Off-the-grid charge algorithm for curve reconstruction in inverse problems *

Bastien Laville[1][0000−0003−2933−8171], Laure Blanc-Féraud[1], and Gilles Aubert[2]

[1] Morpheme project: Université Côte d'Azur, Inria, CNRS, France.
(bastien.laville@inria.fr, laure.blanc_feraud@i3s.unice.fr,)
[2] Université Côte d'Azur, CNRS, LJAD, France.
(gilles.aubert@univ-cotedazur.fr).

Abstract. Several numerical algorithms have been developed in the literature and employed for curves reconstruction. However, these techniques are developed within the discrete setting, namely the super-resolved image is defined on a finer grid than the observed images. Conversely, off-the-grid (or gridless) optimisation does not rely on a fine grid and offer a tractable theoretical and numerical framework. In this work, we present a gridless method accounting for the reconstruction of both open and closed curves, based on the latest theoretical development in off-the-grid curve reconstruction.

Keywords: Off-the-grid variational method · Inverse problem · Frank-Wolfe algorithm · Curve detection.

1 Introduction

This work focuses on the numerical optimisation of the functional CROC designed for inverse problems in order to recover curves in an off-the-grid fashion, by considering the space of vector Radon measures with finite divergence.

Off-the-grid methods is a rather new field of research, introduced a decade ago to overcome some limitations of so-called discrete methods. Indeed, the source estimation problem amounts to recover the source with support lying in some set $\mathcal{X}$, thanks to an altered acquisition on a coarse grid: blurred, noisy, low-passed, *etc.* In a rather classical discrete framework, the recovered source lies on a refined grid *i.e.* it is a matrix. On the contrary, *off-the-grid* (or *gridless*) methods do not rely on a grid: for instance a spike source position is continuously estimated, and cannot be bound to a pixel in a fine grid, thus not bringing any discretisation error. The source is then encoded in a *measure*, lying in a

* The work of Bastien Laville has been supported by the French government, through the UCA DS4H Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-17-EURE-0004. The work of Laure Blanc-Féraud has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

broader set of functions denoted $\mathscr{M}(\mathcal{X})$. Moreover, these gridless methods have several theoretical guarantees [8] and also add up structural information to the optimisation: geometrical information is then used to recover a certain object, on the contrary to discrete methods where it always yields a matrix. The off-the-grid literature built up around the spike reconstruction, but several other structures such as level sets [5], sinusoid [15], dynamic trajectories [3] *etc.* have likewise been explored.

However, to the best knowledge of the authors, the literature does not handle the off-the-grid curve reconstruction problem, let alone provides a tailored numerical algorithm for this specific task. The point/spike reconstruction problem, *i.e.* measure supported on a 0D set, was thoroughly explored [4,7] and the level set reconstruction problem, *i.e.* measure supported on a 2D set, has well-known theoretical results stemming from the geometrical measure theory and was lately successfully adapted to off-the-grid reconstruction [5]. Still, the numerical reconstruction of a measure supported on a 1D set, and more specifically a curve object, was yet to explore. It is all the more unfortunate as the curve structure naturally arise in many inverse problems, such as the super-resolution in biomedical imaging. In this paper, we propose a new algorithm for curve reconstruction in an off-the-grid fashion, based on the latest theoretical results [14] investigating a new regulariser to yield curve minima.

1.1 Notations

In the following, $\mathcal{X}$ is the ambient space where the positions of the objects (e.g. spikes, curves, *etc.*) live, it is a non-empty bounded open set of $\mathbb{R}^d$, hence a submanifold of dimension $d \in \mathbb{N}^*$. $\mathscr{H}_1$ denotes the 1-dimensional Hausdorff measure (see [9] for a definition).

1.2 Related works

This paper makes an extensive use of the last theoretical results on the divergence vector field measure space brought by [14], based on some strong results of [17,12]. As we aim to close the gap between spike (sort of 0D, total variation) and set (somehow 2D, bounded variation) reconstruction, we relate to state-of-art Dirac [7,4] and level set [5] Frank-Wolfe algorithm.

2 Optimisation in the space of divergence vector fields

2.1 The space of charges

We give some useful definitions and properties from the off-the-grid literature, the interested reader can take a look at the review [13].

Definition 1 (Evanescent continuous function on $\mathcal{X}$). *We call $\mathscr{C}_0(\mathcal{X}, \mathcal{Y})$ the set of evanescent continuous functions from $\mathcal{X}$ to a normed vector space $\mathcal{Y}$, namely all the continuous map $\psi : \mathcal{X} \to \mathcal{Y}$ such that :*

$$\forall \varepsilon > 0, \exists K \subset \mathcal{X} \text{ compact}, \quad \sup_{x \in \mathcal{X} \setminus K} \|\psi(x)\|_{\mathcal{Y}} \leq \varepsilon.$$

We write $\mathscr{C}_0(\mathcal{X})$ when $\mathcal{Y} = \mathbb{R}$. We now introduce:

Definition 2 (Set of Radon measures). *We denote by $\mathscr{M}(\mathcal{X})$ the set of real signed Radon measures on $\mathcal{X}$ of finite masses. It is the topological dual of $\mathscr{C}_0(\mathcal{X})$ endowed with* supremum *norm $\|\cdot\|_{\infty,\mathcal{X}}$ by the Riesz-Markov representation theorem [9]. Thus, a Radon measure $m \in \mathscr{M}(\mathcal{X})$ is a continuous linear form on functions $f \in \mathscr{C}_0(\mathcal{X})$, with the duality bracket denoted by $\langle f, m \rangle_{\mathscr{M}(\mathcal{X})} = \int_{\mathcal{X}} f \, \mathrm{d}m$.*

