Applied Mathematics & Optimization (2023) 88:51
https://doi.org/10.1007/s00245-023-10028-y

Boundedness and Unboundedness in Total Variation Regularization

Kristian Bredies[1] · **José A. Iglesias**[2] · **Gwenael Mercier**[3]

Accepted: 2 June 2023 / Published online: 27 June 2023
© The Author(s) 2023

Abstract
We consider whether minimizers for total variation regularization of linear inverse problems belong to L^∞ even if the measured data does not. We present a simple proof of boundedness of the minimizer for fixed regularization parameter, and derive the existence of uniform bounds for sufficiently small noise under a source condition and adequate a priori parameter choices. To show that such a result cannot be expected for every fidelity term and dimension we compute an explicit radial unbounded minimizer, which is accomplished by proving the equivalence of weighted one-dimensional denoising with a generalized taut string problem. Finally, we discuss the possibility of extending such results to related higher-order regularization functionals, obtaining a positive answer for the infimal convolution of first and second order total variation.

Keywords Total variation · Linear inverse problems · Boundedness of minimimizers · Generalized taut string · Vanishing weights · Infimal convolution regularizers

Mathematics Subject Classification 49Q20 · 47A52 · 65J20 · 65J22

✉ José A. Iglesias
jose.iglesias@utwente.nl

Kristian Bredies
kristian.bredies@uni-graz.at

Gwenael Mercier
gwenael.mercier@univie.ac.at

1 Institute of Mathematics and Scientific Computing, University of Graz, Graz, Austria

2 Department of Applied Mathematics, University of Twente, Enschede, The Netherlands

3 Faculty of Mathematics, University of Vienna, Vienna, Austria

1 Introduction

For $\Omega \subset \mathbb{R}^d$ either a bounded Lipschitz domain or the whole $\mathbb{R}^d$ with $d \geq 2$, $\Sigma \subset \mathbb{R}^m$ an arbitrary domain with $m \geqslant 1$, and given a linear bounded operator

$$A : L^{d/(d-1)}(\Omega) \to L^q(\Sigma),$$

we are interested in solutions $u_{\alpha,w}$ of the TV-regularized inverse problem $Au = f$ with noisy data $f + w$ for $w \in L^q(\Sigma)$, that is, solutions of the minimization problem

$$\min_{u \in L^{d/(d-1)}(\Omega)} \frac{1}{\sigma} \left(\int_\Sigma |Au - (f + w)|^q \right)^{\sigma/q} + \alpha \, \mathrm{TV}(u), \tag{1}$$

where $1 < q < \infty$, $\sigma = \min(q, 2)$ and $\mathrm{TV}(u)$ denotes the total variation of u, see (3) below for its definition. This convex minimization leads, at any minimizer $u_{\alpha,w}$, to the optimality condition

$$\begin{aligned} v_{\alpha,w} &:= A^* p_{\alpha,w} \in \partial_{L^{d/(d-1)}} \mathrm{TV}(u_{\alpha,w}) \subset L^d(\Omega), \\ p_{\alpha,w} &:= \frac{1}{\alpha} \|Au_{\alpha,w} - f - w\|_{L^q(\Sigma)}^{\sigma-2} \, j(f + w - Au_{\alpha,w}), \end{aligned} \tag{2}$$

where $j(u) = |u|^{q-2}u$ is the duality mapping of $L^q(\Sigma)$ and $\partial_{L^{d/(d-1)}} \mathrm{TV}(u_\alpha)$ denotes the subgradient (see Definition 1 below). Our main goal is to present a proof of the following uniform boundedness result:

Theorem 1 *Assume that for A, f there is a unique solution $u^\dagger$ for $Au = f$ which satisfies the source condition $\mathrm{Ran}(A^*) \cap \partial \, \mathrm{TV}(u^\dagger) \neq \emptyset$. There is some constant $C(q, \sigma, d, \Omega)$ such that if α_n, w_n are sequences of regularization parameters and perturbations for which*

$$\alpha_n \geqslant C(q, \sigma, d, \Omega) \|A^*\| \|w_n\|_{L^q(\Sigma)}^{\sigma-1} \xrightarrow[n\to\infty]{} 0,$$

then the corresponding sequence u_{α_n,w_n} of minimizers is bounded in $L^\infty(\Omega)$, and (possibly up to a subsequence)

$$u_{\alpha_n,w_n} \xrightarrow[n\to\infty]{} u^\dagger \text{ strongly in } L^{\overline{p}}(\Omega) \text{ for all } \overline{p} \in (1, \infty).$$

This result is a combination of Proposition 1, Proposition 2 and Corollary 1 in Sect. 3, which in turn depend on the preliminaries reviewed in Sect. 2.

In the case of denoising $A = \mathrm{Id}$, it is well known that TV regularization satisfies a maximum principle and $\|u_{\alpha,w}\|_{L^\infty} \leqslant \|f + w\|_{L^\infty}$; a proof can be found in [19, Lem. 2.1] for $L^2(\mathbb{R}^2)$ and [37, Prop. 3.6] for the $L^{d/(d-1)}(\mathbb{R}^d)$ case that most closely resembles the situation considered here. Another related result is the nonexpansiveness in L^p norm for denoising with rectilinear anisotropy found in [41]. In comparison, our results work with unbounded measurements and linear operators, and although

the analysis depends on the noise level and parameter choice, the bound on $\|u_{\alpha,w}\|_{L^\infty}$ can be made uniform in the regime of small noise and regularization parameter. Our proof of Theorem 1 hinges on the L^d summability and stability in L^d norm of the subgradients $v_{\alpha,w}$ appearing in (2) and their relation with the perimeter of the level sets of $u_{\alpha,w}$. To the best of our knowledge, the subgradient of TV was first characterized in [3, 4] (see also [6] for further context) relying on the results of [7]. In Sect. 2 we summarize the parts of these results that we will use in the sequel.

Moreover, we are also interested in finding cases outside the assumptions of the result above in which minimizers fail to be bounded. To this end, in Sect. 4 we consider a one-dimensional denoising problem (21), with weights present in both the fidelity term and the total variation which may vanish at the boundary of the interval. By proving its equivalence to a generalized taut string problem, one obtains that the minimizers can be constructed by gluing a few different types of behavior on finitely many subintervals. This in turn allows us to produce explicit unbounded minimizers for radial data in 3D and boundedness conditions for radial powers. Moreover, we believe that results on weighted taut-string formulations can be of independent interest.

Indeed, the taut string approach to one-dimensional total variation minimization has been studied in many works, either from a continuous [29] or discrete [25, 33] point of view. Versions for more general variants have also been considered, like [46] for higher-order total variation, [30] for 'nonuniform tubes' which can be seen as a weight imposed directly on the taut string formulation, and [40] where the graph setting is considered. Likewise, there is a large number of works considering weighted TV denoising seen as having spatially-dependent regularization parameters, and approaches to their automatic selection. This literature is extensive and the choice of weights for particular tasks is beyond the scope of this work, so we only explicitly mention the analytical studies [34] for the one-dimensional case and [8] for higher dimensions and weights that degenerate at the boundary. In contrast, we have not been able to find material combining both weighted TV functionals and taut string characterizations, which motivates our investigations below.

In Sect. 5 we explore the possibility of extending Theorem 1 or parts thereof to the setting of higher-order regularizers related to the total variation. We obtain a boundedness result for the case of infimal convolution of TV and TV^2 regularizers, for which the optimality conditions are closely related to subgradients of TV. Finally, we also present counterexamples suggesting that L^∞ bounds in the case of total generalized variation (TGV) likely need methods different to those considered in this work.

2 Preliminaries

In this section we collect definitions and preliminary results.

Definition 1 For a convex functional $F : X \to \mathbb{R} \cup \{+\infty\}$, where X is a Banach space, the subgradient or subdifferential of F at some element $u \in X$ is the set defined as

$$\partial_X F(u) = \left\{ v \in X' \,\middle|\, F(\bar{u}) - F(u) \geqslant \langle v, \bar{u} - u \rangle_{(X',X)} \text{ for all } \bar{u} \in X \right\},$$

where X' is the dual Banach space of X and $\langle \cdot, \cdot \rangle_{(X',X)}$ denotes the corresponding duality product.

2.1 Subgradient of TV, Pairings, and Slicing

The total variation is defined as

$$\mathrm{TV}(u;\Omega) := \sup \left\{ \int_\Omega u(x) \operatorname{div} z(x) \,\mathrm{d}x \,\middle|\, z \in C_c^1(\Omega;\mathbb{R}^d) \text{ with } |z(x)| \leqslant 1 \text{ for all } x \in \Omega \right\}, \tag{3}$$

which in turn motivates defining the perimeter in Ω of a Lebesgue measurable subset $E \subset \Omega$ as $\mathrm{Per}(E) := \mathrm{TV}(1_E)$, where 1_E is the indicatrix taking the value 1 on E and 0 on $\mathbb{R}^d \setminus E$.

We notice that (3) is the Fenchel conjugate with respect to the dual pair $(L^d, L^{d/(d-1)})$ of the convex set $K := \operatorname{div}\{z \in C_c^1(\Omega;\mathbb{R}^d), \|z\|_{L^\infty(\Omega)} \leqslant 1\}$, so that

$$(\mathrm{TV})^* = \left((\chi_K)^*\right)^* = \chi_{\overline{K}}, \tag{4}$$

where χ_K is the convex characteristic function with value 0 on K and $+\infty$ elsewhere, and $\overline{K}$ is the strong L^d closure of K. This closure in turn satisfies (as stated in [20, Def. 2.2] and proved in [13, Prop. 7]) the identity

$$\begin{aligned} \partial\,\mathrm{TV}(0) &= \overline{K} \\ &= \left\{ \operatorname{div} z \,\middle|\, z \in L^\infty(\Omega;\mathbb{R}^d),\ \|z\|_{L^\infty(\Omega)} \leqslant 1,\ \operatorname{div} z \in L^d(\Omega),\ z \cdot n_{\partial\Omega} = 0 \text{ on } \partial\Omega \right\}, \end{aligned}$$

where the equality $z \cdot n_{\partial\Omega} = 0$ is understood in the sense of the normal trace in $W^{1,d}(\mathrm{div})$, that is,

$$\int_\Omega \operatorname{div} z\, u = -\int_\Omega z \cdot \nabla u \text{ for all } u \in W^{1,d/(d-1)}(\Omega).$$

Now, since TV is positively one-homogeneous, we have (see [38, Lem. A.1], for example) that

$$v \in \partial\,\mathrm{TV}(u) \text{ whenever } v \in \partial\,\mathrm{TV}(0) \text{ and } \int_\Omega vu = \mathrm{TV}(u). \tag{5}$$

It is natural to ask whether one can use that $v = \operatorname{div} z$ for some z to integrate by parts in the last equality, which would formally lead to $z \cdot Du = |Du|$, or $z = Du/|Du|$. However, since z is only in $L^\infty(\Omega;\mathbb{R}^d)$ and not guaranteed to be continuous, there is no immediate meaning to the action of the measure Du on z, and $Du/|Du|$ can only be defined $|Du|$-a.e. as the polar decomposition of Du. This difficulty is mitigated by

defining (as first done in [7]) the product between Du and z as a distribution (z, Du) given by

$$\big\langle (z, Du), \varphi \big\rangle = -\int_\Omega (z \cdot \nabla\varphi) u - \int_\Omega u\varphi \operatorname{div} z \text{ for } \varphi \in C_c^\infty(\Omega). \tag{6}$$

This definition makes sense whenever $u \in \mathrm{BV}(\Omega) \cap L^{d/(d-1)}(\Omega)$ and

$$z \in X_d(\Omega) = \Big\{ z \in L^\infty(\Omega; \mathbb{R}^d) \;\Big|\; \operatorname{div} z \in L^d(\Omega) \Big\} \supset \partial\,\mathrm{TV}(0).$$

Moreover if $\operatorname{supp}\varphi \subset A$ for some open set A also [7, Thm. 1.5]

$$\big|\big\langle (z, Du), \varphi \big\rangle\big| \leqslant \|z\|_{L^\infty(\Omega)} \|\varphi\|_{L^\infty(\Omega)} |Du|(A),$$

so it can be extended to a measure, which in fact is absolutely continuous with respect to $|Du|$, and whose corresponding Radon-Nikodým derivative we denote by $\theta(z, Du, \cdot) \in L^1(\Omega, |Du|)$. Moreover, there exists a generalized normal trace of z on $\partial\Omega$ denoted by $[z, n_{\partial\Omega}] \in L^\infty(\partial\Omega)$ for which the following Green's formula [7, Thm. 1.9] holds:

$$\int_{\partial\Omega} [z, n_{\partial\Omega}] \tilde{u} \, \mathrm{d}\mathcal{H}^{d-1} = (z, D\tilde{u})(\Omega) + \int_\Omega \operatorname{div} z \, \tilde{u} \text{ for all } \tilde{u} \in \mathrm{BV}(\Omega) \cap L^{d/(d-1)}(\Omega). \tag{7}$$

With these definitions in mind (5) may be written [6, Prop. 1.10] as $v \in \partial\,\mathrm{TV}(u)$ if and only if

$$\begin{aligned} v = -\operatorname{div} z \text{ for some } z \in X_d(\Omega) \text{ satisfying} \\ \|z\|_{L^\infty(\Omega)} \leqslant 1,\ [z, n_{\partial\Omega}] = 0,\ \text{and } (z, Du)(\Omega) = \mathrm{TV}(u). \end{aligned} \tag{8}$$

Since we will make extensive use of level sets, we now fix the notation we use for them:

Definition 2 Given a function $u \in L^{d/(d-1)}(\Omega)$ we denote by E^s the upper level set $\{u > s\}$ if $s > 0$, and the lower level set $\{u < s\}$ for $s < 0$, so that $|E^s| < +\infty$ for all $s \in \mathbb{R}\backslash\{0\}$. These two cases can be summarized as

$$E^s := \big\{ x \mid \operatorname{sign}(s)\, u(x) > |s| \big\},$$

Finally, the following characterization of the subgradient of TV in terms of perimeter of level sets is crucial for our results below:

Lemma 1 *Let u and E^s be as in Definition 2. The following assertions are equivalent:*

- $v \in \partial\,\mathrm{TV}(u) \subset L^d(\Omega)$.
- $v \in \partial\,\mathrm{TV}(0)$ *and* $\int_\Omega vu = \mathrm{TV}(u)$.

