



HAL
open science

Eventual linear convergence rate of an exchange algorithm for superresolution

Frédéric de Gournay, Pierre Weiss, Axel Flinth

► **To cite this version:**

Frédéric de Gournay, Pierre Weiss, Axel Flinth. Eventual linear convergence rate of an exchange algorithm for superresolution. SPARS, Jul 2019, TOULOUSE, France. hal-02374440

HAL Id: hal-02374440

<https://hal.science/hal-02374440>

Submitted on 3 Jan 2021

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.

Eventual linear convergence rate of an exchange algorithm for superresolution.

Axel Flinth Université de Toulouse, UPS
Frédéric de Gournay Université de Toulouse, INSA
Pierre Weiss CNRS, ITAV, IMT Toulouse

Abstract—We study an iterative discretization algorithm for solving optimization problems regularized by the total variation norm. Its design relies on ideas from the Frank-Wolfe algorithm, as well as from semi-infinite programming. For smooth regularity terms, we prove an eventual linear convergence rate guarantee.

I. EXCHANGE ALGORITHMS

For a bounded subset $\Omega \subseteq \mathbb{R}^d$, let $\mathcal{M}(\Omega)$ denote the space of Radon measures of bounded total variation $\|\cdot\|_{\mathcal{M}}$. For an operator $A : \mathcal{M}(\Omega) \rightarrow \mathbb{R}^m$ and a function $f : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\infty\}$, we consider the following optimization problem

$$\min_{\mu \in \mathcal{M}(\Omega)} \|\mu\|_{\mathcal{M}} + f(A\mu). \quad (\mathcal{P})$$

This formulation encompasses for instance the infinite-dimensional Basis Pursuit and the BLASSO, which in particular can be applied in superresolution [1], [2], [11], [4]. The dual program of (\mathcal{P}) is given by

$$\sup_{q \in \mathbb{R}^m, \|A^*q\|_{\infty} \leq 1} -f^*(q). \quad (\mathcal{D})$$

In this work, we present theoretical properties of numerically finding a solution of (\mathcal{P}) provided one follows:

Algorithm A

- 1) Given a grid $\Omega_k \subseteq \Omega$, find a solution pair (μ_k, q_k) of the (finite-dimensional) restrictions of (\mathcal{P}) and (\mathcal{D}) ;

$$\min_{\mu \in \mathcal{M}(\Omega_k)} \|\mu\|_{\mathcal{M}} + f(A\mu) \quad (\mathcal{P}_k)$$

$$\sup_{q \in \mathbb{R}^m, \sup_{x \in \Omega_k} |(A^*q)(x)| \leq 1} -f^*(q). \quad (\mathcal{D}_k)$$

- 2) Calculate the set X_k of local maximizers x of $|A^*q_k|$ with $|(A^*q_k)(x)| > 1$. Define new grid $\Omega_{k+1} := \Omega_k \cup X_k$.
- 3) Repeat until convergence.

This strategy is an example of a so-called *exchange algorithm*, applied to solve the semi-infinite program (\mathcal{D}) [9], [7], [10]. Notably, it also has interesting connections to the so-called Frank-Wolfe or Conditional gradient method [6], [5], [3].

II. EVENTUAL LINEAR CONERGENCE

It is not hard to prove that our exchange algorithm in great generality will converge towards a solution pair of (\mathcal{P}) – (\mathcal{D}) . Our main finding is that this will occur at an eventually linear rate under the following reasonable regularity assumptions:

- 1) The function f is differentiable, with an L -Lipschitz gradient

- 2) The measurement functions a_j are all C_0^2 .
- 3) (\mathcal{P}) has a unique solution $\mu^* = \sum_{i=1}^s \alpha_i^* \delta_{\xi_i}$.
- 4) The associated dual solution q^* obeys the following :
 - a) $|A^*q^*(x)| = 1$ only if $x \in \xi$
 - b) $x \mapsto |A^*q^*(x)|$ is strongly concave around each ξ_i .