The space of integrable equivalent classes $\mathrm{L}^1(\mathcal{X})$ continuously injects into the Radon measure space $\mathrm{L}^1(\mathcal{X}) \hookrightarrow \mathscr{M}(\mathcal{X})$, measures are then a generalisation of functions to a broader set. An example of Radon measure is the Dirac measure δ_x, where $x \in \mathcal{X}$. Also, since $\mathscr{C}_0(\mathcal{X})$ is a normed vector space, $\mathscr{M}(\mathcal{X})$ is complete [4] if endowed with its dual norm called the total variation (TV) norm, defined for $m \in \mathscr{M}(\mathcal{X})$ by $\|m\|_{\mathrm{TV}}$.

Now, we use the latest developments in the off-the-grid literature concerning the curve reconstruction [14]. Consider the space of *vector* Radon measures:

Definition 3. *We define the set of vector Radon measures $\boldsymbol{\mathscr{M}(\mathcal{X})^2}$ as the topological dual of the space of continuous vector functions $\boldsymbol{\mathscr{C}_0(\mathcal{X})^2} \overset{\text{def.}}{=} \boldsymbol{\mathscr{C}_0(\mathcal{X}, \mathbb{R}^2)}$. The properties of the scalar case hold for the vector one, indeed $\boldsymbol{\mathscr{M}(\mathcal{X})^2}$ has a natural TV-norm denoted by $\|\cdot\|_{\mathrm{TV}^2}$, a duality bracket $\langle \cdot, \cdot \rangle_{\boldsymbol{\mathscr{M}(\mathcal{X})^2}}$,* etc.

The following results in this section only hold in dimension $d = 2$, since some tumultuous pathological cases appear when $d > 2$, see [17, Section 1.3]. We denote by div the divergence operator, understood in the distributional sense. Indeed, for all $\boldsymbol{m} \in \boldsymbol{\mathscr{M}(\mathcal{X})^2}$ and $\mathscr{C}_0^\infty(\mathcal{X})$ the space of bump functions:

$$\forall \xi \in \mathscr{C}_0^\infty(\mathcal{X}), \quad \langle \operatorname{div} \boldsymbol{m}, \xi \rangle_{\mathscr{D}'(\mathcal{X}) \times \mathscr{C}_0^\infty(\mathcal{X})} = -\langle \boldsymbol{m}, \boldsymbol{\nabla} \xi \rangle_{\boldsymbol{\mathscr{M}(\mathcal{X})^2}}.$$

A measure $\boldsymbol{m}$ is *of finite divergence* if $\operatorname{div}(\boldsymbol{m}) \in \mathscr{M}(\mathcal{X})$. Let us now introduce the following useful space [17,16].

Definition 4 (Space of charges). *We denote by $\boldsymbol{\mathcal{V}}$ the space of* divergence vector fields *or* charges*, namely the space of vector Radon measures with finite divergence:*

$$\boldsymbol{\mathcal{V}} \overset{\text{def.}}{=} \left\{ \boldsymbol{m} \in \boldsymbol{\mathscr{M}(\mathcal{X})^2}, \operatorname{div}(\boldsymbol{m}) \in \mathscr{M}(\mathcal{X}) \right\}.$$

It is a Banach space with respect to the norm $\|\cdot\|_{\boldsymbol{\mathcal{V}}} \overset{\text{def.}}{=} \|\cdot\|_{\mathrm{TV}^2} + \|\operatorname{div}(\cdot)\|_{\mathrm{TV}}$.

We define in the following the *curve measure* belonging to $\boldsymbol{\mathcal{V}}$ *i.e.* a measure supported on a curve and defined through integration:

Definition 5 (Curve measure). *Let $\boldsymbol{\gamma} : [0,1] \to \mathcal{X}$ a parametrised Lipschitz curve, we say that $\boldsymbol{\mu_\gamma} \in \mathcal{V}$* is a measure supported on the curve $\boldsymbol{\gamma}$ *if:*

$$\forall \boldsymbol{g} \in \mathscr{C}_{\mathbf{0}}(\boldsymbol{\mathcal{X}})^{\mathbf{2}}, \quad \langle \boldsymbol{\mu_\gamma}, \boldsymbol{g} \rangle_{\mathscr{M}(\boldsymbol{\mathcal{X}})^{\mathbf{2}}} \overset{\text{def.}}{=} \int_0^1 \boldsymbol{g}(\boldsymbol{\gamma}(t)) \cdot \dot{\boldsymbol{\gamma}}(t) \, \mathrm{d}t.$$

We denote by $\Gamma \overset{\text{def.}}{=} \boldsymbol{\gamma}([0,1])$ the support of the curve.

The bracket w.r.t. to a curve measure is the circulation [17] of a test vector field function along the curve $\boldsymbol{\gamma}$. The curve has a finite length since its parametrisation is Lipschitz, hence $\mathscr{H}_1(\Gamma) < +\infty$. Some properties that a curve might exhibit are introduced in the following.