- $v \in \partial\,\mathrm{TV}(0)$ *and for a.e. s,*

$$\mathrm{Per}(E^s) = \mathrm{sign}(s) \int_{E^s} v. \tag{9}$$

- *For a.e. s, the level sets* E^s *satisfy*

$$E^s \in \arg\min_{E \subset \Omega} \mathrm{Per}(E) - \mathrm{sign}(s) \int_E v.$$

Proof The equivalence between the first two items follows from TV being positively one-homogeneous, and a proof can be found in [38, Lem. A.1], for example. A proof for the other statements in the $L^2(\mathbb{R}^2)$ setting, but which generalizes without modification to the present case, can be found in [21, Prop. 3]. □

2.2 Dual Solutions and Their Stability in One-Homogeneous Regularization

In [36, Prop. 3.3], the dual problem to (1) is computed to be

$$\sup_{\substack{p \in L^{q'}(\Sigma),\\ A^* p \in \partial\,\mathrm{TV}(0) \subset L^d(\Omega)}} \langle p,\ f + w\rangle_{(L^{q'}, L^q)} - \frac{\alpha^{\frac{1}{\sigma-1}}}{\sigma'} \|p\|^{\sigma'}_{L^{q'}(\Sigma)}, \tag{$D_{\alpha,w}$}$$

where $q' = q/(q-1)$ is the conjugate exponent of q, and analogously for σ', and we notice that owing to the strict concavity of the objective, this problem has a unique maximizer. Also, in [36, Prop. 3.6] it is proved that the solutions of ($D_{\alpha,w}$) satisfy

$$\|p_{\alpha,w} - p_{\alpha,0}\|_{L^{q'}(\Sigma)} \leqslant \rho_{L^q,\sigma}\left(\frac{\|w\|_{L^q(\Sigma)}}{2\alpha^{\frac{1}{\sigma-1}}}\right), \tag{10}$$

where $\rho_{L^q,\sigma}$ is defined as the inverse of the function $\mathbb{R}^+ \to \mathbb{R}^+$ defined by

$$t \mapsto \frac{\delta_{\|\cdot\|^{\sigma'}_{L^{q'}}/\sigma'}(t)}{t}$$

where $\delta_{\|\cdot\|^{\sigma'}_{L^{q'}}/\sigma'}$ is in turn the largest modulus of uniform convexity of the functional $\|\cdot\|^{\sigma'}_{L^{q'}(\Sigma)}/\sigma'$, that is, the largest function ψ satisfying for all $u, v \in L^{q'}(\Sigma)$ and $\lambda \in (0, 1)$:

$$\begin{aligned}\frac{1}{\sigma'}\|\lambda u + (1-\lambda)v\|^{\sigma'}_{L^{q'}(\Sigma)} \leqslant\ & \frac{\lambda}{\sigma'}\|u\|^{\sigma'}_{L^{q'}(\Sigma)} \\ & + \frac{1-\lambda}{\sigma'}\|v\|^{\sigma'}_{L^{q'}(\Sigma)} - \lambda(1-\lambda)\psi(\|u-v\|_{L^{q'}(\Sigma)}).\end{aligned}$$

Now, since we have chosen $\sigma = \min(q, 2)$, we have [11, Thm. 5.4.6, Ex. 5.4.7] and [51, Example 2.47]

$$\delta_{L^{q'}}(\epsilon) \geqslant C\epsilon^{\max(2,q/(q-1))} \implies \delta_{\|\cdot\|_{L^{q'}}^{\sigma'}/\sigma'}(t) \geqslant Ct^{\sigma'}$$

where $\delta_{L^{q'}}$ denotes the modulus of uniform convexity of $L^{q'}(\Sigma)$ (which is independent of Σ) defined using points in the unit ball, that is

$$\begin{aligned}\delta_{L^{q'}}(\epsilon) = \inf \Bigg\{ 1 - \left\| \frac{p_1 + p_2}{2} \right\|_{L^{q'}(\Sigma)} \Bigg| \, \|p_1\|_{L^{q'}(\Sigma)} \\ = \|p_2\|_{L^{q'}(\Sigma)} = 1 \text{ and } \|p_1 - p_2\|_{L^{q'}(\Sigma)} \geqslant \epsilon \Bigg\}.\end{aligned}$$

In this setting and defining $v_{\alpha,w} := A^* p_{\alpha,w} \in \partial\,\mathrm{TV}(u_{\alpha,w})$ as in (2), we can deduce from (10) that

$$\begin{aligned}\|v_{\alpha,w} - v_{\alpha,0}\|_{L^d(\Omega)} \leqslant C\|A^*\| \left(\frac{\|w\|_{L^q(\Sigma)}}{2\alpha^{\frac{1}{\sigma-1}}} \right)^{\frac{1}{\sigma'-1}} = C_{q,\sigma}\|A^*\| \left(\frac{\|w\|_{L^q(\Sigma)}^{\frac{1}{\sigma'-1}}}{\alpha^{\frac{1}{\sigma-1}\frac{1}{\sigma'-1}}} \right) \\ = C_{q,\sigma}\|A^*\| \frac{\|w\|_{L^q(\Sigma)}^{\sigma-1}}{\alpha},\end{aligned} \tag{11}$$

which tells us that the threshold to control the effect of the noise on the dual variables is a direct generalization of the linear parameter choice that would arise in the case $\sigma = q = 2$.

To know more about the asymptotic behavior of $v_{\alpha,w}$ as both α and w vanish one needs an additional assumption. Whenever $u^\dagger$ denotes a solution of $Au = f$ of minimal TV among such solutions, we say that the source condition is satisfied if there exists $p \in L^{q'}(\Omega)$ such that

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

Now, since TV is not strictly convex, $u^\dagger$ may not be unique, but arguing as in [38, Rem. 3] we have that if $\hat{u}^\dagger$ is another such solution and we have (12), then $A^* p \in \partial\,\mathrm{TV}(\hat{u}^\dagger)$ as well, for the same source element p. For our purposes, the significance of this source condition comes from the fact that it guarantees that the formal dual problem $(D_{0,0})$ obtained by setting $\alpha = 0$ and $w = 0$ in $(D_{\alpha,w})$ has at least one maximizer, namely the source element p.

Under this source condition assumption and taking a sequence $\alpha_n \to 0$ we have in the case where noise is not present ($w = 0$) the convergence

$$v_{\alpha_n,0} = A^* p_{\alpha,0} \to A^* p_{0,0} =: v_{0,0} \text{ strongly in } L^d(\Omega), \tag{13}$$

where $p_{0,0}$ is the element of $L^{q'}(\Omega)$ of minimal norm among those satisfying (12), which is unique. This convergence is proved (see [36, Prop. 3.4] for a detailed argument) by testing the optimality of $p_{\alpha,0}$ and $p_{0,0}$ in $(D_{\alpha,0})$ and $(D_{0,0})$ with respect to each other to obtain both $p_{\alpha,0} \rightharpoonup p_{0,0}$ weakly in $L^{q'}(\Sigma)$ and $\|p_{\alpha,0}\|_{L^{q'}(\Sigma)} \leqslant \|p_{0,0}\|_{L^{q'}(\Sigma)}$, which together imply strong convergence by using the Radon-Riesz property of $L^{q'}(\Sigma)$. Finally, combining this convergence with (11), we also have that if $w_n \in L^q(\Sigma)$ are such that

$$\frac{\|w_n\|_{L^q(\Sigma)}^{\sigma-1}}{\alpha_n} \to 0, \text{ then also } v_{\alpha_n,w_n} \to A^* p_{0,0} \text{ strongly in } L^d(\Omega).$$

3 Boundedness, Uniform Boundedness, and Strong Convergence

We start with a direct proof of boundedness of minimizers $u_{\alpha,w}$ of (1) which, using the results cited in the previous section, can be made uniform for small noise and strong enough regularization. It relies on studying the level sets

$$E^s_{\alpha,w} := \left\{x \mid \operatorname{sign}(s)\, u_{\alpha,w}(x) > |s|\right\}.$$

Proposition 1 *Let $\Omega = \mathbb{R}^d$. Then minimizers $u_{\alpha,w}$ of (1) belong to $L^\infty(\mathbb{R}^d)$. Moreover, under the source condition $A^* p_0 \in \partial\,\mathrm{TV}(u^\dagger)$ and a parameter choice satisfying the condition*

$$\|w\|_{L^q(\Sigma)}^{\sigma-1} \leqslant \frac{\eta\alpha}{C_{q,\sigma}\|A^*\|} \tag{14}$$

where $C_{q,\sigma}$ is the constant in (11) and $\eta < \Theta_d$ is the constant of the isoperimetric inequality $\mathrm{Per}(E) \geqslant \Theta_d |E|^{(d-1)/d}$ for E of finite perimeter, we have a bound for $\|u_{\alpha,w}\|_{L^\infty(\mathbb{R}^d)}$ which is uniform in α and w.

Proof We first show the claim for fixed α, w. To do this, we make use of $(D_{\alpha,w})$: the problem has a solution $p_{\alpha,w}$, we have strong duality and the optimality condition

$$v_{\alpha,w} := A^* p_{\alpha,w} \in \partial\,\mathrm{TV}(u_{\alpha,w}).$$

In what follows, we assume without loss of generality that $s > 0$ so that $E^s_{\alpha,w} := \{u_{\alpha,w} > s\}$. From (9) and the Hölder inequality we can derive for a.e. $s > 0$ the estimate

$$\mathrm{Per}(E^s_{\alpha,w}) = \int_{E^s_{\alpha,w}} v_{\alpha,w} \leqslant |E^s_{\alpha,w}|^{(d-1)/d} \left(\int_{E^s_{\alpha,w}} |v_{\alpha,w}|^d\right)^{1/d}.$$

Using the isoperimetric inequality $\mathrm{Per}(E^s_{\alpha,w}) \geqslant \Theta_d |E^s_{\alpha,w}|^{(d-1)/d}$, we obtain

$$\mathrm{Per}(E^s_{\alpha,w}) \leqslant \frac{1}{\Theta_d} \mathrm{Per}(E^s_{\alpha,w}) \left(\int_{E^s_{\alpha,w}} |v_{\alpha,w}|^d \right)^{1/d}. \tag{15}$$

Now, $v_{\alpha,w} \in L^d(\mathbb{R}^d)$ so for any $\varepsilon > 0$, there exists a $\delta > 0$ such that for sets E with $|E| \leqslant \delta$, $\int_E |v_{\alpha,w}|^d \leqslant \varepsilon^d$. In particular, if $|E^s_{\alpha,w}| \leqslant \delta$, it implies

$$\mathrm{Per}(E^s_{\alpha,w}) \leqslant \frac{\varepsilon}{\Theta_d} \mathrm{Per}(E^s_{\alpha,w}),$$

which is not possible for ε too small if $|E^s_{\alpha,w}| > 0$. Note that the bound on δ does not depend on s, only on $v_{\alpha,w}$, which means that we have a uniform positive lower bound on the mass of *every* level set $E^s_{\alpha,w}$. From the layer-cake formula for $L^{d/(d-1)}(\mathbb{R}^d)$ functions [42, Thm. 1.13]

$$\int_{\mathbb{R}^d} (u^+_{\alpha,w})^{d/(d-1)} = \int_0^\infty \frac{d}{d-1} s^{1/(d-1)} |E^s_{\alpha,w}| \, \mathrm{d}s, \text{ where } u^+_{\alpha,w} = \max(u_{\alpha,w}, 0)$$

we conclude that there must exist $s_0 > 0$ such that for $s \geqslant s_0$, $|E^s_{\alpha,w}| = 0$, which means that $u_{\alpha,w} \leqslant s_0$ a.e. in $\mathbb{R}^d$. We prove similarly that $u^-_{\alpha,w} = \max(0, -u_{\alpha,w})$ is bounded.

Finally, let us assume that (α_n, w_n) is a sequence of regularization parameters and noises for which (14) holds. From $v_{\alpha_n,0} \to v_{0,0}$ for any sequence $\alpha_n \to 0$ (see (13)), one infers that the family $\{v_{\alpha,0}\}_{\alpha>0}$ is equi-integrable in $L^d(\mathbb{R}^d)$. We want, as before, to estimate $\int_{E^s_{\alpha,w}} |v_{\alpha,w}|^d$. This can be done by writing

$$\begin{aligned}\left(\int_{E^s_{\alpha,w}} |v_{\alpha,w}|^d \right)^{1/d} &\leqslant \left(\int_{E^s_{\alpha,w}} |v_{\alpha,0}|^d \right)^{1/d} + \left(\int_{E^s_{\alpha,w}} |v_{\alpha,w} - v_{\alpha,0}|^d \right)^{1/d} \\ &\leqslant \left(\int_{E^s_{\alpha,w}} |v_{\alpha,0}|^d \right)^{1/d} + \|v_{\alpha,w} - v_{\alpha,0}\|_{L^d(\mathbb{R}^d)}.\end{aligned}$$

Now, using the equi-integrability of $v_{\alpha,0}$, one can, for any ε, find δ so that as soon as $|E^s_{\alpha,w}| \leqslant \delta$, the first term of the right hand side is bounded by ε. On the other hand, the second term is, independently from δ, bounded by $\eta < \Theta_d$, as a consequence of (11) and (14). We conclude that as soon as $|E^s_{\alpha,w}| \leqslant \delta$, we have

$$\mathrm{Per}(E^s_{\alpha,w}) \leqslant \frac{\varepsilon + \eta}{\Theta_d} \mathrm{Per}(E^s_{\alpha,w})$$

which is still not possible for ε too small (independent of s and α, w satisfying (14)) if $|E^s_{\alpha,w}| > 0$. □

Remark 1 By a similar argument on the $E^s_{\alpha,w}$ (see [38, Lem. 5]), we can actually show that $u_{\alpha,w}$ has compact support, so that $u^+_{\alpha,w} \in L^1(\mathbb{R}^d)$ and we could use the (simpler) layer cake formula in L^1:

$$\int_{\mathbb{R}^d} u^+_{\alpha,w} = \int_0^\infty |E^s_{\alpha,w}| \, \mathrm{d}s$$

which would provide the same contradiction.

Remark 2 The parameter choice condition (14) does not necessarily imply convergence of the dual variables. This is the case when

$$\frac{\|w_n\|^{\sigma-1}_{L^q(\Sigma)}}{\alpha_n} \to 0 \text{ as } n \to \infty$$

so that, as remarked in Sect. 2.2, in fact (13) and (11) imply that

$$v_{\alpha_n,w_n} \to A^* p_0 \text{ strongly in } L^d(\Omega).$$

In particular, this implies that the family $(v_{\alpha_n,w_n})_n$ is equi-integrable in $L^d(\mathbb{R}^d)$, which means that in the reasoning coming after (15), the δ can be chosen independent of n, which would simplify the proof in this more restrictive case.

Remark 3 In the case of denoising in the plane with the Rudin-Osher-Fatemi model [48], for which $d = m = 2$, $q = 2$ and $A = \mathrm{Id} : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$, we have that $v_{\alpha,w} = \alpha^{-1}(f + w - u_{\alpha,w})$ and Proposition 1 applies as soon as $f + w \in L^2(\mathbb{R}^2)$.

Remark 4 The same proof shows that if the source condition (12) is satisfied, then $u^\dagger \in L^\infty(\mathbb{R}^d)$.

We consider now the case when Ω is a bounded Lipschitz domain, which leads to two cases for the functional in (1). First, we can consider the total variation $\mathrm{TV}(u; \mathbb{R}^d)$ of the extension by zero of u from Ω to the whole $\mathbb{R}^d$, which we refer to as homogeneous Dirichlet boundary conditions. Second, we can consider the variation $\mathrm{TV}(u; \Omega)$ which we refer to as homogeneous Neumann boundary conditions, since the test functions in (3) are compactly supported in Ω.

Proposition 2 *Proposition 1 also holds for bounded Lipschitz domains Ω and either homogeneous Dirichlet or Neumann boundary conditions on $u_{\alpha,w}$.*

Proof n the Dirichlet case we work with the functional in (1) but consider $\mathrm{TV}(\cdot; \mathbb{R}^d)$ among functions in $L^{d/(d-1)}(\mathbb{R}^d)$ constrained to vanish on $\mathbb{R}^d \backslash \Omega$. This can be translated to a formulation on $L^{d/(d-1)}(\Omega)$ by considering $\mathcal{E} : L^{d/(d-1)}(\Omega) \to L^{d/(d-1)}(\mathbb{R}^d)$ the extension by zero on $\mathbb{R}^d \backslash \Omega$ and $\mathrm{TV}(\cdot\,; \mathbb{R}^d) \circ \mathcal{E}$. By [2, Cor. 3.89] and using that Ω is a Lipschitz domain we have that

$$\mathrm{TV}(\mathcal{E}u; \mathbb{R}^d) = \mathrm{TV}(u; \Omega) + \int_{\partial\Omega} |u| \, \mathrm{d}\mathcal{H}^{d-1},$$

where we emphasize that the first total variation is computed in $\mathbb{R}^d$ whereas the second one in Ω as in (3). The values of u at the boundary are understood in the sense of traces.

In this situation, since $\mathcal{E}$ is linear we can consider the composition $\mathrm{TV} \circ \mathcal{E}$ as a convex positively homogeneous functional in its own right, which gives us that indeed,

$$v_{\alpha,w} := A^* p_{\alpha,w} \in \partial_{L^{d/(d-1)}(\Omega)}(\mathrm{TV} \circ \mathcal{E})(u_{\alpha,w}) \subset L^d(\Omega),$$

where $p_{\alpha,w}$ is defined in (2). By general properties of one-homogeneous functionals we then have (as in the first two items of Lemma 1) that

$$\mathrm{TV}(\mathcal{E}u_{\alpha,w}; \mathbb{R}^d) = \int_\Omega v_{\alpha,w} u_{\alpha,w}.$$

Moreover, we can argue exactly as in [21, Prop. 3] (which uses more properties of the subgradient than the equality above) to obtain that also in this case

$$\mathrm{Per}(E^s_{\alpha,w}; \mathbb{R}^d) = \mathrm{sign}(s) \int_{E^s_{\alpha,w}} v_{\alpha,w},$$

where the level sets $E^s_{\alpha,w}$ are defined as in Proposition 1, and from which we can follow the rest of the proof exactly, since $E^s_{\alpha,w} \subset \Omega$.

