

Geometric properties of solutions to the total variation denoising problem

Antonin Chambolle¹, Vincent Duval², Gabriel Peyré^{3,4,2},
Clarice Poon^{4,2,5}

¹ CMAP, Ecole Polytechnique, CNRS, 91128 Palaiseau, FRANCE

² INRIA, MOKAPLAN, 2 rue Simone Iff, Paris, FRANCE

³ CNRS

⁴ CEREMADE, Université Paris-Dauphine, CNRS, Place du Marechal De Lattre De Tassigny, 75775 Paris 16, FRANCE

⁵ DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0DZ, UK

E-mail: antonin.chambolle@cmap.polytechnique.fr, vincent.duval@inria.fr
gabriel.peyre@ceremade.dauphine.fr, cmhsp2@cam.ac.uk

Abstract. This article studies the denoising performance of total variation (TV) image regularization. More precisely, we study geometrical properties of the solution to the so-called Rudin-Osher-Fatemi total variation denoising method. The first contribution of this paper is a precise mathematical definition of the “extended support” (associated to the noise-free image) of TV denoising. It is intuitively the region which is unstable and will suffer from the staircasing effect. We highlight in several practical cases, such as the indicator of convex sets, that this region can be determined explicitly. Our second and main contribution is a proof that the TV denoising method indeed restores an image which is exactly constant outside a small tube surrounding the extended support. The radius of this tube shrinks toward zero as the noise level vanishes, and we are able to determine, in some cases, an upper bound on the convergence rate. For indicators of so-called “calibrable” sets (such as disks or properly eroded squares), this extended support matches the edges, so that discontinuities produced by TV denoising cluster tightly around the edges. In contrast, for indicators of more general shapes or for complicated images, this extended support can be larger. Beside these main results, our paper also proves several intermediate results about fine properties of TV regularization, in particular for indicators of calibrable and convex sets, which are of independent interest.

1. Introduction

The total variation (TV) denoising method was introduced by Rudin, Osher and Fatemi in [79]. It is one of the first proposed non-linear image restoration methods, and has had an enormous impact on shaping modern imaging sciences. Despite being quite old, this method is still routinely used today, and its popularity probably stems from both its simplicity and its ability to restore “cartoon-looking” images. While far from being the state of the art for denoising in terms of performance (see Section 1.2 for some

more recent works), it is still featured as a benchmark in most papers being published on image restoration.

1.1. Total Variation Denoising

The total variation of a function $u \in L^2(\mathbb{R}^2)$ is defined as

$$J(u) \stackrel{\text{def.}}{=} \int_{\mathbb{R}^2} |Du| \stackrel{\text{def.}}{=} \sup \left\{ \int_{\mathbb{R}^2} u \operatorname{div} z ; z \in \mathcal{C}_c^1(\mathbb{R}^2, \mathbb{R}^2), \|z\|_\infty \leq 1 \right\}. \quad (1)$$

Given some noisy input function f , following [79], we are interested in the total variation denoising problem

$$\min_{u \in L^2(\mathbb{R}^N)} \lambda J(u) + \frac{1}{2} \|u - f\|_{L^2}^2. \quad (\mathcal{P}_\lambda(f))$$

Here, $\lambda > 0$ is the regularization parameter, and it should be adapted by the user to the noise level.

The goal of this paper is to study the ability to restore the geometrical structures (in particular the edges) of some (typically unknown) noise-free function f by solving $\mathcal{P}_\lambda(f + w)$, i.e. by applying TV regularization to the input noisy image $f + w$. Here w accounts for some additive noise in the image formation process, and is assumed to have a finite L^2 norm $\|w\|_{L^2}$.

Let us first emphasize that our analysis focusses on regimes where the noise and the regularization parameter are small. It is not very difficult to see that, as $\lambda \rightarrow 0^+$ and $\|w\|_{L^2} \rightarrow 0^+$, the solution $u_{\lambda,w}$ to $\mathcal{P}_\lambda(f + w)$ converges towards f in the L^2 topology. Our goal is to describe this convergence more precisely: *is it possible to say that the level lines of $u_{\lambda,w}$ converge to those of f ? In what sense? Moreover, does the support of $Du_{\lambda,w}$ converge towards the support of Df ?* These questions are all the more important as it is widely acknowledged in image analysis theory that the shape information of an image is contained in the level sets of an image [82, 87], determined in particular by their boundary.

1.2. Previous Works

Image restoration. The TV denoising method, often referred to as the Rudin-Osher-Fatemi (ROF) model, was introduced in [79]. Its basic properties (including the existence and uniqueness of the solution) are derived in [38]. We refer to [39] for an introduction to this model and an overview of its numerous applications in image processing. A thorough study of its properties can be found in [3, 4]. It is important to realize that TV is no longer the state of the art in imaging sciences, and we refer to recent works such as [2, 25, 43, 76] that obtain superior denoising performance on natural images by exploiting more complex and involved regularizers and statistical models.

Beyond denoising, TV methods have been used successfully to solve a wide range of ill-posed inverse problems, see for instance [1, 40, 41, 69]. Following the work of Meyer [71], TV regularization in conjunction to a norm dual of TV (favoring oscillations) is used to separate texture from structure, see for instance [16]. In a finite dimensional setting (using a discretization of the gradient operator), TV methods have been used to solve compressed sensing, where the linear operator is randomized [72, 75] to obtain accurate reconstructions when the number of random samples is nearly proportional to the number of the discretized edges.

Jump sets stability. The use of non-smooth (possibly non-convex) regularizations to restore edges and promote sharp features has been advocated by Mila Nikolova. She provided in a series of papers a detailed analysis of a general class of regularization schemes which admit piecewise smooth solutions, see for instance [73]. In the case of the TV regularization, this analysis can be refined. Explicit solutions are known, mostly in 1-D and for radial 2-D functions (see for instance [83]), as well as for indicators of convex sets in the plane [5, 8]. They suggest that TV methods indeed do maintain sharp features. A landmark result is presented in [30] which shows that total variation regularization does not introduce new jumps, i.e. the “jump set” of the solution of $(\mathcal{P}_\lambda(f))$ is included in the one of the input f . A review of this result and extensions can be found in [85].

These results are, however, of little interest when f is replaced by a noisy function $f + w$ (which is the setting of practical use of the method), since the noise w , which is only assumed to be in L^2 , might introduce jumps everywhere. It is actually the presence of this noise which is responsible for the “staircasing” effect, which creates spurious edges in flat area. Properties of this staircasing are studied in 1-D [77] and in higher dimension in [60]. It is the purpose of the present paper to fill this theoretical gap by analyzing the impact of the noise on the jump set of the solution to $\mathcal{P}_\lambda(f + w)$, when both $\|w\|_{L^2}$ and λ are not too large.

Calibrable and Cheeger sets. Of particular importance for the analysis of TV methods are indicator functions of sets, and their behavior under the regularization. Indicator functions which are invariant (up to a rescaling) under TV denoising define so-called “calibrable” sets. These sets play the role of “stable” sets and one expects the corresponding edges to be well restored by TV denoising, a statement which is made precise in the present paper. We refer to section 3.4 for a detailed description of these sets and their basic properties. An important result is the full characterization of convex calibrable sets in [7]. We remark also that the indicators of calibrable sets are in fact eigenvectors of the curvature operator [19]. This operator informally reads $\operatorname{div}(\frac{Du}{|Du|})$, and is also known as the 1-Laplacian, see [63]. These eigenvectors can be used for image processing purposes, as advocated in [21]. The study of fine geometrical properties of TV minimizers is thus deeply linked with geometric measure theory and in particular sets of finite perimeters [13, 68]. In particular, the construction of calibrable sets is

related to minimal surface problems [52] and capillarity problems [66]. Calibrable sets are also related to Cheeger sets, which are subsets of a given set minimizing the ratio of perimeter over area. These Cheeger sets are useful in the construction of solutions to the TV denoising problem. Cheeger sets associated to a given convex sets are unique [6, 31, 61], and can be approximated using either p -Laplacian [62] or strictly convex penalizations [28] to recover an unique maximal set, which can in turn be computed numerically [29].

TV flow. While our paper studies variational problems, a closely related denoising method is obtained by solving the PDE obtained as a gradient flow of J , see [20] for a formal definition. In this setting, the evolution time t plays the role of λ . This TV flow can be shown in 1-D, for characteristic of convex sets and for radial functions to be equivalent to the TV regularization ($\mathcal{P}_\lambda(f)$), see [23, 24, 60, 77]. All the results available for the variational formulation ($\mathcal{P}_\lambda(f)$) have equivalent formulations in the PDE setting, such as explicit solutions for the indicators of convex sets [8] and the evolution of the jump-set [32]. Some of these results have been extended to more general PDE's, see [14].

1-D setting and statistical estimation. 1-D TV denoising, sometimes referred to as the “taut string method” [70], is a method of choice to perform statistical analysis of time series and in particular to detect jumps and transitions. In 1-D, TV flow and TV regularization are known to be equivalent [23, 77]. In the special 1-D case, it is possible to compute exactly the solution on a grid of P points in $O(P^2)$ operations using a dynamic programming method [42, 46, 56, 58]. Similarly to wavelet thresholding estimators, 1-D TV denoising is known to achieve asymptotic optimal estimation results [70]. This optimality is however measured in term of L^2 error, which does not provide geometric information about the location of jumps. A more precise analysis of the distribution of the jumps is provided in [45]. This analysis is however probabilistic and does not extend to higher dimensions, whereas we targets a deterministic geometric analysis in 2-D (although some of our results cover the general N -dimensional case).

Inverse problem and source condition. The systematic study of noise stability of regularization schemes for general regularizing functionals J relies on the so-called source-condition [80], which reads in the simple denoising setting that $\partial J(f)$ should be non-empty (see Section 3.3 for a primer on the total variation sub-differential ∂J). For non-smooth regularizations over Banach spaces, this study started with the seminal paper of Burger and Osher [27] who show that this source condition implies stability of the solution according to the Bregman divergence associated to J . This Bregman measure of stability is however quite weak, and in particular it does not lead to a precise geometric characterization of the restored jump set. Our analysis can be seen as a generalization and refinement of this approach, as highlighted in Section 2. Note that under a non-degeneracy condition, namely that 0 is in the relative interior of $\partial J(f)$, it is possible to state much stronger results, as detailed in the book [80] for ℓ^1 -based

methods. These results however do not cover the TV regularization and can only be applied to discretized versions of TV regularization problems, see [84].

Numerical algorithms. While this is not the topic of this article, let us note that the discretization (often using finite differences) and the numerical resolution of $(\mathcal{P}_\lambda(f))$ is notoriously difficult, in large part because of the non-smoothness of the TV functional J . Early algorithms rely on various smooth approximations of J [38, 86]. The dual projected gradient method proposed by [35] started a wave of activity on the use of first order proximal splitting schemes to solve $(\mathcal{P}_\lambda(f))$ with a provably convergent scheme, see for instance [18] for accelerated first order schemes. Another option is to solve exactly the denoising problem using graph-cuts methods [59], see also [36, 44, 65] and the references therein. These algorithms work however only for the anisotropic total variation, and thus do not cover our J functional, which is the isotropic total variation. Let us also recall that TV methods, and their discretizations, are intimately linked with iterative non-linear filterings, and in particular local median filters, see [26].

2. A brief overview of our results

A common approach to studying the stability properties of a variational problems is by analysis of the source condition. To explain our approach, we first recall a result of Burger and Osher [27] which provides a link between the source condition and the stability of regularized solutions: Given $f \in L^2(\mathbb{R}^2)$ with finite total variation, suppose that the source condition holds, i.e. there exists some $v \in \partial J(f)$ such that $v = -\operatorname{div} z$ for some $z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$ with $\|z\|_{L^\infty} \leq 1$. Note that elements in $\partial J(f)$ are often referred to as *dual certificates*. Let $T \subset \mathbb{R}^2$ and $\delta \in (0, 1)$ be such that $|z(x)| < 1 - \delta$ for a.e. $x \notin T$. Then, $u_{\lambda,w}$, the solution to $\mathcal{P}_\lambda(f + w)$ satisfies

$$(1 - \delta) \int_{\mathbb{R}^2 \setminus T} |Du_{\lambda,w}| \leq \frac{\|w\|_{L^2}^2}{2\lambda} + \frac{\lambda \|v\|_{L^2}^2}{2} + \|w\|_{L^2} \|v\|_{L^2}.$$

While this result informs us that the variation of $u_{\lambda,w}$ is concentrated in the region T , it does not provide any information on the regions where $Du_{\lambda,w}$ is identically zero and no information is given about how the support of $Du_{\lambda,w}$ differs from the support of Df .

Instead of studying *any* $v \in \partial J(f)$, in this paper, we shall study the minimal norm certificate

$$v_{0,0} \stackrel{\text{def.}}{=} \operatorname{argmin} \{ \|v\|_{L^2} ; v \in \partial J(f) \}. \quad (2)$$

The minimal norm certificate was first proposed in [47] for studying the support of solutions to the sparse spikes deconvolution problem using total variation of measures regularization, but in this particular framework of denoising, it is also known as the *minimal section* of $\partial J(f)$ [81]. Although dual certificates have been widely used to derive stability properties of solutions to the sparse spikes deconvolution problem in terms of the L^2 norm, see for instance [57], the novelty of the minimal norm certificate

(which is itself a dual certificate) is that it additionally addresses support stability questions such as the number and the location of the recovered diracs.

In this paper, we follow the same philosophy: Similarly to the problem of sparse spikes deconvolution, we show that the minimal norm certificate naturally gives rise to the definition of an instability region which we refer to as the extended support. This region, in turn, governs the support of the regularized solution in the low noise regime. Unlike previous works, our analysis is carried out for this very specific dual certificate and in doing so, we are able to characterize the support stability of the total variation denoising problem.

We now present a sketch of why the minimal norm certificate (2) governs the support of the regularized solutions in the low noise regime:

2.1. The significance of the minimal norm certificate for support stability

The minimal norm certificate arises naturally as the strong L^2 limit of $v_{\lambda,w}$, the solutions to the dual problems associated with $(\mathcal{P}_\lambda(f+w))$ as λ and $\|w\|_{L^2}/\lambda$ converges to 0. The significance of this is that if $u_{\lambda,w}$ and $v_{\lambda,w}$ are respectively the primal and the dual solutions associated with $(\mathcal{P}_\lambda(f+w))$, then $v_{\lambda,w} \in \partial J(u_{\lambda,w})$ and a consequence of this is that

$$\text{Supp}(Du_{\lambda,w}) = \bigcup \left\{ \partial^* F ; |F| < \infty, P(F) = \pm \int v_{\lambda,w} \right\}, \quad (3)$$

where $\partial^* F$ denotes the reduced boundary of the set F . Thus, one observes that $v_{\lambda,w}$ is intimately linked to the level sets of $u_{\lambda,w}$ and intuitively, one can infer information on the level sets of $u_{\lambda,w}$ since $v_{0,0}$ is the strong L^2 limit of $v_{\lambda,w}$. In particular, in light of (3), one would expect $\text{Supp}(Du_{\lambda,w})$ to be concentrated around the extended support of f , which can be written as

$$\text{Ext}(Df) = \overline{\bigcup \left\{ \partial^* E ; |E| < +\infty \quad \text{and} \quad \pm \int_E v_{0,0} = P(E) \right\}}.$$

The dual problems associated to $(\mathcal{P}_\lambda(f+w))$ and their limit is discussed in Section 4, while the relationship between the subdifferential and the level sets of a function is recalled in Section 3.5 and in particular, Proposition 4. The notion of the extended support is described in Section 6. Examples of the minimal norm certificate and the expected support can also be found, respectively, in Sections 4.4 and 6.2.

2.2. Properties of the level sets

To formalize the link between $v_{0,0}$ and the level sets of $u_{\lambda,w}$, we derive some properties on the level sets of $u_{\lambda,w}$ in Section 5. In particular, we show that for all λ and $\|w\|_{L^2}$ sufficiently small,

- (i) there exists *uniform* upper (c.f. Lemma 2) and lower bounds (c.f. Lemma 3) on the perimeters and measures of the level sets of $u_{\lambda,w}$,

- (ii) there exists some ball which contains all level sets of $u_{\lambda,w}$ (c.f. Lemma 4),
- (iii) there is a uniform weak regularity on the boundaries of the level sets of $u_{\lambda,w}$ (c.f. Proposition 7), so that despite the inevitable lack of strong regularity results (as explained in Section 3.6), this uniform weak regularity essentially prevents the boundaries from exhibiting singularity properties such as cusps.

Using the properties described above, we show in Section 7 that one can exploit compactness properties of the level sets of $u_{\lambda,w}$ to prove our main result, Theorem 1, which shows that $\text{Supp}(Du_{\lambda,w})$ is contained inside a small neighbourhood around $\text{Ext}(Df)$. In particular, given sequences $w_n \in L^2(\mathbb{R}^2)$ and $\lambda_n \in \mathbb{R}_+$ such that $\|w_n\|_{L^2}/\lambda_n \rightarrow 0$ as $n \rightarrow +\infty$, the conclusion of our main result is that

$$\text{Supp}(Df) \subseteq \liminf_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \limsup_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \text{Ext}(Df),$$

where the limits are in the sense of Painlevé-Kuratowski set convergence, which we recall in Section 3.1.

2.3. Further extensions

Stability estimates in the absence of the source condition. Section 8 presents stability analysis for cases where the source condition is not satisfied, i.e. $\partial J(f) = \emptyset$. One important class of functions which this covers are the indicator functions on convex sets with nonsmooth boundary, such as the square. To our knowledge, there were no previous studies on stability analysis in the absence of the source condition and hence, no stability guarantees for even the simple case of denoising the indicator function of a square. Although in this case, the minimal norm certificate is not defined, we show in Theorem 2 that the techniques developed in the analysis of the minimal norm certificate can be adapted to such special cases to derive stability estimates for general convex sets.

Convergence rates. We stress that via the approach of [27], characterization of the regions where the variation of $u_{\lambda,w}$ is small is possible only when the source condition holds *and* there is precise knowledge of the extremal points and decay of some vector field $z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$ for which $v = -\text{div } z \in \partial J(f)$. In general, this vector field is not unique and such precise characterization is a difficult problem. In contrast, via our approach, explicit knowledge of the vector field associated with the minimal norm certificate is not essential, and in fact, the definition of the extended support is dependent only v_0 and not on the vector fields z for which $v_0 = -\text{div } z$.

Nonetheless, we show in Section 9 that in the special cases where the vector field associated with v_0 is known, our main result in Theorem 1 can be further refined. More specifically, we provide in Theorem 3 an explicit upper bound on the rate of shrinkage of the tube around the extended support with respect to the decay of the noise level. For the indicator function on a calibrable set C with \mathcal{C}^2 boundary, we describe an explicit construction of the vector field z_0 associated with the minimal norm certificate with

$|z_0| < 1$ on all compact subsets of $\mathbb{R}^2 \setminus \partial C$. Therefore, our main result can be seen as a refinement of the work of [27] and can be applied in much greater generality. Some numerical examples are presented in Section 10 for the illustration of our theoretical results with respect to the underlying vector field.

3. Preliminaries

The main contributions of our paper are presented in Sections 4 to 9 with some numerical illustrations in Section 10; the purpose of this section is to recall some essential results which are applied throughout this paper.

3.1. Set convergence

We shall use the notion of *Painlevé-Kuratowski* set convergence (see [78] for more detail). Given a sequence of sets $\{S_n\}_{n \in \mathbb{N}}$, $S_n \subseteq \mathbb{R}^2$, let us define the outer (resp. inner) limit of $\{S_n\}_{n \in \mathbb{N}}$ as

$$\limsup_{n \rightarrow +\infty} S_n \stackrel{\text{def.}}{=} \left\{ x \in \mathbb{R}^2 ; \liminf_{n \rightarrow +\infty} \text{dist}(x, S_n) = 0 \right\}, \quad (4)$$

$$(\text{resp.}) \quad \liminf_{n \rightarrow +\infty} S_n \stackrel{\text{def.}}{=} \left\{ x \in \mathbb{R}^2 ; \limsup_{n \rightarrow +\infty} \text{dist}(x, S_n) = 0 \right\}. \quad (5)$$

It is clear that $\liminf_{n \rightarrow +\infty} S_n \subseteq \limsup_{n \rightarrow +\infty} S_n$. Moreover, those two sets are closed. We say that S_n converges towards $S \subseteq \mathbb{R}^2$, *i.e.* $\lim_{n \rightarrow +\infty} S_n = S$, if

$$\liminf_{n \rightarrow +\infty} S_n = S = \limsup_{n \rightarrow +\infty} S_n. \quad (6)$$

If the sequence S_n is bounded (there exists $R > 0$ such that $S_n \subseteq B(0, R)$ for all n large enough), then the *Painlevé-Kuratowski* convergence is equivalent to the so-called Hausdorff convergence, that is,

$$\lim_{n \rightarrow +\infty} \sup_{x \in S \cup S_n} |\text{dist}(x, S_n) - \text{dist}(x, S)| = 0. \quad (7)$$

3.2. Functions with bounded variation and sets with finite perimeter

We briefly recall some properties of functions of bounded variations and sets of finite perimeter. We refer the reader to [12, 68] for a comprehensive treatment of the subject.