Definition 6 (Several characterisation of curves). *A curve is called* simple *if the restriction of $\boldsymbol{\gamma}$ on $[0,1)$ is an injective mapping. A curve is* closed *if $\boldsymbol{\gamma}(0) = \boldsymbol{\gamma}(1)$, it is called a* loop *if it is simple and closed.*

Using Sard's theorem for Lipschitz functions, one can prove that $\boldsymbol{\mu_\gamma}$ does not depend on the way the curve is parametrized. Therefore, it is assumed that the curve $\boldsymbol{\gamma}$ has a constant speed parametrization[3], unless stated otherwise. Finally, we give the expression for the divergence of a curve:

Proposition 1 (Curve divergence). *Let $\boldsymbol{\mu_\gamma}$ be a measure supported on a curve $\boldsymbol{\gamma}$, then $\operatorname{div}(\boldsymbol{\mu_\gamma}) = \delta_{\boldsymbol{\gamma}(0)} - \delta_{\boldsymbol{\gamma}(1)}$. In particular, $\operatorname{div}(\boldsymbol{\mu_\gamma}) = 0$ if $\boldsymbol{\gamma}$ is closed.*

2.2 The CROC functional and its minimiser structure

Similarly to the scalar case, one can define a variational problem on $\mathcal{V}$. The following functional (CROC) standing for *Curves Represented On Charges* [14] implements the curve reconstruction problem:

$$\operatorname*{argmin}_{\boldsymbol{m} \in \mathcal{V}} T_\alpha(\boldsymbol{m}) \overset{\text{def.}}{=} \frac{1}{2} \|y - \Phi \boldsymbol{m}\|_{\mathcal{H}}^2 + \alpha \|\boldsymbol{m}\|_{\mathcal{V}}. \tag{CROC}$$

$\Phi : \mathcal{V} \to \mathcal{H}$ is linear and maps the divergence vector field set to the acquisition space $\mathcal{H}$ supposed to be Hilbert, where the data observation is $\boldsymbol{y} \in \mathcal{H}$. This functional exhibits existence of a solution, see [14] for a proof and a discussion on extremality conditions. The regulariser penalises the length of the curve and the number of curves. Now, consider the following:

Definition 7 (Curve measures set). *We denote by $\mathfrak{S}$ the space of curve measures, supported on either open or closed simple ones, endowed with weak-$*$ topology:*

$$\mathfrak{S} \overset{\text{def.}}{=} \left\{ \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathcal{V}}}, \boldsymbol{\gamma} \text{ is a simple oriented Lipschitz curve} \right\}.$$

It is a (non-complete) metric space for the weak-$*$ topology.

[3] $\dot{\boldsymbol{\gamma}}$ is a.e. equal to a constant.

We stress that the curve measures involved do not encode any variation of amplitude along the support, also that the elements of $\mathfrak{S}$ are normalised. This obviously affects the terms of the $\mathscr{V}$-norm: from now on, and to avoid any ambiguity, we denote by $\boldsymbol{\nu}_{\boldsymbol{\gamma}} \overset{\text{def.}}{=} \boldsymbol{\mu}_{\boldsymbol{\gamma}}/\|\boldsymbol{\mu}_{\boldsymbol{\gamma}}\|_{\mathscr{V}}$ an element of $\mathfrak{S}$. The following result is a corollary of the main theorem of [14], which establishes that the extreme points of the unit ball of the $\mathscr{V}$-norm are precisely the elements of $\mathfrak{S}$, and the celebrated representer theorem [2,1]. The Hilbert acquisition space is now specifically $\mathcal{H} = \mathcal{H}_n \overset{\text{def.}}{=} \mathbb{R}^n$ a finite dimensional space.

Corollary 1 (Minimiser structure). *The problem* (CROC) *admits a minimiser denoted* $\overline{\boldsymbol{u}} \in \mathscr{V}$*:*

$$\overline{\boldsymbol{u}} = \sum_{i=1}^{p} \alpha_i \boldsymbol{u_i}$$

where $p \leq \dim \mathcal{H}_n$*,* $\boldsymbol{u} \in \mathfrak{S}^p$ *and* $\alpha_i > 0$ *for* $0 \leq i \leq p$*, while* $\sum_{i=1}^{p} \alpha_i = T_\alpha(\overline{\boldsymbol{u}})$*.*

The extreme points result of [14] and these latest corollary are the core component of our numerical implementation. Based on the compelling results obtained in [7], we use similarly a greedy algorithm namely the *Frank-Wolfe* algorithm [10], also known as the conditional gradient method. Hopefully, it consists in the iterative reconstruction of the solution with the regulariser atoms, *i.e.* level sets extreme points, yet precisely curve measures here. In the following sections, we present our main contribution, amounting to the numerical optimisation of (CROC) with an instance for a synthetic super-resolution.

3 The Charge (Sliding) Frank-Wolfe for off-the-grid curve reconstruction

The Frank-Wolfe algorithm perfoms the minimisation of a convex differentiable function over a weakly compact convex subset of a Banach space. It relies on the iterative minimisation of a linearised version of the objective function, benefiting from the fact that it uses the directional derivatives and that it does not require any Hilbertian structure, contrary to classical proximal algorithms. It has gained significant attention from data scientists as it produces iterates that are a combination of only a few atoms, specific to the chosen regulariser. Similarly to other off-the-grid implementation, this algorithm is not straightforwardly appliable to CROC: T_α is not differentiable, and the optimisation set $\mathscr{V}$ is not bounded. It is thus necessary to perform an *epigraphical lift* [11,7] to reach a differentiable functional that shares the same *minimum* measures as T_α. Our proposed algorithm is given in Algorithm 1.

