UNSUPERVISED SEPARATION OF SPARSE MULTIVALUED COMPONENTS WITH APPLICATIONS IN ASTROPHYSICS

Keywords. Sparse signal modelling, sparse representations, blind source separation, tensor decomposition, proximal algorithms.

Context. Nowadays, component separation is a key tool to analyze data in a very large range of scientific applications such as sound processing or medical signal and image analysis. Astrophysics is a particularly enlightning example : the forthcoming missions will provide a very large amount of data, which will be key to answer some of the most key questions in this field. Precisely analyzing such data further requires tackling challenging problems in signal/image processing. Amongst these problems, one of these key challenges is unsupervised component separation. In this context, the data are composed of multiple observations $\mathbf{x}_i$ (whether they are time series, images or more complex data), each of which is assumed to be composed of the combination of several elementary components $\mathbf{C_j}$:

$$\text{for each observation : } \mathbf{x_i} = \sum_j \mathbf{C}_j[i]$$

The goal of unsupervised component separation (UCS) is to blindly retrieve the components $\mathbf{C_j}$ for the observations.
Interestingly, this problem is of paramount importance in various astrophysical projects : the Planck space mission, the γ-ray space telescope Fermi, the X-ray telescopes XMM and Chandra, ground-based telescopes such as Hess or CTA, the forthcoming very large radio-interferometric telescopes such as SKA. Despite the differences between these different applications, data analysis are mathematically similar and described by component separation problems.

Method. Tackling unsupervised component separation problems requires facing two challenges :

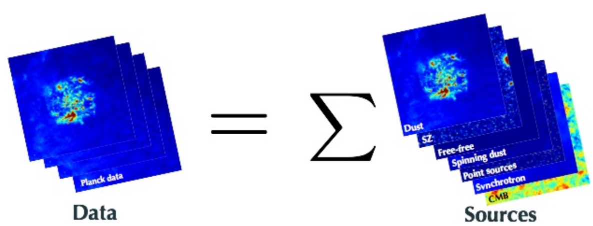

Figure 1. Planck, a component separation example. The Planck telescopes observes 9 channels at different wavelength. Component separation allows retrieving elementary components that are burried in the raw data.

- *Component modelling* an efficient and precise modelling of the components is of paramount importance to design an effective separation process. During the last few years [2,3], sparse signal modelling [?] has been showed to be a very effective framework for component separation.

  Goal 1 : the first step of this thesis will focus on the relevant signal models for astrophysical components. This includes the modeling of sparse tensor models (2D space $\times$ wavelength $\times$ time) using modern-day sparse signal representations (*e.g.* wavelet, curvelets or learned representations). Such models could include physical models or learned signal models.

- *Separation algorithm* unsupervised component separation problems are, by nature, complex non-convex problem. Furthermore, the design of an effective separation algorithm is vital to process the kind of large-scale data (*e.g.* mega-pixel images, etc) that are found in astrophysical applications.

  Goal 2 : the second step of this thesis will focus on the design of effective minimization schemes to tackle component separation problems, especially in the large-scale setting. This algorithm will build upon existing matrix factorization algorithms [6] as well are modern-day optimization techniques such as proximal algorithms [5].

Applications. The methods and algorithms developed during this thesis will be applied to real-world astrophysical applications such as the Planck space mission [4] and the X-ray Chandra data.

The candidate. The candidate must have a Master 2 degree (or equivalent) and should have a good knowledge in signal/image processing. Knowledge in convex optimization is a plus.
The PhD student will acquire an expertise in sparse image processing, multichannel data processing, machine learning and modern optimization techniques.

Contact info. The PhD will take place in the CosmoStat lab (CEA Saclay), which is a joint laboratory at the interface in computer science and astrophysics. The PhD will take an active participation in the European project LENA (http ://lena.cosmostat.org).

— *Contact :* jbobin@cea.fr
— *Lab :* CEA/IRFU in Saclay
— **Applications are expected before the 31th of April 2017.**

Bibliography

[1] P. Comon and C. Jutten, **"Handbook of blind source separation"**, Academic Press, 2010.
[2] J.Bobin *et al*, **"Sparsity and Morphological Diversity in Blind Source Separation"**, IEEE Tr. on Image Processing, 2007.
[3] J.Bobin *et al*, **"Sparsity and adaptivity for the blind separation of partially correlated sources "**, IEEE Tr. on Signal Processing, 2015.
[4] J.Bobin *et al*, **"CMB reconstruction from the WMAP and Planck PR2**

data", A&A, 2016.
[5] N. Parikh and S. Boyd, **"Proximal Algorithms "**, Foundations and Trends in Optimization, 2014.
[6] J.Rapin *et al*, **"NMF with sparse regularizations in transformed domains"**, SIAM Imaging Science, 2014.