Total variation, perimeter. Given $u \in L^1_{loc}(\mathbb{R}^2)$, its total variation is equal to

$$\int_{\mathbb{R}^2} |Du| \stackrel{\text{def.}}{=} \sup \left\{ \int_{\mathbb{R}^2} u \operatorname{div} z ; z \in \mathcal{C}_c^1(\mathbb{R}^2, \mathbb{R}^2), \|z\|_\infty \leq 1 \right\}.$$

If $J(u) < +\infty$, we say that u has bounded variation. The mapping $u \mapsto J(u)$ is lower semi-continuous with respect to the $L^1_{loc}(\mathbb{R}^2)$ topology (hence for the L^2 topology).

If $E \subseteq \mathbb{R}^2$ is a measurable set, we denote by $|E|$ its 2-dimensional Lebesgue measure. The set E is said to be of finite perimeter if $J(\mathbf{1}_E) < +\infty$, where $\mathbf{1}_E$ is the indicator function of E . Its perimeter is defined as $P(E) = \int_{\mathbb{R}^2} |D\mathbf{1}_E|$. For a Borel set $S \subseteq \mathbb{R}^2$, $\mathcal{H}^1 \llcorner S$ denotes the 1-dimensional Hausdorff measure restricted to S , namely $\mathcal{H}^1 \llcorner S(A) = \mathcal{H}^1(A \cap S)$.

The reduced boundary of E is defined as

$$\partial^* E \stackrel{\text{def.}}{=} \left\{ x \in \text{Supp } |D\mathbf{1}_E| ; \nu_E(x) \stackrel{\text{def.}}{=} \lim_{r \rightarrow 0^+} \frac{-D\mathbf{1}_E(B(x, r))}{|D\mathbf{1}_E(B(x, r))|} \text{ exists and } |\nu_E(x)| = 1 \right\}. \quad (8)$$

The vector $\nu_E(x)$ is the measure theoretic outer unit normal to E . When the context is clear, we shall write ν instead of ν_E . Moreover, $D\mathbf{1}_E = -\nu_E \mathcal{H}^1 \llcorner \partial^* E$, and $|D\mathbf{1}_E|(A) = \mathcal{H}^1(\partial^* E \cap A)$ for all open set $A \subseteq \mathbb{R}^2$.

In the following, we use the construction in [52, Prop. 3.1] so as to always consider a Lebesgue representative of E such that for all x in the topological boundary ∂E , $0 < |E \cap B(x, r)| < |B(x, r)|$. Then, with this representative,

$$\text{Supp } D\mathbf{1}_E = \overline{\partial^* E} = \partial E.$$

The area and the perimeter are related by the so-called isoperimetric inequality: for any Lebesgue measurable set $E \subseteq \mathbb{R}^2$,

$$c_2 \min\{|E|, |\mathbb{R}^2 \setminus E|\} \leq (P(E))^2,$$

where $c_2 = 4\pi$ is the isoperimetric constant.

Level sets and the coarea formula. The coarea formula relates the total variation of a function $f \in L^1_{loc}(\mathbb{R}^2)$ and the perimeter of its level sets. Define the level sets of f as

$$\begin{aligned} F^{(t)} &\stackrel{\text{def.}}{=} \{x \in \mathbb{R}^2 ; f(x) \geq t\} \text{ for } t \geq 0, \\ F^{(t)} &\stackrel{\text{def.}}{=} \{x \in \mathbb{R}^2 ; f(x) \leq t\} \text{ for } t < 0. \end{aligned} \quad (9)$$

It is clear that $|F^{(t)}| < +\infty$ except possibly for $t = 0$. Moreover, the family is monotone on $[0, +\infty)$ and $(-\infty, 0)$ with

$$F^{(t)} = \bigcap_{0 < t' < t} F^{(t')} \text{ for } t > 0, \quad F^{(t)} = \bigcap_{0 > t' > t} F^{(t')} \text{ for } t < 0.$$

We handle 0 as a special case with $F^{(0)} = \mathbb{R}^2 \setminus \bigcup_{t' < 0} F^{(t')}$. Now, given an open set $U \subseteq \mathbb{R}^2$, the coarea formula states that if $J(f) < +\infty$ then

$$\int_U |Df| = \int_{-\infty}^{+\infty} P(F^{(t)}; U) dt,$$

where $P(F^{(t)}; U) \stackrel{\text{def.}}{=} |D\mathbf{1}_E|(U)$.

3.3. Subdifferential of J

In the following, unless otherwise stated, we use the $L^2(\mathbb{R}^2)$ topology. The functional $J : L^2(\mathbb{R}^2) \rightarrow \mathbb{R} \cup \{+\infty\}$ is convex, proper lower semi-continuous. It is in fact the support function of the closed convex set

$$\{\operatorname{div} z ; z \in X_2(\mathbb{R}^2), \|z\|_\infty \leq 1\} \subseteq L^2(\mathbb{R}^2),$$

where we defined

$$X_2(\mathbb{R}^N) \stackrel{\text{def.}}{=} \{z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2) ; \operatorname{div} z \in L^2(\mathbb{R}^2)\}.$$

As a result, it is possible to prove that

$$\partial J(0) = \{\operatorname{div} z ; z \in X_2(\mathbb{R}^2), \|z\|_\infty \leq 1\}, \quad (10)$$

$$\partial J(u) = \left\{ v \in \partial J(0) ; \int_{\mathbb{R}^N} uv = J(u) \right\}. \quad (11)$$

Provided that $J(u) < +\infty$, Du is a Radon measure, i.e it is possible to evaluate (z, Du) for all vector field $z \in C_c^0(\mathbb{R}^2; \mathbb{R}^2)$. Following the construction by Anzellotti [15], it is possible to define (z, Du) for less smooth z , namely $z \in X_2(\mathbb{R}^N)$ provided that $u \in L^2(\mathbb{R}^2)$ and $J(u) < +\infty$. Given $\varphi \in \mathcal{C}_c^1(\mathbb{R}^2)$, define

$$\langle (z, Du), \varphi \rangle = - \int_{\mathbb{R}^2} u(x) \varphi(x) \operatorname{div} z(x) dx - \int_{\mathbb{R}^2} u(x) z(x) \cdot \nabla \varphi(x) dx.$$

Then (z, Du) is a Radon measure which is absolutely continuous with respect to $|Du|$, with

$$|\langle (z, Du), \varphi \rangle| \leq \|\varphi\|_\infty \|z\|_{L^\infty(U)} \int_U |Du|,$$

for all $\varphi \in \mathcal{C}_c^1(\mathbb{R}^2)$ and $U \subset \mathbb{R}^2$ open set such that $\operatorname{Supp}(\varphi) \subset U$. Moreover, the following integration by parts holds:

$$\int_{\mathbb{R}^2} u \operatorname{div} z = - \int_{\mathbb{R}^2} (z, Du).$$

If $\theta(z, Du)$ denotes the Radon-Nikodym derivative of (z, Du) with respect to $|Du|$, we may also write $\int_{\mathbb{R}^2} (z, Du) = \int_{\mathbb{R}^2} \theta(z, Du) d|Du|$.

Remark 1. If u is smooth, then (z, Du) can be interpreted as a (defined almost everywhere) pointwise inner product:

$$\int_B (z, Du) = \int_B z(x) \cdot \nabla u(x) dx \quad \text{for any Borel set } B \subseteq \mathbb{R}^2.$$

If u is the characteristic function of set with finite perimeter $E \subset \mathbb{R}^2$, then

$$\int_E \operatorname{div} z(x) dx = \int_E (z, -D\mathbf{1}_E).$$

The question whether it is possible to give a pointwise meaning to $(z, -D\mathbf{1}_E)$ is investigated in [22, 37]. In [22], under some regularity assumption on z (which holds if $\operatorname{div} z \in \partial J(\mathbf{1}_E)$), it is interpreted as $(z, -D\mathbf{1}_E) = Tz \cdot \nu_E \mathcal{H}^1 \llcorner \partial^* E$, where Tz is the full trace of z defined on \mathcal{H}^1 -a.e. on $\partial^* E$ [22]. In [37], it is shown that (in dimension $N = 2$ or 3), if $\operatorname{div} z \in \partial J(\mathbf{1}_E) \cap L^N(\mathbb{R}^N)$, then every point of the reduced boundary $\partial^* E$ is a Lebesgue point of z , with thus $z = \nu_E$ \mathcal{H}^1 -a.e. on $\partial^* E$.

In the general case, recalling that $|D\mathbf{1}_E| = \mathcal{H}^1 \llcorner \partial^* E$, we shall write

$$\int_E \operatorname{div} z(x) dx = \int_{\partial^* E} \theta(z, -D\mathbf{1}_E) d\mathcal{H}^1, \quad (12)$$

keeping in mind that in regular cases this amounts to $\int_{\partial^* E} z \cdot \nu_E d\mathcal{H}^1$.

Remark 2. That enables us to interpret the optimality $\int_{\mathbb{R}^2} uv = J(u)$ as an “optimality $|Du|$ -almost everywhere”:

$$\begin{aligned} \int_{\mathbb{R}^N} u \operatorname{div} z &= \int_{\mathbb{R}^N} |Du| \Leftrightarrow - \int_{\mathbb{R}^N} (z, Du) = \int_{\mathbb{R}^N} |Du| \\ &\Leftrightarrow 0 = \int_{\mathbb{R}^N} (1 + \theta(z, Du)) d|Du|, \end{aligned}$$

where $\theta(z, Du)$ is the Radon-Nikodym derivative of (z, Du) with respect to $|Du|$. Since $|\theta(z, Du)| \leq 1$, this implies that in fact the equality $(z, Du) = -|Du|$ holds $|Du|$ -a.e. Informally, recalling that $\|z\|_\infty \leq 1$, this means that

$$z = -\frac{Du}{|Du|}, \quad |Du| - \text{almost everywhere.}$$

In other words z must be orthogonal to the level lines, and its saturation points contains the support of Du (see also [22, 37] for more rigorous statements).

Examples Let us examine two examples which can be found in [71].

Characteristic function of a disc: Given $R > 0$, consider the vector field

$$z(x) = \begin{cases} \frac{x}{R} & \text{if } |x| \leq R, \\ \frac{R}{|x|^2} x & \text{otherwise.} \end{cases} \quad (13)$$

One may check that $z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$, $\operatorname{div} z = \frac{2}{R} \mathbf{1}_{B(0,R)} \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$, $\|z\|_\infty \leq 1$, and $z \cdot \nu = 1$ on $\{x \in \mathbb{R}^2 ; |x| = R\}$. Hence $\operatorname{div} z \in \partial J(u)$ for $u = \mathbf{1}_{B(0,R)}$.

Characteristic function of a square : Let $u = \mathbf{1}_{[0,1]^2}$ be the characteristic function of the unit square. It turns out that $\partial J(u) = \emptyset$. The argument provided in [71] is the following. Assume that there exists $v \in \partial J(u)$ and let $z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$ be the corresponding vectorfield. We denote by T_ε the triangle

$\{x \in \mathbb{R}^2 ; 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_1 + x_2 \leq \varepsilon\}$ and by ν its outer unit normal (defined \mathcal{H}^1 -a.e.). By the Gauss-Green theorem:

$$\int_{T_\varepsilon} \operatorname{div} z = \int_{\partial T_\varepsilon} \theta(z, -D\mathbb{1}_{T_\varepsilon}) d\mathcal{H}^1.$$

Since $v = \operatorname{div} z \in L^2(\mathbb{R}^2)$, the left term is upper-bounded by $\sqrt{\int_{T_\varepsilon} (\operatorname{div} z)^2} \sqrt{|T_\varepsilon|} = o(\varepsilon)$ whereas the right term is lower-bounded by $(2 - \sqrt{2})\varepsilon$. This is a contradiction. Hence $\partial J(u) = \emptyset$

3.4. Calibrable sets in \mathbb{R}^2

A remarkable family of elements of $\partial J(0)$ is the family of characteristic functions of sets F such that $\lambda_F \mathbb{1}_F \in \partial J(\mathbb{1}_F)$ for some $\lambda_F \in \mathbb{R}$. This family of functions, known as the calibrable sets, will serve as a prime example in the illustration of our theoretical results. In this section, we recall some key results about these functions.

3.4.1. Sets that evolve at constant speed The work [20] presented a study on the total variation flow $\frac{\partial u}{\partial t} \in -\partial J(u)$, namely:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{Du}{|Du|} \right), \quad (14)$$

in $[0, \infty] \times \mathbb{R}^2$ subject to initial data $u(\cdot, 0) = u_0 \in L^2(\mathbb{R}^2)$. They prove existence and uniqueness of a “strong solution” (see [20]) for all initial data $u_0 \in L^2(\mathbb{R}^2)$, and existence and uniqueness of an “entropy solution” for $u_0 \in L^1_{loc}(\mathbb{R}^N)$. In the second part of the paper, they characterize the bounded sets of finite perimeter Ω such that $u = \mathbb{1}_\Omega$ satisfies

$$-\operatorname{div} \left(\frac{Du}{|Du|} \right) = \lambda_\Omega u, \quad \text{where } \lambda_\Omega \stackrel{\text{def.}}{=} \frac{P(\Omega)}{|\Omega|}. \quad (15)$$

Such sets are exactly the sets which evolve with constant boundary, i.e. such that $u(x, t) = \lambda(t) \mathbb{1}_\Omega(x)$, with $\lambda \geq 0$, and they are called calibrable. They are characterized by the fact that $\lambda_\Omega \mathbb{1}_\Omega \in \partial J(\mathbb{1}_\Omega)$:

Definition 1 (Calibrable sets). *A set of finite perimeter $\Omega \subset \mathbb{R}^2$ is said to be calibrable if, writing $v = \mathbb{1}_\Omega$, there exists a vector field $z \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)$ such that $\|z\|_\infty \leq 1$ and*

$$\begin{aligned} \int_{\mathbb{R}^N} (z, Dv) &= \int_{\mathbb{R}^2} |Dv|, \\ -\operatorname{div} z &= \lambda_\Omega v. \end{aligned}$$

In that case, we say that z is a calibration for Ω .

Remark 3. If $\lambda \mathbb{1}_\Omega \in \partial J(\mathbb{1}_\Omega)$ for some $\lambda \in \mathbb{R}$, then necessarily $\lambda = \lambda_\Omega$.

3.4.2. *Characterization in \mathbb{R}^2* The following results characterize convex calibrable sets.

Proposition 1 ([20]). *Let $C \subset \mathbb{R}^2$ be a bounded set of finite perimeter, and assume that C is connected. C is calibrable if and only if the following three conditions hold:*

- (i) C is convex;
- (ii) ∂C is of class $C^{1,1}$;
- (iii) the following inequality holds:

$$\operatorname{ess\,sup}_{p \in \partial C} \kappa_{\partial C}(p) \leq \frac{P(C)}{|C|}. \quad (16)$$

Proposition 2 ([20]). *Let $\Omega \subset \mathbb{R}^2$ be a bounded set of finite perimeter which is calibrable. Then,*

- (i) the following relation holds:

$$\frac{P(\Omega)}{|\Omega|} \leq \frac{P(D)}{|D \cap \Omega|}, \quad \forall D \subseteq \mathbb{R}^2, D \text{ of finite perimeter};$$

- (ii) each connected component of Ω is convex.

3.5. From the subdifferential to the level sets

Let $f \in L^2(\mathbb{R}^2)$, $J(f) < +\infty$, and $v \in \partial J(f)$. By definition of the subdifferential,

$$\forall g \in L^2(\mathbb{R}^2), \quad \int_{\mathbb{R}^2} |Dg| - \int_{\mathbb{R}^2} vg \geq \int_{\mathbb{R}^2} |Df| - \int_{\mathbb{R}^2} vf. \quad (17)$$

In fact, using the coarea formula, one may reformulate that optimality property (see Proposition 3 below) as an optimality property of the level sets. The following result is very similar to [64, Corollary 2.4] but it requires a bit more care in our framework since the domain is \mathbb{R}^2 and $v \in L^2(\mathbb{R}^2)$. The level sets of f (resp. g), are denoted by $\{F^{(t)}\}_{t \in \mathbb{R}}$ (resp. $\{G^{(t)}\}_{t \in \mathbb{R}}$).

Proposition 3. *Let $f \in L^2(\mathbb{R}^2)$, $J(f) < +\infty$, and $v \in L^2(\mathbb{R}^2)$. The following conditions are equivalent.*

- (i) $v \in \partial J(f)$.
- (ii) $v \in \partial J(0)$ and the level sets of f satisfy

$$\forall t > 0, \quad P(F^{(t)}) = \int_{F^{(t)}} v, \quad \forall t < 0, \quad P(F^{(t)}) = - \int_{F^{(t)}} v. \quad (18)$$

- (iii) The level sets of f satisfy

$$\forall t > 0, \quad \forall G \subset \mathbb{R}^2, |G| < +\infty, \quad P(G) - \int_G v \geq P(F^{(t)}) - \int_{F^{(t)}} v, \quad (19)$$

$$\forall t < 0, \quad \forall G \subset \mathbb{R}^2, |G| < +\infty, \quad P(G) + \int_G v \geq P(F^{(t)}) + \int_{F^{(t)}} v. \quad (20)$$

Proof. (iii) \Rightarrow (i) It suffices to use the coarea formula $\int |Dg| = \int_{-\infty}^{\infty} P(G^{(t)}) dt$ and Fubini's theorem in

$$\int_{\mathbb{R}^2} gv = \int_{\mathbb{R}^2} \left(\int_0^{+\infty} \mathbb{1}_{g(x) \geq t} v(x) dt - \int_{-\infty}^0 \mathbb{1}_{g(x) \leq t} v(x) dt \right) dx, \quad (21)$$

and similarly for the level sets of f .

(i) \Rightarrow (ii) Using (11), we see that $v \in \partial J(0)$ and $\int_{\mathbb{R}^2} fv = J(f)$. From $v \in \partial J(0)$, and choosing $\pm \mathbb{1}_F$ (for any $F \subset \mathbb{R}^2$ with $|F| < +\infty$) in the subdifferential inequality, we infer that $P(F) \pm \int_{\mathbb{R}^2} \mathbb{1}_F v \geq 0$. Now, $\int_{\mathbb{R}^2} fv = J(f)$ rewrites

$$0 = \int_0^{+\infty} \left(P(F^{(t)}) - \int_{F^{(t)}} v \right) dt + \int_{-\infty}^0 \left(P(F^{(t)}) + \int_{F^{(t)}} v \right) dt.$$

Since the integrands are nonnegative, we obtain that for a.e. $t \in \mathbb{R}$, $P(F^{(t)}) = \text{sign}(t) \int_{F^{(t)}} v$. In fact, the equality holds for all $t \neq 0$. Indeed, for $t > 0$, we may find a sequence $t_n \nearrow t$ as $n \rightarrow +\infty$ such that $P(F^{(t_n)}) = \int_{F^{(t_n)}} v$. Since $|F^{(t)}| < +\infty$ and by monotonicity, $\mathbb{1}_{F^{(t_n)}}$ converges in $L^2(\mathbb{R}^2)$ towards $\mathbb{1}_{F^{(t)}}$ and we have

$$P(F^{(t)}) = P\left(\bigcap_{n \in \mathbb{N}} F^{(t_n)}\right) \leq \liminf_{n \rightarrow +\infty} P(F^{(t_n)}) = \liminf_{n \rightarrow +\infty} \int_{F^{(t_n)}} v = \int_{F^{(t)}} v.$$

The converse inequality holds from the fact that $v \in \partial J(0)$ so that $\int_{F^{(t)}} v \leq P(F^{(t)})$. In a similar way, we may prove that for all $t < 0$, $P(F^{(t)}) = \int_{F^{(t)}} v$.

(ii) \Rightarrow (iii) From $v \in \partial J(0)$, we infer that $P(G) \pm \int_G v \geq 0$ for any $G \subset \mathbb{R}^2$ with $|G| < +\infty$. Since $P(F^{(t)}) - \text{sign}(t) \int_{F^{(t)}} v = 0$, we obtain the claimed result. \square

As a consequence of Proposition 3, if we are given $v \in \partial J(f)$ rather than f , we may control the localization of the support of Df simply by studying the solutions of (18). The following proposition formalizes this idea.