We precise some notation: we denote $\boldsymbol{\nu}_{\boldsymbol{\gamma}}{}^{[k]} \overset{\text{def.}}{=} (\boldsymbol{\nu}_{\boldsymbol{\gamma}_1}{}^{[k]}, \dots, \boldsymbol{\nu}_{\boldsymbol{\gamma}_{N^{[k]}}}{}^{[k]})$ the vector of reconstructed atoms at the k-th iteration, and $\Phi_{\boldsymbol{\nu}_{\boldsymbol{\gamma}}{}^{[k]}}(a) \overset{\text{def.}}{=} \sum_{i=0}^{k} a_i \Phi \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]}$

where $a \in \mathbb{R}^{N^{[k]}}$ is the vector of the estimated curves weights, see Algorithm 1. We denote by $\mathcal{S}(\mathcal{X})$ the set of curves spanning $\mathfrak{S}$, namely:

$$\mathcal{S}(\mathcal{X}) \overset{\text{def.}}{=} \{\boldsymbol{\gamma}, \boldsymbol{\gamma} \text{ is a simple oriented Lipschitz curve with support in } \mathcal{X}\}.$$

The length of $\boldsymbol{\gamma}$ is given by $\ell(\boldsymbol{\gamma}) \overset{\text{def.}}{=} \mathscr{H}_1(\boldsymbol{\gamma}([0,1]))$ and its $\mathscr{V}$-norm equivalent by $\ell_{\text{div}}(\boldsymbol{\gamma})$:

$$\ell_{\text{div}}(\boldsymbol{\gamma}) = \begin{cases} \ell(\boldsymbol{\gamma}) + 2 \text{ if } \boldsymbol{\gamma} \text{ open.} \\ \ell(\boldsymbol{\gamma}) \qquad \text{otherwise.} \end{cases}$$

Our algorithm benefits from the *sliding* improvement, where the classic Frank-Wolfe is improved by the sliding step in line 8–9. Among others, this non-convex step allows a finite time convergence for off-the-grid spikes reconstruction, while a similar argument for off-the-grid level sets and curves is yet to be found, though observed in practice. The following property derived from Frank-Wolfe algorithm properties [6,7] holds:

Proposition 2. *Let $(m^{[k]})_{k\geq 0}$ a sequence produced by Algorithm 1. Then it has an accumulation point in weak-$*$ topology, the latter being a solution of* (CROC)*. Also, there exists $C > 0$ such that for any minimiser m^* of* (CROC)*:*

$$\forall k \in \mathbb{N}^*, \quad T_\alpha(m^{[k]}) - T_\alpha(m^*) \leq \frac{C}{k}.$$

In practice, a curve measure $\boldsymbol{\nu}_{\boldsymbol{\gamma}}$ is discretised by a polygonal curve with integer $n \geq 2$ segments[4]: it is the set of $x \in (\mathcal{X})^{2n}$ such that the list of vertices is simple. Such a choice of approximation is made for the sake of simplicity, in particular for the numerical implementation. A variant with splines or Bézier curves may be interesting to reach higher accuracy with fewer control points. In further works, we will pursue the theoretical (Γ-)convergence of the discrete approximation towards the continuous one. To help the reader to get a grasp on this numerical implementation, we illustrate the Charge Sliding Frank-Wolfe directly on a practical case of super-resolution.

4 A numerical illustration for super-resolution

The chosen application for these experiments is a super-resolution problem in the context of a Gaussian convolution operator. Let $g \in \mathcal{X}^n$ be the observed image with n pixels. Since our source $\boldsymbol{\nu}_{\boldsymbol{\gamma}}$ is a *vector* measure, a natural choice of vector quantity for the fidelity term from this image would be the gradient of the image g, then we denote $\boldsymbol{y} = \boldsymbol{\nabla} g$. In further works, we plan to justify more thoroughly this choice. In practice, we exploit a trick enabled by the convolution derivative property. Indeed, consider formally the discrete image source I, then

[4] and obviously $n \geq 3$ for closed curves.

Algorithm 1: Charge Sliding Frank-Wolfe

Input: Acquisition $y \in \mathcal{H}$, number of iterations K, regularisation weight $\alpha > 0$.

Initialisation: $m^{[0]} = 0$, $N^{[k]} = 0$.

for k, $0 \leq k \leq K$ **do**

For $m^{[k]} = \sum_{i=1}^{N^{[k]}} a_i^{[k]} \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]}$ such that $a_i^{[k]} \in \mathbb{R}$, $\boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]} \in \mathfrak{S}$, let

$$\boldsymbol{\eta}^{[k]}(x) \stackrel{\text{def.}}{=} \frac{1}{\alpha} \Phi^*(\Phi\, m^{[k]} - y).$$

