Off-the-grid curve reconstruction through divergence regularisation: an extreme point result*

Bastien Laville†, Laure Blanc-Féraud†, and Gilles Aubert‡

Abstract. We propose a new strategy for curve reconstruction in an image through an off-the-grid variational framework, inspired by spike reconstruction in the literature. We introduce a new functional CROC on the space of 2-dimensional Radon measures with finite divergence denoted $\mathscr{V}$, and we establish several theoretical tools through the definition of a certificate. Our main contribution lies in the sharp characterisation of the extreme points of the unit ball of the $\mathscr{V}$-norm: there are exactly measures supported on 1-rectifiable oriented simple Lipschitz curves, thus enabling a precise characterisation of our functional minimisers and further opening a promising avenue for the algorithmic implementation.

Key words. Off-the-grid inverse problem, curve reconstruction, Radon measures, extreme points, imaging, divergence vector field.

AMS subject classifications. 34K29

1. Introduction.

This work focuses on the definition of a functional designed to recover curves in an off-the-grid fashion, by considering the space of vector Radon measures with finite divergence, namely that their divergence in the distributional sense is a Radon measure. Such choice of space is motivated by multiple works [22, 21, 2] pertaining to Jordan curve, or at least Radon measure absolutely continuous with respect to $\mathscr{H}_1$. However, the literature nowadays does not offer a tractable off-the-grid framework for open and closed curves recovery.

Yet such situations arise in several domains such as biomedical imaging (blood vessels, filaments structures), in magnetic imaging through observation of the magnetic field (Scanning Magnetic Microscopy), *etc.* An example of curves encountered in a natural and practical setting is presented in the Figure 1. In the following, we propose a method referred to as *Atomic based Method for Gridless* (AMG), designed for curve reconstruction through off-the-grid variational optimisation. To the best knowledge of the authors, this is the first attempt to recover curves in an off-the-grid manner and one of the few papers to characterise the space of divergence vector field.

1.1. Contributions.

Our main contributions in this article are listed below:

- Propose a new space of measures $\mathscr{V}$ for off-the-grid curve variational analysis, and qualify several of its properties;

---

*Submitted to the editors April 6, 2023.

Funding: 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 Cote d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

†Morpheme project: Université Côte d'Azur, Inria, CNRS, France. (bastien.laville@inria.fr, laure.blanc_feraud@i3s.unice.fr, https://team.inria.fr/morpheme/).

‡Université Côte d'Azur, CNRS, LJAD, France.(gilles.aubert@univ-cotedazur.fr).

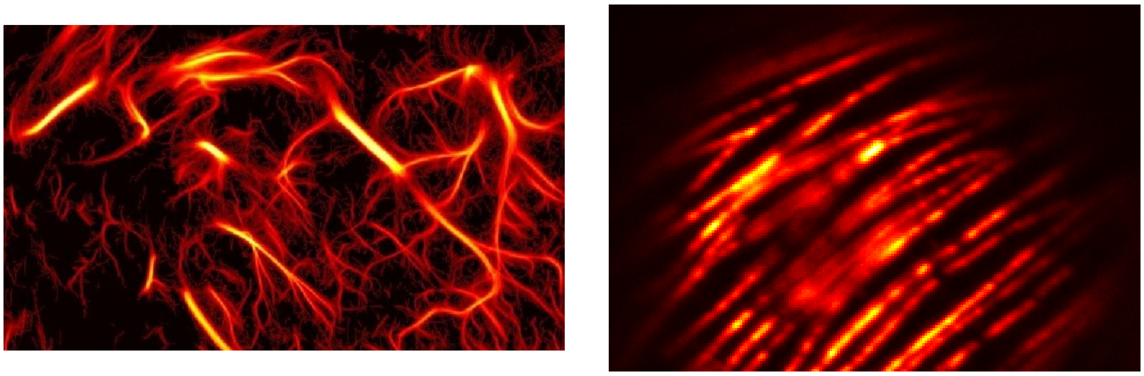

Figure 1: Curves arise genuinely in several biological structures, such as (on the left) blood vessels [1] and (on the right) *Ostreopsis* protein samples.

- Develop a new functional ($\mathcal{Q}_\alpha(y)$) called CROC for curve reconstruction in an off-the-grid variational context;
- Establish the main results, such as the certificate outlining, for theoretical off-the-grid analysis and future numerical implementation;
- Characterise precisely the extreme points of the $\mathscr{V}$-norm unit ball as 1-rectifiable simple oriented Lipschitz curves, thereby describing the structure of the minimisers of our functional and thus proving the interest of AMG.

1.2. Paper outline. The paper is organised as follows. We present the main definitions and tools in section section 2, we propose a functional in section 3, our main theorem lies in section 4 while the conclusions and the outlook follow in section 5.

1.3. Notations. In the following, $\mathcal{X}$ denotes the ambient space where the positions of the curves/spikes live. We suppose that $\mathcal{X}$ is a non-empty bounded open subset of $\mathbb{R}^d$, hence a submanifold of dimension $d \in \mathbb{N}^*$.

Note that this paper makes extensive use of the notion of *measure*, through both functional and set definition. A comprehensive characterisation of measure from a functional standpoint is given in the section 2, and the interested reader can explore the set point-of-view in several celebrated work such as [20]. Let ρ be a measure, $\rho \llcorner A$ denotes the restriction of the measure ρ to the set A, namely for all measurable sets B: $\rho \llcorner A(B) \stackrel{\text{def.}}{=} \rho(A \cap B)$. The support of ρ denoted by $\operatorname{spt}(\rho)$ is a Borel set $B \subset \mathcal{X}$ such that $\operatorname{spt}(\rho) = \{x \in \mathcal{X} \mid \forall U \text{ neighborhood of } x,\, \rho(U) > 0\}$. Let us denote by $\mathscr{H}_1$ the 1-dimensional Hausdorff measure (see [14] for a definition). We say that $E \subset \mathbb{R}^d$ is p-rectifiable if it is the countable union of Lipschitz function images from $\mathbb{R}^p$ to $\mathbb{R}^d$, up to a set of null $\mathscr{H}^p$-measure [18].

2. Off-the-grid framework.

This section is divided into two parts: a recall of the classic off-the-grid framework for spike reconstruction, and the new notions needed for curve reconstruction.

2.1. Classical scalar off-the-grid framework. We recall some definitions and handy properties stemming from the off-the-grid literature, the interested reader may take a deeper look in the review [17] for more insights. In this case, the aim is to reconstruct spikes, *i.e.* dots localised in the space $\mathcal{X}$ with some amplitude information encoded. This problem is encountered in many fields such as astronomy for stars imaging enhancement, fluorescence microscopy where one locates dyes, ultrasound localisation microscopy imaging when one performs the tracking of small air bubbles, *etc.*

2.1.1. Quick digest of measure theory. A spike, not constrained to a finite set of positions, can be accurately modelled by a Dirac measure. Loosely speaking, this map allows us to encode both amplitude and spatial information in the same object. However, since the Dirac measure is not a classic continuous function, one needs to consider a more general class of maps called the Radon measures. From a distributional standpoint, it is a subset of the distribution space $\mathscr{D}'(\mathcal{X})$; the latter being the space of linear forms over the space of test functions $\mathscr{C}_0^\infty(\mathcal{X})$ *i.e.* smooth functions (continuous derivatives of all orders) compactly supported. This functional approach[1] is based on the definition of a measure as a linear form on a function space, namely:

Definition 2.1 (Continuous function on $\mathcal{X}$). *We call $\mathscr{C}(\mathcal{X}, \mathcal{Y})$ the set of continuous functions from $\mathcal{X}$ to a normed vector space $\mathcal{Y}$, endowed with the* supremum *norm $\|\cdot\|_{\infty,\mathcal{X}}$ of functions.*

Definition 2.2 (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 can introduce:

Definition 2.3 (Set of Radon measures). *We denote by $\mathcal{M}(\mathcal{X})$ the set of real signed Radon measures on $\mathcal{X}$ of finite masses. It is the topological dual of $\mathscr{C}_0(\mathcal{X})$ with* supremum *norm $\|\cdot\|_{\infty,\mathcal{X}}$ by the Riesz-Markov representation theorem [14]. Thus, a Radon measure $m \in \mathcal{M}(\mathcal{X})$ is a continuous linear form on functions $f \in \mathscr{C}_0(\mathcal{X})$, with the duality bracket denoted by $\langle f, m \rangle_{\mathcal{M}} = \int_{\mathcal{X}} f \, \mathrm{d}m$.*

A 'signed' measure entails that the quantity $\langle f, m \rangle_{\mathcal{M}}$ can be negative, further generalising the notion of probability (positive) measure. While m is a measure, also remember that it can be evaluated on all Borel sets $A \subset \mathcal{X}$ through:

$$|m|(A) \overset{\text{def.}}{=} \sup \left(\int_A f \, \mathrm{d}m, f \in \mathscr{C}_0(A), \|f\|_{\infty,\mathcal{X}} \leq 1 \right).$$

Classic examples of Radon measures are the Lebesgue measure, the celebrated Dirac measure δ_z centred in $z \in \mathcal{X}$ namely for all $f \in \mathscr{C}_0(\mathcal{X})$ one has $\langle f, \delta_z \rangle_{\mathcal{M}} = f(z)$, *etc.* Eventually,