Proposition 4. *Let $f \in L^2(\mathbb{R}^2)$ with $J(f) < +\infty$, $v \in \partial J(f)$ and let $\text{Supp}(Df)$ denote the support of the Radon measure Df . Then*

$$\text{Supp}(Df) = \overline{\bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R} \setminus \{0\}\}} \quad (22)$$

$$\subseteq \overline{\bigcup \left\{ \partial^* F ; |F| < +\infty \text{ and } P(F) = \pm \int_F v \right\}}. \quad (23)$$

Proof. Let $x \in \mathbb{R}^2 \setminus \text{Supp}(Df)$. There exists $r > 0$ such that $|Df|(B(x, r)) = 0$, hence f is constant in $B(x, r)$, identically equal to some $C \in \mathbb{R}$. Depending the value of $t \in \mathbb{R}$, we see that either $B(x, r) \subseteq F^{(t)}$ or $B(x, r) \cap F^{(t)} = \emptyset$. In any case, $\partial^* F^{(t)} \subseteq \mathbb{R}^2 \setminus B(x, r)$. As a result $x \in \mathbb{R}^2 \setminus \overline{\bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R}\}}$, which proves that $\overline{\bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R}\}} \subseteq \text{Supp}(Df)$.

For the converse inclusion, let $x \in \text{Supp}(Df)$, so that for all $r > 0$,

$$|Df|(B(x, r)) > 0.$$

We apply the coarea formula

$$|Df|(B(x, r)) = \int_{-\infty}^{\infty} P(F^{(t)}, B(x, r)) dt$$

to see that $\mathcal{H}^1(\partial^* F^{(t)} \cap B(x, r)) > 0$ for some $t \in \mathbb{R}$, hence

$$B(x, r) \cap \bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R}\} \neq \emptyset.$$

Since this is true for all $r > 0$, we see that $x \in \overline{\bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R}\}}$.

Now, we prove the last inclusion in (23). First, we observe that

$$\overline{\bigcup \{\partial^* F^{(t)} ; t \in \mathbb{R}\}} = \overline{\bigcup \{\partial^* F^{(t)} ; t \neq 0\}} \cup \overline{\partial^* F^{(0)}}.$$

By Proposition 3, we know that for every $t \neq 0$, $F^{(t)}$ satisfies $|F^{(t)}| < +\infty$ and $\pm \int_{F^{(t)}} v = P(F^{(t)})$. Hence

$$\overline{\bigcup \{\partial^* F^{(t)} ; t \neq 0\}} \subseteq \overline{\bigcup \left\{ \partial^* F ; |F| < +\infty \text{ and } P(F) = \pm \int_F v \right\}},$$

and it is sufficient to prove that

$$\overline{\partial^* F^{(0)}} \subseteq \overline{\bigcup \{\partial^* F^{(t)} ; t \neq 0\}}.$$

Let $x \in \partial^* F^{(0)} = \partial^* (\mathbb{R}^N \setminus \bigcup_{k \in \mathbb{N}^*} F^{(-1/k)}) = \partial^* (\bigcup_{k \in \mathbb{N}^*} F^{(-1/k)})$. Then for all $r > 0$,

$$\left| B(x, r) \cap \bigcup_{k \in \mathbb{N}^*} F^{(-1/k)} \right| > 0 \quad \text{and} \quad \left| B(x, r) \setminus \bigcup_{k \in \mathbb{N}^*} F^{(-1/k)} \right| > 0.$$

In particular, there exists k_0 such that $|B(x, r) \cap F^{(-1/k_0)}| > 0$, and moreover $|B(x, r) \setminus F^{(-1/k_0)}| \geq |B(x, r) \setminus \bigcup_{k \in \mathbb{N}^*} F^{(-1/k)}| > 0$. Hence $\mathbf{1}_{F^{(-1/k_0)}}$ is not constant in $B(x, r)$, so that $\partial^* F^{(-1/k_0)} \cap B(x, r) \neq \emptyset$. As a result, $B(x, r) \cap \bigcup \{\partial^* F^{(t)} ; t \neq 0\} \neq \emptyset$ for all $r > 0$, which proves that $x \in \overline{\bigcup \{\partial^* F^{(t)} ; t \neq 0\}}$. □

3.6. The prescribed mean curvature problem

As a consequence of Propositions 3 and 4, we are led to study the solutions of the *prescribed curvature problem*

$$\min_{\substack{X \subset \mathbb{R}^N \\ |X| < +\infty}} P(X) + \int_X H, \tag{24}$$

for $H = \pm v$, where $v \in \partial J(f)$ is fixed. Following [17], if $E \subset \mathbb{R}^2$ is a solution to (24), we say that v is a *variational mean curvature* \ddagger for E .

\ddagger The careful reader will note that we make a slight abuse in the terminology since in [17, 55] the function H is assumed to be integrable, and the condition $|F| < +\infty$ is not imposed. We make this slight abuse since the local properties of the sets studied in [17, 55] also hold for the solutions of (24).

In the case where $H \in L^p_{loc}(\mathbb{R}^N)$ for some $p \in (N, +\infty]$ and $E \subseteq \mathbb{R}^N$ be a nonempty solution of (24), it can be shown [11] that $\Sigma = \partial E \setminus \partial^* E$ is a closed set of Hausdorff dimension at most $N - 8$, and $\partial^* E$ is a $\mathcal{C}^{1,\alpha}$ hypersurface for all $\alpha < (p - N)/2p$. Furthermore, when H is continuous on an open set A and $N = 2$, $-\frac{1}{N-1}H$ can be shown [11, Th. 1.1.3] to be the mean curvature in the classical sense. However, the integrability of H is crucial of such regularity results. It is easy to show that the level sets of the solution of $(\mathcal{P}_\lambda(f))$ have p -integrable curvature whenever $f \in L^p(\mathbb{R}^N)$ (which yields the same integrability for u), and even continuous curvature if f is continuous [34]. However in general in this paper, we are interested in a limit case $p = N$ where there exist counterexamples [48, 55] for which the Hausdorff dimension of $\partial E \setminus \partial^* E$ is more than $N - 8$.

Nonetheless, when $p = N$, we may rely on a weak regularity theorem [55, Th. 3.6] (see also [54]) which ensures that for all $x \in \partial E$,

$$1 > D_E(x) \stackrel{\text{def.}}{=} \lim_{r \rightarrow 0^+} \frac{|E \cap B(x, r)|}{|B(x, r)|} > 0. \quad (25)$$

Additionally, in the case of $N < 8$, we have that $D_E(x) = 1/2$. In particular, the topological boundary ∂E is equal to the *essential boundary* $\partial^M E$,

$$\partial E = \partial^M E \stackrel{\text{def.}}{=} \{x \in \mathbb{R}^2; \overline{D_E}(x) > 0 \text{ and } \underline{D_E}(x) > 0\},$$

where $\overline{D_E}(x) = \limsup_{r \rightarrow 0^+} \frac{|E \cap B(x, r)|}{|B(x, r)|}$ and $\underline{D_E}(x) = \liminf_{r \rightarrow 0^+} \frac{|E \cap B(x, r)|}{|B(x, r)|}$.

Furthermore, it was shown in [74] that if $H \in L^N(\mathbb{R}^N)$, then $\partial^* E$ is a $C^{0,\alpha}$ hypersurface up to some possible singularities. Thus, in the case where $p = N$, the boundary ∂E does not contain wild singular points such as cusps or points of zero density. In Section 5.2, we apply (25) to show that this weak regularity holds uniformly for the boundaries of the level sets of solutions to $(\mathcal{P}_\lambda(f + w))$ in some low noise regime.

3.7. Decomposition of boundaries into Jordan curves

We shall occasionally rely on the results on the decomposition of sets with finite perimeter provided in [13].

Let E be a set of finite perimeter. By [13, Corollary 1], E can be decomposed into an at most countable union of its M -connected components

$$E = \bigcup_{i \in I \subseteq \mathbb{N}} E_i \quad \text{where} \quad P(E) = \sum_{i \in I} P(E_i), \quad |E_i| > 0,$$

and each M -connected component can be decomposed as

$$E_i = \text{int}(\gamma_i^+) \setminus \bigcup_{j \in L_i} \text{int}(\gamma_j^-) \quad \text{and} \quad \partial^M E_i = \gamma_i^+ \cup \bigcup_{j \in L_i} \gamma_j^- \pmod{\mathcal{H}^1},$$

where each γ_k^\pm is a *rectifiable Jordan curve*, $L_i \stackrel{\text{def.}}{=} \{j \in \mathbb{N}; \text{int}(\gamma_j^-) \subseteq \text{int}(\gamma_i^+)\}$. Here $\text{int}(\gamma)$ denotes the interior of a Jordan curve (but when the context is clear, we shall also use int to denote the topological interior).

Moreover,

$$P(E) = \sum_i \mathcal{H}^1(\gamma_i^+) + \sum_j \mathcal{H}^1(\gamma_j^-) \quad \text{and} \quad P(\text{int } \gamma_i^\pm) = \mathcal{H}^1(\gamma_i^\pm) \quad \text{for all } i.$$

Remark 4. Let $E \subset \mathbb{R}^2$ with $|E| < +\infty$ such that $P(E) = \int_E v$, where $v \in \partial J(0)$. Let us decompose E into its M -connected components, $E = \bigcup_{i \in I} E_i$, where we can assume that I is either \mathbb{N} or of the form $\{0, 1, \dots, n\}$. We observe that $\{E_i\}_{i \in I, i \geq 1}$ yields the decomposition of $E \setminus E_0$ into its M -connected components. Hence,

$$\begin{aligned} 0 &= P(E) - \int_E v \\ &= P(E_0) - \int_{E_0} v + P\left(\bigcup_{i \in I, i \geq 1} E_i\right) - \int_{\bigcup_{i \in I, i \geq 1} E_i} v. \end{aligned}$$

Since $P(E_0) - \int_{E_0} v \geq 0$ and $P\left(\bigcup_{i \in I, i \geq 1} E_i\right) - \int_{\bigcup_{i \in I, i \geq 1} E_i} v \geq 0$, we deduce that these inequalities are in fact equalities. By induction, we deduce that for all $i \in I$,

$$P(E_i) - \int_{E_i} v = 0.$$

Now decomposing, $\partial^M E_i$ into rectifiable Jordan curves, this equivalent to

$$0 = \left(P(\text{int } \gamma_i^+) - \int_{\text{int } \gamma_i^+} v \right) + \sum_{j \in L_i} \left(P(\text{int } \gamma_j^-) + \int_{\text{int } \gamma_j^-} v \right).$$

Since each Jordan curve γ satisfies $P(\text{int } \gamma) - \left| \int_\gamma v \right| \geq 0$, we see that for all i and j in the decomposition,

$$P(\text{int } \gamma_i^+) = \int_{\gamma_i^+} v \quad \text{and} \quad P(\text{int } \gamma_j^-) = - \int_{\gamma_j^-} v.$$

Similarly, we may prove that if $P(E) = - \int_E v$,

$$\begin{aligned} P(E_i) &= - \int_{E_i} v, \\ P(\text{int } \gamma_i^+) &= - \int_{\gamma_i^+} v \quad \text{and} \quad P(\text{int } \gamma_j^-) = \int_{\gamma_j^-} v. \end{aligned}$$

4. Duality for the study of the low noise regime

4.1. Dual problems and “dual certificates”

We are interested in solving:

$$\min_{u \in L^2(\mathbb{R}^2)} J(u) + \frac{1}{2\lambda} \|f - u\|_{L^2}^2, \quad (\mathcal{P}_\lambda(f))$$

where $J(u) = \int_{\mathbb{R}^2} |Du| \in \mathbb{R}_+ \cup \{+\infty\}$.

Using the framework and notations of [49], we set $V = L^2(\mathbb{R}^2)$, $\Lambda = Id$, $Y = L^2(\mathbb{R}^2)$, $F = J$, $G = \frac{1}{2\lambda} \|\cdot - f\|^2$ and we compute the Fenchel-Rockafellar dual problem as

$$\sup_{v \in \partial J(0)} \langle f, v \rangle - \frac{1}{2\lambda} \|v\|_{L^2}^2, \quad (\mathcal{D}'_\lambda(f))$$

$$\text{or equivalently} \quad \inf_{v \in \partial J(0)} \left\| \frac{f}{\lambda} - v \right\|_{L^2}^2. \quad (\mathcal{D}_\lambda(f))$$

It is easy to check that Problem $(\mathcal{P}_\lambda(f))$ is stable in the sense of [49]. In particular, there exists a solution to $(\mathcal{D}'_\lambda(f))$ and strong duality holds between $(\mathcal{P}_\lambda(f))$ and $(\mathcal{D}'_\lambda(f))$, namely $\inf(\mathcal{P}_\lambda(f)) = \sup(\mathcal{D}'_\lambda(f))$. In fact, since $(\mathcal{D}_\lambda(f))$ is a projection problem onto a nonempty closed convex set, it always has a unique solution.

Observe that formally, the limit of $(\mathcal{P}_\lambda(f))$ as $\lambda \rightarrow 0^+$ is the trivial problem

$$\min_{u \in L^2(\mathbb{R}^2)} J(u) \quad \text{s.t.} \quad u = f, \quad (\mathcal{P}_0(f))$$

having $u = f$ as solution. The dual associated with this “exact reconstruction problem” is

$$\sup_{v \in \partial J(0)} \langle f, v \rangle, \quad (\mathcal{D}_0(f))$$

having $\partial J(f)$ as solutions. Here again, strong duality holds, since it is possible to prove that $(\mathcal{D}_0(f))$ is stable. However, a solution to $(\mathcal{D}_0(f))$ does not always exist since it may be that $\partial J(f) = \emptyset$.

The main point in studying the dual problems is that their solutions v_λ are related to the primal solutions u_λ by the extremality relations

$$v_\lambda \in \partial J(u_\lambda) \quad \text{and} \quad v_\lambda = \frac{1}{\lambda}(f - u_\lambda),$$

which enables to study the support of Du_λ (see Section 3). For the noiseless problem, the extremality relation is $v \in \partial J(f)$, for every v solution to $(\mathcal{D}_0(f))$. The term “certificate” stems from the fact that if $u \in L^2(\mathbb{R}^2)$ and $v \in L^2(\mathbb{R}^2)$ satisfy the extremality relations, then u is a solution of the primal problem and v is a solution of the dual problem.

4.2. Low noise regimes and the minimal norm certificate

We shall often consider noisy observations $f + w$, where $w \in L^2(\mathbb{R}^2)$, and from now on, we denote by $u_{\lambda,w}$ (resp. $v_{\lambda,w}$) the unique solution to $\mathcal{P}_\lambda(f + w)$ (resp. $\mathcal{D}_\lambda(f + w)$).

Given $\lambda_0 > 0$, $\alpha_0 > 0$, we consider the low noise regime

$$D_{\lambda_0, \alpha_0} \stackrel{\text{def.}}{=} \left\{ (\lambda, w) \in \mathbb{R}_+ \times L^2(\mathbb{R}^2) ; 0 \leq \lambda \leq \lambda_0 \quad \text{and} \quad \|w\|_{L^2} \leq \alpha_0 \lambda \right\}. \quad (26)$$

The dual solution $v_{\lambda,w}$ being the projection of $(f + w)/\lambda$ onto a convex set, the non-expansiveness of the projection yields

$$\forall (\lambda, w) \in \mathbb{R}_+^* \times L^2(\mathbb{R}^2), \quad \|v_{\lambda,0} - v_{\lambda,w}\|_{L^2} \leq \frac{\|w\|_{L^2}}{\lambda} \leq \alpha_0.$$

As a result, the properties of $v_{\lambda,w}$ are governed by those of $v_{\lambda,0}$, and it turns out that the properties of $v_{\lambda,0}$ are governed, in the low noise regime, by those of a specific solution to $(\mathcal{D}_0(f))$, as the next result hints. The proof is identical to the one in [47].

Proposition 5. *Let $f \in L^2(\mathbb{R}^2)$, $J(f) < +\infty$, and assume that $\partial J(f) \neq \emptyset$. Let $v_{0,0} \in L^2(\mathbb{R}^2)$ be the solution to $(\mathcal{D}_0(f))$ with minimal L^2 norm. Then*

$$\lim_{\lambda \rightarrow 0^+} v_{\lambda,0} = v_{0,0} \quad \text{strongly in } L^2(\mathbb{R}^2).$$

We call $v_{0,0}$ the **minimal norm certificate** for f . It is also known as the *minimal section* in maximal monotone operator theory. The goal of the present paper is to show that $v_{0,0}$ governs the support of the solutions in the low noise regime. In particular, $v_{0,0}$ determines whether the support of $Du_{\lambda,w}$ is close to the support of Df in that regime.

In the next paragraphs, we illustrate the minimal norm certificate in simple cases.

4.3. The minimal norm certificate for calibrable sets

Proposition 6 (Minimal norm certificates for calibrable sets). *Let $C \subseteq \mathbb{R}^2$ be a bounded calibrable set and $f = \mathbb{1}_C$. Then the minimal norm certificate is $v_{0,0} = h_C \mathbb{1}_C$, where $h_C = \frac{P(C)}{|C|}$.*

We provide two different proofs of the above result, each highlighting different aspects of the minimal norm certificate.

Proof ($v_{0,0}$ as a limit). From [20], we know that for a calibrable set $C \subseteq \mathbb{R}^2$, the solution to $(\mathcal{P}_\lambda(f))$ with $f = \mathbb{1}_C$ is given by $u_{\lambda,0} = (1 - \lambda h_C)_+ \mathbb{1}_C$. From the optimality conditions,

$$v_{\lambda,0} = \frac{1}{\lambda}(f - u_{\lambda,0}),$$

we obtain that $v_{\lambda,0} = h_C \mathbb{1}_C$ provided $0 < \lambda \leq \frac{1}{h_C}$. Taking the limit as $\lambda \rightarrow 0^+$, we obtain $v_{0,0} = h_C \mathbb{1}_C$. \square

Another Proof ($v_{0,0}$ as a minimal norm element). Observe that for all $f \in L^2(\mathbb{R}^2)$ with $J(f) < +\infty$, and $v \in L^2(\mathbb{R}^2)$, v is a solution to $(\mathcal{D}_0(f))$ if and only if $v \in \partial J(f)$. For $C \subset \mathbb{R}^2$ bounded calibrable, we obtain that $h_C \mathbb{1}_C$ is a solution to $(\mathcal{D}_0(f))$. It remains to prove that it is the one with minimal norm.

Let $v \in L^2(\mathbb{R}^2)$ be any solution to $(\mathcal{D}_0(f))$. By the Cauchy-Schwarz inequality

$$\sup \mathcal{D}_0(\mathbb{1}_C) = \langle v, \mathbb{1}_C \rangle \leq \|v\|_{L^2} \|\mathbb{1}_C\|_{L^2} = \sqrt{|C|} \|v\|_{L^2}.$$

However, $\|h_C \mathbb{1}_C\|_{L^2} = \frac{P(C)}{\sqrt{|C|}} = \frac{\sup \mathcal{D}_0(\mathbb{1}_C)}{\sqrt{|C|}}$, so that $h_C \mathbb{1}_C$ has minimal norm. \square

4.4. The minimal norm certificate for smooth convex sets

Let C be a nonempty open bounded convex subset of \mathbb{R}^2 . Given $\rho > 0$ we denote by C_ρ the opening of C by open balls with radius ρ , namely $C_\rho = \bigcup_{B(y,\rho) \subseteq C} B(y,\rho)$. For $f = \mathbb{1}_C$, it is proved in [7, 9, 39] that the solution $u_{\lambda,0}$ to $(\mathcal{P}_\lambda(f))$ is

$$u_{\lambda,0} = (1 + \lambda v_C)^+ \mathbb{1}_C,$$

where, by letting R be such that C_R is the maximal calibrable set in C , the function $v_C : \mathbb{R}^2 \rightarrow \mathbb{R}$ is defined by

$$v_C(x) \stackrel{\text{def.}}{=} \begin{cases} 1/R & x \in C_R \\ 1/r & x \in \partial C_r, \ r \in [0, R) \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

Since $v_{\lambda,0} = \lambda^{-1}(f - u_{\lambda,0})$, it follows that

$$v_{\lambda,0}(x) = \begin{cases} v_C(x) & x \in C_\lambda \\ 1/\lambda & x \in C \setminus C_\lambda \\ 0 & \text{otherwise.} \end{cases} \quad (28)$$

Now we assume that $C \subset \mathbb{R}^2$ has $\mathcal{C}^{1,1}$ boundary, and we let $\rho_0 > 0$ such that

$$\kappa_{\partial C}(x) \leq \frac{1}{\rho_0} \quad \text{for } \mathcal{H}^1\text{-a.e. } x \in \partial C,$$

where $\kappa_{\partial C}$ is the curvature of ∂C (defined \mathcal{H}^1 -almost everywhere on ∂C). We shall need the following lemma.

Lemma 1 ([20]). *Let $C \subset \mathbb{R}^2$ be a bounded open convex set. The following conditions are equivalent:*

- there exists $\rho > 0$ such that $C = C_\rho$.
- ∂C is of class $\mathcal{C}^{1,1}$ and $\text{ess sup}_{p \in \partial C} \kappa_{\partial C}(p) \leq \frac{1}{\rho}$.

Since for $0 < r \leq \rho_0$, $C_{\rho_0} \subseteq C_r \subseteq C$, we see that $C_r = C$ for $0 < r \leq \rho_0$.