Find $\boldsymbol{\gamma}_*{}^{[k]} \in \mathcal{S}(\mathcal{X})$ such that :

$$\boldsymbol{\gamma}_* \in \operatorname*{argmax}_{\boldsymbol{\gamma} \in \mathcal{S}(\mathcal{X})} \frac{1}{\ell_{\mathrm{div}}(\boldsymbol{\gamma})} \int_0^1 \boldsymbol{\eta}^{[k]}(\boldsymbol{\gamma}(t)) \cdot \dot{\boldsymbol{\gamma}}(t)\, \mathrm{d}t.$$

if $\left|\eta^{[k]}\left(\boldsymbol{\gamma}_*\right)\right| \leq 1$ **then**

$m^{[k]}$ is the solution of CROC. Stop.

else

Compute $m^{[k+1/2]} = \sum_{i=1}^{N^{[k]}} a_i^{[k+1/2]} \boldsymbol{\nu}_{\boldsymbol{\gamma}_i} + a_{N^{[k]}+1}^{[k+1/2]} \boldsymbol{\nu}_{\boldsymbol{\gamma}_*}{}^{[k+1/2]}$ such that:

$$a_i^{[k+1/2]} \in \operatorname*{argmin}_{a \in \mathbb{R}^{N^{[k]}+1}} T_\alpha \left(\sum_{i=1}^{N^{[k]}} a_i \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]} + a_{N^{[k]}+1} \boldsymbol{\nu}_{\boldsymbol{\gamma}_*}{}^{[k+1/2]} \right)$$

Compute $m^{[k+1]} = \sum_{i=1}^{N^{[k+1]}} a_i^{[k+1]} \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}$, output of the optimisation initialised with $m^{[k+1/2]}$:

$$\left(a^{[k+1]}, \boldsymbol{\nu}_{\boldsymbol{\gamma}}{}^{[k+1]}\right) \in \operatorname*{argmin}_{(a,\boldsymbol{\gamma}) \in \mathbb{R}^{N^{[k]}+1} \times \mathcal{S}(\mathcal{X})^{N^{[k]}+1}} T_\alpha \left(\sum_{i=1}^{N^{[k]}+1} a_i \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]} \right).$$

Set $m^{[k+1]} = \sum_{i=1}^{N^{[k]}+1} a_i^{[k]} \boldsymbol{\nu}_{\boldsymbol{\gamma}_i}{}^{[k]}$. Prune the low amplitude atoms.

end

end

Output: Discrete measure $m^{[k]}$ where k is the stopping iteration.

$\boldsymbol{y} = h * \boldsymbol{\nabla} I = \boldsymbol{\nabla} h * I$. Hence, $\Phi\, \boldsymbol{\nu}_{\boldsymbol{\gamma}}$ will be computed as the support convolved with the gradient of the Gaussian kernel. Then we consider the *vector* forward operator Φ with kernel $\boldsymbol{\varphi}(x)$ for $x \in \mathcal{X}$ writing down:

$$\forall x \in \mathcal{X}, \quad \boldsymbol{\varphi}(x) \stackrel{\text{def.}}{=} \frac{1}{2\pi\sigma} \begin{pmatrix} -x_1\, e^{-\left(\frac{i-1}{K-1} - x_1\right)^2 / 2\sigma^2} \\ -x_2\, e^{-\left(\frac{i-1}{K-1} - x_2\right)^2 / 2\sigma^2} \end{pmatrix}_{1 \leq i \leq n}.$$

Obviously, the gradient $\boldsymbol{y}$ ought to be smoothed up, since it is the vector image $\boldsymbol{y} = \boldsymbol{\nabla} g$ used in practice in the algorithm and all the more as the noise in g has even more impact on $\boldsymbol{y}$.

In the following, we consider a synthetic example where one wants to recover curves from a classic image acquisition y, altered by a convolution with standard deviation $\sigma = 2 \times 10^{-2}$ and white additive noise with spread $\sigma_{\mathrm{b}} = 4 \times 10^{-3}$. Let two measures $\boldsymbol{\mu_1}$ an open curve and $\boldsymbol{\mu_2}$ a closed one, both belonging to $\mathfrak{S}$; consider now the source charge $\boldsymbol{T} = 4\boldsymbol{\mu_1} + \boldsymbol{\mu_2}$, the latter 4 is chosen for the sake of visualisation. The image g and its source $\boldsymbol{T}$ are plotted on the Figure 1. The gradient $\boldsymbol{y}$ of the image is plotted on the Figure 2.

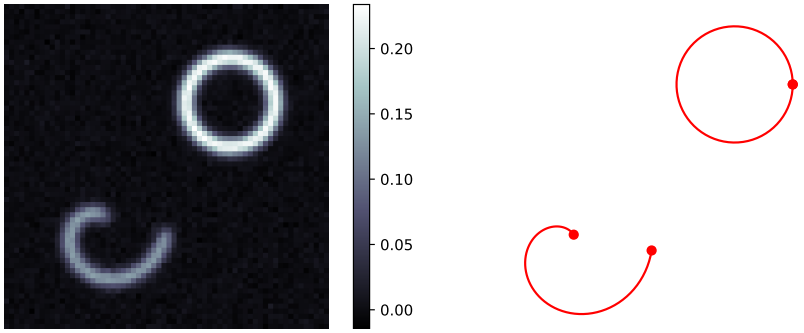

Fig. 1: Left: The observed image g, generated from a charge $\boldsymbol{T}$ composed of a spiral $\boldsymbol{\mu_1}$ and a loop $\boldsymbol{\mu_2}$, which have different intensities. Right: the two curves support. Note the smooth curvature of $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ we aim to recover.