Statements like 4) have been proven to be true in the important special case of $f(v) = \frac{\lambda}{2} \|v - b\|_2^2$ for a $b = A\mu_0$ for an atomic μ_0 with well-separated support, at least for large values of λ , for many classes of measurement operators A (see [4], [8]).

We may now formulate our main result. We say that the algorithm has entered a τ -*regime* if in all future iterations, only points within a distance τ from ξ will be added to the grid.

Theorem II.1. *Under the above assumptions, Algorithm A obeys:*

- 1) For each $\tau > 0$, the algorithm will enter a τ -regime after finitely many iterations.
- 2) After a finite number of iterations N , the algorithm will add exactly one point x_k^i in the proximity of each ξ_i .
- 3) For small $\tau > 0$, the algorithm will enter a τ -regime after at most $N + C \log(\tau^{-1})$ iterations. For such k , we will furthermore have $\inf(\mathcal{P}_k) \leq \inf(\mathcal{P}) + \kappa\tau^2$. That is, the algorithm will eventually converge linearly.

III. STOPPING CRITERION AND MERGING

Since the grid Ω_k will get increasingly fine around the points ξ_i , the discretized problems will inevitably become increasingly ill-conditioned, and we must hence typically stop the algorithm quite early. A computable stopping criterion is given by the following proposition.

Proposition III.1. *Assume that each local maximizer x_k^i is at a distance at most δ from the grid Ω_k . We then have*

$$\sup(\mathcal{D}_k) \leq \sup(\mathcal{D}) + C\delta^2, \quad \|q_k - q^*\|_2 \leq C_2\delta,$$

If f is strongly convex, δ is small and A obeys an additional regularity condition, μ_k will furthermore be supported on clusters K_i with $\sup_{x \in K_i} \|x - \xi_i\|_2 \leq C_3\delta$. If we merge such clusters to their centers of mass, we obtain a measure $\sum_{i=1}^s \tilde{\alpha}_i^0 \delta_{\tilde{x}_i^0}$ with

$$\sup_i \|\tilde{x}_i^0 - \xi_i\|_2 \leq C_3\delta, \quad \|\tilde{\alpha} - \alpha^*\|_2 \leq C_4\delta.$$

The last assertion of Proposition III.1 allows efficient merging operations that pave the way to postprocessing the measure.

REFERENCES

- [1] Emmanuel J Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.
- [2] Yohann De Castro and Fabrice Gamboa. Exact reconstruction using beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 395(1):336–354, 2012.
- [3] Denoyelle, Quentin, Duval, Vincent, Peyré, Gabriel and Soubies, Emmanuel. The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy. *Preprint arXiv:1811.06416*, 2018.
- [4] Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015.
- [5] Armin Eftekhari and Andrew Thompson. A bridge between past and present: Exchange and conditional gradient methods are equivalent. *arXiv preprint arXiv:1804.10243*, 2018.
- [6] Marguerite Frank and Philip Wolfe. An Algorithm for quadratic programming. *Naval Res. Logist. Quart.* 3:95-110, 1956.
- [7] Rainer Hettich and Kenneth O Kortanek. Semi-infinite programming: theory, methods, and applications. *SIAM review*, 35(3):380–429, 1993.
- [8] Clarice Poon, Nicolas Keriven, and Gabriel Peyré. Support localization and the fisher metric for off-the-grid sparse regularization. *arXiv preprint arXiv:1810.03340*, 2018.
- [9] Rembert Reemtsen. Discretization methods for the solution of semi-infinite programming problems. *Journal of Optimization Theory and Applications*, 71(1):85–103, 1991.
- [10] Rembert Reemtsen and Stephan Görner. Numerical methods for semi-infinite programming: A survey. pages 195–262, 1998.
- [11] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *IEEE transactions on information theory*, 59(11):7465–7490, 2013.
- [12] Michael Unser, Julien Fageot, and John Paul Ward. Splines are universal solutions of linear inverse problems with generalized tv regularization. *SIAM Review*, 59(4):769–793, 2017.