As a result, $\lambda \mapsto v_{\lambda,0}$ is constant on $(0, \rho_0]$, and the minimal norm certificate is thus

$$v_{0,0} = v_C. \quad (29)$$

It turns out that $v_{0,0}$ is precisely the subgradient constructed by Alter *et al.* [9] for the evolution of convex sets by the total variation flow. It is instructive to look at the associated vector field z_0 such that $\text{div } z_0 = v_{0,0}$.

For every $x \in \text{int}(C) \setminus \overline{C_R}$, there exists a unique $r(x)$ such that $x \in \partial C_{r(x)}$, and x belongs to an arc of circle of radius $r(x)$. Defining $\nu(x)$ as the outer unit normal to this set, define

$$z_0(x) \stackrel{\text{def.}}{=} \begin{cases} \nu(x) & \text{if } x \in \text{int}(C) \setminus C_R \\ z_{C_R}(x) & \text{if } x \in C_R \\ \bar{z}(x) & \text{if } x \in \mathbb{R}^2 \setminus C, \end{cases}$$

where \bar{z} is a calibration of $\mathbb{R}^2 \setminus C$ (see Section 9.4). As for z_{C_R} , since C_R is calibrable (C_R is then the Cheeger set of C) there exists a vector field z_{C_R} such that $|z_{C_R}| \leq 1$, $\theta(z_{C_R}, -D\mathbf{1}_{C_R}) = 1$, and $\operatorname{div} z_{C_R} = h_{C_R} \mathbf{1}_{C_R}$ with $h_{C_R} = P(C_R)/|C_R|$. It is proved in [9] that $\operatorname{div} z_0 = v_C$ (in the sense of distributions).

It is notable that the construction is quite similar to the one proposed by Barrozzzi et al. in [17] and studied in [55]. In particular, the L^2 -minimality (or even L^p minimality) of the above constructions is already noted in [55].

5. Properties of the level sets in the low noise regime

In this section, we rely on the properties of the minimal norm certificate $v_{0,0}$ to study the solutions of (23) for $v = v_{\lambda,w}$ in a low noise regime. More precisely we study the elements of

$$\mathcal{F}_{\lambda,w} \stackrel{\text{def.}}{=} \left\{ E \subset \mathbb{R}^2 ; |E| < +\infty, \quad \text{and} \quad \pm \int_E v_{\lambda,w} = P(E) \right\}, \quad (30)$$

for $(\lambda, w) \in D_{\lambda_0, \alpha_0}$ with $\lambda_0 > 0$, $\alpha_0 > 0$ small enough. In the following, we denote by E or $E_{\lambda,w}$ any nonempty element of $\mathcal{F}_{\lambda,w}$. Let us emphasize that we allow the case $(\lambda, w) = (0, 0)$, in which case $v_{\lambda,w}$ in (30) is the minimal norm certificate $v_{0,0}$. Typically, from Section 3, one may think of $E_{\lambda,w}$ as a level set of $u_{\lambda,w}$ (or f , for $(\lambda, w) = (0, 0)$), but additional sets may solve (30).

5.1. Upper and lower bounds

In the following lemmas, we prove that there exist uniform upper and lower bounds on the perimeters and the measures of all sets in $\mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$.

Lemma 2. *Let $\alpha_0 \leq \frac{\sqrt{c_2}}{4}$, where $c_2 = 4\pi$ is the isoperimetric constant. Then,*

$$\sup \{P(E) ; E \in \mathcal{F}_{\lambda,w}, \quad \text{and} \quad (\lambda, w) \in D_{1, \alpha_0}\} < +\infty, \quad (31)$$

$$\sup \{|E| ; E \in \mathcal{F}_{\lambda,w}, \quad \text{and} \quad (\lambda, w) \in D_{1, \alpha_0}\} < +\infty. \quad (32)$$

Proof. First, we prove (31). Since $\lim_{\lambda \rightarrow 0^+} v_{\lambda,0} = v_{0,0}$ in $L^2(\mathbb{R}^2)$, the mapping $\lambda \mapsto v_{\lambda,0}$ is continuous on the compact set $[0, 1]$, hence bounded. Moreover, the family $\{v_{\lambda,0}\}_{0 \leq \lambda \leq 1}$ is L^2 -equiintegrable so that given any $\varepsilon > 0$, there exists $R > 0$ such that $\int_{\mathbb{R}^2 \setminus B(0,R)} v_{\lambda,0}^2 \leq \varepsilon^2$ for all $\lambda \in [0, 1]$. Let us also assume that $\alpha_0 \leq \varepsilon$ (so that $\|v_{\lambda,w} - v_{\lambda,0}\|_{L^2} \leq \frac{\|w\|_{L^2}}{\lambda} \leq \varepsilon$).

To simplify the notation, we denote by E (rather than $E_{\lambda,w}$) any nonempty set such that $P(E) = \pm \int_E v_{\lambda,w}$.

Now, the triangle and the Cauchy-Schwarz inequalities yield

$$\begin{aligned}
P(E) &\leq \left| \int_E (v_{\lambda,w} - v_{\lambda,0}) \right| + \left| \int_E v_{\lambda,0} \right| \\
&\leq \varepsilon \sqrt{|E|} + \left| \int_{E \cap B(0,R)} v_{\lambda,0} \right| + \left| \int_{E \setminus B(0,R)} v_{\lambda,0} \right| \\
&\leq \varepsilon \sqrt{|E|} + \sqrt{|B(0,R)|} \|v_{\lambda,0}\|_{L^2} + \sqrt{|E \setminus B(0,R)|} \sqrt{\int_{\mathbb{R}^2 \setminus B(0,R)} v_{\lambda,w}^2} \\
&\leq \left(\varepsilon + \sup_{\lambda \in [0,1]} \|v_{\lambda,0}\|_{L^2} \right) \sqrt{|B(0,R)|} + 2\varepsilon \sqrt{|E \setminus B(0,R)|}
\end{aligned}$$

Recalling that $P(E \setminus B(0,R)) \leq P(E) + P(B(0,R))$, and using the isoperimetric inequality, we obtain

$$\sqrt{|E \setminus B(0,R)|} \leq \frac{1}{\sqrt{c_2}} (P(E) + P(B(0,R))),$$

where c_2 is the isoperimetric constant. We choose $\varepsilon = \frac{\sqrt{c_2}}{4}$ and we define $C = (\varepsilon + \sup_{\lambda \in [0,1]} \|v_{\lambda,0}\|_{L^2}) \sqrt{|B(0,R)|}$ so as to get

$$P(E) - \frac{1}{2} (P(E) + P(B(0,R))) \leq C.$$

We obtain that $P(E)$ is uniformly bounded in $E \in \mathcal{F}_{\lambda,w}$, $(\lambda, w) \in D_{1,\alpha_0}$.

As for (32), the isoperimetric inequality yields

$$|E| \leq \frac{1}{c_2} (P(E))^2.$$

Hence, $|E|$ is uniformly bounded in $E \in \mathcal{F}_{\lambda,w}$, $(\lambda, w) \in D_{1,\alpha_0}$. \square

Conversely, the perimeters and areas of the solutions are also lower bounded, as the next result shows.

Lemma 3. *Let $\alpha_0 \leq \frac{\sqrt{c_2}}{4} = \sqrt{\pi}/2$. Then,*

$$\inf \{P(E) ; E \in \mathcal{F}_{\lambda,w}, E \neq \emptyset \text{ and } (\lambda, w) \in D_{1,\alpha_0}\} > 0, \quad (33)$$

$$\inf \{|E| ; E \in \mathcal{F}_{\lambda,w}, E \neq \emptyset \text{ and } (\lambda, w) \in D_{1,\alpha_0}\} > 0. \quad (34)$$

Moreover, there exists a number $N_0 \in \mathbb{N}$ such that the number of M -connected components E and the number of Jordan curves in the essential boundary $\partial^M E$ is uniformly bounded by N_0 for all $E \in \mathcal{F}_{\lambda,w}$, $(\lambda, w) \in D_{1,\alpha_0}$.

Proof. By the L^2 -equiintegrability of the family $\{v_{\lambda,0}\}_{0 \leq \lambda \leq 1}$, for all $\varepsilon > 0$, there exists δ such that for all $E \subset \mathbb{R}^2$, with $|E| \leq \delta$,

$$\int_E v_{\lambda,0}^2 \leq \varepsilon^2.$$

We choose $\varepsilon = \frac{\sqrt{c_2}}{4}$, $0 < \alpha_0 \leq \varepsilon$, and we consider by contradiction a set $E \in \mathcal{F}_{\lambda,w}$ such that $0 < |E| \leq \delta$. Then,

$$\begin{aligned} P(E) &\leq \left| \int_E (v_{\lambda,w} - v_{\lambda,0}) \right| + \left| \int_E v_{\lambda,0} \right| \\ &\leq \|v_{\lambda,w} - v_{\lambda,0}\|_{L^2} \sqrt{|E|} + \sqrt{\int_E v_{\lambda,0}^2} \sqrt{|E|} \\ &\leq 2\varepsilon \sqrt{|E|} \leq \frac{1}{2} P(E), \end{aligned}$$

by the isoperimetric inequality. Dividing by $P(E) > 0$ yields a contradiction, hence $|E| > \delta$ for all $E \neq \emptyset$, that is (34). We deduce the uniform lower bound on the perimeter (33) by the isoperimetric inequality.

Now, let us decompose the essential boundary of $E \in \mathcal{F}_{\lambda,w}$ into *at most countably* many non trivial Jordan curves $\{\gamma_i^+, \gamma_j^- ; i \in I, j \in J, I \subseteq \mathbb{N}, J \subseteq \mathbb{N}\}$. By Remark 4, we know that for each $\sigma \in \{-1, 1\}$ and $j \in \mathbb{N}$, $\left| \int_{\text{int } \gamma_j^\sigma} v_{\lambda,w} \right| = \mathcal{H}^1(\gamma_j^\sigma)$, that is $(\text{int } \gamma_j^\sigma) \in \mathcal{F}_{\lambda,w}$. As a result $\mathcal{H}^1(\gamma_j^\sigma) \geq \mu$, where μ is the infimum defined in (33) (in I, J , we only consider the non-trivial Jordan curves). Expressing the perimeter of E in terms of these Jordan curves, we get

$$C \geq P(E_{\lambda,w}) = \sum_{i \in I} \mathcal{H}^1(\gamma_i^+) + \sum_{j \in J} \mathcal{H}^1(\gamma_j^-) \geq \mu(\text{Card } I + \text{Card } J),$$

where C is the supremum in (31). Hence the number of Jordan curves is at most C/μ , and the same holds for the number of M-connected components. \square

Additionally, the next result shows that the level sets are uniformly contained in some large ball.

Lemma 4. *Let $\alpha_0 \leq \frac{\sqrt{c_2}}{4} = \sqrt{\pi}$. Then, there exists $R > 0$ such that*

$$\forall (\lambda, w) \in D_{1,\alpha_0}, \quad \forall E \in \mathcal{F}_{\lambda,w}, \quad E \subset B(0, R).$$

Proof. We begin with the same equiintegrability argument as in Lemma 2, choosing again $\varepsilon = \frac{\sqrt{c_2}}{4}$. Now, let $E \in \mathcal{F}_{\lambda,w}$. By the results of Section 3.7, we may further decompose, up to an \mathcal{H}^1 -negligible set, its essential boundary $\partial^M E$ into an at most countable union of Jordan curves γ which satisfy

$$\pm \int_{\text{int } \gamma} v_{\lambda,w} = P(\text{int } \gamma) = \mathcal{H}^1(\gamma).$$

Assume by contradiction that γ is such that $(\text{int } \gamma) \cap B(0, R) = \emptyset$. Then by the isoperimetric inequality,

$$P(\text{int } \gamma) \leq \sqrt{\int_{\mathbb{R}^2 \setminus B(0,R)} v_{\lambda,w}^2} \sqrt{|\text{int } \gamma|} \leq \frac{2\varepsilon}{\sqrt{c_2}} P(\text{int } \gamma).$$

Dividing by $P(\text{int } \gamma)$ yields a contradiction for $\varepsilon = \frac{\sqrt{c_2}}{4}$ if γ is not trivial. Hence $(\text{int } \gamma) \cap B(0, R) \neq \emptyset$. But the uniform bound (31) also holds for γ , hence there is some $C > 0$ (independent from $(\lambda, w) \in D_{1, \alpha_0}$) such that $\mathcal{H}^1(\gamma) \leq C$. As a result, $\text{diam}(\text{int } \gamma) \leq C$ so that $(\text{int } \gamma) \subset B(0, R + C)$, and since this holds for any γ which is involved in the decomposition of $\partial^M E$, it also holds for all $E \in \mathcal{F}_{\lambda, w}$, uniformly in $(\lambda, w) \in D_{1, \alpha_0}$. \square

Remark 5. Let us divide $\mathcal{F}_{\lambda, w}$ into two classes corresponding respectively to the condition $\int_E v_{\lambda, w} = P(E)$ and $-\int_E v_{\lambda, w} = P(E)$ (the empty set being the only element which belongs to both). A consequence of (32) is that *each class is stable by a countable union or intersection*. Indeed, if E and F are two elements of $\mathcal{F}_{\lambda, w}$ such that $\int_E v_{\lambda, w} = P(E)$ (and similarly for F), the submodularity of the perimeter yields

$$P(E \cap F) + P(E \cup F) \leq P(E) + P(F) = \int_E v_{\lambda, w} + \int_F v_{\lambda, w} = \int_{E \cap F} v_{\lambda, w} + \int_{E \cup F} v_{\lambda, w}.$$

Using the subdifferential inequality (on $v_{\lambda, w} \in \partial J(0)$) we obtain that $P(E \cap F) = \int_{E \cap F} v_{\lambda, w}$ and $P(E \cup F) = \int_{E \cup F} v_{\lambda, w}$. Iterating, we get for finite union or intersection $P(\bigcup_{k=1}^n E_k) = \int_{\bigcup_{k=1}^n E_k} v_{\lambda, w}$ and $P(\bigcap_{k=1}^n E_k) = \int_{\bigcap_{k=1}^n E_k} v_{\lambda, w}$. The lower semi-continuity of the perimeter together with $|E_1| < +\infty$ yields

$$P\left(\bigcap_{k=1}^{\infty} E_k\right) \leq \liminf_{n \rightarrow +\infty} P\left(\bigcap_{k=1}^n E_k\right) = \lim_{n \rightarrow +\infty} \int_{\bigcap_{k=1}^n E_k} v_{\lambda, w} = \int_{\bigcap_{k=1}^{\infty} E_k} v_{\lambda, w},$$

and the converse inequality holds by the subdifferential inequality. As for the union, we know from (32) that $|\bigcup_{k=1}^{\infty} E_k| = \sup_{n \in \mathbb{N}} |\bigcup_{k=1}^n E_k| < +\infty$, hence

$$P\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \liminf_{n \rightarrow +\infty} P\left(\bigcup_{k=1}^n E_k\right) = \lim_{n \rightarrow +\infty} \int_{\bigcup_{k=1}^n E_k} v_{\lambda, w} = \int_{\bigcup_{k=1}^{\infty} E_k} v_{\lambda, w},$$

and the opposite inequality also holds, for the same reason as above.

5.2. Weak regularity

In this section, we show that (25) holds uniformly on the boundaries of the sets in $\mathcal{F}_{\lambda, w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$. The proof of Proposition 7 is in fact almost identical to the proof of [54, Lem. 1.2], however, it is included for the sake of completeness, and so as to emphasize the uniformity of this estimate with respect to (λ, w) .

Proposition 7. *There exists $r_0 > 0$ such that for all $r \in (0, r_0]$ and $E_{\lambda, w} \in \mathcal{F}_{\lambda, w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$,*

$$\forall x \in \partial E_{\lambda, w}, \quad \frac{|B(x, r) \cap E_{\lambda, w}|}{|B(x, r)|} \geq \frac{1}{16} \quad \text{and} \quad \frac{|B(x, r) \setminus E_{\lambda, w}|}{|B(x, r)|} \geq \frac{1}{16}. \quad (35)$$

Proof. We give the proof for $P(E_{\lambda,w}) = \int_{E_{\lambda,w}} v_{\lambda,w}$, the other case being similar. Since $\{v_{\lambda,0}\}_{\lambda \in [0,1]}$ is equiintegrable, there exists $r_0 > 0$ such that for all subsets $E \subset \mathbb{R}^2$ with $|E| \leq \pi r_0^2$,

$$\left(\int_E |v_{\lambda,0}|^2 \right)^{1/2} \leq \frac{\sqrt{c_2}}{4}. \quad (36)$$

First observe that by optimality of $E_{\lambda,w}$,

$$P(E_{\lambda,w}) - \int_{E_{\lambda,w}} v_{\lambda,w} \leq P(E_{\lambda,w} \setminus B(x,r)) - \int_{E_{\lambda,w} \setminus B(x,r)} v_{\lambda,w}. \quad (37)$$

For a.e. $r \in (0, r_0]$, $\mathcal{H}^1(\partial^* E_{\lambda,w} \cap \partial B(x,r)) = 0$, so (37) yields

$$\mathcal{H}^1(\partial^* E_{\lambda,w} \cap B(x,r)) - \int_{E_{\lambda,w} \cap B(x,r)} v_{\lambda,w} \leq \mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w}).$$

By adding $\mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w})$ to both sides, it follows that

$$P(E_{\lambda,w} \cap B(x,r)) - \int_{E_{\lambda,w} \cap B(x,r)} v_{\lambda,w} \leq 2\mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w}).$$

By the Cauchy-Schwarz inequality, (36) and since $\|v_{\lambda,w} - v_{\lambda,0}\|_{L^2} \leq \sqrt{c_2}/4$

$$P(E_{\lambda,w} \cap B(x,r)) - \frac{\sqrt{c_2} |E_{\lambda,w} \cap B(x,r)|^{1/2}}{2} \leq 2\mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w}).$$

The isoperimetric inequality then implies that

$$\sqrt{c_2} |E_{\lambda,w} \cap B(x,r)|^{1/2} \leq 4\mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w}).$$

Let $g(r) = |E_{\lambda,w} \cap B(x,r)|$. Then $g(r) > 0$ since $x \in \partial E_{\lambda,w}$, and for a.e. r , $g'(r) = \mathcal{H}^1(\partial B(x,r) \cap E_{\lambda,w})$. Therefore, for a.e. $r \in (0, r_0]$,

$$\sqrt{c_2} \leq 8 \frac{d}{dr} \sqrt{g(r)}.$$

By integrating on both sides,

$$r\sqrt{c_2} \leq 8\sqrt{g(r)}.$$

and the first inequality in (35) follows by recalling that $c_2 = 4\pi$. The proof of $|B(x,r) \setminus E_{\lambda,w}| \geq |B(x,r)|/16$ is similar: instead of comparing $E_{\lambda,w}$ with $E_{\lambda,w} \setminus B(x,r)$ in (37), simply compare $E_{\lambda,w}$ with $E_{\lambda,w} \cup B(x,r)$ and proceed as before. \square

6. The extended support

Let $f \in L^2(\mathbb{R}^2)$, with $J(f) < +\infty$, such that $\partial J(f) \neq \emptyset$, or equivalently that $(\mathcal{D}_0(f))$ has a solution (source condition). Let $v_{0,0}$ be the corresponding minimal norm certificate and let us define the extended support as

$$\text{Ext}(Df) \stackrel{\text{def.}}{=} \overline{\bigcup \{\text{Supp } Dg ; v_{0,0} \in \partial J(g)\}}.$$

As we shall see in Section 7, the extended support governs the location of $\text{Supp}(Du_{\lambda,w})$ for (λ, w) in some low noise regime.

6.1. Properties of the extended support

A first remark in view of Proposition 4 is that we may rewrite the extended support as

$$\text{Ext}(Df) = \overline{\bigcup \left\{ \partial^* E ; |E| < +\infty \quad \text{and} \quad \pm \int_E v_{0,0} = P(E) \right\}}. \quad (38)$$

The inclusion of $\text{Ext}(Df)$ in the right hand-side of the above equation is clear by Proposition 4. The converse inclusion is obtained by considering, for any E in the right hand-side, the function $g = \mathbf{1}_E$, so as to have $g \in L^2$ and $v_{0,0} \in \partial J(g)$ (since $\overline{\partial^* E} = \text{Supp}(Dg)$).

From the above equalities, we see that *all the properties of Section 5 (lower and upper boundedness of the perimeter, uniform boundedness...) hold for the elements of the right hand-side whose union determines the extended support.*

The rest of the section is devoted to examples of extended supports, in the case of indicator function of convex calibrable sets or more general convex sets.

6.2. Convex Calibrable sets

Let $C \subset \mathbb{R}^2$ be a bounded convex calibrable set. We wish to describe the extended support of $f = \mathbf{1}_C$. This may be done by looking at a vector field z with divergence $v_{0,0}$ (see Section 9.2), which is more informative, or by the following approach.