[1]One can then define equivalently the space of Radon measures, either by a set-related approach or by functional analysis approach (thanks to Riesz–Markov theorem). In the more set-related [20] insight, a *measure* is an object which takes sets as an input. A *Borel measure* is a measure defined on all open sets of $\mathcal{X}$, and a *Radon measure* is a Borel measure such that it is finite on all compact sets of $\mathcal{X}$ (by an isomorphism). The functional and the set point-of-views are different approaches to describe the same object.

since $\mathscr{C}_0(\mathcal{X})$ is a normed vector space, $\mathcal{M}(\mathcal{X})$ is complete [6] if endowed with its dual norm called the total variation (TV) norm, defined for $m \in \mathcal{M}(\mathcal{X})$ by:

$$\|m\|_{\mathrm{TV}} \stackrel{\text{def.}}{=} |m|(\mathcal{X}) = \sup\left(\int_{\mathcal{X}} f \, \mathrm{d}m, f \in \mathscr{C}_0(\mathcal{X}), \|f\|_{\infty,\mathcal{X}} \leq 1\right).$$

2.1.2. Observation model. Let us introduce $\mathcal{H}$ the Hilbert space containing the acquired data. In the case of images, we use a finite dimensional space of acquisition $\mathcal{H} = \mathcal{H}_n \stackrel{\text{def.}}{=} \mathbb{R}^n$ for $n \in \mathbb{N}$. Let $m \in \mathcal{M}(\mathcal{X})$ be a source measure, we call *acquisition* $y \in \mathcal{H}$ the result of the *forward/acquisition map* $\Phi : \mathcal{M}(\mathcal{X}) \to \mathcal{H}$ evaluated on m, with measurement kernel $\varphi : \mathcal{X} \to \mathcal{H}$ continuous and bounded [8, 11]:

$$y \stackrel{\text{def.}}{=} \Phi m = \int_{\mathcal{X}} \varphi(x) \, \mathrm{d}m(x). \tag{2.1}$$

Also note that the forward operator Φ incorporates a sampling operation, hence $\mathcal{H} = \mathcal{H}_n$. In the following, we impose $\varphi \in \mathscr{C}^2(\mathcal{X}, \mathcal{H})$ and $\forall q \in \mathcal{H}$, the map $x \mapsto \langle \varphi(x), q \rangle_{\mathcal{H}}$ ought to belong to $\mathscr{C}_0(\mathcal{X})$. Let us also define the adjoint operator of $\Phi : \mathcal{M}(\mathcal{X}) \to \mathcal{H}$ in the weak-$*$ topology, namely the map $\Phi^* : \mathcal{H} \to \mathscr{C}_0(\mathcal{X})$, defined for all $x \in \mathcal{X}$ and $p \in \mathcal{H}$ by $\Phi^*(p)(x) = \langle p, \varphi(x) \rangle_{\mathcal{H}}$. The choice of φ and $\mathcal{H}$ is dictated by the physical process of acquisition, with generic measurement kernels such as convolution, Fourier, Laplace, *etc* [17].

2.1.3. An off-the-grid functional: the BLASSO. Consider the source measure $m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_{0,i}\delta_{x_{0,i}}$ with amplitudes $\mathbf{a_0} \in \mathbb{R}^N$ and positions $\mathbf{x_0} \in \mathcal{X}^N$, the sparse spike problem aims to recover this measure from the acquisition $y \stackrel{\text{def.}}{=} \Phi m_{a_0,x_0} + w$ where $w \in \mathcal{H}$ is an additive noise. In order to tackle this inverse problem, let us introduce the following convex functional called BLASSO [6, 10], standing for Beurling-LASSO:

$$(\mathcal{P}_\lambda(y)) \qquad \underset{m \in \mathcal{M}(\mathcal{X})}{\operatorname{argmin}} \, T_\lambda(m) \stackrel{\text{def.}}{=} \frac{1}{2}\|y - \Phi(m)\|_{\mathcal{H}}^2 + \lambda |m|(\mathcal{X}),$$

where $\lambda > 0$ is the regularisation parameter accounting for the trade-off between fidelity and sparsity of the reconstruction. The BLASSO in a noiseless setting writes down:

$$(\mathcal{P}_0(y_0)) \qquad \underset{\Phi m = y_0}{\operatorname{argmin}} \, |m|(\mathcal{X}) \quad \text{with } y_0 = \Phi m_{a_0,x_0}.$$

BLASSO is genuinely linked with its discrete counterpart, the LASSO [12]: one can formally see BLASSO as the functional limit of LASSO on a finer and finer grid. If the LASSO problem exhibits existence and uniqueness of the solution, is it extendable to its off-the-grid counterpart? Foremost, let us observe that:

- $m \mapsto |m|(\mathcal{X})$ is lower semi-continuous w.r.t. the weak-$*$ convergence;
- Φ is continuous from the weak-$*$ topology of $\mathcal{M}(\mathcal{X})$ to the $\mathcal{H}$ weak topology.

Then thanks to convex analysis results, one can establish the existence of solutions to $(\mathcal{P}_\lambda(y))$ as proved in [6]. The difficulties also lie in the question of uniqueness of the solution and correct support recovery: in order to tackle these concerns, let us introduce several notions of convex analysis in the following subsection.

2.1.4. Dual problems and certificates. The BLASSO problem ($\mathcal{P}_\lambda(y)$) above is convex, thus one can define its dual problem [6, 17] which writes down for $p \in \mathcal{H}$:

$$(\mathcal{D}_\lambda(y)) \qquad \underset{\|\phi^* p\|_{\infty,\mathcal{X}} \leq 1}{\operatorname{argmax}} \ \langle y, p \rangle_{\mathcal{H}} - \frac{\lambda}{2} \|p\|^2_{\mathcal{H}}.$$

Strong duality between ($\mathcal{P}_\lambda(y)$) and ($\mathcal{D}_\lambda(y)$) is proved in [6]. Hence, any solution m_λ of ($\mathcal{P}_\lambda(y)$) is linked [12] to the unique solution p_λ of ($\mathcal{D}_\lambda(y)$) by the extremality conditions:

$$(2.2) \qquad \begin{cases} \Phi^* p_\lambda \in \partial |m_\lambda|(\mathcal{X}), \\ -p_\lambda = \dfrac{1}{\lambda}(\Phi m_\lambda - y), \end{cases}$$

where $\partial|\cdot|(\mathcal{X})$ is the sub-differential of the TV norm. Indeed and similarly to the ℓ^1 norm, the total variation is not differentiable but lower semi-continuous w.r.t. the weak-$*$topology. Thus, we use its sub-differential [12] which identifies to, for $m \in \mathcal{M}(\mathcal{X})$:

$$(2.3) \qquad \partial |m|(\mathcal{X}) = \left\{ \eta \in \mathscr{C}_0(\mathcal{X})\,;\, \|\eta\|_{\infty,\mathcal{X}} \leq 1 \text{ and } \int_{\mathcal{X}} \eta \, \mathrm{d}m = |m|(\mathcal{X}) \right\}.$$

Elements of this subgradient are called *certificate*; then thanks to strong duality, let us define peculiar certificates named the *dual certificates* [7].

Definition 2.4. *We call* $\eta_\lambda \overset{\text{def.}}{=} \Phi^* p_\lambda$ *a* dual certificate *of* m_λ*, where* p_λ *satisfies* (2.2).

η_λ is a certificate since $\Phi^* p_\lambda \in \partial |m_\lambda|(\mathcal{X})$, it is called *dual* since it verifies the second extremality (2.2) condition: indeed it is defined by the dual solution p_λ. Loosely speaking, a dual certificate η_λ is associated to a measure m_λ and it *certifies* that the measure m_λ is a *minimum* of the BLASSO. If there exist for instance solutions of ($\mathcal{P}_\lambda(y)$) of the form $m_\lambda \overset{\text{def.}}{=} \sum_{i=1}^N a_i \delta_{x_i}$, the support satisfies [12] for all integer $1 \leq i \leq N$: $|\eta_\lambda|(x_i) = 1$. The interested reader can take a glance at [12] for uniqueness and support recovery guarantees.

We now extend this classical scalar off-the-grid formulation to the vector case, and present the space of divergence vector field.

2.2. The space of divergence vector fields/charges. Consider the space of *vector* Radon measures:

Definition 2.5. *We define the set of vector Radon measures* $\boldsymbol{\mathcal{M}}(\boldsymbol{\mathcal{X}})^{\mathbf{2}}$ *as the topological dual of the space of continuous vector functions* $\boldsymbol{\mathscr{C}_0}(\boldsymbol{\mathcal{X}})^{\mathbf{2}} \overset{\text{def.}}{=} \boldsymbol{\mathscr{C}_0}(\boldsymbol{\mathcal{X}}, \mathbb{R}^{\mathbf{2}})$*. The properties of the scalar case hold for the vector one, indeed* $\boldsymbol{\mathcal{M}}(\boldsymbol{\mathcal{X}})^{\mathbf{2}}$ *has a natural TV-norm denoted by* $\|\cdot\|_{\mathrm{TV}^2}$*, a duality bracket* $\langle \cdot, \cdot \rangle_{\boldsymbol{\mathcal{M}}^{\mathbf{2}}}$*,* etc.