In the Neumann case we consider TV as $\mathrm{TV}(u; \Omega)$. This parallels what is done in [36, Sec. 6] and with more details in the 2D case in [38, Sec. 4.3]. In this case, one uses the estimate (see [38, Sec. 4.3] for a proof)

$$C_\Omega \, \mathrm{Per}(F; \Omega) \geqslant \frac{|F|^{\frac{d-1}{d}} |\Omega \setminus F|^{\frac{d-1}{d}}}{|\Omega|^{\frac{d-1}{d}}}, \tag{16}$$

where C_Ω is the constant of the Poincaré-Sobolev inequality

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

To see that (16) can play the role of the isoperimetric inequality, notice that we only need to use this inequality for large values of $|s|$ and sets of small measure. Therefore we may assume

$$|E^s_{\alpha,w}| \leqslant \frac{|\Omega|}{2}, \text{ so that } C_\Omega \, \mathrm{Per}(E^s_{\alpha,w}; \Omega) \geqslant \frac{1}{2^{\frac{d-1}{d}}} |E^s_{\alpha,w}|^{\frac{d-1}{d}}.$$

□

Remark 5 Let us assume that A and A^* preserve boundedness, that is, $Au \in L^\infty(\Sigma)$ for all $u \in L^\infty(\Omega)$ and $A^* p \in L^\infty(\Omega)$ whenever $p \in L^\infty(\Sigma)$. Then, if $f, w \in L^\infty(\Sigma)$ and in the situation of Proposition 1, the optimality condition (2) implies that $v_{\alpha,w} \in L^\infty(\Omega)$ as well. One can then use strong regularity results (see [44, Thm. 3.1]

or [28, Thm. 3.6(b)], for example) to obtain that $\partial\{u_{\alpha,w} > s\} \in C^{1,\gamma}$ for all $\gamma < 1/4$, provided that $d \leqslant 7$. Moreover, since our estimates are uniform under the assumptions of Theorem 1, this regularity can also be made uniform along a sequence [53, Secs. 1.9, 1.10] as well.

The assumption of both A and A^* preserving boundedness is easily seen to be satisfied for convolution operators with regular enough kernels, since in this case A^* is of the same type. However, it also holds for other commonly used operators. As an example, let us consider the case of the Radon transform $\mathcal{R}$ of functions on a bounded domain $\Omega \subset \mathbb{R}^d$. In this case one can consider $\Sigma = \mathbb{S}^{d-1} \times (-R, R)$ with $R > 0$ large enough so that $\Omega \subset B(0, R)$ and

$$A : L^{d/(d-1)}(\Omega) \subset L^{2d/(2d-1)}(\mathbb{R}^d) \xrightarrow{\mathcal{R}} L^2(\mathbb{S}^{d-1} \times \mathbb{R}) \xrightarrow{r_{\mathbb{S}^{d-1}\times(-R,R)}} L^2\big(\mathbb{S}^{d-1} \times (-R, R)\big),$$

the last map being restriction, since R is continuous between the middle two spaces [45, Thm. 1]. Now, if $u \in L^\infty(\Omega)$ then clearly $Au \in L^\infty\big(\mathbb{S}^{d-1} \times (-R, R)\big)$ as well, since the integrals on planes in the definition of $\mathcal{R}$ are then on domains of uniformly bounded Hausdorff measure $\mathcal{H}^{d-1}$, that is

$$\begin{aligned} |Au(\theta, t)| &= \left| \int_{\theta^\perp} u(t\theta + y)\, \mathrm{d}\mathcal{H}^{d-1}(y) \right| \\ &\leqslant \|u\|_{L^\infty(\Omega)}\, \mathcal{H}^{d-1}\big(B(0, R) \cap \{x_1 = 0\}\big) \text{ for a.e. } (\theta, t). \end{aligned}$$

Similarly, A^* is then extension by zero composed with the backprojection integral operator, so we have

$$|A^* p(x)| = \left| \int_{\mathbb{S}^{d-1}} p(\theta, x \cdot \theta)\, \mathrm{d}\mathcal{H}^{d-1}(\theta) \right| \leqslant \|p\|_{L^\infty(\mathbb{S}^{d-1}\times(-R,R))} \mathcal{H}^{d-1}(\mathbb{S}^{d-1}) \text{ for a.e. } x.$$

Observe that for the case of $\Omega \subset \mathbb{R}^2$ and convolutions with L^2 kernels boundedness is immediate, since Young's inequality for convolutions used in (2) directly implies $v_{\alpha,w} \in L^\infty(\Omega)$ as soon as $u_{\alpha,w}, f, w \in L^2(\Omega)$.

Corollary 1 *Under the assumptions of Proposition 1, for a sequence of minimizers $\{u_{\alpha_n,w_n}\}_n$ with α_n, w_n satisfying (14) and $\alpha_n \to 0$ and $w_n \to 0$ as $n \to \infty$, we have that, up to a subsequence,*

$$u_{\alpha_n,w_n} \xrightarrow[n\to\infty]{} u^\dagger \text{ strongly in } L^{\overline{p}}(\Omega) \text{ for all } \overline{p} \in (1, \infty), \tag{17}$$

where $u^\dagger$ is an exact solution of $Au = f$ with minimal TV *among such solutions. If there is only one exact solution, then the whole sequence converges to it in the same fashion.*

Proof We first notice that the parameter choice (14) is less restrictive than the one needed in [36, Prop. 3.1] which provides strong convergence to some $u^\dagger$ in $L^{\hat{p}}_{\mathrm{loc}}(\Omega)$ for $\hat{p} \in (1, d/(d-1))$ and up to a subsequence by a basic compactness argument.

Moreover, if $\Omega = \mathbb{R}^d$, we can apply [36, Lem. 5.1] to obtain that all of the u_{α_n,w_n} and $u^\dagger$ are supported inside a common ball $B(0, R)$ for some $R > 0$. If, in contrast, Ω is bounded, then $\Omega \subset B(0, R)$ and we may extend to the latter by zero.

For this subsequence (which we do not relabel) and $\hat{p} \leqslant \overline{p}$ we have

$$\int_\Omega |u_{\alpha_n,w_n} - u^\dagger|^{\overline{p}} \leqslant \|u_{\alpha_n,w_n} - u^\dagger\|^{\hat{p}}_{L^{\hat{p}}(B(0,R))} \|u_{\alpha_n,w_n} - u^\dagger\|^{\overline{p}-\hat{p}}_{L^\infty(\Omega)}, \tag{18}$$

which using Proposition 1 immediately implies (17).

Finally, if the minimal TV solution of $Au = f$ is unique, any subsequence has a further subsequence converging to this unique solution, so the whole sequence must in turn converge to it. □

Remark 6 In the plane, under the same source condition (12), a parameter choice equivalent to the one we use for $d = 2$, and additional assumptions for both A and $u^\dagger$, linear convergence rates are proved in the L^2 setting in [54, Thm. 4.12], that is $\|u_{\alpha_n,w_n} - u^\dagger\|_{L^2(\Omega)} = O(\alpha_n)$ (observe that the parameter choice forces $\|w_n\| = O(\alpha_n)$). Therefore, this result can be combined with ours to obtain also a convergence rate of order $2/\overline{p}$ in $L^{\overline{p}}(\Omega)$. To see this, we just notice that by Remark 4 we have $u^\dagger \in L^\infty(\Omega)$ in addition to the uniform bound on $\|u_{\alpha_n,w_n}\|_{L^\infty(\Omega)}$, so that as in (18) we have

$$\|u_{\alpha_n,w_n} - u^\dagger\|_{L^{\overline{p}}(\Omega)} \leqslant \|u_{\alpha_n,w_n} - u^\dagger\|^{2/\overline{p}}_{L^2(\Omega)} \|u_{\alpha_n,w_n} - u^\dagger\|^{(\overline{p}-2)/\overline{p}}_{L^\infty(\Omega)} = O(\alpha_n^{2/\overline{p}}).$$

One might wonder if strong convergence in $L^\infty(\Omega)$ is possible. In fact, it is not:

Example 1 Consider (as in Remark 3) denoising in the plane $\mathbb{R}^2$ with parameter $\alpha_n = 1/n$ of $f + w_n$, where for arbitrarily small $c > 0$ we define

$$u_{0,0} = f = 1_{B(0,1)} \text{ and } w_n = 1_{B\left(0,\sqrt{1+\frac{c}{n}}\right)} - 1_{B(0,1)},$$
$$\text{so that } \|w_n\|_{L^2(\mathbb{R}^2)} = \frac{\sqrt{\pi}c}{n} \text{ and } \frac{\|w_n\|_{L^2(\mathbb{R}^2)}}{\alpha_n} = \sqrt{\pi}c.$$

Then since $f + w_n = 1_{B\left(0,\sqrt{1+\frac{c}{n}}\right)}$ we know [20, Sec. 2.2.3] that u_{α_n,w_n} is also proportional to $1_{B\left(0,\sqrt{1+\frac{c}{n}}\right)}$, making $L^\infty(\mathbb{R}^2)$ convergence impossible. Notice that this situation does not change for a more aggressive parameter choice.

3.1 Results with Density Estimates

The combination of the source condition $\mathrm{Ran}(A^*) \cap \partial\,\mathrm{TV}(u^\dagger) \neq \emptyset$ and the parameter choice (14) leads to uniform weak regularity estimates for the level sets of $u_{\alpha,w}$, and boundedness may also be deduced from those. More precisely, recalling that

$$E^s_{\alpha,w} = \left\{x \mid \mathrm{sign}(s)\, u_{\alpha,w}(x) > |s|\right\},$$

it is proved in [36, Thm. 4.5] that

$$|B(x,r)\cap E^s_{\alpha,w}| \geqslant C|B(x,r)| \text{ and } |B(x,r)\setminus E^s_{\alpha,w}| \geqslant C|B(x,r)|, \tag{19}$$

for $x \in \partial E^s_{\alpha,w}$ and $r \leqslant r_0$, where r_0 and C are independent of x, n and s, and where we have taken a representative of $E^s_{\alpha,w}$ for which the topological boundary equals the support of the derivative of its indicatrix [43, Prop. 12.19], that is $\partial E^s_{\alpha,w} = \operatorname{supp} D1_{E^s_{\alpha,w}}$. In turn, this support can be characterized [43, Prop. 12.19] by the property

$$\operatorname{supp} D1_{E^s_{\alpha,w}} = \left\{ x \in \mathbb{R}^d \,\middle|\, 0 < \frac{|E^s_{\alpha,w}\cap B(x,r)|}{|B(x,r)|} < 1 \text{ for all } r > 0 \right\},$$

where we note that it could be that these quotients tend to 0 or 1 as $r \to 0$.

We refer to the inequalities in (19) as inner and outer density estimates respectively, and the combination of both as $E^s_{\alpha,w}$ satisfying uniform density estimates. Such estimates are the central tool for the results of convergence of level set boundaries in Hausdorff distance in [21, 36–38].

The proof of (19) is more involved than those in the previous section, but once it is obtained, boundedness of minimizers follows promptly:

Proposition 3 *Assume that the $E^s_{\alpha,w}$ satisfy uniform density estimates at scales $r \leqslant r_0$ with constant C, and that α is chosen in terms of $\|w\|_{L^q(\Sigma)}$ so that $u_{\alpha,w}$ is bounded in $L^{d/(d-1)}(\Omega)$ for some $p \geqslant 1$. Then, in fact $\{u_{\alpha,w}\}$ is uniformly bounded in $L^\infty(\Omega)$. If additionally we have a sequence $u_{\alpha_n,w_n} \to u^\dagger$ strongly in $L^{\hat{p}}(\Omega)$ for some $\hat{p}$, then also $u_{\alpha_n,w_n} \to u^\dagger$ strongly in $L^{\overline{p}}(\Omega)$ for all $\overline{p}$ such that $\hat{p} \leqslant \overline{p} < \infty$.*

Proof If $u_{\alpha,w}$ is not bounded in $L^\infty(\Omega)$, then for every $M > 0$ there are $\alpha := \alpha(M)$ and $w := w(M)$ for which $|E^M_{\alpha,w}| > 0$. The inner density estimate at scale r_0 for some point $x \in \partial E^M_{\alpha,w}$ then reads

$$|B(x,r_0)\cap E^M_{\alpha,w}| \geqslant C|B(x,r_0)|,$$

which implies

$$\int_\Omega |u_{\alpha,w}|^{d/(d-1)} \geqslant \int_{B(x,r_0)\cap E^M_{\alpha,w}} |u_{\alpha,w}|^{d/(d-1)} \geqslant C|B(x,r_0)|M^{d/(d-1)},$$

which, since r_0 is fixed, contradicts the fact that the family $\{u_{\alpha,w}\}$ is bounded in $L^{d/(d-1)}(\Omega)$. For the second statement, we can argue as in (18). □

Remark 7 In the Dirichlet case of Proposition 2 we have assumed only that Ω is a Lipschitz domain without further restrictions. In this context, unless Ω is convex, it is not true that E^s is a minimizer of

$$\operatorname{Per}(E) - \int_E \mathcal{E}v_{\alpha,w} \text{ among } E \subset \mathbb{R}^d,$$

since these may extend beyond Ω, while the functional $u \mapsto (\mathrm{TV} \circ \mathcal{E})(u) - \int_\Omega v_{\alpha,w} u$ is only sensitive to variations supported in Ω. This kind of variational problem is used in [36, 38] to obtain the density estimates (19), but for that one needs (see [38, Lem. 9]) to extend $v_{\alpha,w}$ by a variational curvature of Ω minorized by a function in $L^d(\mathbb{R}^d)$. The existence of such a curvature is an additional restriction on Ω and it is not satisfied for domains with inner corners in $\mathbb{R}^2$, for example.

One can also make the bound slightly more explicit in terms of the constant and scale of the density estimates:

Remark 8 Let u have level sets $E^s := \{\mathrm{sign}(s)u > |s|\}$ satisfying density estimates with constant C at scale r_0. Then, we have

$$\|u\|_{L^\infty} \leqslant \frac{1}{C^{1/p} \cdot |B(0, r_0)|^{1/p}} \|u\|_{L^p}.$$

This is a direct application of the Markov inequality. Indeed, since there is always a boundary point x_0, the level sets E^s which are nonempty must satisfy $|E^s| \geqslant C|B(x_0, r_0)| = C|B(0, r_0)|$, which implies

$$\int_\Omega |u|^p \geqslant |s|^p |E^s| \geqslant |s|^p C|B(0, r_0)|.$$

In particular, this implies

$$\frac{1}{C^{1/p}|B(0, r_0)|^{1/p}} \|u\|_{L^p} \geqslant |s|$$

meaning that as soon as $|s|$ exceeds $\frac{1}{C^{1/p}|B(0,r_0)|^{1/p}} \|u\|_{L^p}$, E^s must be empty.

Corollary 2 *Under the assumptions of Proposition 3 above, let s with $|s| > \|u^\dagger\|_{L^\infty}$. Then, we have*

$$\limsup_{n\to\infty} \partial E^s_{\alpha_n, w_n} = \emptyset,$$

where $\limsup \partial E^s_{\alpha_n, w_n}$ *consists [47, Definition 4.1] of all of limits of subsequences of points in* $\partial E^s_{\alpha_n, w_n}$.