By Proposition 6, we know that the minimal norm certificate associated to $f = \mathbf{1}_C$ is $v_{0,0} = h_C \mathbf{1}_C$, where $h_C = \frac{P(C)}{|C|}$. By (38), we are thus led to solve

$$\inf_{\substack{E \subset \mathbb{R}^2 \\ |E| < +\infty}} P(E) - h_C |E \cap C|, \quad (39)$$

$$\text{and} \quad \inf_{\substack{E \subset \mathbb{R}^2 \\ |E| < +\infty}} P(E) + h_C |E \cap C|. \quad (40)$$

Problem (40) is trivial and its only solution is \emptyset , so that we only focus on (39). By Proposition 2, we see that $E = C$ is a minimizer. Moreover, since C is convex, for all E with finite perimeter $P(C \cap E) \leq P(E)$ with strict inequality whenever $|E \setminus C| > 0$. As a result, any other solution must satisfy $E \subseteq C$. But with this condition, either $E = \emptyset$ or E is a solution to the Cheeger problem

$$\min_{E \subseteq C} \frac{P(E)}{|E|}.$$

The uniqueness of the solution to the Cheeger problem inside any convex set is proved in [6, 53], and we already know that C is optimal. As a result, either $E = \emptyset$ or $E = C$, and eventually

$$\text{Ext}(Df) = \partial C.$$

6.3. Smooth convex sets

Let $C \subset \mathbb{R}^2$ be a bounded open convex set with $\mathcal{C}^{1,1}$ boundary. We describe the extended support of $f = \mathbb{1}_C$ by considering the minimal norm certificate $v_{0,0}$ defined in (27). We need to study the solutions of

$$\inf_{\substack{E \subset \mathbb{R}^2 \\ |E| < +\infty}} P(E) - \int_E v_{0,0}, \quad (41)$$

$$\text{and } \inf_{\substack{E \subset \mathbb{R}^2 \\ |E| < +\infty}} P(E) + \int_E v_{0,0}. \quad (42)$$

Since $v_{0,0} \geq 0$, we see that the only solution to (42) is \emptyset . As for (41), the same convexity argument as above shows that any solution must be included in C .

Now let $r \in [\rho_0, R]$, where $1/\rho_0 \geq \text{ess sup}_{x \in \partial C} \kappa(x)$ and C_r be the opening of C with radius r as defined in Section 4.4. Denoting by ν_{C_r} the outer unit normal to ∂C_r , we have

$$P(C_r) = \int_{\partial C_r} z_0 \cdot \nu_{C_r} d\mathcal{H}^1 = \int_{C_r} \text{div } z_0 = \int_{C_r} v_{0,0},$$

hence C_r is a solution to (41) and $\text{Ext}(Df) \supseteq \overline{\bigcup \{\partial C_r ; \rho_0 \leq r \leq R\}}$.

Let us prove that there is no solution E such that the reduced boundary $\partial^* E$ intersects C_R . By Remark 5, the solutions to (41) are stable by intersection. If a solution E is such that $E \cap C_R \neq \emptyset$, then $P(E \cap C_R) = \int_{E \cap C_R} v_{0,0} = h_{C_R} |E \cap C_R|$ where $h_{C_R} = \frac{P(C_R)}{|C_R|}$ and $E \cap C_R$ is a solution to the Cheeger problem

$$\min_{F \subseteq C_R} \frac{P(F)}{|F|}.$$

By uniqueness of the Cheeger set of C_R , we obtain that $E \cap C_R = C_R$. Eventually, we have proved

$$\text{Ext}(Df) = \overline{\bigcup \{\partial C_r ; \rho_0 \leq r \leq R\}}. \quad (43)$$

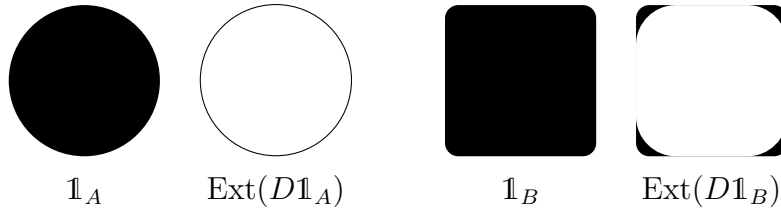


Figure 1: Examples of the extended support for two indicator functions.

7. Support stability outside the extended support

In this section, we prove the main result of this paper, Theorem 1, which shows that, under the source condition $\partial J(f) \neq \emptyset$, as $\lambda \rightarrow 0^+$ and $\|w\|_{L^2}/\lambda$ is small enough,

almost all topological boundaries of the level sets of the solutions to $(\mathcal{P}_\lambda(f+w))$ converge towards the topological boundaries of the corresponding level sets of f in the sense of Hausdorff convergence. If, moreover, $\|w\|_{L^2}/\lambda \rightarrow 0$, the support of $Du_{\lambda,w}$ is contained in arbitrarily small tubular neighborhoods of the extended support $\text{Ext}(Df)$. In Section 9, we show that the width of this tube can be further characterized through the knowledge of a vector field z_0 associated with $v_{0,0}$. An interesting consequence of our main result is that the minimal norm certificate $v_{0,0}$ is constant on each connected component of the extended support (See Corollary 1).

Throughout this section, we denote by $v_{\lambda,w}$ the solution of $(\mathcal{D}_\lambda(f+w))$ and let $E_{\lambda,w}$ be any set of finite perimeter such that $\left| \int_{E_{\lambda,w}} v_{\lambda,w} \right| = P(E_{\lambda,w})$. We also denote by $u_{\lambda,w}$ the solution of $\mathcal{P}_\lambda(f+w)$. Finally, let the level sets of f be denoted by $F^{(t)}$ (refer to (9) for the definition of level sets).

We begin by recalling an elementary result, which holds under very weak assumptions.

Proposition 8. *Let $f \in L^2(\mathbb{R}^2)$ such that $J(f) < +\infty$. Let $\{w_n\}_{n \in \mathbb{N}}$, $\{\lambda_n\}_{n \in \mathbb{N}}$ be sequences such that $w_n \in L^2(\mathbb{R}^2)$, $\|w_n\|_{L^2} \rightarrow 0$, and $\lambda_n \rightarrow 0^+$. Then $\lim_{n \rightarrow +\infty} \|u_{\lambda_n, w_n} - f\|_{L^2} = 0$ and $\text{Supp}(Df) \subseteq \liminf_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n})$.*

Proof. For the sake of simplicity, we shall denote u_{λ_n, w_n} by u_n . From the optimality of u_n ,

$$\lambda_n J(u_n) + \frac{1}{2} \int_{\mathbb{R}^2} (f + w_n - u_n)^2 \leq \lambda_n J(f),$$

we see that $\|f - u_n\|_{L^2} \rightarrow 0$ as $n \rightarrow +\infty$. Together with the fact that $J(u_n) \leq J(f) < +\infty$, that implies that Du_n converges towards Df in the weak-* topology of Radon measures. If $\text{Supp}(Df) = \emptyset$, there is nothing to prove. Otherwise, let $x \in \text{Supp}(Df)$. By weak-* convergence, for all $r > 0$,

$$0 < |Df|(B(x, r)) \leq \liminf_{n \rightarrow +\infty} |Du_n|(B(x, r)).$$

Hence, $\limsup_{n \rightarrow +\infty} \text{dist}(x, \text{Supp}(Du_n)) \leq r$, and since this is true for all $r > 0$, we obtain $x \in \liminf_{n \rightarrow +\infty} \text{Supp}(Du_n)$. This yields $\text{Supp}(Df) \subseteq \liminf_{n \rightarrow +\infty} \text{Supp}(Du_n)$. \square

With the additional assumption that $\partial J(f) \neq \emptyset$, it is possible to describe the behavior of the level lines more precisely. In the following, we denote by $U_n^{(t)}$ the t -level set of u_{λ_n, w_n} .

Theorem 1. *Let $f \in L^2(\mathbb{R}^2)$ such that $J(f) < +\infty$ and $\partial J(f) \neq \emptyset$. Let $\{w_n\}_{n \in \mathbb{N}}$, $\{\lambda_n\}_{n \in \mathbb{N}}$ be sequences such that $w_n \in L^2(\mathbb{R}^2)$, $\lambda_n \rightarrow 0^+$, and $\|w_n\|_{L^2}/\lambda_n \leq \sqrt{c_2}/4$. Then, up to a subsequence, for a.e. $t \in \mathbb{R}$,*

$$\lim_{n \rightarrow +\infty} |U_n^{(t)} \Delta F^{(t)}| = 0, \quad \text{and} \quad \lim_{n \rightarrow +\infty} \partial U_n^{(t)} = \partial F^{(t)}, \quad (44)$$

where the last limit holds in the sense of Hausdorff convergence.

If additionally, $\|w_n\|_{L^2}/\lambda_n \rightarrow 0$ as $n \rightarrow +\infty$, the full sequence satisfies

$$\limsup_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \text{Ext}(Df). \quad (45)$$

Remark 6. It is possible to reformulate (45) in the following way. By Lemma 4, there exists $R > 0$ such that for all n , $\text{Supp}(Du_{\lambda_n, w_n}) \subseteq B(0, R)$ and $\text{Ext}(Df) \subseteq B(0, R)$ so that by [78, Thm. 4.10]), (45) is equivalent to

- (*outer limit inclusion*) for all $r > 0$, there exists $n_0 \in \mathbb{N}$ such that,

$$\forall n \geq n_0, \quad \text{Supp}(Du_{\lambda_n, w_n}) \subseteq T_r \stackrel{\text{def.}}{=} \{x \in \mathbb{R}^2; \text{dist}(x, \text{Ext}(Df)) \leq r\}.$$

- (*inner limit inclusion*) for all $r > 0$, there exists $n_1 \in \mathbb{N}$ such that,

$$\forall n \geq n_1, \quad \text{Supp}(Df) \subseteq \{x \in \mathbb{R}^2; \text{dist}(x, \text{Supp}(Du_{\lambda_n, w_n})) \leq r\}.$$

The second equation of (44) has a similar reformulation.

Remark 7. Note that while [47] considered low noise regimes in which $\lambda_n \rightarrow 0^+$ and $\|w_n\|_{L^2}\lambda_n \leq C$ for some well chosen constant C (for the problem of sparse spikes deconvolution), to obtain the strongest result for the problem of TV denoising, we require that $\|w_n\|_{L^2}/\lambda_n \rightarrow 0$ as $n \rightarrow +\infty$.

Proof. By Lemma 4, there exists some radius $R > 0$ such that for any $t \neq 0$, any $n \in \mathbb{N}$, the level set $U_n^{(t)}$ of u_n is included in $B(0, R)$ (since $(\lambda_n, w_n) \in D_{1, \sqrt{c_2}/4}$). The same also holds for the level sets $F^{(t)}$ of f . As a result, $\text{Supp}(u_n) \cup \text{Supp}(f) \subseteq \overline{B(0, R)}$ and the L^2 convergence of u_n towards f also implies its L^1 convergence. But by Fubini's theorem,

$$0 = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^2} |u_n - f| = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} |U_n^{(t)} \triangle F^{(t)}| dt,$$

so that, up to the extraction of a subsequence $(u_{n'})_{n' \in \mathbb{N}}$, for a.e. $t \in \mathbb{R}$, $\lim_{n \rightarrow +\infty} |U_{n'}^{(t)} \triangle F^{(t)}| = 0$.

Now let us fix such $t \in \mathbb{R}$, and such a subsequence $(u_{n'})_{n' \in \mathbb{N}}$. By L^1 convergence of $\mathbb{1}_{U_{n'}^{(t)}}$ towards $\mathbb{1}_{F^{(t)}}$, and the fact that $|D\mathbb{1}_{U_{n'}^{(t)}}|(\mathbb{R}^2) = P(U_{n'}^{(t)})$ is uniformly bounded (by Lemma (2)), the gradient $D\mathbb{1}_{U_{n'}^{(t)}}$ converges towards $D\mathbb{1}_{F^{(t)}}$ in the weak-* topology. Repeating the same argument as in Proposition 8 above, we obtain that

$$\partial F^{(t)} = \text{Supp}(D\mathbb{1}_{F^{(t)}}) \subseteq \liminf_{n' \rightarrow +\infty} \text{Supp}(D\mathbb{1}_{U_{n'}^{(t)}}) = \liminf_{n' \rightarrow +\infty} \partial U_{n'}^{(t)}.$$

Let us prove that $\limsup_{n' \rightarrow +\infty} \partial U_{n'}^{(t)} \subseteq \partial F^{(t)}$. If $\partial U_{n'}^{(t)} = \emptyset$ for all n' large enough, then $\limsup_{n' \rightarrow +\infty} \partial U_{n'}^{(t)} = \emptyset$ and there is nothing to prove. Otherwise, let $(x_{n'})_{n \in \mathbb{N}}$ such that $x_{n'} \in \partial U_{n'}^{(t)}$ and (up to the additional extraction of a subsequence - that we do not relabel) $\lim_{n' \rightarrow +\infty} x_{n'} = x \in \mathbb{R}^2$. By Proposition 7, for all $r \leq r_0$,

$$|B(x_{n'}, r) \cap U_{n'}^{(t)}| \geq \frac{1}{16} |B(x_{n'}, r)|, \quad \text{and} \quad |B(x_{n'}, r) \setminus U_{n'}^{(t)}| \geq \frac{1}{16} |B(x_{n'}, r)|.$$

By the dominated convergence theorem, we obtain for $n \rightarrow +\infty$,

$$|B(x, r) \cap F^{(t)}| \geq \frac{1}{16} |B(x, r)|, \quad \text{and} \quad |B(x, r) \setminus F^{(t)}| \geq \frac{1}{16} |B(x, r)|.$$

Since this holds for all $r \in (0, r_0]$, we see that $x \in \partial F^{(t)}$, hence $\limsup_{n' \rightarrow +\infty} \partial U_{n'}^{(t)} \subseteq \partial F^{(t)}$.

To prove $\limsup_{n \rightarrow +\infty} \text{Supp}(Du_n) \subseteq \text{Ext}(Df)$, we consider the full sequence again and we now assume that $\|w_n\|_{L^2}/\lambda_n \rightarrow 0$ as $n \rightarrow +\infty$. We denote by v_n the dual certificate v_{λ_n, w_n} . If $\text{Supp}(Du_n) = \emptyset$ for all n' large enough, there is nothing to prove. Otherwise, let $(x_n)_{n \in \mathbb{N}}$ such that $x_n \in \text{Supp}(Du_n)$ and (up to the extraction of a subsequence) $\lim_{n \rightarrow +\infty} x_n = x$ for some $x \in \mathbb{R}^2$. By Proposition 4, it is not restrictive to assume that $x_n \in \partial E_n$ for some $E_n \in \mathcal{F}_{\lambda_n, w_n}$ (otherwise we may replace x_n with $y_n \in \partial E_n$ such that $|x_n - y_n| \leq 1/n$).

By Lemma 2 and 4, the family $\{E_n\}_{n \in \mathbb{N}}$ is relatively compact in the L^1 topology (see [68, Thm. 12.26]), that is, there exists $E \subseteq \mathbb{R}^2$ with finite measure such that, up to the extraction of a subsequence, $\lim_{n \rightarrow +\infty} |E \triangle E_n| = 0$ (we do not relabel the subsequence). Moreover, up to the additional extraction of a subsequence, we may assume that either for all n , $\int_{E_n} v_n = P(E_n)$, or for all n , $\int_{E_n} v_n = -P(E_n)$. We deal with the first case, the other being similar.

Passing to the limit in the optimality equation for E_n , we get

$$P(E) \leq \liminf_{n \rightarrow +\infty} P(E_n) = \lim_{n \rightarrow +\infty} \int_{E_n} v_n = \int_E v_{0,0}, \quad (46)$$

by the lower semi-continuity of the perimeter, and since $\mathbb{1}_{E_n}$ (resp. v_{λ_n, w_n}) converges strongly in $L^2(\mathbb{R}^2)$ towards $\mathbb{1}_E$ (resp. $v_{0,0}$). Since $v_{0,0} \in \partial J(f) \subseteq \partial J(0)$, the converse inequality also holds, so that $P(E) = \int_E v_{0,0}$, and $E \in \mathcal{F}_{0,0}$. By definition of the extended support, this means that $\partial^* E \subseteq \text{Ext}(Df)$, hence $\partial E \subseteq \text{Ext}(Df)$.

Simarly as above, we conclude that $x \in \partial E$ in the following way. By Proposition 7, for all $r \leq r_0$,

$$|B(x_n, r) \cap E_n| \geq \frac{1}{16} |B(x_n, r)|, \quad \text{and} \quad |B(x_n, r) \setminus E_n| \geq \frac{1}{16} |B(x_n, r)|. \quad (47)$$

By the dominated convergence theorem, we obtain for $n \rightarrow +\infty$,

$$|B(x, r) \cap E| \geq \frac{1}{16} |B(x, r)|, \quad \text{and} \quad |B(x, r) \setminus E| \geq \frac{1}{16} |B(x, r)|.$$

Since this holds for all $r \in (0, r_0]$, we see that $x \in \partial E \subseteq \text{Ext}(Df)$. Hence $\limsup_{n \rightarrow +\infty} \text{Supp}(Du_n) \subseteq \text{Ext}(Df)$. \square

Remark 8 (On dimensions $N \geq 3$). There are two key elements to the proof of Theorem 1:

- (i) **Compactness.** Lemma 2 and 4 give the fact that there exists $R, L > 0$ such that $P(E_n) < L$ and the fact that there exists R such that $E_n \subset B(0, R)$. This allows the required compactness result to be applied.

- (ii) **Weak Regularity.** Proposition 7 which ensures that the boundaries of all level sets are uniformly weakly regular.

The difficulty with extending Theorem 1 to higher dimensions is that the second property of weak regularity is no longer true: In dimension N , for weak regularity, we would require that

$$\lim_{\lambda, \|w\|_{L^2}/\lambda \rightarrow 0} \|v_{\lambda,w} - v_{0,0}\|_{L^N} = 0.$$

However, the natural topology for $\{v_{\lambda,w}\}_{\lambda,w}$ is $L^2(\mathbb{R}^N)$ and when $N \geq 3$, there is no guarantee that the boundaries of the level sets of $u_{\lambda,w}$ do not have arbitrarily many singular points such as cusps, and it may be the case that there are level sets of $u_{\lambda,w}$ arbitrarily far out with arbitrarily small measure and perimeter.

When $N \geq 3$, it is still true that there exists L such that $P(E) \leq L$ for all $E \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{e_2}/4}$ and it is possible to adapt the argument in the proof of Theorem 1 to conclude that for each $r > 0$,

$$\lim_{(\lambda_0, \alpha_0) \rightarrow (0,0)} \sup \left\{ \mathcal{H}^1(\partial^* E \setminus T_r) ; E \in \mathcal{F}_{\lambda,w}, (\lambda, \alpha) \in D_{\alpha_0, \lambda_0} \right\} = 0.$$

However, we have no guarantee that there exists $\lambda_0, \alpha_0 > 0$ such that $\mathcal{H}^{N-1}(\partial^* E \setminus T_r) = 0$ for all $E \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{\lambda_0, \alpha_0}$.

As consequence of the support stability theorem, we obtain the following result on the minimal norm certificate.

Corollary 1. *Let Ω be any connected component of $\mathbb{R}^2 \setminus \text{Ext}(f)$. Then $v_{0,0}$ is constant on Ω .*

Proof. For $\delta > 0$, let $\Omega_\delta = \{x \in \Omega : \text{dist}(\partial\Omega, x) > \delta\}$. From Theorem 1, we know that for all δ , there exists $\lambda_\delta > 0$ such that for all $\lambda \in D_{\lambda_\delta, 0}$, $u_{\lambda,0}$ is constant on Ω_δ . Since $v_{\lambda,0} = (f - u_{\lambda,0})/\lambda$, it follows that $v_{\lambda,0}$ is also constant on Ω_δ . So, since $v_{0,0}$ is the L^2 limit of $v_{\lambda,0}$, $v_{0,0}$ must be constant on Ω_δ for all $\delta > 0$. Therefore, $v_{0,0}$ is constant on Ω . \square

8. Support stability for nonsmooth convex sets

The theory developed in Sections 5, 6 and 7 relies on the existence of a subgradient of the total variation for f in the L^2 topology (source condition). As natural as it may seem, this hypothesis does not always hold even for simple functions (like the indicator function of a square). In some cases, however, there is a natural limit for the dual certificates $v_{\lambda,0}$ when considering another topology.

This section studies the case of a union Ω of disjoint convex subsets of \mathbb{R}^2 which are sufficiently far apart. If their boundary is not smooth enough, the source condition is not satisfied. Still, we shall prove that one can guarantee support stability for the solutions of $(\mathcal{P}_\lambda(f + w))$. A notable example is the unit square where $\Omega = [0, 1]^2$.

As usual, throughout this section, we let $u_{\lambda,w}$ be the solution of $(\mathcal{P}_\lambda(f+w))$ and $v_{\lambda,w}$ be the solution of $(\mathcal{D}_\lambda(f+w))$. We also recall the notation from Section 4 where given any bounded open convex set C , C_ρ is the opening of C by open balls of radius ρ and there exists a unique function $r(x)$ such that $x \in \partial C_{r(x)}$ and x belongs to an arc of a circle of radius $r(x)$.

