Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

A note on convergence of solutions of total variation regularized linear inverse problems

J.A. Iglesias, G. Mercier, O. Scherzer

RICAM-Report 2017-39

A note on convergence of solutions of total variation regularized linear inverse problems

José A. Iglesias*, Gwenael Mercier*, Otmar Scherzer*†

Abstract

In a recent paper by A. Chambolle et al. [10] it was proven that if the subgradient of the total variation at the true data is not empty, the level-sets of the total-variation denoised solutions converge to the level-sets of the true solution with respect to the Hausdorff distance. This paper explores a new aspect of total variation regularization theory based on the source condition introduced by Burger and Osher [9] to prove convergence rates results with respect to the Bregman distance. We generalize the results of Chambolle et al. to total variation regularization of general linear inverse problems. As applications we consider denoising in bounded and unbounded, convex and non convex domains, deblurring and inversion of the circular Radon transform. In all these examples we can prove Hausdorff convergence of the level-sets of the total variation regularized solutions.

1 Introduction

In this paper we are concerned with total variation regularization of linear inverse problems

$$Au = f\,, \tag{1}$$

for functions u defined in a domain Ω which is either $\mathbb{R}^2$ or a bounded Lipschitz domain $D \subset \mathbb{R}^2$, and $A : \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Sigma)$ is a linear bounded (typically compact) operator. Since in general the solution of (1) is ill-posed, some sort of regularization needs to be employed.

The method considered in this paper is total variation regularization, in which a regularization parameter $\alpha > 0$ is chosen, and either of the two following minimization problems is solved:

- The Dirichlet (resp. full space) problem consisting of minimization of the functional

$$\mathcal{F}_\alpha(u) := \frac{1}{2}\left\|Au - f\right\|^2_{\mathrm{L}^2(\Sigma)} + \alpha \mathrm{TV}(u) \tag{2}$$

  among $u \in \mathrm{L}^2(D) \cong \left\{u \in \mathrm{L}^2(\mathbb{R}^2) \,\middle|\, \mathrm{supp}(u) \subset \overline{D}\right\}$ (resp. $u \in \mathrm{L}^2(\mathbb{R}^2)$), where the quantity $\mathrm{TV}(u) \in [0, +\infty]$ denotes the total variation, computed in $\mathbb{R}^2$, of the extension of u by zero outside of D (resp. the total variation in $\mathbb{R}^2$ of u).

- The Neumann problem

$$\hat{\mathcal{F}}_\alpha(u) := \frac{1}{2}\left\|Au - f\right\|^2_{\mathrm{L}^2(\Sigma)} + \alpha \mathrm{TV}(u\,;\,\Omega), \tag{3}$$

  where $\mathrm{TV}(u\,;\,\Omega)$ is the total variation of u computed in the open set Ω.

*Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz, Austria (jose.iglesias@ricam.oeaw.ac.at, gwenael.mercier@ricam.oeaw.ac.at)

†Computational Science Center, University of Vienna, Austria (otmar.scherzer@univie.ac.at)

Structure of the paper. The outline of the paper is as follows: In Section 2 we prove existence of minimizers of $\mathcal{F}_\alpha$ and $\hat{\mathcal{F}}_\alpha$ and review dual formulations of these optimization problems, with the goal of exploring the convergence of dual variables (see (8)) and its relation to the source condition. In Section 3, we see that the curvature of level-sets of minimizers is strongly linked to these dual variables and we explain (following [10]) how the convergence of the curvatures implies the main result on Hausdorff convergence of level-sets. In Section 4 we give a proof, for each of the different boundary conditions considered, of the main ingredient needed for the convergence: density estimates (19) for the level-sets. Finally, Section 5 contains some examples where the results of previous sections apply, and the influence of the boundary on the solutions is discussed.

1.1 Notation and spaces

We recall that the total variation in Ω is defined for every function $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ by

$$\mathrm{TV}(u\,;\,\Omega) := |Du|(\Omega) = \sup\left\{ \int_\Omega u \;\operatorname{div} z \; dx \;\middle|\; z \in C_0^\infty(\Omega\,;\,\mathbb{R}^2), \|z\|_{\mathrm{L}^\infty(\Omega)} \leqslant 1 \right\}.$$

If the latter is finite, then the distributional derivative Du is a vector-valued Radon measure on Ω. We also introduce for every Lebesgue measurable $E \subset \Omega$ the perimeter of E in Ω,

$$\mathrm{Per}(E\,;\,\Omega) := \mathrm{TV}(1_E\,;\,\Omega),$$

where 1_E is the indicator function of E. If this quantity is finite, E is said to have finite perimeter. When $\Omega = \mathbb{R}^2$ or when Ω is clear from the context we skip the second argument in the above notations.

In case the domain Ω is bounded, since we have the inclusion $\mathrm{L}^2(\Omega) \subset \mathrm{L}^1(\Omega)$, candidates for minimizers of $\mathcal{F}_\alpha$ are in

$$\mathrm{BV}_\Omega := \left\{ v \in \mathrm{BV}(\mathbb{R}^2) \;\middle|\; v \equiv 0 \text{ on } \mathbb{R}^2 \setminus \Omega \right\},$$

where we adopt the standard definition

$$\mathrm{BV}(\mathbb{R}^2) := \left\{ u \in \mathrm{L}^1(\mathbb{R}^2) \;\middle|\; \mathrm{TV}(u) < +\infty \right\}.$$

Since the total variation in $\mathbb{R}^2$ is used, this corresponds to a homogeneous Dirichlet boundary condition and possible jumps at the boundary are taken into account. On the other hand, if $\Omega = \mathbb{R}^2$, candidates belong only to

$$\left\{ u \in \mathrm{L}^2(\mathbb{R}^2) \,\middle|\, \mathrm{TV}(u) < +\infty \right\}.$$

In the case of $\hat{\mathcal{F}}_\alpha$, which formally requires minimization among functions satisfying homogeneous Neumann boundary conditions, the corresponding space is $\mathrm{BV}(\Omega) \cap \mathrm{L}^2(\Omega)$. The influence of Ω and the boundary conditions on the solutions is discussed in Section 5.

2 Dual certificates and source condition

Proposition 1. *The functional $\mathcal{F}_\alpha$ defined in* (2)*, considered on $\mathrm{L}^2(\Omega)$, has at least one minimizer. If A is injective, the minimizer is unique.*

Proof. Let (u_k) be a minimizing sequence. Since $u_k \in \mathrm{L}^2(\Omega)$ implies $u_k \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ and we work in dimension 2, we can use Sobolev's inequality [3, Theorem 3.47] to get

$$\|u_k\|_{\mathrm{L}^2(\Omega)} \leqslant C\,\mathrm{TV}(u_k).$$

Now, the right hand side is bounded uniformly in k so that the Banach-Alaoglu theorem for $\mathrm{L}^2(\Omega)$ and a compactness result [3, Theorem 3.23] provide a subsequence (u_k) (not relabelled) that converges both weakly in $\mathrm{L}^2(\Omega)$ and strongly in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ to some limit u_α. Since A is a bounded linear operator, Au_k also converges weakly to Au_α in $\mathrm{L}^2(\Sigma)$. Lower semicontinuity of the norm with respect to weak convergence, and of the total variation with respect to strong $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ convergence [3, Remark 3.5] proves that u_α is a minimizer of $\mathcal{F}_\alpha$.

The uniqueness statement is straightforward, since $\|\cdot\|^2_{\mathrm{L}^2(\Sigma)}$ is strictly convex. □

Remark 1. Working in BV_Ω and with the total variation on $\mathbb{R}^2$ can produce markedly different results, compared to considering functions in $\mathrm{L}^2(\Omega)$ and their total variation $\mathrm{TV}(u\,;\,\Omega)$ inside Ω leading to the functional (3). An example is the choice $\Omega = (-1,1)^2$, $\Sigma = (-1+\eta, 1-\eta)^2$ for some $\eta \in (0,1)$, $\alpha = 1$, $f = 0$ and A defined by

$$Au(x) = u(x) - \frac{1}{4\eta^2}\int_{(-\eta,\eta)^2} u(x+y)\,\mathrm{d}y,$$

continuous since $\|Au\|_{\mathrm{L}^2(\Sigma)} \leqslant 2\,\|u\|_{\mathrm{L}^2(\Omega)}$ by the triangle inequality and Young's inequality for convolutions. In this situation, the functional (3) is not coercive: considering the sequence $u_n := n1_\Omega$, we have that $\hat{\mathcal{F}}(u_n) = 0$ for all n, but u_n is not bounded in $\mathrm{L}^2(\Omega)$. The underlying reason is that constant functions are cancelled by A, that is, $A\mathbb{1} = 0$, where $\mathbb{1}$ represents the constant function with value 1 (this situation has also been discussed in [21]). In contrast, when working in BV_Ω we have $\mathrm{TV}(u_n) = 4n$. Note that in the denoising case ($A = \mathrm{Id}$) the data term makes the functional coercive in $\mathrm{L}^2(\Omega)$ even when using $\mathrm{TV}(u\,;\,\Omega)$.