Proof By the convergence $u_{\alpha_n, w_n} \to u^\dagger$ in $L^{\hat{q}}(\Omega)$, if $|s| > \|u^\dagger\|_{L^\infty}$ we have $|E^s_{\alpha_n, w_n}| \to 0$. If we had $x \in \limsup_n \partial E^s_{\alpha_n, w_n}$, we could produce a subsequence $\{x_n\}_n$ in Ω with $x_n \in \partial E^s_{\alpha_n, w_n}$ and $x_n \to x$ for $x \in \overline{\Omega}$. Using the inner density estimate for some $r_0 > 0$, we would end up with

$$|E^s_{\alpha_n, w_n}| \geqslant |B(x_{\alpha_n}, r_0) \cap E^s_{\alpha_n, w_n}| \geqslant C|B(0, r_0)|,$$

which contradicts $|E^s_{\alpha_n, w_n}| \to 0$. □

4 Taut String with Weights and Unbounded Radial Examples

In this section we consider denoising of one-dimensional data with a modified Rudin-Osher-Fatemi (ROF) functional with weights in both terms, obtaining a taut string characterization of the solutions which reduces the problem to finding finitely many parameters. The main difficulty is that we allow weights which may degenerate at the boundary, which forces the use of weighted function spaces in all the arguments.

4.1 One-Dimensional Weighted ROF Problem and Optimality Condition

We start with two weights ϕ, ρ on the interval $(0, 1)$ for which

$$\phi, \rho > 0 \text{ on } (0, 1),\ \rho \in \mathrm{Lip}(0, 1),\ \phi \in C(0, 1) \cap L^\infty(0, 1),\ \text{and } \frac{1}{\phi} \in L^\infty_{\mathrm{loc}}(0, 1), \tag{20}$$

where we notice that it could be that $\phi(x) \to 0$ or $\rho(x) \to 0$ as $x \to 0$ or $x \to 1$. Using these weights we consider the weighted denoising minimization problem

$$\inf_{u \in L^2_\phi(0,1)} \int_0^1 \big(u(x) - f(x)\big)^2 \phi(x)\,\mathrm{d}x + \alpha\,\mathrm{TV}_\rho(u), \tag{21}$$

where we define the weighted total variation with weight ρ in the usual way (see [1, 9], for example) as

$$\mathrm{TV}_\rho(u) := \sup\left\{ \int_0^1 u(x) z'(x)\,\mathrm{d}x \,\middle|\, z \in C^1_c(0, 1) \text{ with } |z(x)| \leqslant \rho(x) \text{ for all } x \right\} \tag{22}$$

and the problem is considered in the weighted Lebesgue space

$$L^2_\phi(0, 1) := \left\{ u \,\middle|\, \|u\|^2_{L^2_\phi} := \int_0^1 u^2 \phi < \infty \right\},$$

which is a Hilbert space when considered with the inner product

$$\langle v, u \rangle_{L^2_\phi} = \int_0^1 \phi\, vu.$$

The predual variables and optimality conditions for (21) will then naturally be formulated on a weighted Sobolev space, namely

$$W^{1,2}_{1,1/\phi}(0, 1) := \left\{ U \in L^2(0, 1) \,\middle|\, U' \in L^2_{1/\phi}(0, 1) \right\}.$$

The first step is to see that under the assumptions (20), the denoising problem is still well posed:

Proposition 4 *Assume that $f \in L^2_\phi(0,1)$. Then, there is a unique minimizer u_α of (21).*

Proof With the assumptions on ϕ, ρ, these weights stay away from zero on any compact subset of $(0,1)$, which means that we have, for an open subset $A \subset (0,1)$, for which in particular $\overline{A} \subset\subset (0,1)$, a constant C_A such that for every $u \in L^2_\phi(0,1)$

$$\int_A \big(u(x) - f(x)\big)^2 \phi(x)\,\mathrm{d}x + \alpha\,\mathrm{TV}_\rho(u;A) \geqslant C_A \left(\int_A \big(u(x) - f(x)\big)^2\,\mathrm{d}x + \alpha\,\mathrm{TV}(u;A) \right).$$

This means that along a minimizing sequence $\{u_n\}_n$ for (21), we have

$$\sup_n \int_A \big(u_n(x) - f(x)\big)^2\,\mathrm{d}x + \alpha\,\mathrm{TV}(u_n;A) < +\infty.$$

Applying Cauchy-Schwarz on the first term and since $\alpha > 0$, we obtain

$$\sup_n \int_A |u_n| + \alpha\,\mathrm{TV}(u_n;A) < +\infty$$

which allows applying [2, Thm. 3.23] to conclude that a not relabeled subsequence of $\{u_n\}_n$ converges in $L^1_{\mathrm{loc}}(0,1)$ to some $u_\infty \in \mathrm{BV}_{\mathrm{loc}}(0,1)$. Now, arguing as in [17, Prop. 1.3.1] we have that TV_ρ is lower semicontinuous with respect to strong L^1_{loc} convergence. To see this, let $z \in C^1_c(0,1)$ with $|z| \leqslant \rho$, so that we have $u_n \xrightarrow{L^1(\mathrm{supp}(z))} u_\infty$ and

$$\int_0^1 u_n z' \to \int_0^1 u_\infty z'.$$

Since the left hand side is bounded by $\mathrm{TV}_\rho(u_n)$, this implies in particular that for any such z,

$$\liminf_n \mathrm{TV}_\rho(u_n) \geqslant \int_0^1 u_\infty z'.$$

Taking the supremum over z as in (22), we obtain the semicontinuity

$$\liminf_n \mathrm{TV}_\rho(u_n) \geqslant \mathrm{TV}_\rho(u_\infty). \tag{23}$$

Moreover, we may take a further subsequence of $\{u_n\}_n$ converging weakly in $L^2_\phi(0,1)$ to a limit which clearly must be again u_∞, and the first term of (21) is lower semi-continuous with respect to this convergence since it involves only the squared norm of $L^2_\phi(0,1)$. This and (23) show that u_∞ is actually a minimizer of (21). □

We have that u realizes the infimum in (21) if and only if

$$\frac{f-u}{\alpha} \in \partial_{L^2_\phi} \mathrm{TV}_\rho(u) \tag{24}$$

where the subgradient means that $v \in \partial_{L^2_\phi} \mathrm{TV}_\rho(u)$ if and only if

$$\mathrm{TV}_\rho(u) + \int_0^1 v(\tilde{u}-u)\phi \leqslant \mathrm{TV}_\rho(\tilde{u}) \text{ for all } \tilde{u} \in \mathrm{BV}_\rho(0,1) \cap L^2_\phi(0,1).$$

This set is characterized (with assumptions on the weights covering ours) in [8, Lem. 2.4], which in one dimension yields the following generalization of (8):

$$\begin{aligned} &\text{There exists } \xi \in L^\infty(0,1) \text{ with } |\xi| \leqslant 1 \text{ and } (\rho\xi)' \in L^2_{1/\phi}(0,1), \\ &\text{for which } \int_0^1 v\tilde{u}\phi = (\rho\xi, D\tilde{u})(0,1) \text{ for all } \tilde{u} \in \mathrm{BV}_\rho(0,1) \cap L^2_\phi(0,1) \\ &\text{and also } \int_0^1 vu\phi = (\rho\xi, Du)(0,1) = \mathrm{TV}_\rho(u), \end{aligned} \tag{25}$$

where it is to be noted that in comparison to (8), we always consider the product $\rho\xi$ together, and generally avoid differentiating ρ alone. Like for (6), the second line of (25) generalizes integration by parts when $\tilde{u} \in C_c^\infty(0,1)$, so we have that $v\phi = -(\rho\xi)'$ in the sense of distributions, and $\rho\xi \in W^{1,2}_{1,1/\phi}(0,1)$. In analogy to (7), it is tempting to interpret this equality in terms of boundary values. However, as remarked in [8, Below Lemma 2.4], to which extent this is possible depends on further properties of ϕ. In our one-dimensional case and noting that because $\phi \in L^\infty(0,1) \cap C(0,1)$ the inverse $1/\phi$ remains bounded away from zero on the closed interval, and we have

$$W^{1,2}_{1,1/\phi}(0,1) \subset W^{1,2}(0,1) \subset C([0,1]), \tag{26}$$

and as we will prove in Theorem 2, in fact we have that $\rho(0)\xi(0) = \rho(1)\xi(1) = 0$, in particular for $v = (f-u)/\alpha$ in (24).

Since TV_ρ is positively one-homogeneous, we recognize in the above that the first statement of (25) characterizes $\partial_{L^2_\phi} \mathrm{TV}_\rho(0)$, and by Fenchel duality in $L^2_\phi(0,1)$ (as in [27, Thm. III.4.2] with $\Lambda = \mathrm{Id}$), we have

$$\begin{aligned} &\min_{u \in L^2_\phi(0,1)} \int_0^1 \big(u(x) - f(x)\big)^2 \phi(x)\,\mathrm{d}x + \alpha\,\mathrm{TV}_\rho(u) \\ &= \max_{v \in \partial_{L^2_\phi} \mathrm{TV}_\rho(0)} \frac{1}{\alpha}\langle v, f\rangle_{L^2_\phi} - \frac{1}{2}\|v\|^2_{L^2_\phi} = \max_{v \in \partial_{L^2_\phi} \mathrm{TV}_\rho(0)} \int_0^1 \left(\frac{vf}{\alpha} - \frac{v^2}{2}\right)\phi, \end{aligned}$$

where the last maximization problem can be written as

$$\max_{v\in\partial_{L^2_\phi}\mathrm{TV}_\rho(0)} \int_0^1 \frac{1}{\alpha} v f \phi - \frac{1}{2}\|v\|^2_{L^2_\phi} = -\left[\min_{v\in\partial_{L^2_\phi}\mathrm{TV}_\rho(0)} \frac{1}{2}\left\|\frac{f}{\alpha} - v\right\|^2_{L^2_\phi} - \frac{1}{2\alpha^2}\|f\|^2_{L^2_\phi}\right],$$

for which we notice that since the last term does not involve v, is the familiar formulation as projection on a convex set.

For explicitly characterizing the minimizers, we will need to give a pointwise meaning to (24), which turns out to be

$$\begin{aligned} \int_0^x \big(f(s) - u(s)\big)\,\phi(s)\,\mathrm{d}s = -U(x) \text{ for all } x \in (0,1), \text{ where} \\ U \in W^{1,2}_{1,1/\phi,0}(0,1) \text{ satisfies } |U| \leqslant \alpha\rho \text{ and} \\ U(x) = \alpha\rho(x)\frac{Du}{|Du|}(x) \text{ for } |Du|\text{-a.e. } x, \end{aligned} \tag{27}$$

as we prove in Theorem 2 below. Here, avoiding density questions and owing to the embedding (26) we have directly defined

$$W^{1,2}_{1,1/\phi,0}(0,1) = \left\{U \in W^{1,2}_{1,1/\phi}(0,1) \;\mid\; U(0) = U(1) = 0\right\}.$$

A first remark about this characterization is that the last equality in the second line of (27) is made possible by the one-dimensional setting, since then functions in $W^{1,2}_{1,1/\phi,0}(0,1)$ are continuous. Whether such a characterization is possible in higher dimensions, even without weights, is a much more delicate question, see [13, 22, 24]. A further remark is that (27) implies in particular

$$\int_0^x \big(f(s) - u(s)\big)\,\phi(s)\,\mathrm{d}s = -\alpha\rho(x)\frac{Du}{|Du|}(x) \text{ for } |Du|\text{-a.e. } x, \tag{28}$$

which is formulated only in terms of the minimizer u and does not involve additional derivatives. However, equation (28) only provides information on $\operatorname{supp}|Du|$, so we cannot immediately conclude that if u satisfies it, then it must be a minimizer of (21).

4.2 The Weighted String

We can pose, for $F \in W^{1,2}_{1,1/\phi}(0,1)$ with $F(0) = 0$, the minimization problem

$$\min\left\{\frac{1}{2}\int_0^1 \frac{1}{\phi(s)}\big(\tilde{U}'(s)\big)^2\,\mathrm{d}s \;\middle|\; \tilde{U} \in W^{1,2}_{1,1/\phi}(0,1),\ |\tilde{U} - F| \leqslant \alpha\rho,\ \tilde{U}(0) = 0,\ \tilde{U}(1) = F(1)\right\}, \tag{29}$$

which can be rewritten, with $U = \tilde{U} - F$, as

$$\min \left\{ \frac{1}{2} \int_0^1 \frac{1}{\phi(s)} \big(U'(s) + F'(s)\big)^2 \, \mathrm{d}s \;\middle|\; U \in W^{1,2}_{1,1/\phi,0}(0,1),\ |U| \leqslant \alpha\rho \right\}. \tag{30}$$

Lemma 2 *There is a unique minimizer $\tilde{U}_0$ of (29).*

Proof Since we have assumed that $\phi \in L^\infty(0,1) \cap C(0,1)$, we have the embedding (26) and the boundary values and constraints are well defined and closed. Moreover, the functional is nonnegative, convex and strongly continuous in $W^{1,2}_{1,1/\phi}(0,1)$, so weakly lower semicontinuous as well. Finally, since $U(0) = 0$ and $U(1) = F(1)$ are fixed, the Poincaré inequality in $W^{1,2}_0(0,1)$ provides us with a bound for $\|U\|_{L^2(0,1)}$, so it is coercive in $W^{1,2}_{1,1/\phi}(0,1)$ and we may apply the direct method of the calculus of variations. □

Theorem 2 *The Fenchel dual of (30), for which strong duality holds, is equivalent to the weighted ROF problem (21). Specifically, if U_0 is the minimizer of (30) and V_0 is optimal in the dual problem, then $u_0 = V_0/\phi = U_0'/\phi + f$ is the minimizer of (21). Moreover, for the pair (u_0, U_0) we have that (27) is satisfied (replacing u, U by u_0, U_0, respectively), and this condition characterizes optimality of this duality pair.*

Proof Following for example [34, Sec. 3.1] and using the notation of [27, Thm. III.4.2], we call

$$\Lambda U = U' \quad \mathcal{G}(\Lambda U) = \frac{1}{2}\|\Lambda U + \Lambda F\|^2_{L^2_{1/\phi}} \quad \mathcal{F} = \chi_{\mathcal{L}},$$

where $\Lambda : W^{1,2}_{1,1/\phi,0}(0,1) \to L^2_{1/\phi}(0,1)$, $\chi_{\mathcal{L}}$ denotes the indicator function of $\mathcal{L}$ and

$$\mathcal{L} = \left\{ U \in W^{1,2}_{1,1/\phi,0}(0,1) \;\middle|\; |U| \leqslant \alpha\rho \right\}, \tag{31}$$

so that (30) can be written as

$$\min_{U \in W^{1,2}_{1,1/\phi,0}(0,1)} \mathcal{F}(U) + \mathcal{G}(\Lambda U).$$

In this situation, the dual problem writes

$$\min_{V \in L^2_{1/\phi}(0,1)} \mathcal{F}^*(-\Lambda^* V) + \mathcal{G}^*(V),$$

where

$$\mathcal{G}^*(V) = \int_0^1 \frac{1}{\phi} \left(\frac{V^2}{2} - \Lambda F V \right),$$

and taking into account $W^{1,2}_{1,1/\phi,0}(0,1) \subset L^2(0,1) \subset \left(W^{1,2}_{1,1/\phi,0}(0,1)\right)^*$ we have

$$\langle V, \Lambda U\rangle_{L^2_{1/\phi}} = \int_0^1 \frac{1}{\phi} V U' = \left\langle -\left(\frac{V}{\phi}\right)', U \right\rangle_{\left(\left(W^{1,2}_{1,1/\phi,0}\right)^*, W^{1,2}_{1,1/\phi,0}\right)}$$

which in turn implies $\Lambda^* V = -(V/\phi)'$ and

$$\begin{aligned}\mathcal{F}^*(-\Lambda^* V) &= \sup_{U\in\mathcal{L}} \left\langle \left(\frac{V}{\phi}\right)', U \right\rangle_{\left(\left(W^{1,2}_{1,1/\phi,0}\right)^*, W^{1,2}_{1,1/\phi,0}\right)} \\ &= \sup_{U\in\mathcal{L}} -\left\langle \frac{V}{\phi}, U' \right\rangle_{L^2} = \sup_{U\in\mathcal{L}} \int_0^1 \frac{1}{\phi} U' V \geqslant \alpha\, \mathrm{TV}_\rho(V/\phi),\end{aligned}$$

where for the last inequality we have used that $\left\{U \in C^1_c(0,1) \,\middle|\, |U| \leqslant \alpha\rho\right\} \subset \mathcal{L}$. We would like to have equality in this last inequality, which holds in particular when

$$\alpha\, \mathrm{TV}_\rho(V/\phi) \geqslant \int_0^1 \frac{1}{\phi} U' V \text{ for every } V \in L^2_{1/\phi}(0,1) \cap \mathrm{BV}_\rho(0,1) \text{ and all } U \in \mathcal{L}. \tag{32}$$

Next, we notice that since

$$(\mathrm{TV}_\rho)^* = \chi_{\overline{D(\mathcal{L}\cap C^1_c(0,1))}^{L^2_{1/\phi}}},$$

where $DU = U'$ for $U \in W^{1,2}_{1,1/\phi,0}$, the statement (32) is in fact equivalent to density of $C^1_c(0,1) \cap \mathcal{L}$ in $\mathcal{L}$ in the strong $W^{1,2}_{1,1/\phi}$ topology. Such a density property can not be obtained by directly mollifying elements of $\mathcal{L}$, because ρ being nonconstant implies that the mollified functions could violate the pointwise constraint; a modified mollifying procedure has been considered in [35], but only for continuous ρ with a positive lower bound, which does not cover our case. Instead, we can replace V/ϕ in (32) by a sequence of smooth approximations for which TV_ρ converges (that is, in strict convergence), and then pass to the limit. This type of approximation of TV_ρ has been proved in [17, Thm. 4.1.6] for Lipschitz ρ but without the lower bound, allowing ρ to vanish at the boundary and thus covering our situation.