8.1. Dual certificates for unions of convex sets

Let C be a bounded open convex subset of \mathbb{R}^2 . The dual certificate $v_{\lambda,0}$ associated with $f = \mathbf{1}_C$ is given in (28).

Now, more generally, if $f = \mathbf{1}_\Omega$, where $\Omega = \cup_{j=1}^M C^{(j)}$ and $\{C^{(j)}\}_{j=1}^M$ are bounded open convex sets such that given any $0 \leq k \leq M$ and any permutation $\{i_1, \dots, i_M\}$ of $\{1, \dots, M\}$,

$$E_{i_1, \dots, i_k} \in \operatorname{argmin} \left\{ P(E) ; P(E) < \infty, \bigcup_{j=1}^k C^{(i_j)} \subset E \subset \mathbb{R}^2 \setminus \bigcup_{j=k+1}^M C^{(i_j)} \right\}$$

implies that $P(E_{i_1, \dots, i_k}) > \sum_{j=1}^k P(C^{(i_j)})$, then, as proved in [7, 9], the solution $u_{\lambda,0}$ to $(\mathcal{P}_\lambda(f))$ is

$$u_{\lambda,0} = \sum_{j=1}^M (1 + \lambda v_{C^{(j)}})^+ \mathbf{1}_{C^{(j)}},$$

and consequently,

$$v_{\lambda,0}(x) = \begin{cases} v_{C^{(j)}}(x) & x \in C_\lambda^{(j)}, j = 1, \dots, M \\ 1/\lambda & x \in C^{(j)} \setminus C_\lambda^{(j)}, j = 1, \dots, M \\ 0 & \text{otherwise.} \end{cases}$$

While $\lim_{\lambda \rightarrow 0} \|v_{\lambda,0}\|_{L^2} = +\infty$, we observe that the function

$$v_{0,0} \stackrel{\text{def.}}{=} \sum_{j=1}^M v_{C^{(j)}} \in L^1(\mathbb{R}^2).$$

Indeed, for each j , by the monotone convergence theorem

$$\|v_{C^{(j)}}\|_{L^1} = \int_{\mathbb{R}^2} v_{C^{(j)}} = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^2} v_{C^{(j)}} \mathbf{1}_{C_{1/n}^{(j)}} = \lim_{n \rightarrow +\infty} P(C_{1/n}^{(j)}) = P(C^{(j)}) < +\infty,$$

where $C_{1/n}^{(j)}$ denotes the opening of $C^{(j)}$ with radius $1/n$. Moreover, since $v_{\lambda,0} \geq v_{\mu,0}$ for $\mu \geq \lambda$ and since for a.e. $x \in \mathbb{R}^2$,

$$\lim_{\lambda \rightarrow 0} v_{\lambda,0}(x) = v_{0,0}(x),$$

it follows by the monotone convergence theorem that

$$\|v_{\lambda,0} - v_{0,0}\|_{L^1} \rightarrow 0, \quad \lambda \rightarrow 0.$$

As before, we may define the extended support of f via $v_{0,0}$ as

$$\begin{aligned} \text{Ext}(Df) &\stackrel{\text{def.}}{=} \overline{\left\{ \partial^* E ; \pm \int_E v_{0,0} = P(E), \quad |E| < \infty \right\}} \\ &= \Omega \setminus \left(\bigcup_{j=1}^M \text{int} \left(C_{R_j}^{(j)} \right) \right), \end{aligned} \quad (48)$$

where $C_{R_j}^{(j)}$ is the maximal calibrable set inside $C^{(j)}$ for each $j = 1, \dots, M$. We remark that $v_{\lambda,0} = v_{0,0}$ on $\mathbb{R}^2 \setminus \text{Ext}(Df)$.

Remark 9. In the limit case where equality may hold in $P(E_{i_1, \dots, i_k}) \geq \sum_{j=1}^k P(C^{(i_j)})$, the extended support of $\mathbf{1}_\Omega$ may be larger than $\partial\Omega$. The case $\Omega = B(x_1, R) \cup B(x_2, R)$ and $|x_1 - x_2| = \pi R$, is shown in Figure 2. In this case, if E is the convex hull of Ω , then $P(E) = P(\Omega)$. In the absence of noise, the support of any TV regularized solution is simply $\partial\Omega$, however, the extended support is strictly larger than $\partial\Omega$. This is essentially reflected in the fact that the presence of any noise which shifts the two balls towards each other will necessarily result in additional level lines. We refer to [5, 33] for a detailed study of this example.

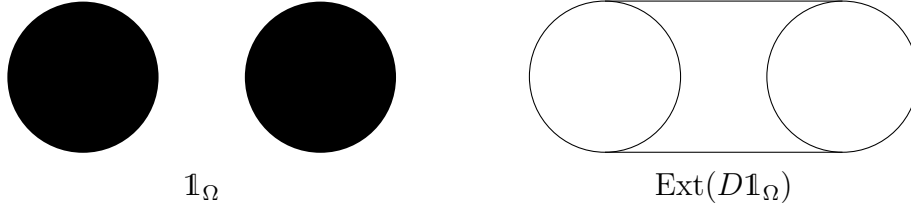


Figure 2: The extended support for the indicator function of a union of two balls.

8.2. Support stability

In this section, we prove that the support of $u_{\lambda,w}$ is stable around the extended support (48), i.e. its support is contained inside some neighborhood of $\text{Ext}(Df)$, whenever λ and $\|w\|_{L^2}/\lambda$ are sufficiently small. We begin by proving some properties of the level sets.

Proposition 9. *The following statements are true.*

- (i) *There exists $\alpha_0, \lambda_0, L > 0$ such that $P(E) \leq L$ for all $E \in \mathcal{F}_{\lambda,w}$ and $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$.*
- (ii) *There exists $R > 0$ such that $E_{\lambda,w} \subset B(0, R)$ for all $E_{\lambda,w} \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$.*

Proof. To prove (i), recall from the discussion in Section 8.1 that $\|v_{\lambda,0}\|_{L^2} \leq \|v_{0,0}\|_{L^1}$. So, for all $E \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$,

$$P(E) = \pm \int_E v_{\lambda,w} \leq \frac{\|w\|_{L^2} |E|^{1/2}}{\lambda} + \|v_{\lambda,0}\|_{L^1} \leq \frac{P(E)}{4} + \|v_{0,0}\|_{L^1}.$$

The proof of (ii) is very similar to the proof of Lemma 4. We first show that there exists $R > 0$ such that $E_{\lambda,w} \cap B(0, R) \neq \emptyset$ for all $E_{\lambda,w} \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$: let R be such that $B(0, R) \supset \Omega$. For a contradiction, suppose that $\emptyset \neq E_{\lambda,w} \subset B(0, R)^c$. Then, since $v_{\lambda,0} = 0$ on Ω^c , we have that

$$P(E_{\lambda,w}) = \int_{E_{\lambda,w}} (v_{\lambda,w} - v_{\lambda,0}) \leq \frac{\|w\|_{L^2} |E_{\lambda,w}|^{1/2}}{\lambda} < \frac{P(E_{\lambda,w})}{4},$$

which is impossible if $P(E_{\lambda,w}) > 0$. So, $E_{\lambda,w} \cap B(0, R) \neq \emptyset$ if $E_{\lambda,w} \neq \emptyset$. Finally, since by (i), there exists $L > 0$ such that $P(E_{\lambda,w}) \leq L$ for all $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$, it follows that $E_{\lambda,w} \subset B(0, R + L)$. □

When the source condition is not satisfied, Proposition 7 cannot be applied directly to the level sets of $u_{\lambda,w}$. However, even if the source condition does not hold, there may still be a subset of V of \mathbb{R}^2 for which

$$\lim_{\lambda \rightarrow 0} \|v_{\lambda,0} - v_{0,0}\|_{L^2(V)} = 0.$$

In this case, one can argue along the lines of Proposition 7 to deduce that there is still weak regularity on a subset of ∂E . Note that for characteristic functions on unions of convex sets as described in Section 8.1, we can let $V = \mathbb{R}^2 \setminus \text{Ext}(Df)$ since $v_{\lambda,0} = v_{0,0}$ on $\mathbb{R}^2 \setminus \text{Ext}(Df)$. The precise regularity statement is given in the following proposition.

Proposition 10. *Let $V \subset \mathbb{R}^2$ be an open set. Suppose that*

$$\lim_{\lambda \rightarrow 0} \|v_{\lambda,0} - v_{0,0}\|_{L^2(V)} = 0.$$

Then, there exists $r_0 > 0$ such that for all $r \in [0, r_0]$ and $E_{\lambda,w} \in \mathcal{F}_{\lambda,w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$, if $x \in \partial E_{\lambda,w}$ is such that $B(x, r_0) \subseteq V$, then

$$\frac{|B(x, r) \cap E_{\lambda,w}|}{|B(x, r)|} \geq \frac{1}{16} \quad \text{and} \quad \frac{|B(x, r) \setminus E_{\lambda,w}|}{|B(x, r)|} \geq \frac{1}{16}.$$

Theorem 2. *Let Ω be a union of convex sets which satisfies the assumptions of Section 8.1. Let $\{w_n\}_{n \in \mathbb{N}}$, $\{\lambda_n\}_{n \in \mathbb{N}}$ be sequences such that $w_n \in L^2(\mathbb{R}^2)$, $\lambda_n \rightarrow 0^+$, and $\|w_n\|_{L^2}/\lambda_n \leq \sqrt{c_2}/4$. Then, up to a subsequence, for a.e. $t \in \mathbb{R}$,*

$$\lim_{n \rightarrow +\infty} |U_n^{(t)} \Delta F^{(t)}| = 0, \quad \text{and} \tag{49}$$

$$\begin{cases} \lim_{n \rightarrow +\infty} \partial U_n^{(t)} = \partial F^{(t)}, & \text{if } 0 < t \leq 1 \\ \limsup_{n \rightarrow +\infty} \partial U_n^{(t)} \subseteq \partial \Omega & \text{otherwise,} \end{cases} \tag{50}$$

where the last limits holds in the sense of Hausdorff convergence.

If additionally, $\|w_n\|_{L^2}/\lambda_n \rightarrow 0$ as $n \rightarrow +\infty$, the full sequence satisfies

$$\text{Supp}(Df) \subseteq \liminf_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \limsup_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \text{Ext}(Df). \tag{51}$$

Proof. By Proposition 9 (ii), the level sets $U_n^{(t)}$ are included in some ball $B(0, R)$. So, by the same argument as in Theorem 1, $\partial F^{(t)} \subseteq \liminf_{n' \rightarrow +\infty} \partial U_{n'}^{(t)}$ and (49) holds.

Now, we prove $\limsup_{n' \rightarrow +\infty} \partial U_{n'}^{(t)} \subseteq \partial \Omega$. Let $(x_{n'})_{n' \in \mathbb{N}}$ be a sequence in \mathbb{R}^2 such that (up to an additional extraction) $x_{n'} \rightarrow x \in \mathbb{R}^2$, and we assume by contradiction that $x \notin \partial \Omega$. We let

$$V \stackrel{\text{def.}}{=} \{y \in \mathbb{R}^2 ; \text{dist}(y, \partial \Omega) > r_1\},$$

where $r_1 > 0$ is such that $r_1 < \text{dist}(x, \partial \Omega)$. We observe that V is open, $x \in V$ and $\lim_{n' \rightarrow +\infty} \|v_{n'} - v_{0,0}\|_{L^2(V)} = 0$.

Applying Proposition 10, we obtain for n' large enough and $r > 0$ small enough,

$$\left| B(x_{n'}, r) \cap U_{n'}^{(t)} \right| \geq \frac{1}{16} |B(x_{n'}, r)|, \quad \text{and} \quad \left| B(x_{n'}, r) \setminus U_{n'}^{(t)} \right| \geq \frac{1}{16} |B(x_{n'}, r)|,$$

and we conclude, passing to the limit as in the proof of Theorem 1 that $x \in \partial F^{(t)} \subseteq \partial \Omega$, which contradicts the hypothesis. Hence $x \in \partial \Omega$, and $\limsup_{n' \rightarrow +\infty} \partial U_{n'}^{(t)} \subseteq \partial \Omega$. Equation (50) follows since $\partial F^{(t)} = \partial \Omega$ for $0 < t \leq 1$.

It remains to prove $\limsup_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \text{Ext}(Df)$. The proof is quite similar to the proof presented for Theorem 1 and we merely sketch it for brevity. We let $x_n \rightarrow x \in \mathbb{R}^2$, where $x_n \in \partial E_n$ for some $E_n \in \mathcal{F}_{\lambda_n, w_n}$ and we assume by contradiction that $x \notin \text{Ext}(Df)$. Arguing as above, and using the compactness property provided by Proposition 9, we see that $x \in \partial E$ where E is the limit of E_n (up to an extraction, for the L^1 topology).

To conclude, we need to prove that $E \in \mathcal{F}_{0,0}$, as in (46). The L^2 convergence of v_{λ_n, w_n} towards $v_{0,0}$ was applied to prove equation (46), but in fact, L^1 convergence $v_{\lambda_n, 0}$ is sufficient: Note that by (i) of Proposition 9 and the isoperimetric inequality, there exists C such that $|E_n| \leq C^2$ for all n . Then by letting $E_n \triangle E = (E_n \setminus E) \cup (E \setminus E_n)$,

$$\begin{aligned} \left| \int_{E_n} v_{\lambda_n, w_n} - \int_E v_{0,0} \right| &\leq \left| \int_{E_n} v_{\lambda_n, w_n} - \int_{E_n} v_{0,0} \right| + \left| \int_{E_n} v_{0,0} - \int_E v_{0,0} \right| \\ &\leq \|v_{\lambda_n, 0} - v_{0,0}\|_{L^1} + \frac{C\|w_n\|_{L^2}}{\lambda_n} + \int_{E_n \triangle E} v_{0,0} \rightarrow 0, \end{aligned}$$

by the L^1 convergence of $(v_{\lambda_n, 0})$ and the absolute continuity of the Lebesgue integral. As a result, $E \in \mathcal{F}_{0,0}$, hence $x \in \partial E \subseteq \text{Ext}(Df)$, and $\limsup_{n \rightarrow +\infty} \text{Supp}(Du_{\lambda_n, w_n}) \subseteq \text{Ext}(Df)$. \square

9. Support stability and calibrations

Theorem 1 shows that as a result of the strong L^2 convergence of $v_{\lambda, w}$ to $v_{0,0}$, one is guaranteed support stability outside a small neighbourhood of the extended support. This section upper bounds the rate of convergence in the outer limit inclusion of (45). In particular, we make explicit the relationship between the width of this neighbourhood, the decay of $\|v_{\lambda, w} - v_{0,0}\|_{L^2}$ and the nondegeneracy of z_0 , the vector field for which $v_{0,0} = -\text{div } z_0$.

9.1. Support stability

In this section, we define

$$T_r \stackrel{\text{def.}}{=} \{x \in \mathbb{R}^2 ; \text{dist}(x, \text{Ext}(Df)) \leq r\},$$

and we make an additional assumption about the decay of z_0 away from the extended support:

$$\delta_r \stackrel{\text{def.}}{=} 1 - \text{ess sup}_{x \in T_r^C} |z_0(x)| > 0. \quad (52)$$

We also let C be such that

$$|E| \leq C^2, \quad \forall E \in \mathcal{F}_{\lambda, w}, \quad (\lambda, w) \in D_{1, \sqrt{c_2}/4}.$$

Recall that the existence of C is guaranteed by Lemma 2. Examples of vector fields whose decay is known outside the extended support are described in Section 9.2.

Proposition 11. *Given any $E \in \mathcal{F}_{\lambda, w}$ with $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$,*

$$\delta_r \mathcal{H}^1(\partial^* E \setminus T_r) \leq \int_E (v_{\lambda, w} - v_{0,0}) \leq C \|v_{\lambda, w} - v_{0,0}\|_{L^2},$$

Proof. Comparing the energy of E with that of the empty set we get,

$$\begin{aligned} P(E) &\leq \int_E v_{\lambda, w} = \int_E (v_{\lambda, w} - v_{0,0}) + \int_E v_{0,0} \\ &\leq \|v_{\lambda, w} - v_{0,0}\|_{L^2} \sqrt{|E|} + \int_{\partial^* E} z_0 \cdot \nu. \end{aligned}$$

Recall from Lemma 2 that there exists $C > 0$ such that $|E| \leq C^2$. So,

$$P(E) - \int_{\partial^* E \cap T_r} z_0 \cdot \nu - \int_{\partial^* E \setminus T_r} z_0 \cdot \nu \leq C \|v_{\lambda, w} - v_{0,0}\|_{L^2}.$$

Since $(\text{ess sup}_{\mathbb{R}^N \setminus T_r} |z_0|) \leq 1 - \delta_r$, and more generally $\|z_0\|_{L^\infty} \leq 1$,

$$\begin{aligned} P(E) - \int_{\partial^* E \cap T_r} z_0 \cdot \nu - \int_{\partial^* E \setminus T_r} z_0 \cdot \nu &\geq \mathcal{H}^1(\partial^* E \setminus T_r) + \mathcal{H}^1(\partial^* E \cap T_r) \\ &\quad - (1 - \delta_r) \mathcal{H}^1(\partial^* E \setminus T_r) - \mathcal{H}^1(\partial^* E \cap T_r) \\ &\geq \delta_r (\mathcal{H}^1(\partial^* E \setminus T_r)), \end{aligned}$$

hence the claimed result. □

Proposition 12. *Let $\lambda > 0$ and $w \in L^2(\mathbb{R}^2)$ be such that $\|v_{\lambda, w} - v_0\|_{L^2} < \delta_{r/2} \sqrt{c_2}$. Then, $E \in \mathcal{F}_{\lambda, w}$ and $P(E) > 0$ implies that*

$$\mathcal{H}^1(\partial^* E \cap T_{r/2}) > 0.$$

Proof. Assume by contradiction that $P(E) > 0$ and $\mathcal{H}^1(\partial^*E \cap T_{r/2}) = 0$ so that $\partial^*E \subset T_{r/2}^c$ up to an \mathcal{H}^1 -negligible set. Then,

$$\begin{aligned} P(E) &= \int_E v_{\lambda,w} \leq \|E\|^{1/2} \|v_{\lambda,w} - v_{0,0}\|_{L^2} + \int_E v_{0,0} \\ &\leq \frac{P(E) \|v_{\lambda,w} - v_{0,0}\|_{L^2}}{\sqrt{c_2}} + (1 - \delta_{r/2}) P(E), \end{aligned}$$

where we have applied the isoperimetric inequality and the fact that $v_{0,0} = \operatorname{div} z_0$ with $|z_0| \leq (1 - \delta_{r/2})$ on $T_{r/2}^c$. Since $P(E) > 0$, this implies that

$$\delta_{r/2} \sqrt{c_2} \leq \|v_{\lambda,w} - v_{0,0}\|_{L^2}.$$

This contradicts the assumption of this proposition. \square

Theorem 3. *Let $r > 0$. If $(\lambda, w) \in D_{1, \sqrt{c_2}/4}$ are such that*

$$\|v_{\lambda,w} - v_0\|_{L^2} \leq \delta_{r/2} \min \left\{ \frac{r}{2C}, \sqrt{c_2} \right\}, \quad (53)$$

then for all level sets E of $u_{\lambda,w}$, $\partial E \subseteq T_r$.

Proof. It is sufficient to show that $\mathcal{H}^1(\partial^*E \setminus T_r) = 0$ for all $E \in \mathcal{F}_{\lambda,w}$ with $\lambda > 0$ and $w \in L^2(\mathbb{R}^2)$ satisfying (53). For, if we have $\mathcal{H}^1(\partial^*E \setminus T_r) = 0$, this means that $\mathbf{1}_E$ is constant on every connected component of the open set $\mathbb{R}^2 \setminus T_r$, hence the topological boundary satisfies $\partial E \subseteq T_r$. Furthermore, by Section 3.7, we may assume that up to an \mathcal{H}^1 -negligible set, ∂^*E is equivalent to a Jordan curve γ .