Proposition 2. *The functional $\hat{\mathcal{F}}_\alpha$ defined in* (3)*, considered in* $\mathrm{L}^2(\Omega)$*, has at least one minimizer. If A is injective, the minimizers is unique.*

Proof. As noticed in Remark 1, the situation is slightly different from Proposition 1. Indeed, if as above, (u_k) is a minimizing sequence, Sobolev inequality gives the existence of a constant m_k such that

$$\|u_k - m_k\|_{\mathrm{L}^2(\Omega)} \leqslant C\,\mathrm{TV}(u_k) \leqslant C. \tag{4}$$

Now, if the constant functions are cancelled by A (that is, $A\mathbb{1} = 0$), then $Au_k - f = A(u_k - m_k) - f$ and $\mathrm{TV}(u_k - m_k\,;\,\Omega) = \mathrm{TV}(u_k\,;\,\Omega)$, so that $v_k := u_k - m_k$ is also a minimizing sequence. Since v_k is bounded in $\mathrm{L}^2(\Omega)$ by (4), it converges weakly to some $v \in \mathrm{L}^2(\Omega)$. Similarly to Proposition 1, one can use compactness and lower semicontinuity to show that v is a minimizer of (3).
On the other hand, if $A\mathbb{1} \neq 0$, the boundedness of A still gives

$$\|A \cdot m_k\|_{\mathrm{L}^2(\Sigma)} = |m_k|\|A\mathbb{1}\|_{\mathrm{L}^2(\Sigma)},$$

and since the left hand side is bounded, the sequence $|m_k|$ is also bounded and therefore (u_k) is bounded in $\mathrm{L}^2(\Omega)$ and converges weakly (up to a subsequence) to some u. The end of the proof works then again as in Proposition 1. □

In the rest of the section, we assume that we are in the case of $\Omega = \mathbb{R}^2$ or Dirichlet boundary conditions, but the results and their proofs are identical for Neumann boundary conditions.

First, we recall some basic results about the convergence of u_α as α vanishes, when some noise is added to the data f.

Lemma 1. *Let $A : \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Sigma)$ be a bounded linear operator. Moreover, assume that there exists a solution $\tilde{u}$ of* (1) *which satisfies* $\mathrm{TV}(\tilde{u}) < \infty$*. Then*

- *There exists a solution $u^\dagger$ of* (1) *with minimal total variation. That is $Au^\dagger = f$ and*
$$\mathrm{TV}(u^\dagger) = \inf\left\{\mathrm{TV}(u) \,\middle|\, u \in \mathrm{L}^2(\Omega),\ Au = f\right\} .$$

- *Given a sequence (α_n) with $\alpha_n \to 0^+$, elements $w_n \in \mathrm{L}^2(\Omega)$ and some positive constant C such that*
$$\frac{\|w_n\|^2_{\mathrm{L}^2(\Sigma)}}{\alpha_n} \leqslant C, \tag{5}$$
there exist a (not relabelled) subsequence (α_n) and minimizers (u_{α_n,w_n}) of
$$u \mapsto \frac{1}{2}\|Au - (f + w_n)\|^2_{\mathrm{L}^2(\Sigma)} + \alpha_n \mathrm{TV}(u)$$
such that $u_{\alpha_n,w_n} \rightharpoonup u^\dagger$ weakly in $\mathrm{L}^2(\Omega)$ for $u^\dagger$ a solution of (1) *with minimal total variation. Additionally, this convergence is also strong with respect to the $\mathrm{L}^1_{loc}(\mathbb{R}^2)$ and $\mathrm{L}^p(\Omega)$ for $1 \leqslant p < 2$, topology, respectively, if Ω is bounded. Furthermore, $\mathrm{TV}(u_{\alpha_n,w_n}) \to \mathrm{TV}(u^\dagger)$.*

Proof. A proof of the first statement can be found in [20, Theorem 3.25]. The second relies on the compactness of the embedding of BV, and may be found in [20, Theorem 3.26] (see also [1, Theorem 5.1]). □

For the results contained in the rest of this section, it will be enough to consider the noiseless case, and therefore we denote a generic minimizer of $\mathcal{F}_\alpha$ with the fixed data f by u_α. The theory of regularization methods is most often concerned with quantitative versions of the convergence of u_α. An essential ingredient of the analysis of convergence rates analysis is the following source condition:

Definition 1. *Let $A^* : \mathrm{L}^2(\Sigma) \to \mathrm{L}^2(\Omega)$ be the adjoint of A. We say that a minimum norm solution $u^\dagger$ satisfies the* source condition *if*
$$\mathcal{R}(A^*) \cap \partial\mathrm{TV}(u^\dagger) \neq \emptyset. \tag{6}$$
Here $\mathcal{R}(A^*)$ denotes the range of the operator A^* and $\partial\mathrm{TV}(u^\dagger)$ denotes the subgradient of $\mathrm{TV}(\cdot)$ at $u^\dagger$ with respect to $\mathrm{L}^2(\Omega)$.

Remark 2.
- This source condition, first introduced in [9], is standard in the inverse problems community. It is the natural condition to obtain convergence rates (with respect to the Bregman distance) of $u_\alpha \to u^\dagger$. See [20, Prop. 3.35, Thm. 3.42].

- Let us notice that the set in (6) does not depend on which minimal variation solution $u^\dagger$ is chosen. Indeed, let $u_1^\dagger, u_2^\dagger$ be such solutions and assume
$$A^*p \in \partial\mathrm{TV}(u_1^\dagger),$$
which means that for every $h \in \mathrm{L}^2(\Omega)$
$$\mathrm{TV}(u_1^\dagger + h) - \mathrm{TV}(u_1^\dagger) \geqslant \langle A^*p, h\rangle_{\mathrm{L}^2(\Omega)} .$$
Now we can write for every k, since $\mathrm{TV}(u_2^\dagger) = \mathrm{TV}(u_1^\dagger)$
$$\begin{aligned}
\mathrm{TV}(u_2^\dagger + k) - \mathrm{TV}(u_2^\dagger) &= \mathrm{TV}(u_1^\dagger + (u_2^\dagger - u_1^\dagger) + k) - \mathrm{TV}(u_1^\dagger) \\
&\geqslant \left\langle A^*p, (u_2^\dagger - u_1^\dagger) + k \right\rangle_{\mathrm{L}^2(\Omega)} \\
&= \left\langle A^*p, u_2^\dagger - u_1^\dagger \right\rangle_{\mathrm{L}^2(\Omega)} + \langle A^*p, k\rangle_{\mathrm{L}^2(\Omega)} \\
&= \left\langle p, Au_2^\dagger - Au_1^\dagger \right\rangle_{\mathrm{L}^2(\Sigma)} + \langle A^*p, k\rangle_{\mathrm{L}^2(\Omega)} \\
&= \langle A^*p, k\rangle_{\mathrm{L}^2(\Omega)} ,
\end{aligned}$$

which means that $A^*p \in \partial \mathrm{TV}(u_2^\dagger)$.

Theorem 1. • *Let $\alpha > 0$. The dual problem (in the sense of [15]) of minimizing the functional $\mathcal{F}_\alpha$, defined in* (2), *on* $\mathrm{L}^2(\Omega)$ *consists in maximizing, among $p \in \mathrm{L}^2(\Sigma)$ such that $A^*p \in \partial\mathrm{TV}(0)$, the quantity*

$$\mathcal{D}_\alpha(p) := \langle f, p\rangle_{\mathrm{L}^2(\Sigma)} - \frac{\alpha}{2}\|p\|^2_{\mathrm{L}^2(\Sigma)}. \tag{7}$$

Moreover,

$$\inf_{u\in\mathrm{L}^2(\Omega)} \mathcal{F}_\alpha(u) = \sup_{A^*p\in\partial\mathrm{TV}(0)} \mathcal{D}_\alpha(p). \tag{8}$$

If these quantities are attained by u_α, p_α, then we have the extremality relations

$$A^*p_\alpha \in \partial\mathrm{TV}(u_\alpha) \tag{9}$$

and

$$p_\alpha \in -\partial\left(\frac{1}{2\alpha}\|A\cdot - f\|^2_{\mathrm{L}^2(\Sigma)}\right)(u_\alpha) = \left\{\frac{1}{\alpha}(f - Au_\alpha)\right\}.$$

- *Similarly, the formal limits of the minimization problems for* (2) *and* (7) *when $\alpha \to 0$ write*

$$l_f := \inf\{\mathrm{TV}(u) \mid u \in \mathrm{L}^2(\Omega),\ Au = f\} \tag{10}$$

and

$$l_d := \sup_{A^*p\in\partial\mathrm{TV}(0)} \langle p,\ f\rangle_{\mathrm{L}^2(\Sigma)} = \sup_{v\in\mathcal{R}(A^*)\cap\partial\mathrm{TV}(0)} \left\langle v,\ u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)}, \tag{11}$$

and satisfy also the strong duality condition $l_f = l_d$. The extremality conditions for (10) *and* (11), *provided the quantities above are attained by some $u^\dagger, p_0$, write*

$$A^*p_0 \in \partial\mathrm{TV}(u^\dagger) \tag{12}$$

and

$$p_0 \in -\left(\partial\chi_{\{f\}}(A\,\cdot)\right)(u^\dagger) = \mathrm{L}^2(\Sigma),$$

where $\chi_{\{f\}}$ is the indicator function of the set $\{f\}$, i.e. $\chi_{\{f\}}(q) = 0$ if $q = f$, and $\chi_{\{f\}}(q) = +\infty$ otherwise.