Summarizing, the dual problem writes

$$\min_{V\in L^2_{1/\phi}(0,1)} \alpha\, \mathrm{TV}_\rho(V/\phi) + \int_0^1 \frac{1}{\phi}\left(\frac{V^2}{2} - F' V\right)$$

that clearly has the same minimizers as

$$\min_{V\in L^2_{1/\phi}(0,1)} \alpha\, \mathrm{TV}_\rho(V/\phi) + \int_0^1 \frac{1}{\phi} \frac{(V-F')^2}{2},$$

which in turn becomes (21) if we define $u := V/\phi$ and $f := F'/\phi$, both of which belong to $L^2_\phi(0,1)$.

The optimality conditions for minimizers U_0, V_0 for this pair of problems are then

$$\begin{aligned} &\Lambda U_0 \in \partial_{L^2_{1/\phi}} \mathcal{G}^*(V_0) = \{V_0 - \Lambda F\} = \{\phi(u_0 - f)\} \text{ for } u_0 = \frac{V_0}{\phi}, \text{ and} \\ &U_0 \in \partial_{W^{1,2}_{1,1/\phi,0}} \mathcal{F}^*(-\Lambda^* V_0). \end{aligned} \tag{33}$$

Note that we use the results from [27], which means that the subgradients of $\mathcal{F}^*$ are elements of the primal space $W^{1,2}_{1,1/\phi,0}(0,1)$ and not of the bidual $\left(W^{1,2}_{1,1/\phi,0}(0,1)\right)^{**}$. In this case, we have (see [13, Prop. 7])

$$\partial \mathcal{F}^*(0) = \overline{\mathcal{L}}^{W^{1,2}_{1,1/\phi}},$$

where the closure is taken in the strong topology of $W^{1,2}_{1,1/\phi,0}(0,1)$. Moreover, we can also extend $\mathcal{F}$ (we denote by $\hat{\mathcal{F}}$ this extension) to $C_0([0,1])$, the space of continuous functions on $[0,1]$ which vanish on the boundary. Then, $\hat{\mathcal{F}}^*$ is defined on Radon measures and we have

$$\partial \hat{\mathcal{F}}^*(0) = \overline{\mathcal{L}}^{C_0},$$

where the closure is now taken with respect to uniform convergence in $[0,1]$. In fact, these two closures satisfy

$$\overline{\mathcal{L}}^{W^{1,2}_{1,1/\phi}} \subset \overline{\mathcal{L}}^{C_0} = \left\{U \in C_0([0,1]) \,\middle|\, |U| \leqslant \alpha\rho\right\}, \tag{34}$$

as shown in Lemma 3 below. This allows us to relate $\partial \mathcal{F}^*(0)$ to subgradients of the weighted total variation norm for measures, which (see for example [34, Lem. 3.1]) satisfy for each $\mu \in \left(C_0([0,1])\right)^*$ that

$$\begin{aligned} &\partial \|\cdot\|_{\mathcal{M}_{\alpha\rho}}(\mu) \cap C_0([0,1]) \\ &\quad = \left\{U \in C_0([0,1]) \,\middle|\, U(x) = \alpha\rho(x)\frac{\mu}{|\mu|}(x) \text{ for } \mu\text{-a.e. } x,\ |U| \leqslant \alpha\rho\right\}, \end{aligned}$$

where

$$\|\cdot\|_{\mathcal{M}_{\alpha\rho}}(\mu) = \sup\left\{\int \varphi \,\mathrm{d}\mu \,\middle|\, \varphi \in C_0([0,1]) \text{ with } |\varphi(x)| \leqslant \alpha\rho(x) \text{ for all } x \in [0,1]\right\}.$$

With this, in the two lines of (33) we have that

$$\begin{aligned}
\Lambda U_0 = \phi(u_0 - f) \text{ if and only if } U_0(x) &= \int_0^x \phi(s)(u_0(s) - f(s))\,\mathrm{d}s \text{ for all } x \in (0,1), \text{ and} \\
\left\langle -\Lambda^* V_0,\ U_0 \right\rangle_{\left(\left(W^{1,2}_{1,1/\phi,0}\right)^*, W^{1,2}_{1,1/\phi,0}\right)} &= \mathrm{TV}_\rho(V_0/\phi) \text{ if and only if} \\
U_0(x) = \alpha\rho(x)\frac{Du_0}{|Du_0|}(x) \text{ for } |Du_0|{-}a.e.\ x, \text{ and } |U_0(x)| &\leqslant \alpha\rho(x) \text{ for all } x \in (0,1),
\end{aligned}$$

with which we arrive at (27). □

Lemma 3 *Let $\mathcal{L}$ be defined as in (31). Then, (34) holds.*

Proof Since $W^{1,2}_{1,1/\phi,0}(0,1)$ is continuously embedded in $C_0([0,1])$, we immediately have

$$\overline{\mathcal{L}}^{W^{1,2}_{1,1/\phi}} \subset \overline{\mathcal{L}}^{C_0}.$$

Moreover, the constraint $|U| \leqslant \alpha\rho$ is closed in $C_0([0,1])$, which shows

$$\overline{\mathcal{L}}^{C_0} \subset \left\{ U \in C_0([0,1]) \,\middle|\, |U| \leqslant \alpha\rho \right\}.$$

For the opposite containment, we consider Pasch-Hausdorff regularizations with respect to the metric induced by the weight ϕ of the positive and negative parts $U^\pm$ of $U = U^+ - U^-$, defined by

$$U_n^\pm(s) := \inf_{t\in(0,1)} \left\{ n\, d_\phi(t,s) + U^\pm(t) \right\},$$

where

$$d_\phi(t,s) := \left| \int_t^s \phi(\omega)\,\mathrm{d}\omega \right|.$$

Defining $U_n(s)$ by $U_n^+(s)$ if $U(s) > 0$ and $-U_n^-(s)$ otherwise, so that again $U_n = U_n^+ - U_n^-$, we obtain regularized functions which are Lipschitz in this metric (this result dates back to [32], see also [31, Thm. 2.1] for a proof), hence belonging to $W^{1,\infty}_{1,1/\phi,0}(0,1) \subset W^{1,2}_{1,1/\phi,0}(0,1)$ and which are by definition pointwise bounded by U, so they remain in the constraint set. Moreover, the U_n converge uniformly to U in $[0,1]$, even if ϕ can vanish at the boundary. To see this, we notice that the functions $U_n^\pm$ are pointwise increasing with respect to n and converging to $U^\pm$ respectively, to use Dini's theorem on both sequences. □

4.3 Switching Behavior

Defining the bijective transformation

$$\Phi : [0, 1] \to \left[0, \int_0^1 \phi(\omega)\, \mathrm{d}\omega\right]$$
$$s \mapsto \int_0^s \phi(\omega)\, \mathrm{d}\omega$$

and denoting $\check{U} := U \circ \Phi^{-1}$ for $U \in W^{1,2}_{1,1/\phi,0}(0, 1)$, we have that $\check{U}' \in L^2(0, \Phi(1))$ and, using the Poincaré inequality in $W^{1,2}_0(0, \Phi(1))$, that $\check{U} \in L^2(0, \Phi(1))$ as well. Likewise, from $\check{U} \in W^{1,2}_0(0, \Phi(1))$ we can conclude $U \in W^{1,2}_{1,1/\phi,0}(0, 1)$ using the inequality

$$\|U\|_{L^2} \leqslant \|U\|_{L^\infty} \leqslant \|U'\|_{L^2} \leqslant \|U'\|_{L^2_{1/\phi}} \|\phi\|_{L^\infty}.$$

This implies that (30) is equivalent to

$$\min \left\{ \frac{1}{2} \int_0^{\Phi(1)} \left(\check{U}'(t) + \check{F}'(t)\right)^2 \mathrm{d}t \;\middle|\; \check{U} \in W^{1,2}_0\big(0, \Phi(1)\big),\ |\check{U}| \leqslant \alpha \left(\rho \circ \Phi^{-1}\right) \right\}, \tag{35}$$

where similarly $\check{F} = F \circ \Phi^{-1}$. In turn, arguing as in [50, Thm. 4.46] the unique minimizer of (35) is the same for any strictly convex integrand, so in particular it can be obtained from

$$\min \left\{ \int_0^{\Phi(1)} \left[\left(\hat{U}'(t)\right)^2 + 1\right]^{\frac{1}{2}} \mathrm{d}t \;\middle|\; \hat{U} \in W^{1,2}_0\big(0, \Phi(1)\big), \right.$$
$$\left. \check{F} - \alpha \left(\rho \circ \Phi^{-1}\right) \leqslant \hat{U} \leqslant \check{F} + \alpha \left(\rho \circ \Phi^{-1}\right) \right\}, \tag{36}$$

by setting $\check{U} = \hat{U} - \check{F}$. This is now a 'generalized' taut string formulation that fits in those considered in [30, Lem. 5.4]. One can then argue as in [10, Lem. 9] (which is directly based on the previously cited result) to conclude that for any $\delta > 0$ the interval $(\delta, \Phi(1) - \delta)$ can be partitioned into finitely many subintervals on which one or neither of the constraints is active. The reasoning behind such a result is that in the form (36), one is minimizing the Euclidean length of the graph of a continuous function with constraints from above and below, so as long as these constraints are at some positive distance ϵ apart, switching from one constraint to the other being active must cost at least ϵ and enforce a subinterval in which neither is active. In our case we need to restrict the interval to avoid the endpoints, because in our setting the weights are potentially degenerate and $\rho \circ \Phi^{-1}$ may vanish at 0 or $\Phi(1)$, and with it the distance between the two constraints. For such degenerate weights, it is enough to

assume that each of the two functions $\check{F} \pm \alpha\big(\rho \circ \Phi^{-1}\big)$ is either convex or concave on a neighborhood of 0 and 1 to ensure only finitely many changes of behavior.

With this property and taking into account Theorem 2 we can go back to (27) to see that the minimizer of (21) alternates between the three behaviors

$$u(x) = f(x) - \alpha \frac{\rho'(x)}{\phi(x)}, \quad u'(x) = 0, \quad u(x) = f(x) + \alpha \frac{\rho'(x)}{\phi(x)} \quad \text{for a.e. } x \tag{37}$$

finitely many times on $(\delta, 1-\delta)$ for any $\delta > 0$, and on $(0, 1)$ for data f which is not oscillating near 0 and 1, and a subinterval corresponding to either the first or the last case is always followed by another on which $u' = 0$.

4.4 Denoising of Unbounded Radial Data

We can apply the results above to find minimizers in the case of radially symmetric data $f \in L^2(B(0,1))$ of

$$\min_{u \in L^2(B(0,1))} \int_{B(0,1)} |u - f|^2 \, \mathrm{d}x + \alpha \, \mathrm{TV}(u), \tag{38}$$

which considered in $\mathbb{R}^d$ corresponds to (21) with $\rho(r) = \phi(r) = r^{d-1}$,

$$\Phi(s) = \frac{1}{d} s^d, \quad (\rho \circ \Phi^{-1})(t) = (td)^{(d-1)/d}, \quad \text{and} \quad \frac{\rho'(r)}{\phi(r)} = \frac{d-1}{r}, \tag{39}$$

where we remark that (37) for this particular case was claimed without proof in [39]. Indeed, it is easy to guess the form of this term by formally differentiating (27), but since this is not possible, we prefer to rigorously justify the behavior of minimizers by the arguments of the previous subsection. The earlier work [52] also contains some results about nearly explicit minimizers for piecewise linear or piecewise constant radial data.

In particular, let us compute explicitly the solution in the case $d = 3$ and

$$f = \frac{1}{|\cdot|} \in L^2_{\mathrm{loc}}(\mathbb{R}^3).$$

In this case we have

$$F(s) = \int_0^s f(r)\phi(r) \, \mathrm{d}r = \frac{s^{d-1}}{d-1} = \frac{s^2}{2} \quad \text{and}$$
$$\check{F}(t) = (F \circ \Phi^{-1})(t) = \frac{(dt)^{(d-1)/d}}{d-1} = \frac{(3t)^{2/3}}{2},$$

and using (39) we conclude that $\check{F} + \alpha\big(\rho \circ \Phi^{-1}\big)$ is concave for all $\alpha > 0$ while $\check{F} - \alpha\big(\rho \circ \Phi^{-1}\big)$ is strictly concave for $0 < \alpha < 1/(d-1) = 1/2$ and convex for

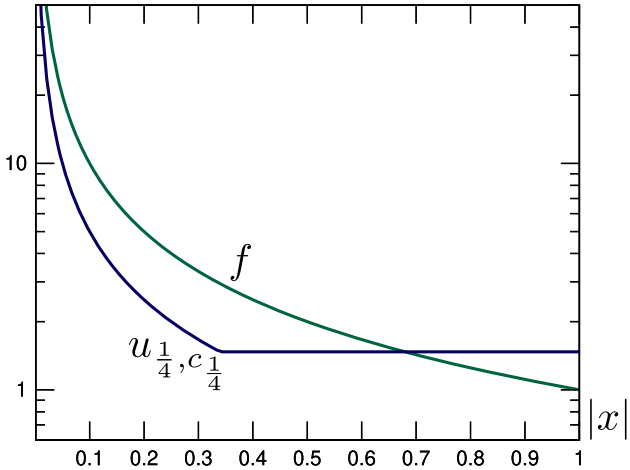

Fig. 1 Logarithmic plot of the radial profile of the unbounded minimizer (40) of the denoising problem in $\mathbb{R}^3$ with data $f(x) = 1/|x| \in L^2\big(B(0,1)\big)$, and $\alpha = 1/4$ resulting in $c \approx 0.34$

