with Applications to Nonnegative Tensor Factorization and Completion. *SIAM Journal on Imaging Sciences*, 6(3):1758–1789, 2013.