We only carry out dimension $d = 2$, since some (wild) pathological cases appear when $d > 2$, see [22, Section 1.3]. Let us denote by div the divergence operator, which ought to be understood in the distributional sense. Indeed, for all $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}(\boldsymbol{\mathcal{X}})^{\mathbf{2}}$:

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

We say that a measure $\boldsymbol{m}$ is *of finite divergence* if $\operatorname{div}(\boldsymbol{m}) \in \mathcal{M}(\mathcal{X})$. Let us now introduce the following useful space [22, 21].

Definition 2.6. *We denote by* $\mathscr{V}$ *the space of* divergence vector fields *or* charges*, namely the space of vector Radon measures with finite divergence:*

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

It is a Banach space with respect to the norm $\|\cdot\|_{\mathscr{V}} \stackrel{\text{def.}}{=} \|\cdot\|_{\mathrm{TV}^2} + \|\operatorname{div}(\cdot)\|_{\mathrm{TV}}$ *(see Appendix A). A charge with null divergence is called* a solenoid*.*

Obviously this topology is not adequate for convergence statement, indeed it does not provide any property for bounded set or sequence. Therefore, one needs to choose a topology with fewer open sets but more compact sets, thus enabling compactness properties:

Definition 2.7 (Weak-$*$ convergence). $\mathscr{V}$ *can be endowed with the weak-$*$ topology of the distributions; we say that a sequence of charges* $(\boldsymbol{m}_n)_{n\in\mathbb{N}}$ *weakly-$*$ converges to* $\boldsymbol{m} \in \mathscr{V}$*, denoted by* $\boldsymbol{m}_n \stackrel{*}{\rightharpoonup} \boldsymbol{m}$*, if:*

$$\forall \boldsymbol{g} \in \mathscr{C}_0^\infty(\mathcal{X}), \quad \int_{\mathcal{X}} \boldsymbol{g} \, \mathrm{d}\boldsymbol{m}_n \xrightarrow[n \to +\infty]{} \int_{\mathcal{X}} \boldsymbol{g} \, \mathrm{d}\boldsymbol{m}.$$

While we have investigated several algebraic facts concerning $\mathscr{V}$, let us precise which maps live in this space.

Claim 2.8. *The* vector *Dirac measure* $\boldsymbol{\delta} \stackrel{\text{def.}}{=} (\delta, \delta)$ *does not belong to* $\mathscr{V}$*, an interesting observation since Dirac measures are extreme points of the TV-norm unit ball.*

Proof. Suppose by contradiction that $\boldsymbol{\delta}$ belongs to $\mathscr{V}$, hence $\operatorname{div} \boldsymbol{\delta} \in \mathcal{M}(\mathcal{X})$. Let $\xi \in \mathscr{C}_0^\infty(\mathcal{X})$ such that $\partial_1 \xi(0) = \partial_2 \xi(0) = 1$ and the sequence:

$$\forall n \in \mathbb{N}^*, \forall x \in \mathcal{X}, \quad \xi_n(x) = \frac{1}{n} \xi(xn).$$

Then, by the dominated convergence theorem, for all measure $\mu \in \mathcal{M}(\mathcal{X})$ one yields $\lim_{n \to +\infty} \int_{\mathcal{X}} \xi_n \, \mathrm{d}\mu = 0$. However, for all $n \in \mathbb{N}^*$ one has:

$$\int_{\mathcal{X}} \xi_n \, \mathrm{d}(\operatorname{div} \boldsymbol{\delta}) = -\int_{\mathcal{X}} \boldsymbol{\nabla} \xi_n \, \mathrm{d}\boldsymbol{\delta} = -\partial_1 \xi(0) - \partial_2 \xi(0) = -2,$$

thus reaching a contradiction. ■

Here is now an example of an element of $\mathscr{V}$, namely the *curve measure i.e.* a measure supported on a curve and defined through integration:

Definition 2.9 (Curve measure). *Let* $\boldsymbol{\gamma} : [0,1] \to \mathcal{X}$ *a parametrised Lipschitz*[2] *curve, we say that* $\boldsymbol{\mu_\gamma} \in \mathscr{V}$ is a measure supported on the curve $\boldsymbol{\gamma}$ *if:*

$$\forall \boldsymbol{g} \in \boldsymbol{\mathscr{C}_0(\mathcal{X})^2}, \quad \langle \boldsymbol{\mu_\gamma}, \boldsymbol{g} \rangle_{\boldsymbol{\mathcal{M}^2}} \stackrel{\text{def.}}{=} \int_0^1 \boldsymbol{g}(\boldsymbol{\gamma}(t)) \cdot \dot{\boldsymbol{\gamma}}(t) \, \mathrm{d}t.$$

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

Loosely speaking, the bracket w.r.t. to a curve measure is the circulation [22] of a test vector field function along the curve $\boldsymbol{\gamma}$. The rectifiability hypothesis can be understood as a finite length enforcement, indeed it is equivalent to $\mathscr{H}_1(\Gamma) < +\infty$. Please note that for the sake of conciseness, we indifferently refer to the curve measure $\boldsymbol{\mu_\gamma}$, or to the curve $\boldsymbol{\gamma}$ the measure is supported on. Let us introduce some properties that a curve might enjoy:

Definition 2.10 (Several characterisation of curves). *A curve is called* simple *if the restriction of* $\boldsymbol{\gamma}$ *to* $[0,1)$ *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 (it is then homotopic to the unit circle of* $\mathbb{R}^2$*).*

By Sard's theorem for Lipschitz function, one can show that $\boldsymbol{\mu_\gamma}$ is independent of the curve parametrisation [2]; thus we assume that $\boldsymbol{\gamma}$ has a constant speed parametrisation (unless stated otherwise). Let us precise that a curve measure can either be seen through a rather functional standpoint (in the latter section), or through a set angle [2]. Indeed, for every Borel set $B \subset \mathcal{X}$, one has:

$$\boldsymbol{\mu_\gamma}(B) = \int_{\Gamma \cap B} \left(\sum_{t \in \boldsymbol{\gamma}^{-1}(x)} \dot{\boldsymbol{\gamma}}(t) \right) \mathrm{d}\, \mathscr{H}_1(x). \tag{2.4}$$

Eventually, here is the simple expression of the divergence of a curve.

Proposition 2.11 (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 a closed curve.*

Proof. Let $\xi \in \mathscr{C}_0^\infty(\mathcal{X})$, then:

$$\begin{aligned}
\langle \operatorname{div} \boldsymbol{\mu_\gamma}, \xi \rangle_{\mathscr{D}'(\mathcal{X}) \times \mathscr{C}_0^\infty(\mathcal{X})} &= - \langle \boldsymbol{\mu_\gamma}, \boldsymbol{\nabla} \xi \rangle_{\mathcal{M}} \\
&= \int_0^1 \boldsymbol{\nabla} \xi(\boldsymbol{\gamma}(t)) \cdot \dot{\boldsymbol{\gamma}}(t) \, \mathrm{d}t \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} (\xi(\boldsymbol{\gamma}(t))) \, \mathrm{d}t \\
&= \xi(\boldsymbol{\gamma}(0)) - \xi(\boldsymbol{\gamma}(1)) \\
&= \langle \delta_{\boldsymbol{\gamma}(0)} - \delta_{\boldsymbol{\gamma}(1)}, \xi \rangle_{\mathscr{D}'(\mathcal{X}) \times \mathscr{C}_0^\infty(\mathcal{X})}.
\end{aligned}$$

Thus $\operatorname{div}(\boldsymbol{\mu_\gamma}) = \delta_{\boldsymbol{\gamma}(0)} - \delta_{\boldsymbol{\gamma}(1)}$ in the sense of distributions. ■

[2] hence 1-rectifiable. The latter property will be recalled to remind the reader of the charge (somehow) finite length.

Remark 2.12 (An insight for optimal transport connoisseurs). A natural question arises: *'is a charge supported on a curve always defined through integration, such as Definition* (2.9)*?'*. This concern is addressed in the De Rham theory [18] through the notion of *rectifiable integrable currents*. Shorthand: a charge has regularity thanks to its finite divergence thus evacuating pathological (non-integrable) cases.