Proof. In the L^2 setting we can make use of classical duality theorems. In the notation of [15, Theorem 4.2] our situation corresponds to:

$$V = \mathrm{L}^2(\Omega), Y = \mathrm{L}^2(\Sigma), \Lambda = A, F(\cdot) = \mathrm{TV}(\cdot) \text{ and } G(\cdot) = \frac{1}{2\alpha}\|\cdot - f\|^2_{\mathrm{L}^2(\Sigma)}.$$

In the formulas that arise from this theorem, we then use the identity (also used in [10])

$$\partial\mathrm{TV}(u^\dagger) = \left\{v \in \partial\mathrm{TV}(0) \,\middle|\, \left\langle v, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} = \mathrm{TV}(u^\dagger)\right\}, \tag{13}$$

to obtain the statement. This identity, which holds for any 1-homogeneous convex functional, can be derived taking advantage of the 0-homogeneity of the subgradient and noting that for such a functional, we have the triangle inequality. □

The assumption that there exists a maximizer of (11) is in fact related to the source condition (6):

Lemma 2. *There exists p_0 maximizing the functional defined in* (11) *over p such that $A^*p \in \partial\mathrm{TV}(0)$ if and only if the source condition* (6) *is satisfied.*

Proof. First, we note that

$$\langle p, f\rangle_{\mathrm{L}^2(\Sigma)} = \left\langle p,\ Au^\dagger\right\rangle_{\mathrm{L}^2(\Sigma)} = \left\langle A^* p, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} . \tag{14}$$

- The source condition (6) implies the existence of $p_0 \in \mathrm{L}^2(\Sigma), v_0 \in \mathrm{L}^2(\Omega)$ such that

  $$v_0 = A^* p_0 \in \partial \mathrm{TV}(u^\dagger).$$

  Then, we note that $v \in \partial\mathrm{TV}(0)$ implies that $\left\langle v, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} \leqslant \mathrm{TV}(u^\dagger)$. From (13) it then follows that v_0 maximizes $\left\langle \cdot, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)}$ in $\partial\mathrm{TV}(0)$, so that together with (14), we get that p_0 is a maximizer of (11).

- Conversely, if $p_0 \in \mathrm{L}^2(\Sigma)$ maximizes $\langle \cdot,\ f\rangle_{\mathrm{L}^2(\Sigma)}$ among p such that $A^* p \in \partial\mathrm{TV}(0)$, then the extremality condition (12) ensures that

  $$A^* p_0 \in \partial\mathrm{TV}(u^\dagger),$$

  and thus the source condition is satisfied.

□

Remark 3. The minimizers of the primal functionals (2), (10) as well as the maximizers of the limit dual functional (11) are not unique in general. However, the dual functional $\mathcal{D}_\alpha$ has a unique maximizer. The existence follows directly since $\partial\mathrm{TV}(0)$ is weakly closed (subgradients of lower semicontinuous convex functions are convex and strongly closed [7, Proposition 16.4], hence weakly closed). Uniqueness follows by the strict convexity of the squared L^2 norm and convexity of $\partial\mathrm{TV}(0)$.

The following proposition is a key result explaining the importance of source condition. In fact boundedness of the maximizers of the dual problems is closely related to the source conditions. The arguments are similar as proving convergence of the Augmented Lagrangian Method (see [17]), which have also been used to prove convergence rates results for dual variables [16] and to prove existence of Bregman TV-flows [8]. The proof of the first part follows [14].

Proposition 3. *Let the source condition* (6) *be satisfied and let* p_α *be the maximizer of* (7)*. Then, we have that*

$$\lim_{\alpha\to 0^+} p_\alpha = p^* \quad \textit{strongly in } \mathrm{L}^2(\Sigma),$$

where p^* *is the maximizer of* (11) *with minimal* $\mathrm{L}^2(\Sigma)$ *norm. Conversely, if* (p_α) *is bounded in* $\mathrm{L}^2(\Sigma)$*, then the source condition is satisfied.*

Proof. Let p_0 be a maximizer of (11), which exists by Lemma 2. We have that

$$\langle p_0, f\rangle_{\mathrm{L}^2(\Sigma)} = \left\langle A^* p_0, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} \geqslant \left\langle A^* p_\alpha, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)}$$

and analogously, since p_α maximizes $\mathcal{D}_\alpha(\cdot)$ that

$$\left\langle A^* p_\alpha, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} - \frac{\alpha}{2}\left\|p_\alpha\right\|^2_{\mathrm{L}^2(\Sigma)} \geqslant \left\langle A^* p_0, u^\dagger\right\rangle_{\mathrm{L}^2(\Omega)} - \frac{\alpha}{2}\left\|p_0\right\|^2_{\mathrm{L}^2(\Sigma)} . \tag{15}$$

Summing these inequalities, we see that (p_α) is bounded and therefore converges weakly (up to a subsequence) to some $p^* \in \mathrm{L}^2(\Sigma)$. Passing to the limit in the two previous equations gives

$$\langle p_0, f\rangle_{\mathrm{L}^2(\Sigma)} = \langle p^*, f\rangle_{\mathrm{L}^2(\Sigma)}$$

where $A^*p^* \in \partial\mathrm{TV}(0)$, the latter being weakly closed. Equation (15) and weak convergence imply that

$$\|p^*\|_{\mathrm{L}^2(\Sigma)} \leqslant \liminf \|p_\alpha\|_{\mathrm{L}^2(\Sigma)} \leqslant \|p_0\|_{\mathrm{L}^2(\Sigma)}$$

which implies that p^* is actually the minimal norm maximizer of the functional $\langle \cdot, f\rangle_{\mathrm{L}^2(\Sigma)}$ over p such that $A^*p \in \partial\mathrm{TV}(0)$, and that the convergence is strong (and for every subsequence).

Let us now assume that (p_α) is bounded in $\mathrm{L}^2(\Sigma)$. Then by weak compactness for the p_α and applying Lemma 1 (with $w_n = 0$), there exist $\alpha_n \to 0$, $u^\dagger$ a solution of $Au = f$ with minimal total variation, and $\overline{p}$ such that

$$\begin{aligned} p_{\alpha_n} &\rightharpoonup \overline{p} \text{ in } \mathrm{L}^2(\Sigma), \text{ and} \\ u_{\alpha_n} &\to u^\dagger \text{ in } \mathrm{L}^1_{\mathrm{loc}}(\Omega). \end{aligned}$$

The extremality conditions (9) and (13) imply that

$$\langle A^* p_{\alpha_n}, u_{\alpha_n}\rangle_{\mathrm{L}^2(\Omega)} = \mathrm{TV}(u_{\alpha_n}).$$

Since $\mathrm{TV}(\cdot)$ is lower semi-continuous on $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$, we have $\mathrm{TV}(u^\dagger) \leqslant \liminf_n \mathrm{TV}(u_{\alpha_n})$. On the other hand, one can write

$$\begin{aligned} \mathrm{TV}(u_{\alpha_n}) &= \left\langle A^* p_{\alpha_n}, u^\dagger \right\rangle_{\mathrm{L}^2(\Omega)} + \left\langle A^* p_{\alpha_n}, u_{\alpha_n} - u^\dagger \right\rangle_{\mathrm{L}^2(\Omega)} \\ &= \left\langle A^* p_{\alpha_n}, u^\dagger \right\rangle_{\mathrm{L}^2(\Omega)} + \langle p_{\alpha_n}, Au_{\alpha_n} - f\rangle_{\mathrm{L}^2(\Sigma)}, \end{aligned}$$

where the first term of the right hand side converges to $\langle A^*\overline{p}, u^\dagger\rangle_{\mathrm{L}^2(\Omega)}$ whereas the second term goes to zero because (p_{α_n}) is uniformly bounded in $\mathrm{L}^2(\Omega)$ and because of the strong $\mathrm{L}^2(\Sigma)$ convergence $Au_{\alpha_n} \to f$. Moreover, $A^*\overline{p} \in \partial\mathrm{TV}(0)$ because $\partial\mathrm{TV}(0)$ is weakly closed. Hence $A^*\overline{p} \in \partial\mathrm{TV}(u^\dagger)$, which is the source condition. □

3 Convergence of Level-Sets

From Proposition 3 it follows that the family

$$(v_\alpha) := (A^* p_\alpha) \text{ converges strongly in } \mathrm{L}^2(\Omega)$$

if the source condition (6) it satisfied, so the family is bounded and equi-integrable. That is the basis of the proof in [10] of convergence of level-sets.

Before stating the result, we introduce $u_{\alpha,w}$ a minimizer of $\mathcal{F}_\alpha$ or $\hat{\mathcal{F}}_\alpha$ for the data $f+w$, where w corresponds to a noise, as in Lemma 1. For every $t \in \mathbb{R}$, we denote by $U^{(t)}_{\alpha,w}$ the t level-set of $u_{\alpha,w}$, that is

$$\begin{aligned} U^{(t)}_{\alpha,w} &:= \{x \in \Omega \mid u_{\alpha,w}(x) \geqslant t\} && \text{for } t \geqslant 0, \\ U^{(t)}_{\alpha,w} &:= \{x \in \Omega \mid u_{\alpha,w}(x) \leqslant t\} && \text{for } t < 0, \end{aligned}$$

this choice being made to ensure that the volume of the level-sets considered are always finite (except the zero one that should be considered separately, see [10]). Similarly, we call $U^{(t)}_\dagger$ the level-sets of $u^\dagger$. Then, we have

Theorem 2. *Assume that either:*

- *For Dirichlet boundary conditions or $\Omega = \mathbb{R}^2$, let $(w_n) \subset \mathrm{L}^2(\Sigma)$ and $\alpha_n \to 0^+$ such that*
$$\frac{\|w_n\|_{\mathrm{L}^2(\Sigma)}\|A^*\|}{\alpha_n} \leqslant \eta < \sqrt{\pi}. \tag{16}$$
If Ω is bounded, assume further that it admits a variational curvature $\kappa_\Omega \in \mathrm{L}^1(\mathbb{R}^2)$ such that $\kappa_\Omega \geqslant g$ with $g \in \mathrm{L}^2(\mathbb{R}^2 \setminus \Omega)$.

- *For Neumann boundary conditions, with $(w_n) \subset \mathrm{L}^2(\Sigma)$ and $\alpha_n \to 0^+$ such that*
$$\frac{\|w_n\|_{\mathrm{L}^2(\Sigma)}\|A^*\|}{\alpha_n} \leqslant \eta < \frac{1}{C(\Omega)}, \tag{17}$$
with $C(\Omega)$ is some Sobolev-Poincaré constant, to be specified later.