First observe that by Proposition 12, $\mathcal{H}^1(\partial^*E \cap T_{r/2}) > 0$. Now, for a contradiction, suppose that $\mathcal{H}^1(\partial^*E \setminus T_r) > 0$. Then since this implies that $\mathcal{H}^1(\gamma \cap T_{r/2}) > 0$, $\mathcal{H}^1(\gamma \setminus T_r) > 0$ and γ is a continuous curve, it follows that

$$\mathcal{H}^1(\gamma \setminus T_{r/2}) = \mathcal{H}^1(\partial^*E \setminus T_{r/2}) > r/2.$$

However, this is a contradiction Proposition 11 implies that

$$\lim_{(\lambda_0, \alpha_0) \rightarrow (0,0)} \sup \left\{ \mathcal{H}^1(\partial^*E \setminus T_r) ; E \in \mathcal{F}_{\lambda,w}, (\lambda, \alpha) \in D_{\alpha_0, \lambda_0} \right\} = 0.$$

Indeed, by our choice of (λ, w) in (53), if $\mathcal{H}^1(\partial^*E \setminus T_r) > 0$, then the combination of Proposition 11 and Proposition 12 yields

$$\frac{r\delta_{r/2}}{2} < \mathcal{H}^1(\partial^*E \setminus T_{r/2}) \leq C \|v_{\lambda,w} - v_{0,0}\|_{L^2} \leq \frac{r\delta_{r/2}}{2}.$$

\square

Example In the case where $f = \mathbf{1}_{B(0,R)}$, by the construction of z_0 from (13), $\delta_r \leq r/R$. Furthermore, since $v_{\lambda,w} = \frac{2}{R}f$, for each $E \in \mathcal{F}_{\lambda,w}$,

$$2\sqrt{\pi} |E|^{1/2} \leq P(E) \leq \int_E v_{\lambda,w} \leq 2\pi R + \|w\|_{L^2} |E|^{1/2},$$

and $|E|^{1/2} \leq 2\pi R / (2\sqrt{\pi} - \|w\|_{L^2})$ provided that $2\sqrt{\pi} > \|w\|_{L^2}$. So, Theorem 3 implies that for all $\lambda > 0$, and $w \in L^2(\mathbb{R}^2)$ such that

$$\|w\|_{L^2} \leq \min \left\{ 2\sqrt{\pi} - \pi, \frac{\lambda r^2}{8R^2} \right\},$$

any level set E of $u_{\lambda,w}$ satisfies $\partial^* E \subset T_r$ up to an \mathcal{H}^1 -negligible set.

9.2. Non-degeneracy of calibrable sets

The aim of this section is to show that if $C \subset \mathbb{R}^2$ is a convex calibrable set, the minimal norm certificate $v_{0,0} = h_C \mathbf{1}_C$ (where $h_C = \frac{P(C)}{|C|}$) can be written as $v_{0,0} = \operatorname{div} z_0$ where $z_0 \in X_2(\mathbb{R}^2)$, $(z, D\mathbf{1}_C) = -|D\mathbf{1}_C|$ and for every compact set $K \subset \mathbb{R}^2 \setminus \partial C$,

$$\operatorname{ess\,sup}_K |z_0| < 1,$$

with an estimation on that inequality. We do not aim at full generality, and we assume that ∂C is of class \mathcal{C}^2 for the sake of simplicity. Reducing the hypotheses is the subject of future work.

The proof relies on the notion of inner and outer calibrations described in [20], which amounts to constructing vector fields “inside” and “outside” the studied set, and then “glue” the two constructions.

Definition 2. Let $C \subseteq \mathbb{R}^2$ be a set of finite perimeter. We say that C is $-$ calibrable if there exists a vector field $z_C^- : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that

- (i) $z_C^- \in L_{loc}^2(\mathbb{R}^2, \mathbb{R}^2)$ and $\operatorname{div} z_C^- \in L_{loc}^2(\mathbb{R}^2)$;
- (ii) $|z_C^-| \leq 1$ almost everywhere in C ;
- (iii) $\operatorname{div} z_C^-$ is constant on C ;
- (iv) $\theta(z^-, -D\mathbf{1}_C)(x) = -1$ for \mathcal{H}^1 -almost every $x \in \partial^* C$.

Similarly C is $+$ calibrable if 1), 2), 3) hold and $\theta(z^+, -D\mathbf{1}_C)(x) = +1$ in 4).

The following lemma tells that one may “glue” calibrations:

Lemma 5 ([20]). Let C be a bounded set of finite perimeter. Then $v = \mathbf{1}_C$ is calibrable if and only if C is $-$ calibrable with $-\operatorname{div} \xi_C^- = h_C$ in C , and $\mathbb{R}^2 \setminus C$ is $+$ calibrable with $\operatorname{div} z_{\mathbb{R}^2}^+ = 0$ in $\mathbb{R}^2 \setminus C$, defining

$$z \stackrel{\text{def.}}{=} \begin{cases} z_C^- & \text{on } C, \\ z_C^+ & \text{on } \mathbb{R}^2 \setminus C. \end{cases}$$

9.3. Inner calibrations

Let $C \subset \mathbb{R}^2$ be a bounded open convex set of class \mathcal{C}^2 , and $h_C \stackrel{\text{def.}}{=} \frac{P(C)}{|C|}$. Following [10] in order to build the calibration, we consider the following auxiliary problem:

$$\operatorname{div} \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = h_C \quad (54)$$

and we define

$$z \stackrel{\text{def.}}{=} \begin{cases} \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} & \text{on } C \\ \nu_C & \text{on } \partial C. \end{cases} \quad (55)$$

Giusti proved the following result in [53] (see also [10, Prop.6.2])

Theorem 4 ([53]). *There exists a solution $u \in \mathcal{C}^2(C)$ to (54) if and only if*

$$\forall B \subsetneq C, B \neq \emptyset, \quad h_C < \frac{P(B)}{|B|}. \quad (56)$$

Furthermore, the solution u is unique up to an additive constant, bounded from below in C , and its graph is vertical at the boundary of C , in the sense that

$$\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \rightarrow \nu_C \text{ uniformly on } \partial C.$$

The consequence is that z defined in (55) is a \mathcal{C}^1 vector field in C , (in fact analytic, see [10]), continuous in \overline{C} .

In fact, Giusti also proved that the condition (56) is equivalent to (16), namely the calibrability of C (this result was extended to \mathbb{R}^N in [6]). As a result, for a calibrable set C , one may choose the calibration given by the vectorfield z such that

$$\forall x \in C, \quad z(x) = \frac{\nabla u(x)}{\sqrt{1 + |\nabla u(x)|^2}}$$

and $z|_{\mathbb{R}^2 \setminus C}$ is a vectorfield such that $\|z\|_\infty \leq 1$, $\operatorname{div} z = h_C$ and $\theta(z, D\mathbf{1}_C) = -1$. A first step in proving that $|z| < 1$ inside C is the following theorem by Giusti.

Theorem 5 ([53]). *For every compact set $K \subset \operatorname{int} C$, there exist exists $Q > 0$ such that for any solution of (54) in $\operatorname{int} C$,*

$$\sup_K |\nabla u| \leq Q.$$

This implies that $\sup_K \frac{|\nabla u|}{\sqrt{1 + |\nabla u|^2}} < 1$. In the next proposition, we study further its decay inside C , which yields a non-degenerate inner calibration for C .

Proposition 13. *Let $C \subset \mathbb{R}^2$ be a bounded strictly convex calibrable set such that ∂C is of class \mathcal{C}^2 and $h_C > \sup_{\partial C} |\kappa_{\partial C}|$. Assume moreover that the solution to (54) is continuous up to the boundary, i.e. $u \in \mathcal{C}(\overline{C})$. Then, there exists $\alpha > 0$, there exists a vector field $z \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^1(\Omega)$ such that $\operatorname{div} z = h_C$, $z \cdot \nu = 1$ on ∂C , and*

$$\forall x \in C, \quad |z(x)| \leq \frac{\alpha}{\sqrt{d(x)^2 + \alpha^2}},$$

where $d(x) \stackrel{\text{def.}}{=} \operatorname{dist}(x, \partial C)$.

Proof. By Theorem 4, there exists a \mathcal{C}^2 solution u to (54) which is vertical at the boundary, and the inequality $h_C > \operatorname{ess\,sup}_{\partial C} |\kappa_{\partial C}|$ implies that u is bounded (see [53, Th. 3.1]). We define $z(x) \stackrel{\text{def.}}{=} \frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}}$ for all $x \in C$.

Let us prove that $|\nabla u(x)| > 0$ for a.e. $x \in C$. First, we assume that C is strictly convex. Since $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$, by [67, Th. 2.2] u is a convex function. As a result, $\{x \in C ; \nabla u(x) = 0\} = \operatorname{argmin}_C u$, and it is thus a closed convex set. Assume by contradiction that the dimension of $\operatorname{argmin}_C u$ is 2, i.e. $\operatorname{argmin}_C u$ contains an open ball $B(x_0, r) \subset \operatorname{int}(C)$ for some $x_0 \in \operatorname{int}(C)$, $r > 0$. Let T denote the operator $T : u \mapsto \frac{\nabla u}{\sqrt{1+|\nabla u|^2}}$, and let w be the constant function $x \mapsto \min_C u$. We have $u \leq w$ in $\partial B(x_0, r)$ (in fact equality holds), and $0 = \operatorname{div} Tw < \operatorname{div} Tu = h_C$ in $B(x_0, r)$. By Theorem 5, Problem (54) is locally uniformly elliptic, and the comparison principle [51, Th. 10.1] yields that $u < w$ in $B(x_0, r)$, which is a contradiction. As a result, the dimension of $\operatorname{argmin}_C u$ is strictly less than 2 and $\operatorname{argmin}_C u$ is Lebesgue-negligible.

Now, for a.e. $x \in \operatorname{int}(C)$, we may define $y \stackrel{\text{def.}}{=} x + d(x) \frac{\nabla u(x)}{|\nabla u(x)|}$, and we observe that $y \in C$. By convexity of u , $u(y) - u(x) \geq \nabla u(x) \cdot (y - x) = d(x) |\nabla u(x)|$. As a result, $|\nabla u(x)| \leq \frac{2\|u\|_\infty}{d(x)}$, and

$$|Tu(x)| \leq \frac{2\|u\|_\infty}{\sqrt{d(x)^2 + 4\|u\|_\infty^2}}.$$

The claimed result holds by a density argument. \square

9.4. Outer calibrations

It is proved in [20, Th. 5] (see also [7, Th. 13] in dimension N) that sets which satisfy a geometric condition (namely convex sets that are far enough from one another) have a complement which is +calibrable. That condition holds for $\mathcal{C}^{1,1}$ convex sets.

However, it is not clear from the proof that the corresponding vector field has norm < 1 in compact sets of $\mathbb{R}^2 \setminus C$. We provide below an explicit construction when the set has \mathcal{C}^2 boundary. Admittedly the hypothesis is quite restrictive but we think that this construction gives some insight on the geometric properties involved.

Proposition 14. *Let $C \subset \mathbb{R}^2$ be a nonempty bounded open convex subset with \mathcal{C}^2 boundary. There exists a vector field $z \in L^\infty \cap \mathcal{C}(\mathbb{R}^2 \setminus \overline{C})$ such that $z = \nu$ on ∂C , $\operatorname{div} z = 0$ in the sense of distributions and $|z| < 1$ on every compact subset of $\mathbb{R}^2 \setminus C$.*

The decay of z is discussed in Remark 10 below.

Proof. We choose an arclength parametrization of ∂C , $s \mapsto y(s)$ defined on $S \stackrel{\text{def.}}{=} \mathbb{R}/(P(C)\mathbb{Z})$, and we consider a basis $(\tau(s), \nu(s))$ such that $\tau(s) = y'(s)$, $\nu(s) = R_{-\pi/2}\tau(s)$, where $R_{-\pi/2}$ the rotation with angle $-\pi/2$. We assume that the parametrization is such that $\nu(s)$ is the outer unit normal to C .

The mapping $\varphi : (s, d) \mapsto y(s) + d\nu(s)$ is a \mathcal{C}^1 -diffeomorphism from $S \times \mathbb{R}_+^*$ onto $\mathbb{R}^2 \setminus \overline{C}$, with

$$\frac{\partial \varphi}{\partial s}(s, d) = \tau(s) + d\kappa(s)\tau(s), \quad \text{and} \quad \frac{\partial \varphi}{\partial d}(s, d) = \nu(s),$$

where $\kappa(s) \geq 0$ is the curvature of ∂C at $y(s)$.

In order to define a vector field $\bar{z} : \mathbb{R}^2 \setminus \overline{C} \rightarrow \mathbb{R}^2$ such that $\text{div } \bar{z} = 0$, it is sufficient to define a vector field $z : S \times \mathbb{R}_+^* \rightarrow \mathbb{R}^2$ such that $\bar{z}(x) = z(\varphi^{-1}(x))$ and

$$\text{Tr}(DzD\varphi^{-1}) = 0.$$

In other words, we shall build a vector field z such that

$$\frac{1}{1 + \kappa(s)d} \partial_s z_1(s, d) + \frac{\kappa(s)}{1 + \kappa(s)d} z_2(s, d) + \partial_d z_2(s, d) = 0. \quad (57)$$

Here, for the sake of brevity, we have denoted by ∂_s (resp. ∂_d) the derivatives with respect to s (resp. d), and by (z_1, z_2) the coordinates of z in the basis $(\tau(s), \nu(s))$.

Given $\alpha > 0$ (to be fixed later), and the function $\eta : t \mapsto \min(t, 2 - t)$, we define

$$z_1(s, d) = -\alpha \left(\int_0^s (\kappa(s') - \frac{2\pi}{P(C)}) ds' \right) \eta(d), \quad (58)$$

$$z_2(s, d) = \frac{1}{1 + \kappa(s)d} \left(1 + \alpha \left(\int_0^d \eta \right) \left(\kappa(s) - \frac{2\pi}{P(C)} \right) \right). \quad (59)$$

Observe that $\lim_{(s,d) \rightarrow (s_0,0)} z(s, d) = \nu(s_0)$, and that z is continuous in $\mathbb{R}^2 \setminus \overline{C}$ since $\int_0^{P(C)} \kappa(s') ds' = 2\pi$. Moreover, it is not difficult to check that z satisfies (57) as well. As a result, $\text{div } \bar{z} = 0$ pointwise in $\mathbb{R}^2 \setminus \overline{C}$, and since \bar{z} is continuous we see by approximation that it also holds in the sense of distributions.

It remains to prove that $|z|^2 - 1 < 0$.

$$\begin{aligned} z_1^2 + z_2^2 - 1 &= \alpha^2 \left(\int_0^s (\kappa - \frac{2\pi}{P(C)}) \right)^2 \eta^2 \\ &\quad + \frac{1}{(1 + \kappa d)^2} \left[\left(1 + \alpha \left(\int_0^d \eta \right) \left(\kappa - \frac{2\pi}{P(C)} \right) \right)^2 - 1 - (\kappa d)^2 - 2\kappa d \right]. \end{aligned}$$

There is a constant $M > 0$ which only depends on $\sup_{\partial C} \kappa$ and $P(C)$ such that $\left(\int_0^s (\kappa - \frac{2\pi}{P(C)}) \right)^2 \leq M$ and $\left| \kappa - \frac{2\pi}{P(C)} \right|^2 \leq M$.

The term inside brackets is equal to

$$\begin{aligned}
& \alpha^2 \left(\int_0^d \eta \right)^2 \left(\kappa - \frac{2\pi}{P(C)} \right)^2 + 2\alpha \left(\int_0^d \eta \right) \left(\kappa - \frac{2\pi}{P(C)} \right) - (\kappa d)^2 - 2\kappa d \\
& \leq \alpha^2 M \left(\int_0^d \eta \right)^2 + 2\kappa \left(\alpha \left(\int_0^d \eta \right) - d \right) - \frac{4\pi\alpha}{P(C)} \int_0^d \eta - (\kappa d)^2 \\
& \leq 2\kappa \left(\alpha \left(\int_0^d \eta \right) - d \right) + \left(\alpha^2 M - \frac{4\pi\alpha}{P(C)} \right) \int_0^d \eta - (\kappa d)^2,
\end{aligned}$$

since $\int_0^d \eta \leq 1$. Hence, for $d \leq 1$, we obtain that for α small enough (depending on M and $P(C)$), that term is less than or equal to

$$-\kappa d - \frac{2\pi\alpha}{P(C)} \frac{d^2}{2} - (\kappa d)^2 \leq 0,$$

which yields (writing $K \stackrel{\text{def.}}{=} \sup_{\partial C} \kappa$)

$$\begin{aligned}
z_1^2 + z_2^2 - 1 & \leq \alpha^2 M d^2 + \frac{1}{1+K} \left(-\kappa d - \frac{2\pi\alpha}{P(C)} \frac{d^2}{2} - (\kappa d)^2 \right) \\
& \leq -\frac{1}{1+K} \left(\kappa d + \frac{\pi\alpha}{P(C)} \frac{d^2}{2} + (\kappa d)^2 \right) < 0,
\end{aligned} \tag{60}$$

for $\alpha > 0$ small enough (depending on M , K and $P(C)$).

As for $d > 1$, we may assume that α is small enough so that $\alpha \int_0^{+\infty} \eta \leq 1/2 \leq d/2$. Moreover, $\int_0^d \eta \geq \int_0^1 \eta = 1/2$, so that the term inside brackets is less than or equal to

$$-\kappa d - \frac{\pi\alpha}{P(C)} - (\kappa d)^2.$$

So,

$$\begin{aligned}
z_1^2 + z_2^2 - 1 & \leq \alpha^2 M - \frac{1}{1+\kappa d} \left(\kappa d + \frac{\pi\alpha}{P(C)} + (\kappa d)^2 \right) \\
& \leq \alpha^2 M - \frac{1}{1+\kappa d} \left(\kappa d + \frac{\pi\alpha}{P(C)} \right).
\end{aligned}$$

For $a \stackrel{\text{def.}}{=} \frac{\pi\alpha}{P(C)} < 1$, the mapping $x \mapsto -\frac{x+a}{x+1}$ is (strictly) decreasing on $[0, +\infty)$, hence upper bounded by $-a$, and we obtain that $z_1^2 + z_2^2 - 1 \leq \alpha^2 M - \frac{\pi\alpha}{P(C)} < 0$ for α small enough. \square

Remark 10. A more straightforward construction would have been to construct z parallel to the normals to C , or equivalently set $\alpha = 0$ in (58) and (59). However, such a vector field would not decay in front of flat areas, where $\kappa(s) = 0$, and we would have $|z(s, d)| = 1$ for all $d > 0$. The above construction “twists” the field lines so as to obtain some decay of the norm.

Still, the resulting upper bound (60) for small d depends on the local curvature of ∂C . If $\kappa(s) > 0$, then, as $d \rightarrow 0^+$,

$$|z(s, d)|^2 \leq 1 - \frac{\kappa(s)}{1+K}d + o(d).$$

On the other hand, if $\kappa(s) = 0$, then

$$|z(s, d)|^2 \leq 1 - \frac{\pi\alpha}{(1+K)P(C)} \frac{d^2}{2}.$$

10. Numerical Illustrations

In order to illustrate our theoretical findings, we have performed numerical computations on a discretized version of the denoising problem $(\mathcal{P}_\lambda(f))$. Let us stress that this section does not provide any theoretical guarantees concerning the geometrical faithfulness of these approximations, and a careful study of the impact of discretization is an interesting avenue for future works. The code to reproduce these results can be found online§.

10.1. Problem discretization

The problem is discretized on an uniform grid $((i/n, j/n))_{i,j=0}^{n-1}$ of n^2 points in $[0, 1]^2$. For simplicity, we use periodic boundary conditions. The input image f is represented on this grid as $(f_{i,j})_{i,j=1}^n$ and is normalized so that $f_{i,j} \in [0, 1]$. The recovered image $(u_{i,j})_{i,j}$ is defined on the same grid. Denoting $y \stackrel{\text{def}}{=} f + w$, the problem $\mathcal{P}_\lambda(y)$ is then approximated as

$$\min_{u \in \mathbb{R}^{n^2}} \sum_{i,j=1}^n |u_{i,j} - y_{i,j}|^2 + \lambda \sum_{i,j} \|\nabla_{i,j} u\| \quad (\mathcal{P}_\lambda^d(y))$$

where $\|\cdot\|$ is the Euclidean norm in \mathbb{R}^4 . Here, $\nabla_{i,j} u \in \mathbb{R}^4$ is a 4-fold discretization of the gradient operator, defined as

$$\nabla_{i,j} u = n(u_{i+1,j} - u_{i,j}, u_{i,j+1} - u_{i,j}, u_{i+1,j+1} - u_{i+1,j}, u_{i+1,j+1} - u_{i,j+1}) \in \mathbb{R}^4$$

so that the discrete gradient operator is $\nabla : u \in \mathbb{R}^{n^2} \mapsto \nabla u \in \mathbb{R}^{n^2 \times 4}$. We also define the discrete divergence as

$$\text{div} \stackrel{\text{def}}{=} -\nabla^\top : \mathbb{R}^{n^2 \times 4} \rightarrow \mathbb{R}^{n^2}.$$

Note that this differs from the more usual forward finite-difference approximation (used for instance in [35]), and we found numerically that this improves the isotropy (rotation invariance) of the scheme.