3. A variational problem on the space of charges.

Similarly to the scalar case, one would want to define a BLASSO on $\mathscr{V}$. Let $\alpha > 0$, we present the following functional we coined CROC (standing for *Curves Represented On Charges*):

$$(\mathcal{Q}_\alpha(y)) \qquad \underset{\boldsymbol{m} \in \mathscr{V}}{\operatorname{argmin}}\, T_\alpha(\boldsymbol{m}) \overset{\text{def.}}{=} \frac{1}{2}\|y - \Phi\, \boldsymbol{m}\|_{\mathcal{H}}^2 + \alpha \|\boldsymbol{m}\|_{\mathscr{V}}.$$

$\Phi : \mathscr{V} \to \mathcal{H}$ is linear and maps the divergence vector field set to the acquisition space $\mathcal{H}$. We recall that $\|\boldsymbol{m}\|_{\mathscr{V}} \overset{\text{def.}}{=} \|\boldsymbol{m}\|_{\mathrm{TV}^2} + \|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}}$, the several terms of our energy may be interpreted in the following sense:

- $\frac{1}{2}\|y - \Phi\, \boldsymbol{m}\|_{\mathcal{H}}^2$ is the data term;
- $\|\boldsymbol{m}\|_{\mathrm{TV}^2}$ amounts to the length of the curve, indeed it is valued $\mathscr{H}_1(\Gamma)$ with Γ the image of the map $\boldsymbol{\gamma}$. It may discard high frequency oscillating behaviour and smooth over the curvature of $\boldsymbol{\gamma}$;
- $\|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}}$ is the (open) curve's counting term. Indeed, this term brings two convenient properties: first and foremost, it regularises the problem and allows yielding curves as CROC minimisers as proved in the latter section. On the other hand, if $\boldsymbol{\gamma}(0) \neq \boldsymbol{\gamma}(1)$ hence $\boldsymbol{\gamma}$ an open curve:

$$\|\operatorname{div} \boldsymbol{\mu}_{\boldsymbol{\gamma}}\|_{\mathrm{TV}} = \left\|\delta_{\boldsymbol{\gamma}(0)} - \delta_{\boldsymbol{\gamma}(1)}\right\|_{\mathrm{TV}} = 2.$$

A convenient improvement for $(\mathcal{Q}_\alpha(y))$ would be the penalisation of curve length and curve count with different weights, such as $\alpha, \beta > 0$ and the regularisation $\alpha\|\boldsymbol{m}\|_{\mathrm{TV}^2} + \beta\|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}}$. This enhancement is out of the scope of this article, but it may be handy for future numerical implementation.

Theorem 3.1. *The problem* $(\mathcal{Q}_\alpha(y))$ *admits solutions.*

Proof. The functional T_α is proper and coercive on $\mathscr{V}$. Let $(u_n)_{n\in\mathbb{N}}$ be a minimising sequence of T_α. Since T_α is coercive, $(u_n)_{n\in\mathbb{N}}$ is bounded in $\mathscr{V}$ so it converges in the weak-$*$ topology to some $\overline{u} \in \mathscr{V}$, up to a subsequence (see Appendix B). Observe also that:

- $\|\cdot\|_{\mathscr{V}}$ is lower semi-continuous w.r.t. the weak-$*$ convergence;
- Φ is continuous from $\boldsymbol{\mathcal{M}}\left(\boldsymbol{\mathcal{X}}\right)^{\mathbf{2}}$ weak-$*$ topology to $\mathcal{H}$ weak topology, since for all $q \in \mathcal{H}$ one has the mapping $x \mapsto \langle \varphi(x), q \rangle$ continuous.

Hence, by the direct method of the calculus of variations, it is clear that $\overline{u}$ is a minimiser of T_α, thus proving the existence of a minimiser. ■

Remark 3.2. The problem $(\mathcal{Q}_\alpha(y))$ enjoys uniqueness if Φ is injective, since it compels T_α strictly convex. Note that this constitutes a rather strong hypothesis, hardly fulfilled in applications, such as super-resolution for instance.

We have then proved the existence of a minimiser. Also consider:

Theorem 3.3. *Let $\boldsymbol{m}$ be a minimiser of* ($\mathcal{Q}_\alpha(y)$). *The dual problem admits a minimiser $q^* \in \mathscr{C}_0(\mathcal{X})$ optimal, strong duality holds, and extremality conditions read:*

$$\begin{cases} -\boldsymbol{\nabla} q^* & \in \partial \|\boldsymbol{m}\|_{\mathrm{TV}^2} + \boldsymbol{\Phi}^*(\Phi\, \boldsymbol{m} - y) \\ -q^* & \in \partial \|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}}. \end{cases} \tag{3.1}$$

Proof. Consider the Ekeland-Temam duality from [13, Remark 4.2], with a little caveat: the Banach space $\mathscr{V}$ has to be reflexive, which is clearly not the case here. However, the reflexive hypothesis is only needed for the sake of the existence proof. Since we already proved the solution existence, this reflexivity hypothesis is not required in our case. Back to the Remark 4.2 of [13] stating, for $\Lambda : \mathscr{V} \to Y$ linear, $F : \mathscr{V} \to \mathbb{R}$ and $G : Y \to \mathbb{R}$ convex, that the primal problem

$$\inf_{\boldsymbol{u} \in \mathscr{V}} F(\boldsymbol{u}) + G(\Lambda \boldsymbol{u})$$

has a dual problem writing down:

$$\sup_{p^* \in Y^*} -F^*(\Lambda^* p^*) - G^*(-p^*). \tag{3.2}$$

If $\boldsymbol{m}$ and p^* are respectively solutions of the primal and dual problem, the extremality conditions are:

$$\begin{cases} \Lambda^* p^* & \in \partial F(\boldsymbol{m}) \\ -p^* & \in \partial G(\Lambda \boldsymbol{m}). \end{cases}$$

Here we set the map

$$\begin{array}{rccc} \Lambda : & \mathscr{V} & \longrightarrow & \mathcal{M}(\mathcal{X}) \\ & \boldsymbol{u} & \longmapsto & \operatorname{div} \boldsymbol{u} \end{array}$$

and its dual

$$\begin{array}{rccc} \Lambda^* : & \mathscr{C}_0(\mathcal{X}) & \longrightarrow & \mathscr{V}^* \\ & p^* & \longmapsto & -\boldsymbol{\nabla} p^*. \end{array}$$

We also set for $\boldsymbol{u} \in \mathscr{V}$, $F(\boldsymbol{u}) \stackrel{\text{def.}}{=} \|\boldsymbol{u}\|_{\mathrm{TV}^2} + \frac{1}{2}\|y - \Phi\, \boldsymbol{u}\|_{\mathcal{H}}^2$, and $G(q) \stackrel{\text{def.}}{=} \|q\|_{\mathrm{TV}}$ for $q \in \mathcal{M}(\mathcal{X})$. The reader might be aware that G ought to be continuous at some Λu_0 where $u_0 \in \mathscr{V}$: this condition is fulfilled for the trivial null measure. Then the dual problem writes down:

$$\operatorname*{argmax}_{q^* \in \mathscr{C}_0(\mathcal{X})} -F^*(-\boldsymbol{\nabla} q^*) - G^*(-q^*).$$

Extremality conditions then boils down to q^* solution of the dual problem (3.2):

$$\begin{cases} \Lambda^* q^* & \in \partial F(\boldsymbol{m}) \\ -q^* & \in \partial G(\Lambda \boldsymbol{m}). \end{cases}$$ ■

In the off-the-grid literature [12, 6, 11], the dual problem ($\mathcal{D}_\lambda(y)$) of the BLASSO ($\mathcal{P}_\lambda(y)$) is useful to establish both the existence of the solutions and to characterise the so-called certificate. The latter is crucial for the sake of numerical implementation, since it is involved in both measure support estimation step and stopping condition of state-of-the-art greedy algorithms [6, 11].

Corollary 3.4. *Let $\boldsymbol{m} \in \mathscr{V}$ be a minimiser of ($\mathcal{Q}_\alpha(y)$). Then, there exist certificates $\boldsymbol{\eta_1} \in \partial \|\boldsymbol{m}\|_{\mathrm{TV}^2}$ and $\boldsymbol{\eta_2} = \boldsymbol{\nabla} q^*$ where $q^* \in \partial \|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}}$ with $\boldsymbol{\eta_1}, \boldsymbol{\eta_2} \in \boldsymbol{\mathscr{C}_0(\mathcal{X})^2}$ such that:*

$$\langle \boldsymbol{m}, \boldsymbol{\eta_1} + \boldsymbol{\eta_2} \rangle_{\boldsymbol{\mathcal{M}^2}} = \|\boldsymbol{m}\|_{\mathscr{V}}.$$

Proof. Let us investigate the two optimality conditions by probing a bit more the extremality conditions:

- let us consider the former optimality condition $-\boldsymbol{\nabla} q^* \in \partial F(\boldsymbol{m})$. In this case the sub-differential of the sum is the sum of the sub-differential[3], meaning:

$$\partial F(\boldsymbol{m}) = \partial \|y - \Phi(\cdot)\|^2_{\mathcal{H}}(\boldsymbol{m}) + \partial \|\cdot\|_{\mathrm{TV}^2}(\boldsymbol{m}),$$

by [13]. We reach, using the closed form of the least squares norm:

$$\partial \|y - \Phi(\cdot)\|^2_{\mathcal{H}}(\boldsymbol{m}) = \{- \boldsymbol{\Phi^*}(y - \Phi\, \boldsymbol{m})\}\,.$$

Then:

$$\begin{cases} -\boldsymbol{\nabla} q^* = \boldsymbol{\eta_1} - \boldsymbol{\Phi^*}(y - \Phi\, \boldsymbol{m}) \\ \boldsymbol{\eta_1} \in \partial \|\boldsymbol{m}\|_{\mathrm{TV}^2}. \end{cases}$$

The latter $\boldsymbol{\eta_1} \in \partial \|\boldsymbol{m}\|_{\mathrm{TV}^2}$, with the result in equation (2.3) adapted to the vector case, amounts to:

$$\begin{cases} \|\boldsymbol{\eta_1}\|_{\infty,\mathcal{X}} \leq 1 \\ \|\boldsymbol{m}\|_{\mathrm{TV}^2} = \langle \boldsymbol{m}, \boldsymbol{\eta_1} \rangle_{\boldsymbol{\mathcal{M}^2}}. \end{cases}$$

[3] Since the former sub-differential is single-valued, see [13].