Then up to a subsequence and for almost all $t \in \mathbb{R}$, denoting U_{α_n,w_n} by U_n, we have that
$$\lim_{n\to\infty} |U_n^{(t)} \Delta U_\dagger^{(t)}| = 0, \quad \textit{and} \quad \lim_{n\to\infty} \partial U_n^{(t)} = \partial U_\dagger^{(t)},$$
the second limit being understood in the sense of Hausdorff convergence.

In what follows, we use the notion of variational curvature that we make precise now.

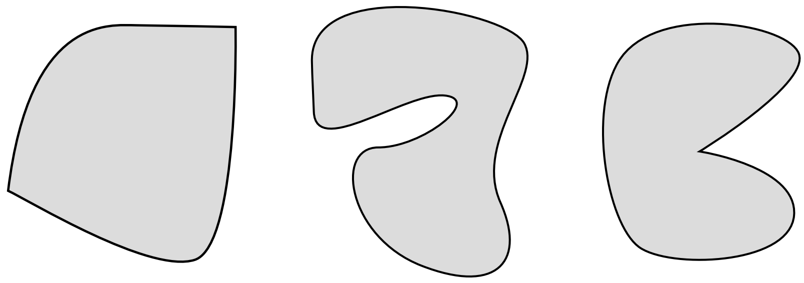
Definition 2. *Let $v \in \mathrm{L}^1(\Omega)$. Then a set E is said to have variational curvature v in Ω if*

- *The perimeter in Ω of E finite.*
- *E minimizes the functional*
$$F \to Per(F) - \int_F v$$
among compact perturbations, that is, for every F such that $F\Delta E$ is compactly supported in Ω we have
$$\mathrm{Per}(E) - \int_E v \leqslant \mathrm{Per}(F) - \int_F v. \tag{18}$$

Remark 4. For smooth sets, this notion is strongly related with the differential notion of curvature. Indeed, assuming that the boundary of E is smooth and that v is also smooth, one may consider diffeomorphic deformations $\phi^s : \Omega \to \Omega$ applied to E such that each boundary point $x \in \partial E$ is mapped to $x + sh(x)\nu(x)$, where ν is the outer unit normal vector and $h : \Omega \to \mathbb{R}$ a smooth function. We obtain at $s = 0$ [19, Section 17.3]
$$\frac{d}{ds} \mathrm{Per}\left(\phi^s(E)\right) = \int_{\partial E} h(x)\kappa(x)\, d\mathcal{H}^1(x) \text{ and}$$
$$\frac{d}{ds} \int_{\phi^s(E)} v = \int_{\partial E} h(x)v(x)\, d\mathcal{H}^1(x),$$
where κ is the curvature of ∂E and $\mathcal{H}^1$ is the 1-dimensional Hausdorff measure. Since h was arbitrary and using the minimality (18) of E, we must have $\kappa\big|_{\partial E} = v$.

In [5], the authors show that every set with finite perimeter has a variational curvature in $\mathrm{L}^1(\mathbb{R}^2)$, so such a quantity will exist for every set considered in this paper.

The restriction $\kappa_\Omega \geqslant g$ that is put on Ω in Theorem 2 roughly means that inside corners (where the curvature is negative, see Figure 1 (c)) are not allowed. Indeed, corners are known not to have a curvature in $\mathrm{L}^2(\mathbb{R}^2)$ [18, Theorem 1.1]. However, many interesting domains satisfy $\kappa_\Omega \geqslant g$ with $g \in \mathrm{L}^2(\mathbb{R}^2 \setminus \Omega)$ (see Figure 1 (a) and (b)), in particular:



(a) (b) (c)

Figure 1: Three Lipschitz domains. Domains (a) and (b) have a variational curvature with lower bound in L^2 of their complements, whereas (c), because of the inside corner, does not.

- Any convex domain, even with corners, has a variational curvature κ_Ω such that $\kappa_\Omega = 0$ on $\mathbb{R}^2 \setminus \Omega$, since a convex set minimizes perimeter among outer perturbations. Indeed, if $F \supset E$, this writes $\operatorname{Per}(F) \geqslant \operatorname{Per}(E)$ which is precisely (18) after using $\int_F \kappa_\Omega = \int_E \kappa_\Omega + \int_{F\setminus E} \kappa_\Omega$.
- Any $C^{1,1}$ domain has a curvature $\kappa_\Omega \in \mathrm{L}^\infty(\mathbb{R}^2)$. To see this, first notice that at boundary points of a $C^{1,1}$ set one can place balls of radius bounded below and completely inside or outside Ω [13, Theorems 7.8.2 (ii) and 7.7.3]. Moreover, it is proved in [4, Remark 1.3 (ii)] that the variational curvature constructed in [18] for a set with the mentioned property is bounded.

We outline the main ingredients of the proof of Theorem 2 below, along the lines of [10], highlighting the differences that appear because of the operator A and the different boundary conditions.

3.1 Structure of the proof of Theorem 2

Denoting by u_n the functions u_{α_n,w_n}, minimizers of either the Functional $\mathcal{F}_\alpha$ or $\hat{\mathcal{F}}_\alpha$ with data $f + w_n$ and parameter α_n, and since the condition (16) (resp. (17)) is stronger than (5), we know by Lemma 1 that $u_n \to u^\dagger$ strongly in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$.

The proof is accomplished in two steps. The first is to improve the convergence of u from $\mathrm{L}^1_{\mathrm{loc}}$ to L^1, which is the convergence in mass of its level-sets. In the second step, the L^1 convergence of the level-sets is improved to Hausdorff convergence.

First step. We show in Section 3.4 that actually the u_n have a compact support (that does not depend on n) so that the convergence of u_n also holds in L^1 strong, which implies that, up to a subsequence that we still denote by n, the level-sets $U_n^{(t)}$ of u_n converge to the level-sets $U_\dagger^{(t)}$ of $u^\dagger$ in L^1, and for almost every t.

Second step. An improvement of mode of convergence like the one we need can, in general, only be accomplished through a regularity result. In our case, the adequate property is termed *weak-regularity* in [10], and relates to the well-known density estimates for Λ-minimizers of perimeter [19, Theorem 21.11]. Thanks to (20) below and the parameter choice (16,17), we will see that the curvature of level-sets of $u_{\alpha,w}$ is controlled which roughly speaking implies that it cannot contain sharp tips, which in turn will imply that such a level-set should locally have significant mass at both sides of its boundary. In symbols, E and its complement E^c have a significant fraction of mass around each $x \in \partial E$.

That is, one considers for small r the sets $E \setminus B(x,r)$ and $E \cup B(x,r)$, and obtains lower bounds of the type

$$\frac{|B(x,r) \cap E|}{|B(x,r)|} \geqslant C \text{ and } \frac{|B(x,r) \setminus E|}{|B(x,r)|} \geqslant C. \tag{19}$$

We prove this property for all the different boundary conditions in Section 4. Let us now explain how (19) implies that for collections of sets which lie inside a common ball (we will see that the level-sets of $u_{\alpha,w}$ do), L^1 convergence and Hausdorff convergence are equivalent. To see this, just consider E_n converging to F in L^1 such that the E_n satisfy (19) with C and $r \leqslant r_0$ uniform in n, and the definition of Hausdorff distance:

$$\begin{aligned} d_H(E_n, F) &= \max \left\{ \sup_{x \in E_n} d(x,F), \sup_{y \in F} d(y, E_n) \right\} \\ &= \max \left\{ \sup_{x \in E_n} \inf_{y \in F} |x-y|, \sup_{y \in F} \inf_{x \in E_n} |x-y| \right\}, \end{aligned}$$

and suppose without loss of generality that the first term of the right hand side does not converge to 0. This would imply that there is a $\delta > 0$ (we can take $\delta < r_0$) and $x_n \in E_n$ such that $d(x_n, F) \geqslant \delta$, and in particular $B(x_n, \delta) \cap F = \emptyset$. This implies using the density estimate (19) that

$$|E_n \,\Delta\, F| \geqslant |(E_n \cap B(x_n,\delta)) \setminus F| = |E_n \cap B(x_n, \delta)| \geqslant C\delta^2,$$

contradicting the L^1 convergence.

3.2 The level-set problem

We turn our attention to the function

$$v_{\alpha,w} = A^* p_{\alpha,w} = \frac{1}{\alpha} A^*(f + w - Au_{\alpha,w}),$$

derived from the dual certificates for the minimization of (2) (resp. (3)). The reason is that this function is the variational curvature of $U^{(t)}_{\alpha,w}$. Indeed, it is proved in [10, Prop. 3] that the extremality relation (9) is equivalent to the statement that for every $F \subset \Omega$ and every $t \neq 0$,

$$\mathrm{Per}(F) - \mathrm{sgn}(t) \int_F v_{\alpha,w} \geqslant \mathrm{Per}(U^{(t)}_{\alpha,w}) - \mathrm{sgn}(t) \int_{U^{(t)}_{\alpha,w}} v_{\alpha,w}, \tag{20}$$

which implies that $U^{(t)}_{\alpha,w}$ has a variational curvature $\mathrm{sgn}(t) v_{\alpha,w}$. Furthermore, it is also shown in [10, Prop. 3] that (13) implies that the $U^{(t)}_{\alpha,w}$ satisfy

$$\mathrm{Per}(U^{(t)}_{\alpha,w}) = \mathrm{sgn}(t) \int_{U^{(t)}_{\alpha,w}} v_{\alpha,w}.$$

These formulas are consequences of slicing equations (9) and (13) using the coarea and layer cake formulas.