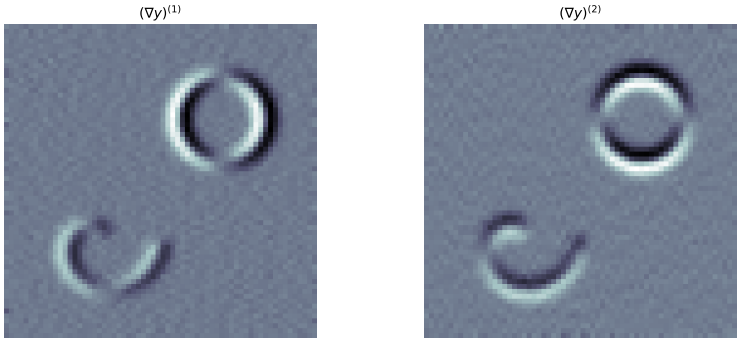

Fig. 2: The two components of the gradient $\boldsymbol{y} = \boldsymbol{\nabla} g$, is the relevant quantity for the fidelity term of (CROC). Note that the noise on the image has an even greater impact on its gradient, such that a gradient denoising strategy must be adopted, especially for experimental images.

As we stated before, the first step of the algorithm lies in the support estimation, this linear step bears an original approach in each off-the-grid regulariser.

The equivalent case for the classic TV-norm consists in a simple grid search of the greatest value pixel of $\boldsymbol{\eta}^{[k]}$, similarly defined as in line 3 of Algorithm 1. Here we have to solve:

$$\operatorname*{argmax}_{\boldsymbol{m}\in\mathscr{V}} \langle \boldsymbol{\eta}^{[k]}, \boldsymbol{m} \rangle_{\mathscr{M}(\boldsymbol{\mathcal{X}})^2}$$

equivalent to the following problem:

$$\operatorname*{argmax}_{\boldsymbol{\gamma}\in\mathcal{S}(\mathcal{X})} \frac{1}{\ell_{\mathrm{div}}(\boldsymbol{\gamma})} \int_0^1 \boldsymbol{\eta}^{[k]}(\boldsymbol{\gamma}(t)) \cdot \dot{\boldsymbol{\gamma}}(t)\,\mathrm{d}t.$$

This can be interpreted as the length of the curve $\boldsymbol{\gamma}$ weighted by a 'metric' $\boldsymbol{\eta}^{[k]}$. Since the kernel in Φ is the gradient of the 2D Gaussian kernel, the reader might be aware that the *certificate* $\boldsymbol{\eta}^{[k]}$ in Algorithm 1, see [14] for more insights, is the Laplacian of the image g. Then, the support of the estimated curve $\boldsymbol{\gamma}_*$ appears naturally in the Laplacian, as it is the maximum of $\left|\boldsymbol{\eta}^{[k]}\right|$. Concerning the heuristic to determine if $\boldsymbol{\gamma}_*$ is open or closed, note that our involved curves are simple. Therefore, we exploit that a loop separates $\mathcal{X}$ onto two connected components, on the contrary to an open curve whose support complement has only one connected set. The algorithm then selects a simple chain of pixels to give an approximate support estimation. The Figure 3 shows the magnitude of the certificate at iteration 0, an initialisation support for the added atom is successfully found.

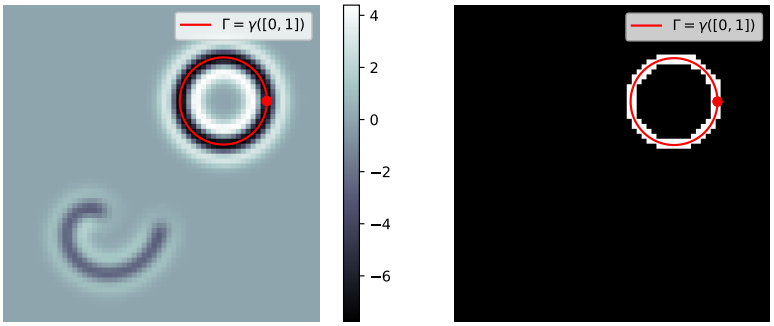

Fig. 3: Left: magnitude of $\Phi^* \boldsymbol{y}$, the support of the estimated curve $\boldsymbol{\gamma}_*$ lies in the (near-)optima pixels. Right: threshold to reach a rough estimate of the support. The ground-truth curve $\boldsymbol{\gamma}$ with support Γ is traced in red.

The convex step in line 7 is a fairly run-of-the-mill routine, the amplitudes of the atoms are estimated; this is a rather classical LASSO up to some tweaks, tackled in partice with a L-BFGS optimiser. The non-convex step in lines 8-9 on the contrary is way more challenging: the *sliding* performed here adjusts

the atoms both in amplitudes and support. The chosen discretisation of curve measure here amounts to a chain, this optimisation can be understood then as a gradient descent over both amplitudes (common for one chain) and positions of numerous discretisation points. This is dealt with in our implementation with a flavour of stochastic gradient descent namely an ADAM optimiser, empirically shown here to outperform other non-convex solvers. which have shown to be empirically . In Figure 4 one can see the output of the convex step, then sharing the same estimated support as the crude one from the linear step; to be compared with the non-convex output on the right.