- let us consider the latter optimality condition $-q^* \in \partial G(\operatorname{div} \boldsymbol{m})$. We note $\nu \stackrel{\text{def.}}{=} \operatorname{div} \boldsymbol{m}$. Then, as stated before in equation (2.3) and since $G \stackrel{\text{def.}}{=} \|\cdot\|_{\mathrm{TV}}$:

$$-q^* \in \partial G(\nu) \iff \begin{cases} q^* \in \mathscr{C}_0(\mathcal{X}) \\ \|q^*\|_{\infty,\mathcal{X}} \leq 1 \\ \langle -q^*, \nu \rangle_{\mathcal{M}} = \|\nu\|_{\mathrm{TV}}. \end{cases}$$

But, while denoting $\boldsymbol{\eta_2} \stackrel{\text{def.}}{=} \boldsymbol{\nabla} q^*$:

$$\|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}} = \langle -q^*, \operatorname{div} \boldsymbol{m} \rangle_{\mathcal{M}} = \langle \boldsymbol{\nabla} q^*, \boldsymbol{m} \rangle_{\boldsymbol{\mathcal{M}^2}} = \langle \boldsymbol{\eta_2}, \boldsymbol{m} \rangle_{\boldsymbol{\mathcal{M}^2}}.$$

Finally, if we merge these two results, we yield:

$$\begin{aligned} \|\boldsymbol{m}\|_{\mathrm{TV}^2} + \|\operatorname{div} \boldsymbol{m}\|_{\mathrm{TV}^2} &= \langle \boldsymbol{m}, \boldsymbol{\eta_1} \rangle_{\boldsymbol{\mathcal{M}^2}} + \langle \boldsymbol{\eta_2}, \boldsymbol{m} \rangle_{\boldsymbol{\mathcal{M}^2}} \\ &= \langle \boldsymbol{m}, \boldsymbol{\eta_1} + \boldsymbol{\eta_2} \rangle_{\boldsymbol{\mathcal{M}^2}}, \end{aligned}$$

thus ensuring $\|\boldsymbol{m}\|_{\mathscr{V}} = \langle \boldsymbol{m}, \boldsymbol{\eta_1} + \boldsymbol{\eta_2} \rangle_{\boldsymbol{\mathcal{M}^2}}$. Eventually, note that one has:

$$\boldsymbol{\eta_1} + \boldsymbol{\eta_2} = \Phi^*(y - \Phi\, \boldsymbol{m}).$$ ■

We have then specified the optimality conditions and defined the certificate $\boldsymbol{\eta_\alpha} = \boldsymbol{\eta_1} + \boldsymbol{\eta_2} \in \boldsymbol{\mathscr{C}_0(\mathcal{X})^2}$. This vector map has several applications from both theoretical and numerical standpoint: indeed, as one works out an algorithm based on the Frank-Wolfe algorithm [15], one needs for instance a stopping condition based on the criterion *'$\|\boldsymbol{\eta_\alpha}\|_{\infty,\mathcal{X}}$ less than a quantity'*. In the 'classic' weighted Dirac sum case, the dual certificate η simply interpolates the sign of the Dirac [12], and the numerical criterion amounts to $\|\eta\|_{\infty,\mathcal{X}} \leq 1$. The equivalent property for curves is not as clear-cut; indeed consider the BV case where the solution is a sum of indicator sets χ_E for E a simple set [9]: the criterion reads $\langle \chi_E, m \rangle_{\mathcal{M}} \leq P(E)$. We feel somehow that the stopping condition will be related to $\langle \boldsymbol{\eta_\alpha}, \boldsymbol{m} \rangle_{\boldsymbol{\mathcal{M}^2}} \leq \|\boldsymbol{m}\|_{\mathscr{V}}$, thereby ensuring that $\boldsymbol{\eta_\alpha}$ is in the sub-differential of the energy, namely the $\mathscr{V}$-norm. Also, let us caution that the precise characterisation of a curve measure dual certificate is not obvious, and is not a copy-and-paste of the Dirac case. However, one could note that:

Remark 3.5. If $\boldsymbol{m}$ is a measure supported on a curve $\boldsymbol{\gamma}$ such as in Definition 2.9 with support $\Gamma = \boldsymbol{\gamma}([0,1])$, then:

- since $\boldsymbol{m}$ is optimal, $\boldsymbol{\eta_1}$ realises the sup through $\|\boldsymbol{m}\|_{\mathrm{TV}^2} = \langle \boldsymbol{m}, \boldsymbol{\eta_1} \rangle_{\boldsymbol{\mathcal{M}^2}}$, and using the coarea formula in (2.4) coupled with a Cauchy-Schwarz equality, we reach with $x \in \mathcal{X}$ $\mathscr{H}_1$–a.e.:

$$\begin{cases} \boldsymbol{\eta_1}(x) = \dfrac{\sum_{t \in \boldsymbol{\gamma}^{-1}(x)} \dot{\boldsymbol{\gamma}}(t)}{\left|\sum_{t \in \boldsymbol{\gamma}^{-1}(x)} \dot{\boldsymbol{\gamma}}(t)\right|} & \text{if } \sum_{t \in \boldsymbol{\gamma}^{-1}(x)} \dot{\boldsymbol{\gamma}}(t) \neq 0 \\ |\boldsymbol{\eta_1}(x)| \leq 1 & \text{otherwise.} \end{cases}$$

Moreover, if $\boldsymbol{\gamma}$ is simple, it simply amounts to $|\boldsymbol{\eta_1}(x)| = 1$ if x belongs to the curve support Γ.

- As for $\boldsymbol{\eta_2}$, the analysis is quite tricky let alone intuitive. First, assume $\boldsymbol{\gamma}$ open ($\boldsymbol{\eta_2}$ is zeroing-out otherwise) and note that $\boldsymbol{\eta_2}$ is the gradient of function q^*. This function q^* lies in the sub-differential of G hence the TV-norm, evaluated at the point $\operatorname{div} \boldsymbol{m}$ which is a sum of Dirac. As stated before, a certificate associated to the TV-norm on a Dirac weighted sum is interpolating the sign of the Dirac [12]. Then q^* relates to a function interpolating the sign of the divergence $\operatorname{div} \boldsymbol{m}$, and $\boldsymbol{\eta_2}$ is the gradient of this function.

The precise study of the certificates is out of scope of this article and will be pursued in further numerical works.

However, we still lack a more precise description of ($\mathcal{Q}_\alpha(y)$) minimisers. To tackle this issue, let us introduce the following representer theorem.

We strongly advise the reader to notice that the choice of F and G in the further section will be different. Indeed, the choice in the latter section was only made for the sake of easing the dual problem computation. Let $n \in \mathbb{N}^*$, $\mathcal{H}_n$ be the finite dimensional n space of acquisitions, $G : \mathcal{H}_n \to \mathbb{R}$ an arbitrary data fitting term; suppose that $\Lambda : \mathscr{V} \to \mathcal{H}_n$ is linear and $F : \mathscr{V} \to \mathbb{R}$ is convex. Consider the general setting for $\boldsymbol{m} \in \mathscr{V}$:

$$J(\boldsymbol{m}) \stackrel{\text{def.}}{=} G(\Lambda \boldsymbol{m}) + F(\boldsymbol{m}). \tag{3.3}$$

The unit ball of the regulariser F is denoted $\mathcal{B}_F \stackrel{\text{def.}}{=} \{\boldsymbol{u} \in \mathscr{V}, F(\boldsymbol{u}) \leq 1\}$. The following theorem, due to [4, 5], establishes (up to some hypotheses) the link between the minimisers of the functional J in (3.3) and the extreme points (see section 4 and Definition 4.1) of $\mathcal{B}_F$ denoted $\operatorname{Ext}(\mathcal{B}_F)$:

Theorem 3.6 (Representer theorem). *If F is semi-norm, there exists $\overline{u} \in \mathscr{V}$, a minimiser of (3.3) with the representation:*

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

where $p \leq \dim \mathcal{H}_n$, $\boldsymbol{u_i} \in \operatorname{Ext}(\mathcal{B}_F)$ and $\alpha_i > 0$ with $\sum_{i=1}^{p} \alpha_i = F(\overline{u})$.

It amounts to the characterisation of minimisers' structural properties, without actually needing to solve the problem [4]. Moreover, the description of the extreme points of the regulariser unit ball is fundamental for the numerical implementation. Indeed, the Frank-Wolfe algorithm recovers a solution by iteratively adding F-unit ball extreme points to the reconstructed measure: at each step, the algorithm adds a new charge to the reconstructed measure, this new charge has a prescribed structure since it has to be an extreme point of the regulariser.