3.3 Parameter choice

First, we will see that the noise does not significantly modify the geometrical properties of the level-sets, as long as α is chosen relatively to the noise level. Indeed, in view of (7) one can note that $p_{\alpha,w}$ is the $\mathrm{L}^2(\Sigma)$ orthogonal projection of $\frac{f+w}{\alpha}$ onto the convex set

$$\left\{ p \in \mathrm{L}^2(\Sigma) \;\middle|\; A^* p \in \partial \mathrm{TV}(0) \right\}.$$

The non-expansiveness of the projection operator leads to

$$\|p_\alpha - p_{\alpha,w}\|_{\mathrm{L}^2(\Sigma)} \leqslant \frac{\|w\|_{\mathrm{L}^2(\Sigma)}}{\alpha},$$

which, together with the boundedness of A^*, means that

$$\|v_\alpha - v_{\alpha,w}\|_{\mathrm{L}^2(\Omega)} \leqslant \frac{\|w\|_{\mathrm{L}^2(\Sigma)}\|A^*\|}{\alpha} \leqslant \eta,$$

so that the curvatures of the level-sets of $u_{\alpha,w}$ and u_α, for noisy and noiseless data respectively, can be forced to be as close as needed when $w \to 0$ through the choice of α. The necessity of the restrictions for η of (16) and (17) is made apparent in Sections 3.4 and 4 below.

3.4 Upper bounds and compact support

We now prove that sets E satisfying

$$\mathrm{Per}(E) = \int_E v_{\alpha,w} \tag{21}$$

(in particular, the level-sets of the minimizers $u_{\alpha,w}$) have uniformly bounded perimeter, and their support is contained in a common ball. The latter completes the first step of Section 3.1.

If we work with bounded Ω, a bound on the perimeter follows easily from (21) and (16) (resp. (17)):

$$\begin{aligned}\mathrm{Per}(E) \leqslant \left|\int_E (v_{\alpha,w} - v_\alpha)\right| + \left|\int_E v_\alpha\right| &\leqslant \eta\sqrt{|\Omega|} + \sqrt{|\Omega|}\|v_\alpha\|_{\mathrm{L}^2(\Omega)} \\ &\leqslant \left(\eta + \sup_\alpha \|v_\alpha\|_{\mathrm{L}^2(\Omega)}\right)\sqrt{|\Omega|}.\end{aligned}$$

The $\mathbb{R}^2$ case is very close to what is done in [10]. We sketch now the arguments given in [10], that apply directly to this case. The same parameter choice (16) is required.

Here, by Proposition 3, we have that $v_\alpha \to v_0$ strongly in $\mathrm{L}^2(\Omega)$, and therefore the family (v_α) is L^2-equiintegrable, which means that one can find a ball $B(0,R)$ such that

$$\int_{\mathbb{R}^2 \setminus B(0,R)} v_\alpha^2 \leqslant \eta.$$

Then, for every E with finite mass that satisfies (21) and provided α and w satisfy (16),

$$\begin{aligned}\mathrm{Per}(E) &\leqslant \left|\int_E (v_{\alpha,w} - v_\alpha)\right| + \left|\int_{E\cap B(0,R)} v_\alpha\right| + \left|\int_{E\setminus B(0,R)} v_\alpha\right| \\ &\leqslant \eta\sqrt{|E|} + \sqrt{|B(0,R)|}\|v_\alpha\|_{\mathrm{L}^2(\Omega)} + \sqrt{|E \setminus B(0,R)|}\eta \\ &\leqslant \left(\eta + \sup_\alpha \|v_\alpha\|_{\mathrm{L}^2(\Omega)}\right)\sqrt{|B(0,R)|} + 2\eta\sqrt{|E \setminus B(0,R)|}.\end{aligned}$$

Now, isoperimetric inequality (that is, $4\pi|E| \leq \mathrm{Per}(E)^2$) and sub-additivity of the perimeter lead to

$$\sqrt{|E \setminus B(0,R)|} \leqslant \frac{1}{\sqrt{4\pi}}\,\mathrm{Per}(E \setminus B(0,R)) \leqslant \frac{1}{\sqrt{4\pi}}\big(\mathrm{Per}(E) + \mathrm{Per}(B(0,R))\big),$$

which when used in the previous equation, since $\eta < \sqrt{\pi}$, implies that $\mathrm{Per}(E)$ is bounded uniformly in α. Once again using the isoperimetric inequality yields the boundedness of $|E|$ independently of α, as long as (16) is satisfied.

We now prove that the mass and perimeter of level-sets of $u_{\alpha,w}$ are bounded away from zero. The equiintegrability of (v_α) ensures that there is no concentration of mass for v_α: $\int_E v_\alpha^2$ is small if $|E|$ is small. Then, if E satisfies (21), Cauchy Schwarz inequality provides an inequality of the type

$$\mathrm{Per}(E) \leqslant \varepsilon\sqrt{|E|},$$

which together with the isoperimetric inequality, implies $\mathrm{Per}(E) \leqslant C\varepsilon\,\mathrm{Per}(E)$, which is not possible for ε too small. Therefore, $|E|$ must be bounded away from zero (and $\mathrm{Per}(E)$ as well thanks to the isoperimetric inequality).

One can then show, using [2], that if E has a finite mass and satisfies (21), it can be split into connected components which also satisfy (21). Therefore, the perimeter and mass of such components are bounded from above and below, which implies that there can only be finitely many of them. Since their perimeter is bounded, their diameter is bounded too, which implies that they all lie in a ball $B(0, R)$. So does E.

Remark 5. As a byproduct of the previous proof, one can notice that all level-sets of $u^\dagger$ belong to some ball $B(0, R)$, which means that $\partial\mathrm{TV}(u^\dagger) \neq \emptyset$ implies that $u^\dagger$ has a compact support. To our knowledge, this property was never stated before, although it is implicit in [10]. Since it is a result on the subgradient, it applies whether $A = \mathrm{Id}$ or not.

4 Proof of the density estimates

In this section, we derive the density estimates (19) in each of the three boundary frameworks that are mentioned in this article. The proof follows the usual strategy for this kind of estimates (see [19], for example), but the appearance of different boundary conditions requires a closer examination.

The general strategy of the proof is to use minimality of a set in problem (20) and compare it with the sets obtaining by adjoining or substracting pieces of balls centered at a point of its boundary, leading to the first and second parts of (19) respectively.

In what follows we consider only the first estimate, since the second one can be derived analogously. We emphasize that the bounds obtained need to be uniform in α in order to obtain the desired convergence.

For a Lebesgue set $E \subset \mathbb{R}^2$, we use the notations $E^{(1)}$ and $E^{(0)}$ for the points where the density of E is 1 and 0 respectively. That is, for $s \in \{0, 1\}$ we have

$$E^{(s)} = \left\{ x \in \mathbb{R}^2 \;\middle|\; \lim_{r\to 0} \frac{|B(x,r) \cap E|}{|B(x,r)|} = s \right\}.$$

Furthermore, we note that by the Lebesgue differentiation theorem

$$|E^{(0)} \Delta (\mathbb{R}^2 \setminus E)| = 0 \text{ and } |E^{(1)} \Delta E| = 0.$$

Since $D1_E$ is a Radon measure, one can consider the perimeter of E in every Borel subset F, which we denote by $P(E\,;\,F) := |D1_E|(F)$. We can then prove

Lemma 3. *Let $\kappa \in \mathrm{L}^2(\Omega)$ (with Ω bounded or $\Omega = \mathbb{R}^2$) and $E \subset \Omega$ minimize*

$$F \mapsto \mathrm{Per}(F) - \int_F \kappa,$$

where the perimeter is either understood in Ω (Neumann) or in $\mathbb{R}^2$ (Dirichlet). Then, one has for almost every r

$$\mathrm{Per}(E \cap B(x,r)) - \int_{E\cap B(x,r)} \kappa \leqslant 2\,\mathrm{Per}(B(x,r)\,;\,E^{(1)}) \tag{22}$$

Remark 6. One can note that

$$\operatorname{Per}(B(x,r)\,;\,E^{(1)}) = \mathcal{H}^1(\partial B(x,r)\cap E) \text{ for almost every } r, \tag{23}$$

for $\mathcal{H}^1$ the 1-dimensional Hausdorff measure. In fact, (22) can be proved for all r by keeping track of extra terms in (24) and (25) that appear when the sets have tangential contact.

Proof. We use the following inequality, valid for every finite perimeter sets $F, G \subset \Omega$,

$$\operatorname{Per}(F\setminus G) + \operatorname{Per}(G\setminus F) \leqslant \operatorname{Per}(F) + \operatorname{Per}(G), \tag{24}$$

which can be proved by using (16.11) of [19, Theorem 16.3] twice. We will also apply the following equality, which holds for almost every $r > 0$

$$\begin{aligned}
&\operatorname{Per}((B(x,r)\cap\Omega)\setminus E)\\
&\quad= \operatorname{Per}(B(x,r)\cap\Omega\,;\,E^{(0)}\cap\Omega) + \operatorname{Per}(E\,;\,B(x,r)\cap\Omega)\\
&\quad= \operatorname{Per}(B(x,r)\cap\Omega\,;\,E^{(0)}\cap\Omega) + \operatorname{Per}(E\cap B(x,r)\,;\,\Omega) - \operatorname{Per}(B(x,r)\cap\Omega\,;\,E^{(1)}\cap\Omega).
\end{aligned} \tag{25}$$

This equality can be deduced using relations (16.11) and (16.10) of [19] and noting that for almost every r,

$$\mathcal{H}^1\left[(\partial(B(x,r)\cap\Omega)\cap\partial^*E)\cap\Omega\right] = 0, \tag{26}$$

since $\mathcal{H}^1(\partial^*E) < \infty$, where ∂^*E is the reduced boundary of E [19, Chapter 15].