§ <https://github.com/gpeyre/2016-IP-tv-denoising/>

10.2. Discrete dual problem and iterative algorithm

We solve the finite dimensional convex optimization problem $(\mathcal{P}_\lambda^d(y))$ using the dual projected gradient descent of [35]. It minimizes a discrete counterpart of $(\mathcal{D}_\lambda(f))$, which reads

$$\min_{z \in \mathbb{R}^{n^2 \times 4}} \left\{ \left\| \frac{y}{\lambda} + \operatorname{div}(z) \right\|_{\ell^2} ; z \in \mathcal{C}_\infty \right\} \quad (\mathcal{D}_\lambda^d(y))$$

where $\mathcal{C}_\infty \stackrel{\text{def.}}{=} \left\{ z \in \mathbb{R}^{n^2 \times 4} ; \forall (i, j), \|z_{i,j}\| \leq 1 \right\}.$

The solutions z_λ of $(\mathcal{D}_\lambda^d(y))$ are in general non-unique because the problem is not strictly convex, but the primal-dual relationship allow one to recover the unique solution u_λ of the primal problem $(\mathcal{P}_\lambda^d(y))$ as

$$u_\lambda \stackrel{\text{def.}}{=} \operatorname{div}(z_\lambda) + \frac{y}{\lambda}.$$

Starting by some initial $z^{(0)} \in \mathbb{R}^{n^2 \times 4}$, the projected gradient descent reads

$$z^{(\ell+1)} \stackrel{\text{def.}}{=} \operatorname{Proj}_{\mathcal{C}_\infty} \left(z^{(\ell)} + \tau \nabla(\operatorname{div}(z) + \frac{y}{\lambda}) \right),$$

where the step size τ should satisfy $\tau < 2/\|\nabla\|^2$ where $\|\nabla\|^2 \leq 16n^2$ is the operator norm of the gradient. The orthogonal projection on \mathcal{C}_∞ is computed as

$$\tilde{z} = \operatorname{Proj}_{\mathcal{C}_\infty}(z) \quad \text{where} \quad \forall (i, j), \quad \tilde{z}_{i,j} = \frac{z_{i,j}}{\max(\|z_{i,j}\|, 1)}.$$

The $z^{(\ell)}$ converge toward a solution z_λ of $(\mathcal{D}_\lambda^d(y))$, while the primal iterates

$$u^{(\ell)} \stackrel{\text{def.}}{=} \operatorname{div}(z^{(\ell)}) + \frac{y}{\lambda}$$

converge toward the unique solution u_λ of $(\mathcal{P}_\lambda^d(y))$ with a speed $\|u_\lambda - u^{(\ell)}\| = O(1/\ell)$ as shown in [50].

10.3. Denoising results

Figure 3 displays the solution u_λ of $(\mathcal{P}_\lambda^d(y))$ for a set of increasing values of λ . We use here $n = 512$, and the noise w is a realization of a white noise where each pixel is Gaussian distributed with a variance σ^2 for $\sigma = 0.2$. As predicted by Theorem 1, this shows how the level sets of the solution progressively clusters around the extended support $\operatorname{Ext}(Df)$ as λ increases. In order to get some insight about the geometry of this extended support, we display the saturation points of $\|z_0\| = (\|z_{0,i,j}\|)_{i,j}$. Note that since z_0 is non-unique and we used the one output by the discrete minimization scheme, we do not claim any theoretical guarantee about this procedure. In practice, we observed that starting the algorithm with $z^{(0)} = 0$ leads to meaningful results about this extended support, as shown on Figure 3.



Figure 3: Display of the discretized solution u_λ of the discretized problem $(\mathcal{P}_\lambda^d(y))$ for several value of λ . The blue curves on top u_λ of indicate the level sets of u_λ (computed using bilinear interpolation on the grid). The blue curves on top of $\|z_0\|$ indicate the obtained approximation of the boundary of the extended support $\text{Ext}(Df)$.

Conclusion

In this paper, we have characterized the regions in which the solutions to the two-dimensional TV denoising problem are geometrically stable under L^2 additive noise. In particular, via the minimal norm certificate, we introduced the notion of an extended support and although the support of TV regularized solutions are in general not stable,

we have proved that the support instabilities are confined to a neighbourhood of the extended support. We have also provided explicit examples of the extended support in the case of indicators of convex sets. Within the low noise regime, for the indicator set of a calibrable set C , the support of the solutions was shown to cluster around ∂C . While for indicator functions of general convex sets (including convex sets for which the source condition is not satisfied), the support of the solutions was shown to cluster around the domain $C \setminus \text{int}(C_{R^*})$, where C_{R^*} is the maximal calibrable set inside C .

Acknowledgements

The authors thank Matteo Novaga for enlightening discussions about calibrations and the capillary problem.

Antonin Chambolle was partially supported by the ANR, programs ANR-12-BS01-0014-01 “GEOMETRYA” and ANR-12-IS01-0003 “EANOI” (joint with FWF No. I1148). He acknowledges the hospitality of DAMTP and Churchill College (U. Cambridge) for the year 2015-2016. Vincent Duval and Clarice Poon acknowledge support from the CNRS (Défi Imag’in de la Mission pour l’Interdisciplinarité, project CAVALIERI). The work of Gabriel Peyré has been supported by the European Research Council (ERC project SIGMA-Vision). Clarice Poon acknowledges support from the Fondation Sciences Mathématiques de Paris.

References

- [1] R. Acar and C.R. Vogel. “Analysis of bounded variation penalty methods for ill-posed problems”. In: *Inverse Problems* 10 (1994), 1217–1229.
- [2] M. Aharon, M. Elad, and A. Bruckstein. “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation”. In: *Signal Processing, IEEE Transactions on* 54.11 (2006), pp. 4311–4322.
- [3] W.K. Allard. “Total variation regularization for image denoising I: Geometric theory”. In: *SIAM Journal on Mathematical Analysis* 39.4 (2007), pp. 1150–1190.
- [4] W.K. Allard. “Total variation regularization for image denoising II: Examples”. In: *SIAM Journal on Imaging Sciences* 1.4 (2008), pp. 400–417.
- [5] W.K. Allard. “Total variation regularization for image denoising III: Examples”. In: *SIAM Journal on Imaging Sciences* 2.2 (2009), pp. 532–568.
- [6] F. Alter and V. Caselles. “Uniqueness of the Cheeger set of a convex body”. In: *Nonlinear Analysis: Theory, Methods & Applications* 70.1 (2009), pp. 32–44.
- [7] F. Alter, V. Caselles, and A. Chambolle. “A characterization of convex calibrable sets in R^N ”. In: *Math. Ann.* 7 (2005), pp. 29–53.
- [8] F. Alter, V. Caselles, and A. Chambolle. “Evolution of characteristic functions of convex sets in the plane by the minimizing total variation flow”. In: *Interfaces and free boundaries* 7.1 (2005), p. 29.

- [9] F. Alter, V. Caselles, and A. Chambolle. “Evolution of Convex Sets in the Plane by Minimizing the Total Variation Flow”. In: *Interfaces and Free Boundaries* 332 (2005), pp. 329–366.
- [10] S. Amato, G. Bellettini, and L. Tealdi. *Anisotropic mean curvature on facets and relations with capillarity*. 2015.
- [11] L. Ambrosio. *Corso introduttivo alla teoria geometrica della misura ed alle superfici minime*. Quad. Scuola Norm. Sup. Pisa. Pantograph, Genova, 1997.
- [12] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford mathematical monographs. Oxford University Press, 2000.
- [13] L. Ambrosio, V. Caselles, S. Masnou, and J-M. Morel. “Connected Components of Sets of Finite Perimeter and Applications to Image Processing”. In: *Journal of EMS*. 3 (2001), pp. 213–266.
- [14] F. Andreu-Vaillo, V. Caselles, and J.M. Mazon. *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*. Progress in Mathematics. Springer Verlag NY, 2004. ISBN: 9783764366193.
- [15] G. Anzellotti. “Pairing Between Measures and Bounded Functions and Compensated Compactness”. In: *Ann. di Matematica Pura ed Appl.* IV.135 (1983), pp. 293–318.
- [16] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. “Image decomposition into a bounded variation component and an oscillating component”. In: *Journal of Mathematical Imaging and Vision* 22 (2005), pp. 71–88.
- [17] E. Barozzi, E. Gonzales, and I. Tamanini. “The mean curvature of a set of finite perimeter”. In: *AMS* 99 (1987), pp. 313–316.
- [18] A. Beck and M. Teboulle. “Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems”. In: *Image Processing, IEEE Transactions on* 18.11 (2009), pp. 2419–2434. ISSN: 1057-7149.
- [19] G. Bellettini, V. Caselles, and M. Novaga. “Explicit Solutions of the Eigenvalue Problem $\operatorname{div}(\frac{D}{u}|Du|) = u$ in R^2 ”. In: *SIAM Journal on Mathematical Analysis* 36.4 (2005), pp. 1095–1129.
- [20] G. Bellettini, V. Caselles, and M. Novaga. “The Total Variation Flow in \mathbb{R}^N ”. In: *J. Differential Equations* 184.2 (2002), pp. 475–525.
- [21] M. Benning and M. Burger. “Ground states and singular vectors of convex variational regularization methods”. In: *Meth. Appl. Analysis* 20 (2013), pp. 295–334.
- [22] K. Bredies and M. Holler. “A pointwise characterization of the subdifferential of the total variation functional”. 2012.

- [23] A. Briani, A. Chambolle, M. Novaga, and G. Orlandi. “On the Gradient Flow of a one-homogeneous functional”. In: *Confluentes Mathematici* 03.04 (2011), pp. 617–635.
- [24] T. Brox, M. Welk, G. Steidl, and J. Weickert. “Equivalence Results for TV Diffusion and TV Regularisation”. English. In: *Scale Space Methods in Computer Vision*. Ed. by L. D. Griffin and M. Lillholm. Vol. 2695. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2003, pp. 86–100. ISBN: 978-3-540-40368-5.
- [25] A. Buades, B. Coll, and J-M. Morel. “A Review of Image Denoising Algorithms, with a New One”. In: *Multiscale Modeling & Simulation* 4.2 (2005), pp. 490–530.
- [26] A. Buades, B. Coll, and J-M. Morel. “Neighborhood filters and PDE’s”. In: *Numerische Mathematik* 105.1 (2006), pp. 1–34.
- [27] M. Burger and S. Osher. “Convergence rates of convex variational regularization”. In: *Inverse problems* 20.5 (2004), p. 1411.
- [28] G. Buttazzo, G. Carlier, and M. Comte. “On the selection of maximal Cheeger sets”. In: *Differential and Integral Equations* 20.9 (2007), pp. 991–1004.
- [29] G. Carlier, M. Comte, and G. Peyré. “Approximation of maximal Cheeger sets by projection”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 43.1 (2009), pp. 131–150.
- [30] V. Caselles, A. Chambolle, and M. Novaga. “The discontinuity set of solutions of the TV denoising problem and some extensions”. In: *Multiscale Modeling and Simulation* 6.3 (2007), pp. 879–894.
- [31] V. Caselles, A. Chambolle, and M. Novaga. “Uniqueness of the Cheeger set of a convex body”. In: *Pacific J. Math.* 232.1 (2007), pp. 77–90.
- [32] V. Caselles, K. Jalalzai, and M. Novaga. “On the jump set of solutions of the total variation flow”. In: *Rendiconti del Seminario Matematico della Università di Padova* 130 (2013), pp. 155–168.
- [33] V. Caselles, M. Novaga, and C. Pöschl. “TV denoising of two balls in the plane”. In: *arXiv preprint arXiv:1605.00247* (2016).
- [34] Vicent Caselles, Antonin Chambolle, and Matteo Novaga. “Regularity for solutions of the total variation denoising problem”. In: *Rev. Mat. Iberoam.* 27.1 (2011), pp. 233–252. ISSN: 0213-2230.
- [35] A. Chambolle. “An algorithm for total variation minimization and applications”. In: *Journal of Mathematical Imaging and Vision* 20.1 (2004), pp. 89–97.
- [36] A. Chambolle and J. Darbon. “On Total Variation Minimization and Surface Evolution Using Parametric Maximum Flows”. In: *International Journal of Computer Vision* 84.3 (2009), pp. 288–307. ISSN: 1573-1405.
- [37] A. Chambolle, M. Goldman, and M. Novaga. “Fine properties of the subdifferential for a class of one-homogeneous functionals”. In: *Adv. Calc. Var.* 8.1 (2015), pp. 31–42.

- [38] A. Chambolle and P-L. Lions. “Image recovery via total variation minimization and related problems”. In: *Numerische Mathematik* 76.2 (1997), pp. 167–188.
- [39] A. Chambolle, M. Novaga, D. Cremers, and T. Pock. “An introduction to total variation for image analysis”. In: *Theoretical Foundations and Numerical Methods for Sparse Recovery*, De Gruyter. 2010.
- [40] T. Chan, A. Marquina, and P. Mulet. “High-Order Total Variation-Based Image Restoration”. In: *SIAM Journal on Scientific Computing* 22.2 (2000), pp. 503–516.
- [41] G. Chavent and K. Kunisch. “Regularization of linear least squares problems by total bounded variation”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 2 (Jan. 1997), pp. 359–376. ISSN: 1262-3377.
- [42] L. Condat. “A Direct Algorithm for 1-D Total Variation Denoising”. In: *IEEE Signal Process. Lett* 20.11 (2013), pp. 1054–1057.
- [43] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian. “Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering”. In: *IEEE Trans. Image Processing* 16.8 (Aug. 2007), pp. 2080–2095.
- [44] J. Darbon and M. Sigelle. “Exact Optimization of Discrete Constrained Total Variation Minimization Problems”. English. In: *Combinatorial Image Analysis*. Ed. by R. Klette and J. Aunias. Vol. 3322. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, pp. 548–557. ISBN: 978-3-540-23942-0.
- [45] P. L. Davies and A. Kovac. “Local Extremes, Runs, Strings and Multiresolution”. In: *Ann. Statist.* 29.1 (Feb. 2001), pp. 1–65.
- [46] L. Dumbgen and A. Kovac. “Extensions of smoothing via taut strings”. In: *Electron. J. Statist.* 3 (2009), pp. 41–75.
- [47] V. Duval and G. Peyré. “Exact Support Recovery for Sparse Spikes Deconvolution”. In: *Foundations of Computational Mathematics* 15.5 (2015), pp. 1315–1355.
- [48] Barozzi E. and Massari U. “On sets with variational mean curvature in $L^n(\mathbb{R}^n)$ ”. In: *Annali Dell’Universita’Di Ferrara* 59.2 (2013), pp. 259–268.
- [49] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. Classics in Applied Mathematics. Philadelphia: Siam, 1999. ISBN: 0-89871-450-8.
- [50] M. J. Fadili and G. Peyré. “Total Variation Projection with First Order Schemes”. In: *IEEE Transactions on Image Processing* 20.3 (2011), pp. 657–669.
- [51] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Grundlehren der mathematischen Wissenschaften. Springer, 1977.
- [52] E. Giusti. *Minimal surfaces and functions of bounded variation*. Notes on pure mathematics. Dept. of Pure Mathematics, 1977. ISBN: 9780708112946.
- [53] E. Giusti. “On the equation of surfaces of prescribed mean curvature”. In: *Invent. Math.* 46.2 (1978), pp. 111–137.

- [54] E. HA Gonzales, U. Massari, and I. Tamanini. “Boundaries of prescribed mean curvature”. In: *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni* 4.3 (1993), pp. 197–206.
- [55] E. H. A. Gonzalez and U. Massari. “Variational mean curvatures”. In: *Rend. Sem. Mat. Univ. Pol. Torino* 52 (1994), pp. 1–28.
- [56] M. Grasmair. “The Equivalence of the Taut String Algorithm and BV-Regularization”. In: *Journal of Mathematical Imaging and Vision* 27.1 (2006), pp. 59–66. ISSN: 1573-7683.
- [57] M. Grasmair, O. Scherzer, and M. Haltmeier. “Necessary and sufficient conditions for linear convergence of ℓ_1 -regularization”. In: *Communications on Pure and Applied Mathematics* 64.2 (2011), pp. 161–182.
- [58] W. Hinterberger, M. Hintermuller, K. Kunisch, M. von Oehsen, and O. Scherzer. “Tube Methods for BV Regularization”. In: *Journal of Mathematical Imaging and Vision* 19.3 (Nov. 2003), pp. 219–235.
- [59] D. S. Hochbaum. “An Efficient Algorithm for Image Segmentation, Markov Random Fields and Related Problems”. In: *J. ACM* 48.4 (2001), pp. 686–701.
- [60] K. Jalalzai. “Some Remarks on the Staircasing Phenomenon in Total Variation-Based Image Denoising”. In: *Journal of Mathematical Imaging and Vision* 54.2 (2015), pp. 256–268.
- [61] B. Kawohl and T. Lachand-Robert. “Characterization of Cheeger sets for convex subsets of the plane”. In: *Pacific J. Math.* 225.1 (2006), pp. 103–118.
- [62] B. Kawohl and M. Novaga. “The p -Laplace eigenvalue problem as $p \rightarrow 1$ and Cheeger sets in a Finsler metric”. In: *Journal of Convex Analysis* 15.3 (2008), p. 623. ISSN: 0944-6532.
- [63] B. Kawohl and F. Schuricht. “Dirichlet Problems for the 1-Laplace Operator, Including the Eigenvalue Problem”. In: *Communications in Contemporary Mathematics* 09.04 (2007), pp. 515–543.
- [64] S. Kindermann, S. Osher, and J. Xu. “Denoising by BV-duality”. In: *Journal of Scientific Computing* 28.2-3 (2006), pp. 411–444.
- [65] V. Kolmogorov and R. Zabini. “What energy functions can be minimized via graph cuts?” In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.2 (2004), pp. 147–159. ISSN: 0162-8828.
- [66] N. J. Korevaar. “Capillary Surface Convexity above Convex Domains”. In: *Indiana Univ. Math. J.* 32 (1 1983), pp. 73–81. ISSN: 0022-2518.
- [67] N. J. Korevaar. “Convex Solutions to Nonlinear Elliptic and Parabolic Boundary Value Problems”. In: *Indiana University Mathematics Journal* 32.4 (1983).
- [68] F. Maggi. *Sets of finite perimeter and geometric variational problems: an introduction to Geometric Measure Theory*. Vol. 135. Cambridge University Press, 2012.

- [69] F. Malgouyres. “Minimizing the total variation under a certain convex constraint for image restoration”. In: *IEEE Trans. Image Proc.* 11 (2002), pp. 1450–1456.
- [70] E. Mammen and S. van de Geer. “Locally adaptive regression splines”. In: *Ann. Statist.* 25.1 (Feb. 1997), pp. 387–413.
- [71] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*. Vol. 22. University Lecture Series. The fifteenth Dean Jacqueline B. Lewis memorial lectures. Providence, RI: American Mathematical Society, 2001. ISBN: 0-8218-2920-3.
- [72] D. Needell and R. Ward. “Stable image reconstruction using total variation minimization”. In: *CoRR* abs/1202.6429 (2012).
- [73] M. Nikolova. “Local Strong Homogeneity of a Regularized Estimator”. In: *SIAM Journal on Applied Mathematics* 61.2 (2000), pp. 633–658.
- [74] E. Paolini. “Regularity for minimal boundaries in R^n with mean curvature in L^n ”. In: *Manuscripta Mathematica* 97.1 (1998), pp. 15–35. ISSN: 1432-1785. DOI: 10.1007/s002290050082. URL: <http://dx.doi.org/10.1007/s002290050082>.
- [75] C. Poon. “On the Role of Total Variation in Compressed Sensing”. In: *SIAM Journal on Imaging Sciences* 8.1 (2015), pp. 682–720.
- [76] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. “Image denoising using scale mixtures of gaussians in the wavelet domain”. In: *IEEE Trans. Image Processing* 12.11 (Nov. 2003), pp. 1338–1351.
- [77] W. Ring. “Structural Properties of Solutions to Total Variation Regularization Problems”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 34 (04 July 2000), pp. 799–810. ISSN: 1290-3841.
- [78] R. T. Rockafellar, R. J-B. Wets, and M. Wets. *Variational analysis*. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg, New York: Springer, 1998.
- [79] L. I. Rudin, S. Osher, and E. Fatemi. “Nonlinear total variation based noise removal algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1 (1992), pp. 259–268.
- [80] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. 1st. Applied Mathematical Sciences. Springer, 2009. ISBN: 0387309314.
- [81] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational Methods in Imaging*. Vol. 167. Springer Science & Business Media, 2008.
- [82] J. Serra. *Image analysis and mathematical morphology*. Academic Press, London, 1982.
- [83] D. M. Strong and T. F. Chan. *Exact Solutions to Total Variation Regularization Problems*. Tech. rep. UCLA, 1996.

- [84] S. Vaiter, G. Peyré, C. Dossal, and J. Fadili. “Robust Sparse Analysis Regularization”. In: *IEEE Transactions on Information Theory* 59.4 (2013), pp. 2001–2016.
- [85] T. Valkonen. “The Jump Set under Geometric Regularization. Part 1: Basic Technique and First-Order Denoising”. In: *SIAM J. Math. Analysis* 47.4 (2015), pp. 2587–2629.
- [86] C. R. Vogel and M. E. Oman. “Iterative Methods for Total Variation Denoising”. In: *SIAM Journal on Scientific Computing* 17.1 (1996), pp. 227–238.
- [87] M. Wertheimer. “Untersuchungen zur Lehre der Gestalt, II”. In: *Psychologische Forschung* 4 (1923), pp. 301–350.