$\alpha \geq 1/(d-1) = 1/2$. Therefore, (37) tells us that either $u(r) = (1-2\alpha)/r$ or $u'(r) = 0$ on each of finitely many subintervals. We consider the simplest case of two such intervals and check that we can satisfy the optimality condition for it. First, the Neumann boundary condition forces $u'(1) = 0$ since the radial weight equals one there. Moreover, if $f \geqslant 0$, the minimizer u of (38) satisfies $u \geqslant 0$ as well, because in this case, replacing u with u^+ cannot increase either term of the energy. Taking this into account, as well as the fact that u should be continuous [18], we consider for $c \in (0,1)$ the family of candidates

$$u_{\alpha,c}(r) := \begin{cases} \frac{1-2\alpha}{r} & \text{for } r \in (0,c) \\ \frac{1-2\alpha}{c} & \text{for } r \in (c,1), \end{cases} \tag{40}$$

where $\alpha \leqslant 1/2$ is required so that $u \geqslant 0$ is satisfied. Moreover, we have also excluded segments with value $(1+2\alpha)/r$, since in that case the fidelity cost is the same as for $(1-2\alpha)/r$, but with a higher variation. To certify optimality of (40), we attempt to satisfy the pointwise characterization (27). The condition $U(1) = 0$, for U defined by the first line of (27) forces

$$\int_0^1 (u(r) - f(r))r^2 \,\mathrm{d}r = -\int_0^c \frac{2\alpha}{r} r^2 \,\mathrm{d}r + \int_c^1 r^2 \left(\frac{1-2\alpha}{c}\right) - r \,\mathrm{d}r = 0,$$

which is equivalent to

$$\alpha c^2 + \frac{1-c^2}{2} + \frac{1-c^3}{3}\frac{2\alpha-1}{c} = 0,$$

which simplifies to

$$c^3 + \frac{3c}{2\alpha - 1} + 2 = 0, \tag{41}$$

which if $0 < \alpha < 1/2$ can be solved to find a $c \in (0, 1)$, as can be easily seen using the intermediate value theorem. Now, since for the family $u_{\alpha,c}$, one has $\frac{Du}{|Du|} = -1$ on $(0, c] = \mathrm{supp}(|Du|)$, one can check that for r on this interval, we have

$$U(r) = \alpha\rho(r)\frac{Du}{|Du|}(r).$$

Now, it remains to check that the inequality $|U| \leqslant \alpha\rho$ holds on $(c, 1)$ (it holds on $(0, c)$ thanks to the previous argument). Since at c, both functions involved are equal, it is enough to show

$$\mathrm{sign}\big(U(r)\big)(f(r) - u_{\alpha,c}(r))\phi(r) = \frac{d}{dr}|U(r)| \leqslant \alpha\rho'(r) \text{ for } r > c,$$

which, using the definition of $u_{\alpha,c}$, is satisfied in particular if

$$\left|\frac{1}{r} - \frac{1-2\alpha}{c}\right| r^2 \leqslant 2\alpha r, \text{ or } 1 \leqslant \frac{r}{c} \leqslant \frac{1+2\alpha}{1-2\alpha}$$

which, as long as $\frac{1}{4} \leqslant \alpha < \frac{1}{2}$ and taking into account (41), is satisfied as soon as $c \leqslant r \leqslant 1$. Equation (41) can be solved directly, for example for $\alpha = \frac{1}{4}$, and we obtain a solution $c_{\frac{1}{4}} \in (0, 1)$, which guarantees that (40) is indeed a minimizer of (38). Thus, we have obtained an explicit example of an unbounded minimizer of (38), whose graph is depicted in Fig. 1.

Remark 9 We note here (as kindly pointed out by an anonymous reviewer) that using techniques of [5] it is also possible to find an explicit vector field $\tilde{z}$ satisfying the characterization (8) to guarantee that (40) is a solution of the denoising problem with data $1/r$. Denoting $e_r = x/|x|$ and

$$\tilde{z}(x) = z(r)e_r, \text{ so that } \operatorname{div} z(x) = z'(r) + \frac{2}{r}z(r) \text{ for } d = 3,$$

one can define

$$z(r) = \begin{cases} -1 & \text{for } r \in (0, c) \\ \frac{1}{2\alpha}\left(\frac{1}{r^2} - 1\right) + \frac{1-2\alpha}{3c\alpha}\left(r - \frac{1}{r^2}\right) & \text{for } r \in [c, 1) \\ 0 & \text{for } r \in [1, +\infty) \end{cases}$$

for which continuity as well as L^∞ boundedness by 1 holds for c solving (41). Moreover

$$z'(r) + \frac{2}{r} z(r) = \begin{cases} -\frac{2}{r} & \text{for } r \in (0, c) \\ \frac{1-2\alpha}{\alpha c} - \frac{1}{\alpha r} & \text{for } r \in (c, 1), \end{cases}$$

so for such c we have

$$\operatorname{div} \tilde{z} \in \partial \operatorname{TV}(u_{\alpha,c}) \text{ and } \frac{u_{\alpha,c} - f}{\alpha} = \operatorname{div} \tilde{z},$$

in which case $u_{\alpha,c}$ indeed solves (38). Furthermore, a construction along these lines would also be possible for Lemma 4 below, for which our proof is based on the same weighted taut string technique.

For other cases such an explicit computation may be impractical or impossible, but we can still use the taut string characterization to derive some conditions for local boundedness and unboundedness of radial minimizers:

Lemma 4 *Let $f : (0, 1) \to \mathbb{R}$ be such that $f\big|_{(\delta,1)} \in L^2(\delta, 1)$ for all $\delta \in (0, 1)$, and such that for some $\delta_0 \in (0, 1)$ we have*

$$f(r) = \frac{1}{r^\beta} \text{ for all } r \in (0, \delta_0), \text{ so that } f \in L^2_{r^{d-1}}(0, 1) \text{ for } 2\beta < d,$$

and consider the corresponding minimizer u of (21) with $\rho(r) = \phi(r) = r^{d-1}$. Then, if u is unbounded on any neighborhood of 0, then necessarily $\beta \geqslant 1$, and if $\beta > 1$, then u is unbounded on any neighborhood of 0. In particular, for $d = 2$ and with f of this form belonging to $L^2_{r^{d-1}}(0, 1)$, we have that u is always bounded on a neighborhood of 0.

Proof First, as above we notice that $f \geqslant 0$ implies $u \geqslant 0$, since otherwise truncating u to $u^+ = \max(u, 0)$ would produce a lower energy in (21). Then, the switching behavior (37) implies that there is $\delta \in (0, \delta_0)$ such that either

$$u(r) = f(r) - \alpha \frac{d-1}{r}, \text{ or } u'(r) = 0 \text{ for all } r \in (0, \delta), \tag{42}$$

and one of the two choices must hold for the whole interval $(0, \delta)$ by the results cited in Sect. 4.3, since once again the behavior of u can change only finitely many times away from 1. To see this, notice that in this case

$$F(s) = \int_0^s f(r)\phi(r)\,\mathrm{d}r = \frac{s^{d-\beta}}{d-\beta} \qquad \text{for } s < \delta_0, \text{ and}$$
$$\check{F}(t) = (F \circ \Phi^{-1})(t) = \frac{(dt)^{(d-\beta)/d}}{d-\beta} \qquad \text{for } t < (d\delta_0)^{1/d},$$

so that $\check{F} + \alpha\big(\rho \circ \Phi^{-1}\big)$ is concave and $\check{F} - \alpha\big(\rho \circ \Phi^{-1}\big)$ is strictly concave around zero for $\beta > 1$ but convex for $\beta < 1$.

Now, we notice that if $\beta < 1$ we must have the second case of (42) because $u \geqslant 0$ would be violated in the first case, which proves the first part of the statement. Let us now assume that $\beta > 1$. Comparing in the energy (21) the first case of (42) holds exactly when for all $0 < \nu \leqslant \delta$ we simultaneously have

$$\int_0^\nu \left[\left| f'(r) + \alpha \frac{d-1}{r^2} \right| + \alpha^2 \left(\frac{d-1}{r} \right)^2 \right] r^{d-1}\,\mathrm{d}r < \int_0^\nu \left(f(\nu) - \alpha \frac{d-1}{\nu} - f(r) \right)^2 r^{d-1}\,\mathrm{d}r, \tag{43}$$

which encodes that switching to the second case $u' = 0$ should not be advantageous at any $\nu \leqslant \delta$, and where equality is not possible since it would contradict uniqueness of minimizers of (21). Moreover, we also notice that the last term of the left hand integral can only be finite if $d \geqslant 3$, proving boundedness in case $d = 2$. Using the specific form of $f(r) = r^{-\beta}$, inequality (43) becomes

$$\int_0^\nu \left[\left| -\frac{\beta}{r^{\beta+1}} + \alpha \frac{d-1}{r^2} \right| + \alpha^2 \left(\frac{d-1}{r} \right)^2 \right] r^{d-1}\,\mathrm{d}r < \int_0^\nu \left(\frac{1}{\nu^\beta} - \alpha \frac{d-1}{\nu} - \frac{1}{r^\beta} \right)^2 r^{d-1}\,\mathrm{d}r. \tag{44}$$

Using the notation $I(\nu) < J(\nu)$ in (44), since the inequality is either true or false for every $\nu < \delta$ simultaneously, it holds if we had that

$$\ell := \lim_{\nu \to 0} \frac{I(\nu)}{J(\nu)} < 1.$$

But for ν small enough we have that $I(\nu)$ is dominated (up to a constant factor) by $\nu^{d-\beta-1}$ and $\nu^{d-2\beta}$ is dominated (also up to a factor) by $J(\nu)$, which, since $\beta > 1$, implies $\ell = 0$. □

This lemma can be used to derive a boundedness criterion for more general radial data, without requiring local convexity or concavity assumptions for the functions appearing in the constraints of (36):

Proposition 5 *Assume that the radial data f satisfies $f \geqslant 0$ and $f \in L^2_{r^{d-1}}(0,1)$. Then, if*

$$\lim_{r \to 0} r^\beta f(r) = +\infty \text{ for some } \beta > 1, \tag{45}$$

then the corresponding minimizer u of (21) with $\rho(r) = \phi(r) = r^{d-1}$ is unbounded on any neighborhood of 0, but bounded on any compact interval excluding zero.

Conversely, if

$$\lim_{r\to 0} r^{\beta} f(r) = 0 \textit{ for some } \beta < 1 \tag{46}$$

then u is bounded.

Proof First, we notice that we must have

$$u \in L^{\infty}(\delta, 1) \text{ for all } \delta > 0, \tag{47}$$

because on the interval $(\delta, 1)$ we have $\rho \geq \delta^{d-1}$, which ensures that the restriction of u to it belongs to $\mathrm{BV}(\delta, 1)$, and in one dimension we have the embedding $\mathrm{BV}(\delta, 1) \subset L^{\infty}(\delta, 1)$. Given (47) and using the equivalence with (38), we can conclude by applying the pointwise comparison principle for minimizers of TV denoising in $B(0, 1) \subset \mathbb{R}^d$ (a proof of which can be found in for example [37, Prop. 3.6]) against inputs of the form treated in Lemma 4 above. Assuming (45), we choose δ_0 such that $f(r) \geq r^{-\beta}$ for $r \in (0, \delta_0)$, while for the case (46) we take δ_0 such that $f(r) \leq r^{-\beta}$ for $r \in (0, \delta_0)$, and in both cases compare with $\tilde{f}$ defined by

$$\tilde{f}(r) = \begin{cases} \frac{1}{r^{\beta}} & \text{if } r \in (0, \delta_0) \\ f(r) & \text{if } r \in [\delta_0, 1). \end{cases}$$

□

We can also see this criterion in light of the results of Sect. 3, in particular Propositions 1 and 2. There, the main requirement for boundedness of u is that there should be $v \in \partial_{L^{d/(d-1)}} \mathrm{TV}(u) \subset L^d(\Omega)$. In the case of radial data on $\Omega = B(0, 1) \subset \mathbb{R}^d$, we have worked with the ROF denoising problem for which

$$v_{\alpha} := -\frac{1}{\alpha}(u_{\alpha} - f) \in \partial_{L^2} \mathrm{TV}(u_{\alpha}),$$

and by (37) we know that $u_{\alpha} - f$ switches between $\pm\alpha(d-1)/|\cdot|$ and zero on annuli. Moreover, the power growth $1/|\cdot|$ is precisely the threshold for a function on $\mathbb{R}^d$ to belong to $L^d(B(0, 1))$, assuming it is in $L^{\infty}\big(B(0, 1) \setminus B(0, \delta)\big)$ for all δ. That is, for slower power growth as in (46) we have $v_{\alpha} \in L^d(B(0, 1))$ and the minimizer is bounded, which we could have proved using the techniques of Sect. 3. In contrast, for faster power growth as in (45) we have that $v_{\alpha} \notin L^d(B(0, 1))$, so the methods of Sect. 3 are not applicable, and by Proposition 5 in this case TV denoising produces an unbounded minimizer.

5 Higher-Order Regularization Terms

Here we treat problems regularized with two popular approaches combining derivatives of first and higher orders, with the goal of extending our boundedness results

from Sect. 3 to these more involved settings. In this section, we limit ourselves to the case of functions defined on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$.

The first approach, for which we are able to prove an analog of Theorem 1 in dimension 2, is infimal convolution of first and second order total variation first introduced in [23] (see also [14]), that is

$$\min_{u \in L^{d/(d-1)}(\Omega)} \frac{1}{\sigma} \left(\int_\Sigma |Au - (f + w)|^q \right)^{\sigma/q} + \left[\inf_{g \in \mathrm{BV}^2(\Omega)} \alpha_1 \, \mathrm{TV}(u - g) + \alpha_2 \, \mathrm{TV}^2(g) \right], \tag{48}$$

where $\mathrm{TV}^2(g)$ for $g \in \mathrm{BV}^2(\Omega) \subset W^{1,1}(\Omega)$ is the norm of the distributional Hessian as a matrix-valued Radon measure, that is

$$\mathrm{TV}^2(g) := \sup \left\{ \int_\Omega \nabla g(x) \cdot \operatorname{div} M(x) \, \mathrm{d}x \;\middle|\; M \in C_c^1(\Omega; \mathbb{R}^{d \times d}) \text{ with } |M(x)|_F \leqslant 1 \text{ for all } x \right\},$$

with $|M|_F$ denoting the Frobenius norm, that is $|M|_F^2 = \sum_{ij} (M_{ij})^2$. Here and in the rest of this section, A and the exponents σ, q are as in Sects. 1 and 3, i.e. $A : L^{d/(d-1)}(\Omega) \to L^q(\Sigma)$ is a bounded linear operator, and $\sigma = \min(q, 2)$.

The second approach, widely used in applications but harder to treat analytically, is total generalized variation of second order, introduced in [15] and for which we use the characterization of [16, Thm. 3.1] to write

$$\min_{u \in L^{d/(d-1)}(\Omega)} \frac{1}{\sigma} \left(\int_\Sigma |Au - (f + w)|^q \right)^{\sigma/q} + \left[\inf_{z \in \mathrm{BD}(\Omega)} \alpha_1 |Du - z|(\Omega) + \alpha_2 \, \mathrm{TD}(z) \right], \tag{49}$$

for which we denote the second term as $\mathrm{TGV}(u)$, not making the regularization parameters α_1, α_2 explicit in the notation. Here, $\mathrm{TD}(z)$ denotes the 'total deformation' of the vector field z defined in terms of its distributional symmetric gradient, that is

$$\mathrm{TD}(z) := \sup \left\{ \int_\Omega z(x) \cdot \operatorname{div} M(x) \, \mathrm{d}x \;\middle|\; M \in C_c^1(\Omega; \mathbb{R}^{d \times d}_{\mathrm{sym}}) \text{ with } |M(x)|_F \leqslant 1 \text{ for all } x \right\}.$$

Since we are working on a bounded Lipschitz domain, the inner infima are attained in both cases and the regularization functionals (48) and (49) have minimizers. For a proof, see [14, Prop. 4.8, Prop. 4.10] for (48) and [16, Thm. 3.1, Thm. 4.2] [12] [14, Thm. 5.9, Prop. 5.17] for the TGV case (49).

Lemma 5 *Assuming that u is a minimizer of (48) and the inner infimum is attained by some g_u, we have the necessary and sufficient optimality condition:*

$$\begin{aligned} &v \in \alpha_1 \partial \operatorname{TV}(u - g_u) \cap \alpha_2 \partial \operatorname{TV}^2(g_u), \text{ for} \\ &v := \|Au - f - w\|_{L^q(\Sigma)}^{\sigma-2} A^* \left(|f + w - Au|^{q-2}(f + w - Au) \right). \end{aligned} \tag{50}$$

If (49) is minimized by u and the inner infimum is attained for some z_u, we also have the necessary condition

$$v \in \alpha_1 \partial J_{z_u}(u), \tag{51}$$

where the functional J_{z_u} is defined by $J_{z_u}(\tilde{u}) = |D\tilde{u} - z_u|(\Omega)$, and v is as in (50).