In the following, it rather makes sense to consider $F : \boldsymbol{m} \mapsto \|\boldsymbol{m}\|_{\mathscr{V}}$, and obviously the data term $G : p \mapsto \|y - p\|_{\mathcal{H}}^2$ with $\Lambda = \Phi$.

4. Extreme points of the $\mathscr{V}$-norm's unit ball are curves.

First, we recall some definitions and convenient results concerning extreme points.

Definition 4.1. *Let X be a topological vector space and $K \subset X$. An* extreme point x of K *is a point such that:*

$$\forall y, z \in K, \quad \forall \lambda \in (0,1), \qquad x = \lambda y + (1-\lambda) z \implies x = y = z.$$

The set of extreme points of K is denoted by $\operatorname{Ext} K$.

We further introduce the celebrated Krein-Milman theorem, stating that if K convex and compact, the closed convex hull of the set of the extreme points of K coincides with K.

Theorem 4.2 (Krein-Milman theorem). *If $K \subset X$ is a non-empty, compact and convex subset of a Hausdorff locally convex space, then $K = \overline{\operatorname{co}}(\operatorname{Ext}(K))$.*

Also note the following refinement of Krein-Milman, stating the decomposition of any point of the space onto a 'combination' of extreme points, through a concept [19] called the Choquet integral:

Theorem 4.3 (Choquet theorem). *Let X be a metrisable topological vector space and K a non-empty, convex and compact subset. Then for any $x \in K$ there exists a Borel measure ρ on X, concentrated on $\operatorname{Ext} K$ and satisfying:*

$$x = \int_{\operatorname{Ext} K} r \, \mathrm{d}\rho(r)$$

namely for every $\omega : X \to \mathbb{R}$ linear and continuous, $\omega(x) = \int_{\operatorname{Ext} K} \omega(r) \, \mathrm{d}\rho(r)$.

These notions lie in a very general framework, let us precise some theoretical tools pertaining to the space of charges for our main theorem. Consider the following [22, 16]:

Definition 4.4. *Let $\boldsymbol{T} \in \mathscr{V}$ and $J \subset \mathscr{V}$, we say that $\boldsymbol{T}$* decomposes *into charges lying in J if there exists a finite positive Borel measure ρ on J such that:*

$$\boldsymbol{T} = \int_J \boldsymbol{R} \, \mathrm{d}\rho(\boldsymbol{R})$$

$$\|\boldsymbol{T}\|_{\mathrm{TV}^2} = \int_J \|\boldsymbol{R}\|_{\mathrm{TV}^2} \, \mathrm{d}\rho(\boldsymbol{R}).$$

We say that $\boldsymbol{T}$ completely decomposes *into charges lying in J if the latter conditions hold and moreover if:*

$$\|\operatorname{div} \boldsymbol{T}\|_{\mathrm{TV}} = \int_J \|\operatorname{div} \boldsymbol{R}\|_{\mathrm{TV}} \, \mathrm{d}\rho(\boldsymbol{R}).$$

In our case we can use the following sets:

Definition 4.5 (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:*

$$\mathfrak{S} \stackrel{\text{def.}}{=} \left\{ \frac{\boldsymbol{\mu}_{\boldsymbol{\gamma}}}{\|\boldsymbol{\mu}_{\boldsymbol{\gamma}}\|_{\mathscr{V}}}, \boldsymbol{\gamma} \textit{ is a simple oriented Lipschitz curve} \right\}.$$

It is a (non-complete) metric space for the weak-$*$ topology. We also denote by $\mathfrak{S}_{\text{loop}}$ the space of curve measures supported on loops:

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

Eventually we note $\mathfrak{S}_{\text{open}} \overset{\text{def.}}{=} \mathfrak{S} \backslash \mathfrak{S}_{\text{loop}}$ the space of open curves.

We now recall the following fundamental results of [22, 16]. Remember that a solenoid is a charge of $\mathscr{V}$ with null divergence:

Theorem 4.6 (Solenoid decomposition theorem [16, Proposition 1.17]). *Any solenoid $\boldsymbol{T} \in \mathscr{V}$ can be decomposed into elements of $\mathfrak{S}_{\text{loop}}$.*

Let us stress out that this latter theorem only holds in the $d = 2$ case; indeed there exists several counterexamples for $d > 2$, see [22] for various illustrations. The following theorem holds in any dimension:

Theorem 4.7 (Smirnov's Theorem C [22], refined with [16, Proposition 1.17]). *Any charge $\boldsymbol{T} \in \mathscr{V}$ is a sum of two charges $\boldsymbol{P}$ and $\boldsymbol{Q}$ such that $\operatorname{div} \boldsymbol{P} = 0$ (Theorem (4.6) can then be applied) and $\boldsymbol{Q}$ is completely decomposable into measures supported on 1-rectifiable Lipschitz simple oriented curves.*

We now present the main result of this paper, namely the theorem establishing a link between curve measures and extreme points of the unit ball of the regulariser (*i.e.* the $\mathscr{V}$-norm).

Theorem 4.8 (Main result). *Let us denote by $\mathcal{B}^1_{\mathscr{V}} \overset{\text{def.}}{=} \{\boldsymbol{m} \in \mathscr{V}, \|\boldsymbol{m}\|_{\mathscr{V}} \leq 1\}$ the unit ball of the norm of $\mathscr{V}$, weakly-$*$ compact. Then one has:*

$$\operatorname{Ext}(\mathcal{B}^1_{\mathscr{V}}) = \mathfrak{S} = \left\{ \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}, \boldsymbol{\gamma} \textit{ is a 1-rectifiable simple oriented Lipschitz curve} \right\}.$$

Proof. We start by proving

$$\operatorname{Ext}(\mathcal{B}^1_{\mathscr{V}}) \supset \left\{ \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}, \boldsymbol{\gamma} \text{ is a 1-rectifiable simple oriented Lipschitz curve} \right\}.$$

Let $\boldsymbol{\gamma} : [0,1] \to \mathcal{X}$ be a Lipschitz 1-rectifiable simple oriented curve of length $\ell > 0$, with image $\Gamma = \boldsymbol{\gamma}([0,1])$, and $\boldsymbol{\mu_\gamma}$ the measure supported on this curve. Note that it is either an open or a closed curve. By contradiction, let us suppose that there exists $\boldsymbol{u_1}, \boldsymbol{u_2} \in \mathscr{V}$ such that $\|\boldsymbol{u_1}\|_{\mathscr{V}} \leq 1$, $\|\boldsymbol{u_2}\|_{\mathscr{V}} \leq 1$ and for $\lambda \in (0,1)$:

$$\frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} = \lambda \boldsymbol{u_1} + (1-\lambda)\boldsymbol{u_2}. \tag{4.1}$$

For every $A \subset \mathcal{X}$ Borel, while denoting $|\cdot|_{\mathscr{V}}(A) \overset{\text{def.}}{=} |\cdot|(A) + |\operatorname{div} \cdot|(A)$, one has:

$$\frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A) = \lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A). \tag{4.2}$$

Indeed, following a proof from [5, Theorem 4.7], if there exists $A \subset \mathscr{X}$ such that $\lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A) > \frac{|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A)}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}$ then we get

$$\begin{aligned}
1 &= \frac{|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(\mathscr{X})}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \\
&= \frac{|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A)}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} + \frac{|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A^c)}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \\
&< \lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A) + \frac{|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A^c)}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \\
&< \lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A) + \lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A^c) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A^c) \\
&< 1.
\end{aligned}$$

So we reach $1 < 1$ which yields a contradiction, then:

$$\lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A) \leq \frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}|\boldsymbol{\mu_\gamma}|_{\mathscr{V}}(A).$$

Since we always have the opposite (triangle) inequality, we get (4.2). We further deduce from equation (4.2) that:

$$\|\boldsymbol{u_1}\|_{\mathscr{V}} = \|\boldsymbol{u_2}\|_{\mathscr{V}} = 1. \tag{4.3}$$

Note that the same argument can be used on both terms of the $\mathscr{V}$–norm to reach for all Borel set A:

$$\frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}|\boldsymbol{\mu_\gamma}|_{\mathrm{TV}^2}(A) = \lambda|\boldsymbol{u_1}|_{\mathrm{TV}^2}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathrm{TV}^2}(A) \tag{4.4}$$

$$\frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}|\operatorname{div}\boldsymbol{\mu_\gamma}|_{\mathrm{TV}}(A) = \lambda|\operatorname{div}\boldsymbol{u_1}|_{\mathrm{TV}}(A) + (1-\lambda)|\operatorname{div}\boldsymbol{u_2}|_{\mathrm{TV}}(A). \tag{4.5}$$

Let $A \subset \mathscr{X}$ be a Borel set such that $A \cap \Gamma = \varnothing$, then from (4.2) we deduce:

$$0 = \lambda|\boldsymbol{u_1}|_{\mathscr{V}}(A) + (1-\lambda)|\boldsymbol{u_2}|_{\mathscr{V}}(A) \implies \boldsymbol{u_1}(A) = \boldsymbol{u_2}(A) = 0. \tag{4.6}$$

By setting $A = \Gamma^c$ in equation (4.6), we yield $\operatorname{spt}\boldsymbol{u_i} \subset \Gamma$. Now, using Theorems 4.6 and 4.7, there exists two finite positive measures ρ_i[4] measures on $\mathfrak{S}$ such that:

[4]Since $\|\boldsymbol{u_i}\|_{\mathscr{V}} = \int_{\mathfrak{S}} \underbrace{\|\boldsymbol{R}\|_{\mathscr{V}}}_{=1} \,\mathrm{d}\rho_i(\boldsymbol{R}) = \rho_i(\mathfrak{S})$ for $\|\boldsymbol{u_i}\|_{\mathscr{V}} = 1$, it is clear-cut that $\rho_i(\mathfrak{S}) = 1$.

$$\boldsymbol{u_i} = \int_{\mathfrak{S}} \boldsymbol{R} \,\mathrm{d}\rho_i(\boldsymbol{R}).$$

We denote $\rho \overset{\text{def.}}{=} \lambda\rho_1 + (1-\lambda)\rho_2$, hence we yield:

$$\lambda \boldsymbol{u_1} + (1-\lambda)\boldsymbol{u_2} = \int_{\mathfrak{S}} \boldsymbol{R} \,\mathrm{d}\rho(\boldsymbol{R}).$$

First, each curve of the decomposition is included in the support of the charge, *i.e.* ρ–a.e. one has $\operatorname{spt} \boldsymbol{R} \subseteq \operatorname{spt} \boldsymbol{\mu_\gamma}$ as pointed out by Smirnov [22, Remark 5]. Also using equations (4.4), note that:

$$\begin{aligned}
\frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2} &= \|\lambda \boldsymbol{u_1} + (1-\lambda)\boldsymbol{u_2}\|_{\mathrm{TV}^2} \\
&= \lambda\|\boldsymbol{u_1}\|_{\mathrm{TV}^2} + (1-\lambda)\|\boldsymbol{u_2}\|_{\mathrm{TV}^2} \\
&= \lambda \int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho_1(\boldsymbol{R}) + (1-\lambda)\int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho_2(\boldsymbol{R}) \\
&= \int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho(\boldsymbol{R}).
\end{aligned}$$

Hence,

$$\frac{1}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2} = \int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho(\boldsymbol{R}) \tag{4.7}$$

Then, $\boldsymbol{R}$ ought to be maximal in length, *i.e.* $\operatorname{spt} \boldsymbol{R} = \operatorname{spt} \boldsymbol{\mu_\gamma}$. Suppose elseways that $\operatorname{spt} \boldsymbol{R} \subsetneq \operatorname{spt} \boldsymbol{\mu_\gamma}$ on a set of positive ρ measure. The curve associated with $\boldsymbol{R}$ is necessarily open, then:

$$\|\boldsymbol{R}\|_{\mathrm{TV}^2} = \frac{\mathscr{H}_1(\operatorname{spt} \boldsymbol{R})}{\mathscr{H}_1(\operatorname{spt} \boldsymbol{R}) + 2} < \frac{\mathscr{H}_1(\operatorname{spt} \boldsymbol{\mu_\gamma})}{\mathscr{H}_1(\operatorname{spt} \boldsymbol{\mu_\gamma}) + 2} \leq \frac{\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}.$$

Hence $\|\boldsymbol{R}\|_{\mathrm{TV}^2} < \dfrac{\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}$, therefore:

$$\begin{aligned}
\int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho(\boldsymbol{R}) &< \frac{\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \int_{\mathfrak{S}} \mathrm{d}\rho \\
&< \frac{\|\boldsymbol{\mu_\gamma}\|_{\mathrm{TV}^2}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \underbrace{\rho(\mathfrak{S})}_{=1} = \int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2} \,\mathrm{d}\rho(\boldsymbol{R}),
\end{aligned}$$

implying

$$\int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2}\, \mathrm{d}\rho(\boldsymbol{R}) < \int_{\mathfrak{S}} \|\boldsymbol{R}\|_{\mathrm{TV}^2}\, \mathrm{d}\rho(\boldsymbol{R}),$$

thus reaching a contradiction. We denote by $\boldsymbol{\gamma_R}$ the curve parametrisation with constant speed associated to $\boldsymbol{R}$, let us now sum up the properties we gathered:

- $\operatorname{spt} \boldsymbol{R} = \operatorname{spt} \boldsymbol{\mu_\gamma}$, ρ-a.e.;
- $\boldsymbol{\gamma_R}$ is Lipschitz simple.

Let $x \in \operatorname{spt} \boldsymbol{R}$. Then there exists $t \in [0, 1)$ such that $\boldsymbol{\gamma_R}(t) = x$. Since the supports of $\boldsymbol{R}$ and $\boldsymbol{\mu_\gamma}$ are shared, observe that there also exists $\tilde{t} \in (0, 1)$ such that $\boldsymbol{\gamma}(\tilde{t}) = x$, then $\boldsymbol{\gamma_R}(t) = \boldsymbol{\gamma}(\tilde{t})$. Since both curves are simple, obviously $t = \boldsymbol{\gamma_R}^{-1}(\boldsymbol{\gamma}(\tilde{t}))$, implying that $\boldsymbol{\gamma_R}$ is a reparametrisation of $\boldsymbol{\gamma}$. Hence, with $\|\boldsymbol{R}\|_{\mathscr{V}} = \|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}} = 1$, one has $\boldsymbol{R} = \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}$ ρ-a.e. therefore yielding:

$$\boldsymbol{u_i} = \int_{\mathfrak{S}} \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}\, \mathrm{d}\rho_i = \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \underbrace{\rho_i(\mathfrak{S})}_{=1} = \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}.$$

It ensues that either $\lambda = 0$ or $\lambda = 1$, $\boldsymbol{u_1} = \boldsymbol{u_2} = \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}$; thus concluding this part.

We now prove the converse, namely:

$$\operatorname{Ext}(\mathcal{B}^1_{\mathscr{V}}) \subset \left\{ \frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}},\ \boldsymbol{\gamma} \text{ is a 1-rectifiable simple oriented Lipschitz curve} \right\}.$$

Let $\boldsymbol{T} \in \operatorname{Ext}(\mathcal{B}^1_{\mathscr{V}})$ be an extreme point of the unit ball norm of $\mathscr{V}$, one has $\|\boldsymbol{T}\|_{\mathscr{V}} = 1$ and there exists[5] a finite (probability) Borel measure ρ such that:

$$\boldsymbol{T} = \int_{\mathfrak{S}} \boldsymbol{R}\, \mathrm{d}\rho(\boldsymbol{R}),$$

by Smirnov's decomposition in Theorem (4.7) [22, 16, 3]. Since $\boldsymbol{T}$ is an extreme point, by [19, Proposition 1.4] ρ is an atomic measure, supported on a singleton of $\mathfrak{S}$ and the proof is achieved: there exists a 1-rectifiable simple Lipschitz curve $\boldsymbol{\gamma}$ such that $\frac{\boldsymbol{\mu_\gamma}}{\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}} \in \mathfrak{S}$ and $\rho = \delta_{\boldsymbol{\mu_\gamma}/\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}}$ hence $\boldsymbol{T} = \boldsymbol{\mu_\gamma}/\|\boldsymbol{\mu_\gamma}\|_{\mathscr{V}}$, therefore concluding the proof. ■

Remark 4.9. One can see Smirnov's theorems A and C of [22] as a refined Choquet integral result. Indeed, the article yields the same conclusions (with some extensions) as the application of Choquet theorem, *i.e.* every charge decomposes as a weighted average of the set of extreme points $\mathfrak{S}$.

[5] Actually, this is allowed by the decomposition $\boldsymbol{T} = \boldsymbol{P} + \boldsymbol{Q}$ from Theorem 4.7 and $d = 2$. Hence, both $\boldsymbol{P}$ and $\boldsymbol{Q}$ decompose on $\mathfrak{S}$ and their respective Borel measures can be summed up to reach a unique Borel measure ρ concentrated on $\mathfrak{S}$.

We have then successfully established a result on extreme points, thus enabling a promising avenue for numerical implementation. Indeed, the state-of-the-art algorithms in the off-the-grid literature boil down to the iterative reconstruction of a linear sum of extreme points of the energy regulariser.

5. Outlook. This article performed the analysis of the space of divergence vector fields $\mathscr{V}$, and proposed a new optimisation problem $(\mathcal{Q}_\alpha(y))$ based on an off-the-grid functional called CROC. Existence of CROC minimisers and characterisation of the associated certificates have been established. Moreover, the extreme point theorem 4.8, coupled with the representer theorem 3.6, allows a precise analysis of the structure of $(\mathcal{Q}_\alpha(y))$ minimisers.

In further works, we plan to exploit this theoretical result through an adaptation of the (Sliding) Frank-Wolfe algorithm [6, 11] suited for curves. Indeed, this greedy algorithm reconstructs an iterated linear combination of extreme points, hence offering a tractable approach for curve reconstruction in an off-the-grid fashion.