Using the minimality of E, then the two formulas above, and the additivity of perimeter, we get

$$\begin{aligned}
\operatorname{Per}(E) - \int_E \kappa &\leqslant \operatorname{Per}(E\setminus B(x,r)) - \int_{E\setminus B(x,r)} \kappa\\
&\leqslant \operatorname{Per}(E) + \operatorname{Per}(B(x,r)\cap\Omega) - \operatorname{Per}((B(x,r)\cap\Omega)\setminus E) - \int_{E\setminus B(x,r)} \kappa\\
&= \operatorname{Per}(E) + \operatorname{Per}(B(x,r)\cap\Omega) - \operatorname{Per}((B(x,r)\cap E) - \operatorname{Per}(B(x,r)\cap\Omega\,;\,E^{(0)}\cap\Omega)\\
&\quad + \operatorname{Per}(B(x,r)\cap\Omega\,;\,E^{(1)}\cap\Omega) - \int_{E\setminus B(x,r)} \kappa\\
&= \operatorname{Per}(E) - \operatorname{Per}(B(x,r)\cap E) + 2\operatorname{Per}(B(x,r)\,;\,E^{(1)}\cap\Omega) - \int_{E\setminus B(x,r)} \kappa,
\end{aligned}$$

where in the last equality we use [19, Theorem 16.2] and again (26). Since $E^{(1)} \subset \Omega$, the above is the statement of (22). □

4.1 The $\mathbb{R}^2$ case.

Here, $\Omega = \mathbb{R}^2$ and the proof is then the one presented in [10] up to making more explicit the constants involved. We denote by E a level-set (which we assume without loss of generality to be positive) of $u_{\alpha,w}$ that therefore minimizes

$$F \mapsto \operatorname{Per}(F) - \int_F v_{\alpha,w},$$

and $x \in \partial E$. Thanks to the equiintegrability of v_α (which, as noted before, follows from the strong convergence in L^2 showed in Proposition 3), for every $\delta > 0$ and $|F| \leqslant \pi r_0^2$ with r_0 small enough (independent of α but dependent of δ) one has

$$\left(\int_F |v_\alpha|^2\right)^{1/2} \leqslant \delta. \tag{27}$$

Then, (16) and the above imply that for $r \leqslant r_0$,

$$\begin{aligned}\left|\int_{E\cap B(x,r)} v_{\alpha,w}\right| &\leqslant |E\cap B(x,r)|^{1/2}\|v_{\alpha,w}\|_{\mathrm{L}^2(B(x,r))}\\ &\leqslant |E\cap B(x,r)|^{1/2}\left(\|v_\alpha\|_{\mathrm{L}^2(B(x,r))}+\eta\right)\\ &\leqslant |E\cap B(x,r)|^{1/2}(\delta+\eta).\end{aligned}$$

Using the above in (22), we obtain

$$\mathrm{Per}(E\cap B(x,r)) - |E\cap B(x,r)|^{1/2}(\delta+\eta) \leqslant 2\,\mathrm{Per}(B(x,r)\,;\,E^{(1)}),$$

which combined with the isoperimetric inequality in $\mathbb{R}^2$ and finally (23) yields

$$|E\cap B(x,r)|^{1/2}(2\sqrt{\pi}-\delta-\eta) \leqslant 2\mathcal{H}^1(E\cap\partial B(x,r)). \tag{28}$$

Now, denoting by

$$g(r) := |E\cap B(x,r)|,$$

we have that for a.e. r, $g'(r) = \mathcal{H}^1(E\cap\partial B(x,r))$. As a result, (28) reads

$$(2\sqrt{\pi}-\delta-\eta)\sqrt{g} \leqslant 2g'.$$

Now, if η and δ are chosen such that $\delta+\eta < 2\sqrt{\pi}$, one can integrate on both sides and use $g(0)=0$ to get $(2\sqrt{\pi}-\delta-\eta)r \leqslant 4\sqrt{g(r)}$, which reads

$$\frac{|B(x,r)\cap E|}{|B(x,r)|} \geqslant \frac{(2\sqrt{\pi}-\delta-\eta)^2r^2}{16\pi r^2} = \frac{(2\sqrt{\pi}-\delta-\eta)^2}{16\pi},$$

which is uniform in α. Since δ was arbitrary and the parameter choice (16) implies $\eta < \sqrt{\pi}$, we obtain (19).

4.2 The Dirichlet case

In this subsection, we consider the case of Dirichlet conditions in a bounded domain, and see that it can be treated through a variational problem formulated in $\mathbb{R}^2$:

Lemma 4. *Assume that Ω admits a variational curvature κ_Ω such that $\kappa_\Omega \geqslant g$ with $g \in \mathrm{L}^2(\mathbb{R}^2)$, and let $E \subset \Omega$ satisfy* (20) *(we assume that $t>0$). Then, E satisfies the following variational problem among sets $F \subset \mathbb{R}^2$ such that $F\Delta E$ is bounded:*

$$\mathrm{Per}(E) - \int_E \kappa_{\alpha,w} \leqslant \mathrm{Per}(F) - \int_F \kappa_{\alpha,w}, \quad \textit{where } \kappa_{\alpha,w} = v_{\alpha,w}1_\Omega + g1_{\mathbb{R}^2\setminus\Omega}.$$

Proof. Similarly to [6, Lemma, p. 132], we consider the constraint as an obstacle and we write

$$\begin{aligned}\mathrm{Per}(E) - \int_E \kappa_{\alpha,w} = \mathrm{Per}(E) - \int_E v_{\alpha,w} &\leqslant \mathrm{Per}(F\cap\Omega) - \int_{F\cap\Omega} v_{\alpha,w}\\ &\leqslant \mathrm{Per}(F) + \mathrm{Per}(\Omega) - \mathrm{Per}(F\cup\Omega) - \int_{F\cap\Omega} v_{\alpha,w}\\ &= \mathrm{Per}(F) + \mathrm{Per}(\Omega) - \mathrm{Per}(F\cup\Omega) - \int_{F\cap\Omega} \kappa_{\alpha,w}.\end{aligned} \tag{29}$$

On the other hand, the assumption on Ω implies

$$\operatorname{Per}(\Omega) - \int_\Omega \kappa_\Omega \leqslant \operatorname{Per}(F \cup \Omega) - \int_{F\cup\Omega} \kappa_\Omega,$$

which we can use in (29) to get

$$\begin{aligned}
\operatorname{Per}(E) - \int_E \kappa_{\alpha,w} &\leqslant \operatorname{Per}(F) - \int_{F\cap\Omega} \kappa_{\alpha,w} - \int_{F\setminus\Omega} \kappa_\Omega \\
&\leqslant \operatorname{Per}(F) - \int_{F\cap\Omega} \kappa_{\alpha,w} - \int_{F\setminus\Omega} g \\
&= \operatorname{Per}(F) - \int_{F\cap\Omega} \kappa_{\alpha,w} - \int_{F\setminus\Omega} \kappa_{\alpha,w} \\
&= \operatorname{Per}(F) - \int_F \kappa_{\alpha,w}.
\end{aligned}$$

□

Now, for $x \in \partial E$, we may perturb E with balls $B(x, r)$ not necessarily contained in Ω. Since the $\kappa_{\alpha,w}$ are, as the $v_{\alpha,w}$, equiintegrable, we can apply the $\mathbb{R}^2$ density estimates of Section 4.1 to obtain (19) for the Dirichlet boundary conditions.

4.3 The Neumann case

We have assumed that Ω is such that its boundary can be locally represented as the graph of a Lipschitz function (so it is in particular an extension domain, see [3, Definition 3.20, Proposition 3.21]). Therefore, as a replacement for the isoperimetric inequality, we can use the following Poincaré-Sobolev inequality [3, Remark 3.50] valid for $u \in \mathrm{BV}(\Omega)$:

$$\left\| u - \frac{1}{|\Omega|}\int_\Omega u \right\|_{\mathrm{L}^2(\Omega)} \leqslant C(\Omega)\, \mathrm{TV}(u\,;\,\Omega).$$

With $u = 1_F$ the indicator function of some $F \subset \Omega$, the left hand side reads

$$\int_\Omega \left| 1_F - \frac{|F|}{|\Omega|} \right|^2 = |F| \left(\frac{|\Omega \setminus F|}{|\Omega|} \right)^2 + |\Omega \setminus F| \left(\frac{|F|}{|\Omega|} \right)^2$$

and the inequality yields

$$C(\Omega) \operatorname{Per}(F) \geqslant \left(\frac{|F|\ |\Omega \setminus F|}{|\Omega|^2} \right)^{1/2} (|\Omega \setminus F| + |F|)^{1/2} \geqslant \left(\frac{|F|\ |\Omega \setminus F|}{|\Omega|} \right)^{1/2}. \tag{30}$$

As before, let E satisfy (20). Applying (30) to $E \cap B(x, r)$, we get

$$\begin{aligned}
|E \cap B(x,r)|^{1/2} &\leqslant C(\Omega) \left(\frac{|\Omega|}{|\Omega \setminus (E \cap B(x,r))|} \right)^{1/2} \operatorname{Per}(E \cap B(x,r)) \\
&\leqslant C(\Omega) \left(\frac{|\Omega|}{|\Omega \setminus B(x,r)|} \right)^{1/2} \operatorname{Per}(E \cap B(x,r)).
\end{aligned} \tag{31}$$

Now, the parameter choice (17) implies that one can choose r_0 independent of x such that for every $r \leqslant r_0$,

$$\eta < \frac{1}{C(\Omega)} \left(\frac{|\Omega| - |B(r)|}{|\Omega|} \right)^{1/2}$$

and such that (27) holds for some δ that satisfies

$$\frac{1}{C(\Omega)}\left(\frac{|\Omega|-|B(r_0)|}{|\Omega|}\right)^{1/2}-\delta-\eta>0.$$

We can then use (31) in (22) (which holds since E satisfies (20)) and perform the same remaining steps as in Section 4.1 to get the estimate

$$\frac{|B(x,r)\cap E|}{|B(x,r)|}\geqslant\frac{\left(\frac{1}{C(\Omega)}\left(\frac{|\Omega|-|B(r_0)|}{|\Omega|}\right)^{1/2}-\delta-\eta\right)^2}{16\pi},$$

where the right hand side is uniform in r and x. This is the first part of (19). In this case, the other estimate of (19) reads

$$\frac{|(B(x,r)\cap\Omega)\setminus E|}{|B(x,r)|}\geqslant C,$$

which is still enough for the Hausdorff convergence of $\partial U^{(t)}_{\alpha,w}$ of Section 3.1.