Proof We assume for the sake of simplicity that $\alpha_1 = \alpha_2 = 1$, and first prove (51). We notice that the first term in (49) is continuous with respect to $u \in L^{d/(d-1)}(\Omega)$, so we have [27, Prop. I.5.6] for an optimal u that

$$v \in \partial \left[\inf_{z \in \operatorname{BD}(\Omega)} |D \cdot - z|(\Omega) + \operatorname{TD}(z) \right](u), \tag{52}$$

and similarly for (48). Now, one would like to use the characterization of the subgradient of exact infimal convolutions as intersections of subgradients at the minimizing pair (see [55, Cor. 2.4.7], for example). Since the complete result cannot be applied directly to (52), we make explicit the parts of its proof that are applicable. To this end, let $u^* \in L^d(\Omega)$ belong to the right hand side of (52). We have for all $\tilde{u} \in L^{d/(d-1)}(\Omega)$, noticing that $\operatorname{TGV}(u) < +\infty$ implies $u \in \operatorname{BV}(\Omega)$ since we may just consider $z = 0$ in the inner minimization, that

$$\begin{aligned} &|D(u + \tilde{u}) - z_u|(\Omega) - |Du - z_u|(\Omega) - \langle u^*, \tilde{u} \rangle_{L^d \times L^{d/(d-1)}} \\ &\quad = |D(u + \tilde{u}) - z_u|(\Omega) + \operatorname{TD}(z_u) - |Du - z_u|(\Omega) - \operatorname{TD}(z_u) - \langle u^*, \tilde{u} \rangle_{L^d \times L^{d/(d-1)}} \\ &\quad = |D(u + \tilde{u}) - z_u|(\Omega) + \operatorname{TD}(z_u) - \left[\inf_{z \in \operatorname{BD}(\Omega)} |Du - z|(\Omega) + \operatorname{TD}(z) \right] \\ &\qquad - \langle u^*, \tilde{u} \rangle_{L^d \times L^{d/(d-1)}} \\ &\quad \geqslant \left[\inf_{z \in \operatorname{BD}(\Omega)} |D(u + \tilde{u}) - z|(\Omega) + \operatorname{TD}(z) \right] \\ &\qquad - \left[\inf_{z \in \operatorname{BD}(\Omega)} |Du - z|(\Omega) + \operatorname{TD}(z) \right] - \langle u^*, \tilde{u} \rangle_{L^d \times L^{d/(d-1)}} \\ &\quad \geqslant 0, \end{aligned}$$

which combined with (52) implies (51). To prove that (50) is necessary for optimality, assuming that u now minimizes (48) we have analogously

$$v \in \partial \left[\inf_{g \in L^2(\Omega)} \operatorname{TV}(\cdot - g) + \operatorname{TV}^2(g) \right](u), \tag{53}$$

and the analogous computation leads to $v \in \partial\,\mathrm{TV}(u - g_u)$. To see that we also have $v \in \partial\,\mathrm{TV}^2(g_u)$, we consider $u^* \in L^d(\Omega)$ in the right hand side of (53) so that for $\tilde{u} \in L^{d/(d-1)}(\Omega)$,

$$\begin{aligned}
&\mathrm{TV}^2(g_u + \tilde{u}) - \mathrm{TV}^2(g_u) - \langle u^*, \tilde{u}\rangle_{L^2} \\
&\quad = \mathrm{TV}^2(g_u + \tilde{u}) + \mathrm{TV}(u - g_u) - \mathrm{TV}^2(g_u) - \mathrm{TV}(u - g_u) - \langle u^*, \tilde{u}\rangle_{L^d \times L^{d/(d-1)}} \\
&\quad = \mathrm{TV}^2(g_u + \tilde{u}) + \mathrm{TV}(u - g_u) - \left[\inf_{g \in L^{d/(d-1)}(\Omega)} \mathrm{TV}(u - g) + \mathrm{TV}^2(g)\right] \\
&\qquad - \langle u^*, \tilde{u}\rangle_{L^d \times L^{d/(d-1)}} \\
&\quad \geqslant \left[\inf_{g \in L^{d/(d-1)}(\Omega)} \mathrm{TV}(u + \tilde{u} - g) + \mathrm{TV}^2(g)\right] \\
&\qquad - \left[\inf_{g \in L^{d/(d-1)}(\Omega)} \mathrm{TV}(u - g) + \mathrm{TV}^2(g)\right] - \langle u^*, \tilde{u}\rangle_{L^d \times L^{d/(d-1)}} \geqslant 0.
\end{aligned}$$

Notice that this last part is not a priori possible for (49) due to the lack of symmetry in the variables involved. That (50) is also sufficient can be seen from [55, Cor. 2.4.7], since in this case we do have an infimal convolution. □

Proposition 6 *Whenever $d = 2$, minimizers of (48) are in $L^\infty(\Omega)$. For $d > 2$ they belong to $L^{d/(d-2)}(\Omega)$.*

Proof Using (50) we notice as in Remark 4 that the proof of Proposition 1 implies that $u - g_u \in L^\infty(\Omega)$. Now since $\mathrm{TV}^2(g_u) < +\infty$, for $d > 2$ one gets $g_u \in L^{d/(d-2)}(\Omega)$ so that also $u \in L^{d/(d-2)}(\Omega)$, and in fact for $d = 2$ the critical embedding $\mathrm{BV}^2(\Omega) \subset L^\infty(\Omega)$ holds, see [26, Thm. 3.2] for domains with C^2 boundary and [49, Rem. 0.3] for Lipschitz domains. Therefore, in that case $u \in L^\infty(\Omega)$. □

More generally one can prove the following restricted analogue of Theorem 1:

Theorem 3 *Let $d = 2$ and Ω be a bounded Lipschitz domain. Assume that for $A : L^2(\Omega) \to L^q(\Sigma)$ linear bounded, $f \in L^q(\Sigma)$ and a fixed $\gamma > 0$ there is a unique solution $u^\dagger$ for $Au = f$ which satisfies the source condition* $\mathrm{Ran}(A^*) \cap \partial(\mathrm{TV} \,\square\, \gamma\mathrm{TV}^2)(u^\dagger) \neq \emptyset$*, where we denote*

$$(\mathrm{TV} \,\square\, \gamma\mathrm{TV}^2)(u) := \inf_{g \in \mathrm{BV}^2(\Omega)} \mathrm{TV}(u - g) + \gamma\,\mathrm{TV}^2(g).$$

In this case, there is some constant $C(\gamma, q, \sigma, \Omega)$ such that if α_n, w_n are sequences of regularization parameters and perturbations for which

$$\alpha_n \geqslant C(\gamma, q, \sigma, \Omega)\|A^*\|\|w_n\|_{L^q(\Sigma)}^{\sigma-1} \xrightarrow[n\to\infty]{} 0$$

then the corresponding sequence u_n of solutions of

$$\min_{u \in L^2(\Omega)} \frac{1}{\sigma}\left(\int_\Sigma |Au - (f + w_n)|^q\right)^{\sigma/q} + \alpha_n(\mathrm{TV} \,\square\, \gamma\mathrm{TV}^2)(u) \tag{54}$$

is bounded in $L^\infty(\Omega)$, and (possibly up to a subsequence)

$$u_n \xrightarrow[n\to\infty]{} u^\dagger \text{ strongly in } L^{\overline{p}}(\Omega) \text{ for all } \overline{p} \in (1,\infty).$$

Proof Since the proof follows by using the methods of Sect. 3 essentially verbatim, we provide just an outline highlighting the steps that pose significant differences.

The first ingredient is to repeat the results on convergence and stability of dual variables summarized in Sect. 2.2, which in fact only depend on the regularization term being positively one-homogeneous and lower semicontinuous. Once these are obtained, from the characterization (50) and Proposition 1 one gets uniform L^∞ bounds for $u_n - g_{u_n}$. To finish, one notices that by comparing with $u^\dagger$ in the minimization problem (54) we get using $Au^\dagger = f$ that

$$\gamma\,\mathrm{TV}^2(g_{u_n}) \leqslant \frac{1}{\sigma}\|w_n\|^\sigma_{L^q(\Sigma)} + \alpha_n\left[\inf_{g\in \mathrm{BV}^2(\Omega)} \mathrm{TV}(u^\dagger - g) + \gamma\,\mathrm{TV}^2(g)\right],$$

in which the first term of the right hand side is bounded above in terms of α_n by the parameter choice and $(\mathrm{TV}\,\square\,\gamma\mathrm{TV}^2)(u^\dagger) < +\infty$ by the source condition, so the bound on g_{u_n} obtained by the embedding $\mathrm{BV}^2(\Omega) \subset L^\infty(\Omega)$ is not just uniform in n but in fact vanishes as $n \to \infty$. Since we have assumed that Ω is bounded, the strong convergence in $L^{\bar{p}}(\Omega)$ follows directly along the lines of the proof of Corollary 1. □

In comparison with Theorem 1, the above result is weaker in two aspects. The first is that we need to impose that the two regularization parameters maintain a constant ratio between them, in order to formulate a source condition in terms of a subgradient of a fixed functional. Moreover, the result is limited to $d = 2$ and bounded domains. Without the latter assumption, in particular the upgrade to strong convergence in $L^{\bar{p}}(\Omega)$ would require a common compact support for all u_n. However, even assuming attainment of the inner infimum, using Lemma 5 we would only get a common support for $u_n - g_{u_n}$, and in the absence of coarea formula for TV^2 it is not clear how to control the one of g_{u_n}.

For the case of TGV regularization it is unclear if one can also use the boundedness results of Sect. 3, which are based on subgradients of TV. A first observation is that we cannot use the TGV subgradients directly, except in trivial cases:

Proposition 7 *If for $u \in L^{d/(d-1)}(\Omega)$ and $v \in \partial\,\mathrm{TGV}(u)$ we have that also $v \in \alpha_1\partial\,\mathrm{TV}(u)$, then necessarily $\alpha_1\,\mathrm{TV}(u) = \mathrm{TGV}(u)$.*

Proof Assuming without loss of generality that $\alpha_1 = 1$, we have by testing with $z = 0$ that

$$\inf_{z\in \mathrm{BD}(\Omega)} |Du - z|(\Omega) + \alpha_2\,\mathrm{TD}(z) \leqslant |Du|(\Omega) = \mathrm{TV}(u) \text{ for all } u \in L^{d/(d-1)}(\Omega),$$

which since the Fenchel conjugate is order reversing means that

$$\mathrm{TV}^*(v) \leqslant \mathrm{TGV}^*(v) \text{ for all } v \in L^d(\Omega).$$

Because both these functionals are convex positively one-homogeneous we have as in (4) that

$$\mathrm{TV}^* = \chi_{\partial\,\mathrm{TV}(0)} \text{ and } \mathrm{TGV}^* = \chi_{\partial\,\mathrm{TGV}(0)},$$

so that

$$\chi_{\partial\,\mathrm{TV}(0)} \leqslant \chi_{\partial\,\mathrm{TGV}(0)}, \text{ or } \partial\,\mathrm{TGV}(0) \subset \partial\,\mathrm{TV}(0).$$

Now, again since these functionals are positively one-homogeneous we have

$$v \in \partial\,\mathrm{TGV}(u) \text{ exactly when } v \in \partial\,\mathrm{TGV}(0) \text{ and } \int_\Omega vu = \mathrm{TGV}(u)$$

and similarly for TV, as stated in (5). But since we assumed that $v \in \partial\,\mathrm{TV}(u)$ as well, we must then have

$$\mathrm{TGV}(u) = \int_\Omega vu = \mathrm{TV}(u).$$

□

In Appendix A we explore whether it is possible to use a more refined approach by trying to find elements of $\partial\,\mathrm{TV}(u)$ from those of $\partial J_{z_u}(u)$ appearing in (51) for TGV regularization, with a negative conclusion (at least without using further properties of z_u) in the form of an explicit counterexample.

Acknowledgements We wish to thank Martin Holler for some enlightening discussions about higher-order regularization terms in the early stages of this work. Moreover, we also would like to thank the anonymous reviewers for their careful reading of this manuscript, and for the explicit construction in Remark 9 provided by one of them.

Funding A large portion of this work was completed while the second-named author was employed at the Johann Radon Institute for Computational and Applied Mathematics (RICAM) of the Austrian Academy of Sciences (ÖAW), during which his work was partially supported by the State of Upper Austria. The Institute of Mathematics and Scientific Computing of the University of Graz, with which the first-named author is affiliated, is a member of NAWI Graz (https://www.nawigraz.at/en/).

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

A A Subgradient of TGV Does Not Easily Induce One for TV

To attempt to infer that TGV-regularized minimizers are bounded, an alternative approach to the arguments of Sect. 5 would be to try and use that the inner minimization problem is attained. Proposition 6 above for the infimal convolution case uses only that $\partial\,\mathrm{TV}(u - g_u) \neq \emptyset$ and $g_u \in \mathrm{BV}^2(\Omega)$, and not any information about subdifferentials of TV^2. It is natural to wonder whether it is possible to follow a similar approach for TGV starting from (51). Unlike in the infimal convolution case, clearly this cannot follow just from embeddings.

What (51) provides is an element in the subdifferential of the functional $J_z = |D \cdot - z|(\Omega)$ at u, and Du is only a measure but $z \in \mathrm{BD}(\Omega) \subset L^{d/(d-1)}(\Omega; \mathbb{R}^d)$ is much more regular. We could then ask ourselves if this means that $\partial\,\mathrm{TV}(u) \neq \emptyset$ or, equivalently, if u being irregular enough to not have a subgradient of TV would imply $\partial J_z(u) = \emptyset$ as well, making z just a 'more regular perturbation'. Unfortunately this is not the case, which prevents using (51) and knowledge of subgradients of TV to infer whether minimizers of TGV regularization are bounded.

For our counterexample we will use the following characterization:

Lemma 6 *For arbitrary $u \in \mathrm{BV}(\Omega)$ and $z \in \mathcal{M}(\Omega; \mathbb{R}^d)$ (which could be $z = 0$ for the* TV *case) we have that $\partial J_z(u) = \emptyset$ if and only if there are v_n with $\|v_n\|_{L^{d/(d-1)}(\Omega)} = 1$ and $t_n \to 0^+$ such that*

$$\lim_{n\to\infty} \frac{1}{t_n}\Big(|Du + t_n Dv_n - z|(\Omega) - |Du - z|(\Omega)\Big) = -\infty. \tag{55}$$

Proof This is a general fact for proper convex functionals whose domain is a linear subspace. Let F be such a functional on a separable Banach space X and let $x \in \operatorname{dom} F$. For any $h \in X$, the directional derivative

$$D_h F(x) := \lim_{t \searrow 0} \frac{F(x + th) - F(x)}{t}$$

in the direction h exists (it can be $\pm\infty$) since the difference quotient

$$t \mapsto \frac{F(x + th) - F(x)}{t}$$

is either constant $+\infty$ or nondecreasing with $t > 0$. Moreover, we notice that $D_{(\cdot)} F(x)$ is sublinear and positively homogeneous, in particular convex. And since the domain of F is a linear subspace, it is its own relative interior, so $D_{(\cdot)} F(x)$ is also proper [55, Thm. 2.1.13].

Let us assume that there is $v \in X^*$ such that $v \in \partial F(x)$. Then, we have $\frac{F(x+th)-F(x)}{t} \geqslant \langle v, h\rangle$, which implies $D_h F(x) \geqslant \langle v, h\rangle$. If in particular, $\|h\| = 1$, we have, for all t,

$$\frac{F(x + th) - F(x)}{t} \geqslant -\|v\|.$$

In consequence, if $\partial F(x) \neq \emptyset$ the limit in (55) cannot be $-\infty$.

For the opposite direction, let us now assume that $\partial F(x) = \emptyset$ for $x \in \operatorname{dom} F$. We want to show that in this case

$$\left(D_{(\cdot)} F(x)\right)^* = +\infty.$$

For that, let us assume the contrary, that is, at some $v \in X^*$,

$$\left(D_{(\cdot)} F(x)\right)^*(v) = \sup_{h \in X} \langle v,\, h\rangle - D_h F(x) = M < +\infty.$$

Noticing that since $D_{(\cdot)} F(x)$ is positively one-homogeneous, $\left(D_{(\cdot)} F(x)\right)^*$ is a characteristic function, so we must have that $M = 0$. This means that for all h, we have

$$D_h F(x) \geqslant \langle v,\, h\rangle\,,$$

which implies that for all $t > 0$,

$$\frac{F(x + th) - F(x)}{t} \geqslant \langle v,\, h\rangle$$

or again, using $t = 1$

$$F(x + h) - F(x) \geqslant \langle v,\, h\rangle$$

and $v \in \partial F(x)$, which is a contradiction.