Appendix A. $\mathscr{V}$ is a Banach space . Let $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{V}$. By definition, it is also a Cauchy sequence in $\mathcal{M}(\mathcal{X})^2$, which is complete therefore it admits a limit $u \in \mathcal{M}(\mathcal{X})^2$. Since $(u_n)_{n\in\mathbb{N}}$ is bounded (as a Cauchy sequence) for the norm topology in $\mathscr{V}$ and by lower semi-continuity of $\|\mathrm{div}(\cdot)\|_{\mathrm{TV}}$, then:

$$\|\mathrm{div}\, u\|_{\mathrm{TV}} \leq \liminf_{n\to+\infty} \|\mathrm{div}\, u_n\|_{\mathrm{TV}} < +\infty$$

and since $u \in \mathcal{M}(\mathcal{X})^2$, one has $\|u\|_{\mathscr{V}} = \|\mathrm{div}\, u\|_{\mathrm{TV}} + \|u\|_{\mathrm{TV}^2} < +\infty$. Hence $u \in \mathscr{V}$, let us now study the convergence in $\mathscr{V}$. Since $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, there exists $\nu \in \mathcal{M}(\mathcal{X})$ such that $\mathrm{div}\, u_n \xrightarrow[n\to+\infty]{} \nu$ for the TV-topology. Then up to the density argument, for all $\phi \in \mathscr{C}_0^\infty(\mathcal{X})$, one has:

$$\langle \mathrm{div}\, u_n, \phi\rangle \xrightarrow[n\to+\infty]{} \langle \nu, \phi\rangle$$

but by derivative of distributions $\langle \mathrm{div}\, u_n, \phi\rangle = \langle -u_n, \boldsymbol{\nabla}\phi\rangle$ and since $\langle u_n, \boldsymbol{\nabla}\phi\rangle \to \langle u, \boldsymbol{\nabla}\phi\rangle$ one has $\nu = \mathrm{div}\, u = \lim_{n\to+\infty} \mathrm{div}\, u_n$ in the TV-topology. Then,

$$\|u_n - u\|_{\mathscr{V}} = \|u_n - u\|_{\mathrm{TV}^2} + \|\mathrm{div}\, u_n - \mathrm{div}\, u\|_{\mathrm{TV}} \xrightarrow[n\to+\infty]{} 0,$$

thus ensuring that $\mathscr{V}$ is a Banach space, and therefore concluding the proof.

Appendix B. Every bounded sequence of $\mathscr{V}$ has a weak-$*$ convergent subsequence in $\mathscr{V}$. Let $(u_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathscr{V}$, hence in $\mathcal{M}(\mathcal{X})^2$. By the Banach-Alaoglu theorem, there exists a weak-$*$ convergent subsequence, namely a point $u \in \mathcal{M}(\mathcal{X})^2$ and a strictly increasing positive function $\varphi : \mathbb{N} \to \mathbb{N}$ such that:

$$u_{\varphi(n)} \overset{*}{\rightharpoonup} u \quad \text{in}\, \mathcal{M}(\mathcal{X})^2.$$

Since $(u_n)_{n\in\mathbb{N}}$ is bounded for the norm topology in $\mathscr{V}$ and by lower semi-continuity of $\|\mathrm{div}(\cdot)\|_{\mathrm{TV}}$, then:

$$\|\mathrm{div}\, u\|_{\mathrm{TV}} \leq \liminf_{n\to+\infty} \left\|\mathrm{div}\, u_{\varphi(n)}\right\|_{\mathrm{TV}} < +\infty,$$

and one might prove that $u \in \mathscr{V}$, while $\mathrm{div}\, u_{\varphi(n)} \overset{*}{\rightharpoonup} \mathrm{div}\, u$ using the same argument as the latter appendix. Hence $u_{\varphi(n)} \overset{*}{\rightharpoonup} u$ in $\mathscr{V}$ and this completes the proof.

Acknowledgments. We would like to thank both Ms Vasiliki Stergiopoulou and Mr Sébastien Schaub for their help on the *Ostreopsis* SOFI imaging; Mr Théo Bertrand for his helpful remarks on the draft and his advice on the blood vessel ultrasound imaging. We also would like to express our gratitude to the reviewers, whose relevant remarks and careful reviews helped to greatly improve the overall quality, clarity and correctness of the paper. Eventually, we would like to thank the `EUR-DS4H` and `3IA` fundings for their financial support.

REFERENCES

[1] C. Arthur, B. Heiles, H. Vincent, Lopez Pauline, Eliott Teston, and C. Olivier, *Opulm pala*, 2020, https://doi.org/10.5281/ZENODO.4343435.

[2] L. Baratchart, D. Hardin, and C. Villalobos-Guillén, *Divergence-free measures in the plane and inverse potential problems in divergence form*, https://arxiv.org/abs/2006.09072.

[3] P. Bonicatto and N. A. Gusev, *On the structure of divergence-free measures on* $\mathbb{R}^2$, arXiv:1912.10936 [math], (2020), http://arxiv.org/abs/1912.10936 (accessed 2021-12-13). arXiv: 1912.10936.

[4] C. Boyer, A. Chambolle, Y. D. Castro, V. Duval, F. de Gournay, and P. Weiss, *On representer theorems and convex regularization*, SIAM Journal on Optimization, 29 (2019), pp. 1260–1281, https://doi.org/10.1137/18m1200750, https://arxiv.org/abs/1806.09810.

[5] K. Bredies and M. Carioni, *Sparsity of solutions for variational inverse problems with finite-dimensional data*, Calculus of Variations and Partial Differential Equations, 59 (2019), https://doi.org/10.1007/s00526-019-1658-1, https://arxiv.org/abs/1809.05045.

[6] K. Bredies and H. K. Pikkarainen, *Inverse problems in spaces of measures*, ESAIM: Control, Optimisation and Calculus of Variations, 19 (2012), pp. 190–218, https://doi.org/10.1051/cocv/2011205.

[7] E. J. Candès and C. Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics, 67 (2013), pp. 906–956, https://doi.org/10.1002/cpa.21455, https://arxiv.org/abs/1203.5871.

[8] L. Chizat and F. Bach, *On the global convergence of gradient descent for over-parameterized models using optimal transport*, Advances in Neural Information Processing Systems (NIPS), (2018), https://arxiv.org/abs/1805.09545.

[9] Y. de Castro, V. Duval, and R. Petit, *Towards off-the-grid algorithms for total variation regularized inverse problems*, in Lecture Notes in Computer Science, Springer International Publishing, 2021, pp. 553–564, https://arxiv.org/abs/2104.06706.

[10] Y. de Castro and F. Gamboa, *Exact reconstruction using beurling minimal extrapolation*, Journal of Mathematical Analysis and Applications, 395 (2012), pp. 336–354, https://doi.org/10.1016/j.jmaa.2012.05.011, https://arxiv.org/abs/1103.4951.

[11] Q. Denoyelle, V. Duval, G. Peyré, and E. Soubies, *The sliding frank–wolfe algorithm and its application to super-resolution microscopy*, Inverse Problems, 36 (2019), p. 014001, https://doi.org/10.1088/1361-6420/ab2a29, https://arxiv.org/abs/1811.06416.

[12] V. Duval and G. Peyré, *Exact support recovery for sparse spikes deconvolution*, Foundations of Computational Mathematics, 15 (2014), pp. 1315–1355, https://doi.org/10.1007/s10208-014-9228-6, https://arxiv.org/abs/1306.6909.

[13] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Society for Industrial and Applied Mathematics, Jan 1999, https://doi.org/10.1137/1.9781611971088, https://doi.org/10.1137/1.9781611971088.
[14] H. Federer, *Geometric measure theory*, Springer, Berlin, 1996.
[15] M. Frank and P. Wolfe, *An algorithm for quadratic programming*, Naval Research Logistics Quarterly, 3 (1956), pp. 95–110, https://doi.org/10.1002/nav.3800030109.
[16] V. P. Khavin and S. K. Smirnov, *Approximation and extension problems for some classes of vector fields*, St. Petersburg Department of Steklov Institute of Mathematics, Russian Academy of Sciences, 10 (1998), pp. 507–528, https://mathscinet.ams.org/mathscinet-getitem?mr=1628034 (accessed 2022-02-07).
[17] B. Laville, L. Blanc-Féraud, and G. Aubert, *Off-The-Grid Variational Sparse Spike Recovery: Methods and Algorithms*, Journal of Imaging, 7 (2021), p. 266, https://doi.org/10.3390/jimaging7120266, https://www.mdpi.com/2313-433X/7/12/266 (accessed 2021-12-07).
[18] F. Morgan, *Geometric measure theory : a beginner's guide*, Academic Press, Boston, 1988.
[19] R. R. Phelps, ed., *Lectures on Choquet's Theorem*, Springer Berlin Heidelberg, 2001, https://doi.org/10.1007/b76887.
[20] W. Rudin, *Real and Complex Analysis*, MCGRAW HILL BOOK CO, 1986, https://doi.org/10.2307/2348852, https://www.ebook.de/de/product/4028964/walter_rudin_real_and_complex_analysis.html.
[21] M. Šilhavỳ, *Divergence measure vectorfields: their structure and the divergence theorem*, Mathematical modelling of bodies with complicated bulk and boundary behavior, 20 (2008), pp. 214–237.
[22] S. K. Smirnov, *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 (1993), pp. 206–238, https://mathscinet.ams.org/mathscinet-getitem?mr=1246427 (accessed 2022-02-07).