5 Examples and discussion

We first consider two particular examples of operators A where the above results apply.

5.1 The circular Radon transform

We review from [20] the problem of inverting the *circular Radon transform*

$$\mathrm{R}_{\mathrm{circ}}\, u=v \tag{32}$$

in a stable way, where

$$\begin{aligned}\mathrm{R}_{\mathrm{circ}}:\mathrm{L}^2(\mathbb{R}^2)&\to\mathrm{L}^2\big(\Sigma=\mathbb{S}^1\times(0,2)\big)\,,\\ u&\mapsto(\mathrm{R}_{\mathrm{circ}}\,u)(\vec{z},t):=t\int_{\mathbb{S}^1}u(\vec{z}+t\vec{\omega})\,d\mathcal{H}^1(\vec{\omega})\,.\end{aligned} \tag{33}$$

In the following let $\Omega:=B(0,1)$ be an open Ball of Radius 1 with center 0 in $\mathbb{R}^2$ and let $\varepsilon\in(0,1)$. We are considering the spherical Radon transform defined on the subspace of functions supported in $\overline{B(0,1-\varepsilon)}$, that is on

$$\mathrm{L}^2(B(0,1-\varepsilon)):=\left\{u\in\mathrm{L}^2(\mathbb{R}^2)\;\middle|\;\mathrm{supp}(u)\subseteq\overline{B(0,1-\varepsilon)}\right\}.$$

Some properties [20, Prop. 3.80 and 3.81] of the circular Radon transform are:

- The circular Radon transform, as defined in (33), is well-defined, bounded, and satisfies $\|\mathrm{R}_{\mathrm{circ}}\|\leq 2\pi$.
- There exists a constant $C_\varepsilon>0$, such that

  $$C_\varepsilon^{-1}\,\|\mathrm{R}_{\mathrm{circ}}\,u\|_2\leq\|i^*(u)\|_{1/2,2}\leq C_\varepsilon\,\|\mathrm{R}_{\mathrm{circ}}\,u\|_2\,,\quad u\in\mathrm{L}^2(B(0,1-\varepsilon))\,,$$

  where i^* is the adjoint of the embedding $i:\mathrm{W}^{1/2,2}(B(0,1))\to\mathrm{L}^2(B(0,1))$ of the standard Sobolev space of differentiation of order 1/2 on Ω.

- For every $\varepsilon \in (0,1)$ we have

$$\mathcal{R}(\mathrm{R}_{\mathrm{circ}}^*) \cap \mathrm{L}^2(B(0,1-\varepsilon)) = \mathrm{W}^{1/2,2}(B(0,1-\varepsilon))\,,$$

  where

$$\begin{aligned}&\mathrm{W}^{1/2,2}(B(0,1-\varepsilon))\\ &\qquad := \left\{ u \in \mathrm{L}^2(\mathbb{R}^2) \;\middle|\; \operatorname{supp} u \subset \overline{B(0,1-\varepsilon)} \text{ and } u|_{B(0,1)} \in \mathrm{W}^{1/2,2}(B(0,1)) \right\}.\end{aligned}$$

  Note that $\mathrm{W}^{1/2,2}$ is not the standard definition of a Sobolev space because we associate with each function of the space $\mathrm{W}^{1/2,2}(B(0,1-\varepsilon))$ an extension to $\mathbb{R}^2$ by 0 outside. We could also say, in the terminology of this paper, that these functions satisfy zero Dirichlet boundary condition on $B(0,1-\varepsilon)$.

It was shown in [20, Prop. 3.82 and 3.83] that minimization of the functional (2) with $A = \mathrm{R}_{\mathrm{circ}}$:

- is well-posed, stable, and convergent.

- Moreover, the following result holds: Let $\varepsilon \in (0,1)$ and $u^\dagger$ be the solution of (32). Then we have the following convergence rates result for TV-regularization: If $\xi \in \partial \mathrm{TV}(u^\dagger) \cap \mathrm{W}^{1/2,2}(B(0,1-\varepsilon))$, then

$$\mathrm{TV}(u_{\alpha(\delta)}^\delta) - \mathrm{TV}(u^\dagger) - \left\langle \xi,\ u_{\alpha(\delta)}^\delta - u^\dagger \right\rangle = \mathcal{O}(\delta) \qquad \text{for } \alpha(\delta) \sim \delta.$$

  In the last equation, the left hand side is called *Bregman distance* of TV at $u^\dagger$ and ξ.

  With the results of this paper, if the parameter α is chosen finer, meaning satisfying (16), we not only have convergence rates of the Bregman distance, but also convergence of the level-sets.

  There are particular examples for which the source condition is satisfied:

  – Let $\rho \in C_0^\infty(\mathbb{R}^2)$ be an adequate mollifier and ρ_μ the scaled function of ρ. Moreover, let $x_0 = (0.2, 0)$, $a = 0.1$, and $\mu = 0.3$. Then

$$u^\dagger := 1_{B(x_0, a+\mu)} * \rho_\mu$$

    satisfies the source condition.

  – Let $u^\dagger := 1_F$ be the indicator function of a bounded subset of $\mathbb{R}^2$ with smooth boundary. Then, the source condition is satisfied as well [20, Example 3.74].

5.2 A numerical deblurring example

The second situation we consider is a numerical deconvolution example, in which an indicator function has been blurred with a Gaussian kernel and subsequently corrupted by additive Gaussian noise. Both the convolution kernel and the variance of the noise are assumed known, and Dirichlet boundary conditions on a rectangle are used. These choices lead directly to the minimization of (2) and enable the use of a parameter choice according to (16), so that the the results of Section 3 provide convergence of level-lines.

The discretization of choice is the 'upwind' finite difference scheme of [11], and the resulting discrete problem is solved with a primal-dual algorithm with the convolutions implemented through Fourier transforms as in [12]. The boundary conditions were imposed by extending the computational domain and projection onto the corresponding constraint. The results and parameter choices are shown in Figure 2.

Finally, we make some remarks on the influence of the boundary conditions in the qualitative properties of the corresponding solutions.

5.3 Denoising in $\mathbb{R}^2$ or in Ω with Dirichlet conditions

In this subsection, we consider only denoising ($A = \mathrm{Id}$). If $u^\dagger = f$ has a bounded support in $\mathbb{R}^2$, one can minimize (2) either in $\mathbb{R}^2$ or in a bounded domain Ω containing the support of f. In general, these minimizations yields different results. Nevertheless, when Ω is convex, we can easily show

Proposition 4. *Let f have compact support included in an open convex set Ω. Then, minimizing* (2) *on Ω with Dirichlet homogeneous boundary conditions or $\mathbb{R}^2$ lead to the same solution.*

Proof. We just need to show that the minimizer u of (2) in $\mathbb{R}^2$ has a support in Ω. If it were not the case, just note that replacing u by $u \cdot 1_\Omega$ decreases both terms of the functional. For the total variation part, this result uses the convexity of Ω. $\square$

If Ω is not convex, it is easy to construct examples where this result is no longer true, even for denoising. See Figure 3. Nevertheless, the direct application of Theorem 2 show that as $\alpha \to 0$, the level-sets of these two minimizers concentrate around the ones of f.

5.4 Denoising with Neumann boundary conditions

As in the Dirichlet case, there are some configurations where solving in a bounded domain does not correspond to solving for $\mathbb{R}^2$. For example, if $A = \mathrm{Id}$, $f = 1_{B(0,1)}$ and $\Omega = B(0, R)$ with $R > 1$, the minimizer of (3) is

$$u_\alpha = \left(1 - \alpha - \frac{2\alpha}{R^2 - 1}\right) 1_{B(0,1)} + \frac{2\alpha}{R^2 - 1} 1_{B(0,R)},$$

whereas the minimizer of (2) in $\mathbb{R}^2$ is clearly $1_{B(0,1)}$.

One can also see lower left image in Figure 3, which contains the denoising of the $\mathcal{C}$ in a rectangle.

Acknowledgments

The third author was supported by the Austrian Science Fund (FWF) through the National Research Network "Geometry+Simulation" (NFN S11704) and the project "Interdisciplinary Coupled Physics Imaging" (FWF P26687).

References

[1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.*, 10(6):1217–1229, 1994.

[2] L. Ambrosio, V. Caselles, S. Masnou, and J.-M. Morel. Connected components of sets of finite perimeter and applications to image processing. *J. Eur. Math. Soc. (JEMS)*, 3(1):39–92, 2001.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs. Oxford University Press, New York, 2000.

[4] E. Barozzi, E. Gonzalez, and U. Massari. The mean curvature of a Lipschitz continuous manifold. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 14(4):257–277 (2004), 2003.

[5] E. Barozzi, E. Gonzalez, and I. Tamanini. The mean curvature of a set of finite perimeter. *Proc. Amer. Math. Soc.*, 99(2):313–316, 1987.