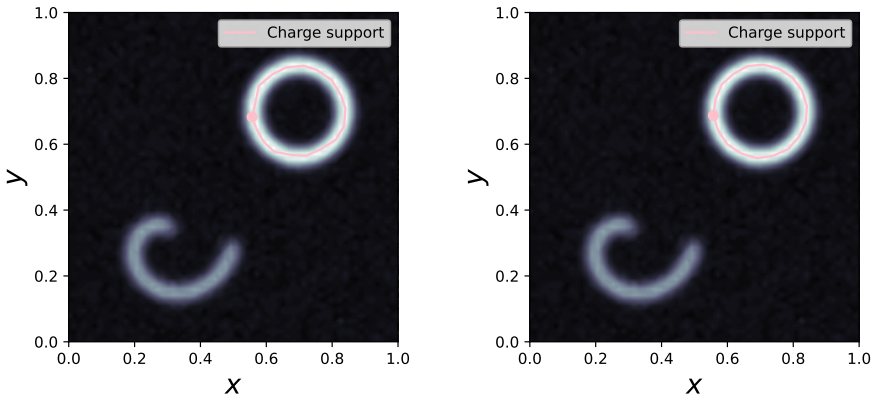

Fig. 4: Left: crude estimation from the oracle step. Right: the estimate at the end of the first iteration. Hopefully the sliding step smooths over the curvature and corrects the rough support estimation.

Interestingly enough, the loop endpoints *i.e.* $\boldsymbol{\gamma}(0)$ and $\boldsymbol{\gamma}(1)$, by definition equal, tends in fact to move a bit away from each other with the non-convex iteration. This is maybe related to the faces and broadly speaking to the geometric structure of the unit ball of the $\mathcal{V}$-norm: to the best knowledge of the authors there is no such work investigating this curious change of topology since the $\mathcal{V}$ space is a rather new concept for the off-the-grid community. Our observation is at this point purely empirical, and we only implemented a strategy to merge really closed endpoints. Our algorithm has then reconstructed one curve, and we loop over so forth to yield a reconstruction of the source charge $\boldsymbol{T}$, see Figure 5.

Finally, the reconstruction is plotted in Figure 6. The curves supports are greatly recovered[5], with of course some small differences due to the noise. A theoretical bound of this support estimation error is not trivial: it was investigated

[5] 20 seconds on an *Intel Xeon E5-2687W v3* for a 64×64 image. A CUDA implementation is enabled for larger images, still ensuring reconstruction for a 512×512 image in less than 4 minutes on a *NVIDIA Quadro K2200*. See `https://gitlab.inria.fr/blaville/amg`.

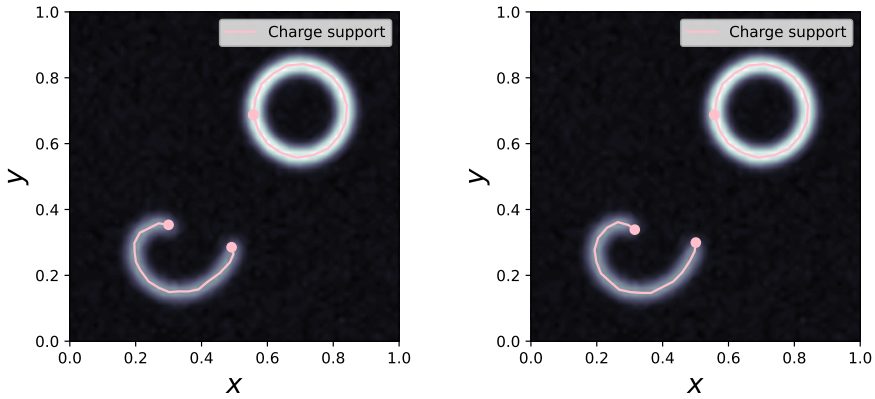

Fig. 5: Left: crude estimation from the oracle step. Right: the estimate at the end of the first iteration. Estimates matched the structure of the curves but exhibits some 'thorns': due to the noise, the reconstruction seems not as smoothed as the ground-truth. The non-convex step better recover the spiral folding.

for spikes [8] and recently a similar result for level sets was advertised. Still, an equivalent for the off-the-grid curve reconstruction is yet to be explored and is clearly out of the scope of the present paper.

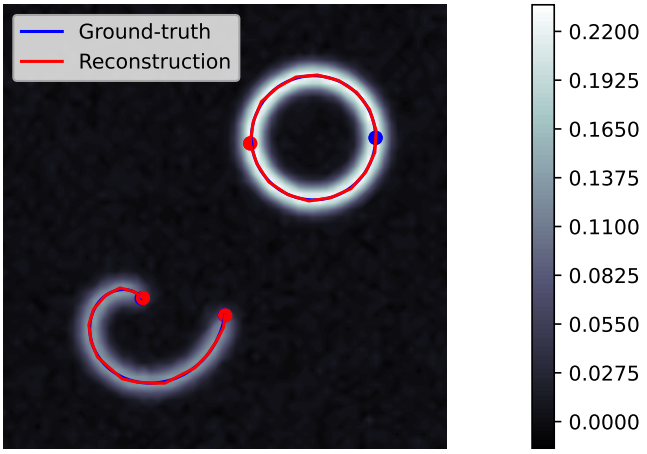

Fig. 6: The final reconstruction captures well the curvature of the source curves.

