The Sliding Frank-Wolfe Algorithm for the BLASSO
Quentin Denoyelle, Vincent Duval, Gabriel Peyré, Emmanuel Soubies

To cite this version:
Quentin Denoyelle, Vincent Duval, Gabriel Peyré, Emmanuel Soubies. The Sliding Frank-Wolfe Algorithm for the BLASSO. Workshop on Signal Processing with Adaptative Sparse Structured Representations - SPARS 2019, Sep 2019, Toulouse, France. hal-03012568

HAL Id: hal-03012568
https://hal.science/hal-03012568
Submitted on 24 Nov 2020

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.
The Sliding Frank-Wolfe Algorithm for the BLASSO

Quentin Denoyelle
EPFL, Biomedical Imaging Group
Email: quentin.denoyelle@epfl.ch

Vincent Duval
INRIA Paris, MOKAPLAN
Email: vincent.duval@inria.fr

Gabriel Peyré
CNRS, ENS
Email: gabriel.peyre@ens.fr

Emmanuel Soubies
CNRS, IRIT
Email: emmanuel.soubies@irit.fr

Abstract—This paper showcases the Sliding Frank-Wolfe (SFW), which is a novel optimization algorithm to solve the BLASSO sparse spikes super-resolution problem. The BLASSO is the continuous (i.e. off-the-grid or grid-less) counterpart of the well-known $\ell^1$ sparse regularisation method (also known as LASSO or Basis Pursuit). Our algorithm is a variation on the classical Frank-Wolfe (also known as conditional gradient) method which follows a recent trend of interleaving convex optimization updates (corresponding to adding new spikes) with non-convex optimization steps (corresponding to moving the spikes). We prove theoretically that this algorithm terminates in a finite number of steps under a mild non-degeneracy hypothesis.

I. INTRODUCTION

A. Sparse Spikes Super-Resolution

The sparse spikes super-resolution problem aims at recovering an unknown point source signal from noisy measurements. For Fourier measurements, it is possible to solve the BLASSO by convex regularization. The problem is a way to solve it in a stable way by introducing a sparsity-enforcing regularization on the classical Frank-Wolfe (also known as conditional gradient) method (also known as LASSO or Basis Pursuit). Our algorithm is a variation on the classical Frank-Wolfe (also known as conditional gradient) method which follows a recent trend of interleaving convex optimization updates (corresponding to adding new spikes) with non-convex optimization steps (corresponding to moving the spikes). We prove theoretically that this algorithm terminates in a finite number of steps under a mild non-degeneracy hypothesis.

Algorithm 1 Sliding Frank-Wolfe (SFW) Algorithm

1: Initialize with $m^{[0]} = 0$ and $n = 0$.
2: for $k = 0, \ldots, n$ do
3: \hspace{1cm} $m^{[k]} = \sum_{i=1}^{N} a_i^{[k]} \delta_{x_i^{[k]}}, a_i^{[k]} \in \mathbb{R}, x_i^{[k]}$ pairwise distincts,
4: \hspace{1cm} find $x_i^{[k]} \in \mathbb{R}$ s.t.: $x_i^{[k]} = \arg \max_{x \in \mathbb{R}} |y_i^{[k]}(x)|$ where $y_i^{[k]} = \frac{1}{X} \Phi^*(y - \Phi(m^{[k]}))$.
5: \hspace{1cm} if $|y_i^{[k]}(x_i^{[k]})| \leq 1$ then
6: \hspace{2cm} $m^{[k]}$ is a solution of $\mathcal{P}_\lambda(y)$, Stop.
7: \hspace{1cm} else
8: \hspace{2cm} Get $m^{[k+1/2]} = \sum_{i=1}^{N} a_i^{[k+1/2]} \delta_{x_i^{[k+1/2]}}, a_i^{[k+1/2]} \in \mathbb{R}$ s.t. $a_i^{[k+1/2]} = \min_{a \in \mathbb{R}^N} \frac{1}{2} \|\Phi(m_{a,x}^{[k+1/2]}) - y\|_2^2 + \lambda \|a\|_1$
9: \hspace{2cm} where $x_i^{[k+1/2]} = (x_i^{[k]}, \ldots, x_i^{[k]}, x_i^{[k+1]})$.
10: \hspace{1cm} end if
11: end for

III. CONVERGENCE RESULT

Definition 1 (Admissible kernels $\varphi$). We denote by $KER^d$, the set of admissible kernels of order $k$. A function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ belongs to $KER^d$ if:

$\varphi \in C^k(\mathbb{R})$

$\varphi$ vanishes at infinity.

The next theorem gives a finite time convergence guarantee under mild assumptions that significantly improves the previously known convergence for this kind of algorithm (weak-* convergence).

Theorem 1. Suppose that $\varphi \in KER^d$, then $m_{\varphi,x} = \sum_{i=1}^{N} a_i \delta_{x_i}$, is the unique solution of $\mathcal{P}_\lambda(y)$, and that $\eta_{\varphi} = \frac{1}{X} \Phi^*(y - \Phi(m_{\varphi,x}))$ is nondegenerate, i.e. for all $x \in \mathbb{R} \setminus \bigcup_{i=1}^{N} \{x_i\}$

$|\eta_{\varphi}(x)| < 1$ and $\forall i \in \{1, \ldots, N\}$, $\eta_{\varphi}(x_i) \neq 0$. (1)

Then Algorithm 1 recovers $m_{\varphi,x}$ after a finite number of steps (i.e. there exists $k \in \mathbb{N}$ such that $m^{[k]} = m_{\varphi,x}$).

Acknowledgement

The work of Gabriel Peyré has been supported by the European Research Council (ERC project NORIA). The work of Emmanuel Soubies has been supported by the European Research Council (ERC project GlobalBioId).
Fig. 1: Values of the objective function throughout the SFW algorithm (cumulative iterations of the BFGS used to find a critical point of the non-convex problem of Step 8). The vertical black lines separate the main outer iterations of the algorithm. The measure $m_*$ represents the unique solution of the BLASSO. It is the same example as in Figure 2.

Fig. 2: Main steps of the SFW algorithm.

REFERENCES