[6] E. Barozzi and U. Massari. Regularity of minimal boundaries with obstacles. *Rend. Sem. Mat. Univ. Padova*, 66:129–135, 1982.

[7] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, second edition, 2011.

[8] M. Burger, K. Frick, S. Osher, and O. Scherzer. Inverse total variation flow. *Multiscale Model. Simul.*, 6(2):365–395 (electronic), 2007.

[9] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Prob.*, 20(5):1411–1421, 2004.

[10] A. Chambolle, V. Duval, G. Peyré, and C. Poon. Geometric properties of solutions to the total variation denoising problem. *Inverse Prob.*, 33(1):015002, 2017.

[11] A. Chambolle, S. E. Levine, and B. J. Lucier. An upwind finite-difference method for total variation-based image smoothing. *SIAM J. Imaging Sci.*, 4(1):277–299, 2011.

[12] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40(1):120–145, 2011.

[13] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries*, volume 22 of *Advances in Design and Control.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2011.

[14] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *Found. Comput. Math.*, 15(5):1315–1355, 2015.

[15] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999.

[16] K. Frick and O. Scherzer. Regularization of ill-posed linear equations by the non-stationary augmented Lagrangian method. *J. Integral Equations Appl.*, 22(2):217–257, June 2010.

[17] R. Glowinski. *Numerical methods for nonlinear variational problems.* Springer Series in Computational Physics. Springer-Verlag, New York, 1984.

[18] E. Gonzales, U. Massari, and I. Tamanini. Boundaries of prescribed mean curvature. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 4(3):197–206, 1993.

[19] F. Maggi. *Sets of finite perimeter and geometric variational problems*, volume 135 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 2012.

[20] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging.* Number 167 in Applied Mathematical Sciences. Springer, New York, 2009.

[21] L. Vese. A study in the BV space of a denoising-deblurring variational problem. *Appl. Math. Optim.*, 44(2):131–161, 2001.

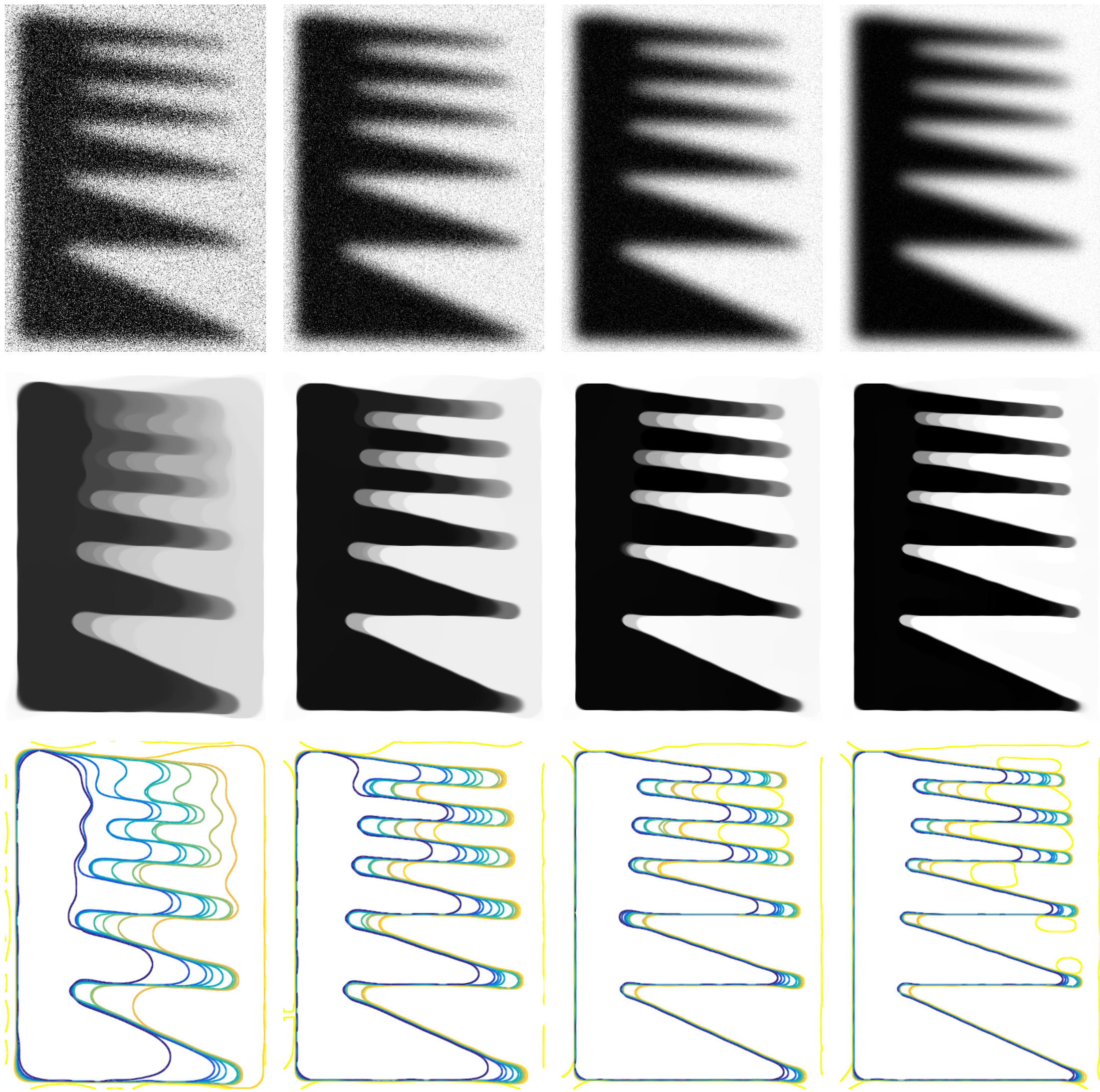


Figure 2: Deblurring of a characteristic function by total variation regularization with Dirichlet boundary conditions. First row: Input image blurred with a known kernel and with additive noise. Second row: numerical deconvolution result, corresponding to minimizers of (2). Third row: some level lines of the results. The regularization parameters are $\alpha = 1, 0.25, 0.0625, 0.0156$ and the variance of the Gaussian noise used is $\alpha/10$.

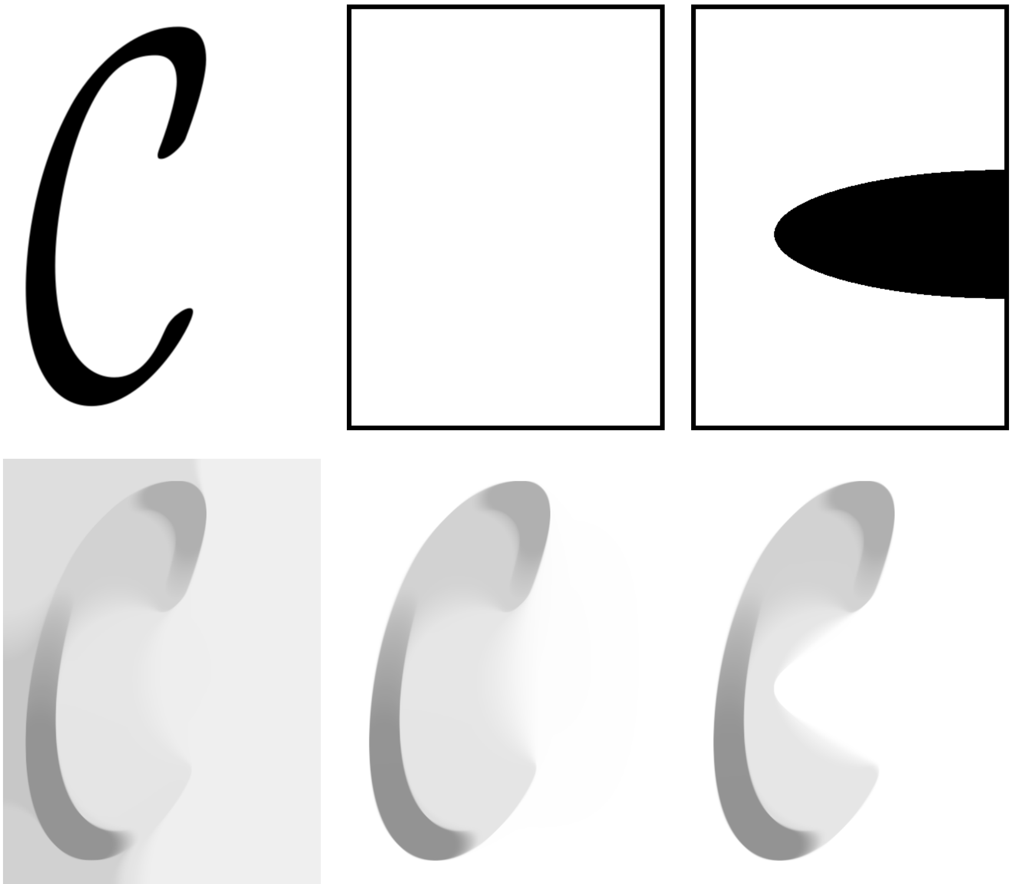


Figure 3: Denoising of a $\mathcal{C}$ with different boundaries and boundary conditions. Upper left: original image. Upper middle: convex Dirichlet domain. Upper right: nonconvex domain. Lower left: result with Neumann boundary. Lower middle: Dirichlet result in the convex domain Lower right: Dirichlet result in the nonconvex domain. The Neumann solution reflects the fact that for level-sets that reach the boundary of Ω, part of their perimeter is not penalized. For the rightmost solution, since the domain is not convex, the solution is different to that of the $\mathbb{R}^2$ case.