The proposed results might seem a bit simple. However, we emphasize that we can replace the convolution kernel with Fourier or Laplace measurements to handle more difficultly interpretable acquisition encountered in inverse problems. We wish to apply our algorithm to experimental data, but it requires a precise

post-processing of the gradient such as denoising or a basic super-resolution, as the algorithm needs a (mediocre but at least usable) first guess to estimate a point of $\mathscr{V}$ in the basin of attraction of the global minimiser.

5 Conclusion

Based on the latest developments in off-the-grid theoretical results, we proposed a new algorithm called *Charge Sliding Frank-Wolfe* to perform a gridless reconstruction of curves while successfully implementing and reaching first results on a synthetic example. In further works, we plan to carry out the experiments with real data in localisation microscopy, while digging some numerical properties of our algorithm: an equivalence of the finite-time convergence rate of the *Sliding Frank-Wolfe* for instance would be a quite convenient property.

Acknowledgements The work of Bastien Laville has been supported by the French government, through the UCA DS4H Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-17-EURE-0004. The work of Laure Blanc-Féraud has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. This work was partially funded by the ANR Micro-Blind with the reference number ANR-21-CE48-0008. BL would like to warmly thanks Mr. Romain Petit for the genuinely constructive discussion they shared in Évian-les-Bains and later remote conference, and Mr. Théo Bertrand for his helpful advices on some of the numerous quirkiness in the numerical implementation.

References

1. Boyer, C., Chambolle, A., Castro, Y.D., Duval, V., de Gournay, F., Weiss, P.: On representer theorems and convex regularization. SIAM Journal on Optimization **29**(2), 1260–1281 (jan 2019). https://doi.org/10.1137/18m1200750
2. Bredies, K., Carioni, M.: Sparsity of solutions for variational inverse problems with finite-dimensional data. Calculus of Variations and Partial Differential Equations **59**(1) (dec 2019). https://doi.org/10.1007/s00526-019-1658-1
3. Bredies, K., Carioni, M., Fanzon, S., Romero, F.: On the extremal points of the ball of the benamou–brenier energy. Bulletin of the London Mathematical Society (jun 2021). https://doi.org/10.1112/blms.12509
4. Bredies, K., Pikkarainen, H.K.: Inverse problems in spaces of measures. ESAIM: Control, Optimisation and Calculus of Variations **19**(1), 190–218 (mar 2012). https://doi.org/10.1051/cocv/2011205
5. de Castro, Y., Duval, V., Petit, R.: Towards off-the-grid algorithms for total variation regularized inverse problems. In: Lecture Notes in Computer Science, pp. 553–564. Springer International Publishing (2021)
6. Demianov Vladimir Fedorovich, M.: Approximate methods in optimization problems. American Elsevier Pub. Co. (1970)

7. Denoyelle, Q., Duval, V., Peyré, G., Soubies, E.: The sliding frank–wolfe algorithm and its application to super-resolution microscopy. Inverse Problems **36**(1), 014001 (dec 2019). https://doi.org/10.1088/1361-6420/ab2a29
8. Duval, V., Peyré, G.: Exact support recovery for sparse spikes deconvolution. Foundations of Computational Mathematics **15**(5), 1315–1355 (oct 2014). https://doi.org/10.1007/s10208-014-9228-6
9. Federer, H.: Geometric measure theory. Springer, Berlin (1996)
10. Frank, M., Wolfe, P.: An algorithm for quadratic programming. Naval Research Logistics Quarterly **3**(1-2), 95–110 (mar 1956). https://doi.org/10.1002/nav.3800030109
11. Harchaoui, Z., Juditsky, A., Nemirovski, A.: Conditional gradient algorithms for norm-regularized smooth convex optimization. Mathematical Programming **152**(1-2), 75–112 (apr 2014). https://doi.org/10.1007/s10107-014-0778-9
12. Khavin, V.P., Smirnov, S.K.: Approximation and extension problems for some classes of vector fields. St. Petersburg Department of Steklov Institute of Mathematics, Russian Academy of Sciences **10**(3), 507–528 (1998), https://mathscinet.ams.org/mathscinet-getitem?mr=1628034
13. Laville, B., Blanc-Féraud, L., Aubert, G.: Off-The-Grid Variational Sparse Spike Recovery: Methods and Algorithms. Journal of Imaging **7**(12), 266 (Dec 2021). https://doi.org/10.3390/jimaging7120266, https://www.mdpi.com/2313-433X/7/12/266
14. Laville, B., Blanc-Féraud, L., Aubert, G.: Off-the-grid curve reconstruction through divergence regularisation: an extreme point result. Preprint (2023), https://hal.science/hal-03658949/document
15. Parhi, R., Nowak, R.D.: On continuous-domain inverse problems with sparse superpositions of decaying sinusoids as solutions. In: ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE (may 2022). https://doi.org/10.1109/icassp43922.2022.9746165
16. Šilhavỳ, M.: Divergence measure vectorfields: their structure and the divergence theorem. Mathematical modelling of bodies with complicated bulk and boundary behavior **20**, 214–237 (2008)
17. Smirnov, S.K.: Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. St. Petersburg Department of Steklov Institute of Mathematics, Russian Academy of Sciences **5**(4), 206–238 (1993), https://mathscinet.ams.org/mathscinet-getitem?mr=1246427