Conjugating again, we get $\left(D_{(\cdot)} F(x)\right)^{**} = -\infty$. This implies that the lower semicontinuous envelope of the convex function $D_{(\cdot)} F(x)$ is not proper, since otherwise [55, Thm. 2.3.4(i)] it would equal $\left(D_{(\cdot)} F(x)\right)^{**}$, giving a contradiction. And since $D_{(\cdot)} F(x)$ is not everywhere $+\infty$ (just consider $D_0 F(x) = 0$), there must be some $h \in X$ at which this lower semicontinuous envelope takes the value $-\infty$. This means that there is a sequence of directions $h_n \to h$ for which $D_{h_n} F(x) \to -\infty$, and by removing some elements if necessary, we may also assume that $D_{h_n} F(x) < +\infty$ and $h_n \neq 0$ for all n (again since $D_0 F(x) = 0$).

By definition of the directional derivative, there exists t_n (we can choose $t_{n+1} < t_n/2$ to enforce $t_n \to 0$) such that

$$\left| \frac{F(x + t_n h_n) - F(x)}{t_n} - D_{h_n} F(x) \right| \leqslant 1,$$

where the left hand is well defined since $D_{h_n} F(x) < +\infty$.

We obtain then

$$\frac{F(x + t_n h_n) - F(x)}{t_n} \to -\infty.$$

Finally, setting $\tilde{h}_n = \frac{h_n}{\|h_n\|}$ and $\tilde{t}_n = \|h_n\| t_n$, we have $\tilde{t}_n \to 0$ and

$$\frac{F(x + \tilde{t}_n \tilde{h}_n) - F(x)}{\tilde{t}_n} \to -\infty,$$

which was to be shown. □

Example 2 We start with some nonzero function $S \in W^{1,1}(0, 1)$ with $S(0) = S(1) = 0$. For illustration purposes one can think for example of the hat function

$$S(x) = 1 - |2x - 1|.$$

We extend it by zero outside of $(0, 1)$ and define the functions, for $n \geqslant 0$,

$$S_n(x) := S\left[2^n x - \left(2^{n+1} - 2\right)\right]$$

which have support in

$$\left[2 - 2^{-n+1}, 2 - 2^{-n}\right] \subset [0, 2].$$

Note that these supports have an intersection of measure zero. We then introduce $\Omega = (0, 2) \times (0, 1)$ and define $u_n \in W^{1,1}(\Omega)$ by

$$u_n(x_1, x_2) := S_n(x_1).$$

Now, we have that

$$\begin{aligned} \|\nabla u_n\|_{L^1} &= \int_\Omega |\nabla u_n(x_1, x_2)| \, \mathrm{d}x = \int_{(0,2)} |S_n'(t)| \, \mathrm{d}t \\ &= \int_{(0,1)} |S'(t)| \, \mathrm{d}t = \|S'\|_{L^1(0,1)}, \end{aligned}$$

while

$$\|u_n\|_{L^2}^2 = \int_\Omega |u_n(x_1, x_2)|^2 \, \mathrm{d}x = \frac{1}{2^n} \int_{(0,1)} |S(t)|^2 \, \mathrm{d}t = \frac{1}{2^n} \|S\|_{L^2(0,1)}^2.$$

We define then

$$u := \sum_{n=0}^{\infty} 2^{-2n} u_n \in W^{1,1}(\Omega).$$

For it we can use Lemma 6 with $z = 0$,

$$v_n := -\frac{u_n}{\|u_n\|_{L^2}}, \quad t_n := 2^{-2n} \|u_n\|_{L^2} \to 0, \quad |Dv_n|(\Omega) = 2^{n/2} \frac{\|S'\|_{L^1(0,1)}}{\|S\|_{L^2(0,1)}} \to \infty.$$

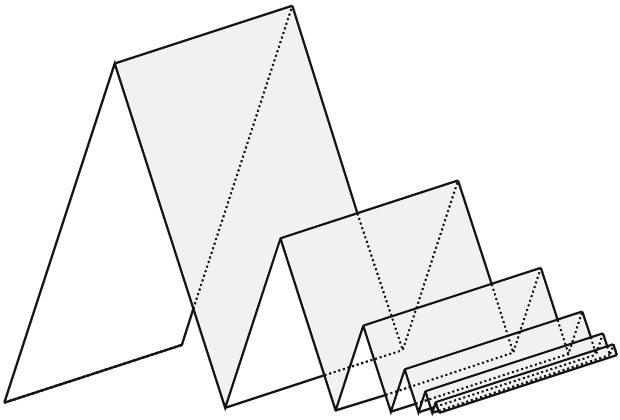

Fig. 2 Sketch of the graph of the function u in Example 2

Moreover by construction $|Du + t_n Dv_n| = |Du| - t_n|Dv_n|$, so we conclude that $|D \cdot |\big((0,2) \times (0,1)\big)$ is not subdifferentiable at u. In particular, for the particular choice $z = Du$ we have $0 \in \partial J_z(u)$, but $\partial\,\mathrm{TV}(u) = \emptyset$. Notice that $Du \in L^1(\Omega; \mathbb{R}^2)$, indicating that lack of subdifferentiability of TV arises not only from the singular part of the derivative of BV functions. We also observe that, extending the v_n to $\mathbb{R}^2$, we have

$$\frac{|Dv_n|}{|Dv_n|(\Omega)} \overset{*}{\rightharpoonup} \mathcal{H}^1 \llcorner \big(\{2\} \times (0,1)\big),$$

and in particular the Lebesgue measure of the support of Dv_n vanishes in the limit.

Remark 10 From the example above we also get a counterexample of the statement we wanted to prove in the first place. The idea is just to take u as above, and for z consisting of the tail of the series starting from some $N > 0$ to remove the small amplitude oscillations in $Du - z$, that is

$$z := Dw, \text{ for } w = \sum_{n=N}^{\infty} 2^{-2n} u_n.$$

Then $\partial\,\mathrm{TV}(u) = \emptyset$ as above, but $\partial J_z(u) \neq \emptyset$ for $J_z = |D \cdot -z|(\Omega)$. Say that $N = 1$ and in the construction of the example we choose

$$S = \rho_{1/8} * 1_{[1/4,3/4]},$$

where $\rho_{1/8}$ is a standard mollifier supported on $(-1/8, 1/8)$. Then the function defined by

$$g(t) = \begin{cases} 8t & \text{if } t \in (0, 1/8) \\ +1 & \text{if } t \in (1/8, 3/8) \\ -8t + 4 & \text{if } t \in (3/8, 5/8) \\ -1 & \text{if } t \in (5/8, 7/8) \\ 8t - 8 & \text{if } t \in (7/8, 1) \end{cases}$$

satisfies $|g(t)| \leqslant 1$, $g(t) = \operatorname{sign}(S'(t))$ whenever $S'(t) \neq 0$, and $g' \in L^\infty(0, 1)$. This implies that the negative divergence of the vector field $\tilde{g}$ defined by $\tilde{g}(x_1, x_2) = (g(x_1), 0)^T$ is in $\partial \operatorname{TV}(u_0) = \partial \operatorname{TV}(u - w) = \partial J_z(u)$, since $-\operatorname{div} \tilde{g}(x_1, x_2) = -g'(x_1)$ and

$$\begin{aligned} \int_\Omega -\operatorname{div} \tilde{g}(x)\, u(x)\, \mathrm{d}x &= \int_\Omega \tilde{g}(x) \cdot \nabla u(x)\, \mathrm{d}x \\ &= \int_0^1 g(t)\, S'(t)\, \mathrm{d}t = \int_0^1 |S'(t)|\, \mathrm{d}t = \operatorname{TV}(u). \end{aligned}$$

Remark 11 The Du and z of the example above are not in BD. However, if one starts with a smoother $S \in W^{2,1}(0, 1)$, the construction works just the same. We would get

$$\|\nabla^2 u_n\|_{L^1} = 2^n \|S''\|_{L^1(0,1)} \text{ but } \sum_{n=1}^{\infty} 2^{-2n} \cdot 2^n < +\infty,$$
$$\text{so } u \in W^{2,1}(\Omega) \text{ and } z \in \left[W^{1,1}(\Omega)\right]^d \subset \operatorname{BD}(\Omega).$$

To summarize, the subdifferential of the functional $J_z = |D \cdot -z|(\Omega)$ being nonempty at u does not automatically imply that $\partial \operatorname{TV}(u) \neq \emptyset$, which prevents us from easily applying the techniques of Sect. 3, in particular Remark 4, to the TGV case. In any case, we note that our counterexamples are built with vector fields z which are not optimal for the inner minimization of (49) defining $\operatorname{TGV}(u)$, so it could still be that $\partial \operatorname{TGV}(u) \neq \emptyset$ implies (beyond the trivial case in Proposition 7) that $\partial \operatorname{TV}(u) \neq \emptyset$.

References

1. Amar, M., Bellettini, G.: A notion of total variation depending on a metric with discontinuous coefficients. Ann. Inst. H. Poincaré C Anal. Non Linéaire **11**(1), 91–133 (1994)
2. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs. Oxford University Press, New York (2000)
3. Andreu, F., Ballester, C., Caselles, V., Mazón, J.M.: The Dirichlet problem for the total variation flow. J. Funct. Anal. **180**(2), 347–403 (2001)
4. Andreu, F., Ballester, C., Caselles, V., Mazón, J.M.: Minimizing total variation flow. Differential Integral Equations **14**(3), 321–360 (2001)

5. Andreu, F., Caselles, V., Díaz, J.I., Mazón, J.M.: Some qualitative properties for the total variation flow. J. Funct. Anal. **188**(2), 516–547 (2002)
6. Andreu-Vaillo, F., Caselles, V., Mazón, J.: Parabolic quasilinear equations minimizing linear growth functionals. Progress in Mathematics, vol. 223. Birkhäuser Verlag, Basel (2004)
7. Anzellotti, G.: Pairings between measures and bounded functions and compensated compactness. Ann. Mat. Pura Appl. (4) **135**(293–318), 1983 (1984)
8. Athavale, P., Jerrard, R.L., Novaga, M., Orlandi, G.: Weighted TV minimization and applications to vortex density models. J. Convex Anal. **24**(4), 1051–1084 (2017)
9. Baldi, A.: Weighted BV functions. Houston J. Math. **27**(3), 683–705 (2001)
10. Bauer, U., Munk, A., Sieling, H., Wardetzky, M.: Persistence barcodes versus Kolmogorov signatures: detecting modes of one-dimensional signals. Found. Comput. Math. **17**(1), 1–33 (2017)
11. Borwein, J.M., Vanderwerff, J.D.: Convex functions: constructions, characterizations and counterexamples. Encyclopedia of Mathematics and its Applications, vol. 109. Cambridge University Press, Cambridge (2010)
12. Bredies, K., Holler, M.: Regularization of linear inverse problems with total generalized variation. J. Inverse Ill-Posed Probl. **22**(6), 871–913 (2014)
13. Bredies, K., Holler, M.: A pointwise characterization of the subdifferential of the total variation functional. Preprint arXiv:1609.08918 [math.FA], (2016)
14. Bredies, K., Holler, M.: Higher-order total variation approaches and generalisations. Inverse Problems, 36(12):123001, 128, (2020)
15. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM J. Imaging Sci. **3**(3), 492–526 (2010)
16. Bredies, K., Valkonen, T.: Inverse problems with second-order total generalized variation constraints. In: Proceedings of the 9th International Conference on Sampling Theory and Applications (SampTA), Singapore, (2011). arXiv:2005.09725
17. Camfield, C. S.: Comparison of BV norms in weighted Euclidean spaces and metric measure spaces. Ph.D. Thesis, University of Cincinnati, (2008)
18. Caselles, V., Chambolle, A., Novaga, M.: Regularity for solutions of the total variation denoising problem. Rev. Mat. Iberoam. **27**(1), 233–252 (2011)
19. Chambolle, A.: An algorithm for mean curvature motion. Interfaces Free Bound. **6**(2), 195–218 (2004)
20. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation for image analysis. In: Theoretical foundations and numerical methods for sparse recovery, volume 9 of *Radon Ser. Comput. Appl. Math.*, pp. 263–340. Walter de Gruyter, Berlin, (2010)
21. Chambolle, A., Duval, V., Peyré, G., Poon, C.: Geometric properties of solutions to the total variation denoising problem. Inverse Prob. **33**(1), 015002 (2017)
22. Chambolle, A., Goldman, M., Novaga, M.: Fine properties of the subdifferential for a class of one-homogeneous functionals. Adv. Calc. Var. **8**(1), 31–42 (2015)
23. Chambolle, A., Lions, P.-L.: Image recovery via total variation minimization and related problems. Numer. Math. **76**(2), 167–188 (1997)
24. Comi, G. E., Crasta, G., De Cicco, V., Malusa, A.: Representation formulas for pairings between divergence-measure fields and BV functions. Preprint arXiv:2208.10812 [math.FA], (2022)
25. Condat, L.: A direct algorithm for 1-D total variation denoising. IEEE Signal Process. Lett. **20**(11), 1054–1057 (2013)
26. Demengel, F.: Fonctions à hessien borné. Ann. Inst. Fourier (Grenoble) **34**(2), 155–190 (1984)
27. Ekeland, I., Témam, R.: Convex analysis and variational problems, volume 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, (1999)
28. Gonzalez, E., Massari, U.: Variational mean curvatures. Rend. Sem. Mat. Univ. Politec. Torino **52**(1), 1–28 (1994)
29. Grasmair, M.: The equivalence of the taut string algorithm and BV-regularization. J. Math. Imaging Vision **27**(1), 59–66 (2007)
30. Grasmair, M., Obereder, A.: Generalizations of the taut string method. Numer. Funct. Anal. Optim. **29**(3–4), 346–361 (2008)
31. Gutev, V.: Lipschitz extensions and approximations. J. Math. Anal. Appl. **491**(1), 124242 (2020)
32. Hausdorff, F.: Über halbstetige Funktionen und deren Verallgemeinerung. Math. Z. **5**(3–4), 292–309 (1919)

33. Hinterberger, W., Hintermüller, M., Kunisch, K., von Oehsen, M., Scherzer, O.: Tube methods for BV regularization. J. Math. Imaging Vision **19**(3), 219–235 (2003)
34. Hintermüller, M., Papafitsoros, K., Rautenberg, C.N.: Analytical aspects of spatially adapted total variation regularisation. J. Math. Anal. Appl. **454**(2), 891–935 (2017)
35. Hintermüller, M., Rautenberg, C.N.: On the density of classes of closed convex sets with pointwise constraints in Sobolev spaces. J. Math. Anal. Appl. **426**(1), 585–593 (2015)
36. Iglesias, J.A., Mercier, G.: Influence of dimension on the convergence of level-sets in total variation regularization. ESAIM Control Optim. Calc. Var. **26**, 52 (2020)
37. Iglesias, J.A., Mercier, G.: Convergence of level sets in total variation denoising through variational curvatures in unbounded domains. SIAM J. Math. Anal. **53**(2), 1509–1545 (2021)
38. Iglesias, J.A., Mercier, G., Scherzer, O.: A note on convergence of solutions of total variation regularized linear inverse problems. Inverse Prob. **34**(5), 055011 (2018)
39. Jalalzai, K.: Some remarks on the staircasing phenomenon in total variation-based image denoising. J. Math. Imaging Vision **54**(2), 256–268 (2016)
40. Kirisits, C., Scherzer, O., Setterqvist, E.: Invariant φ-minimal sets and total variation denoising on graphs. SIAM J. Imaging Sci. **12**(4), 1643–1668 (2019)
41. Kirisits, C., Setterqvist, E.: Universal minimality in TV regularization with rectilinear anisotropy. Preprint arXiv:1910.05186 [math.OC], (2019)
42. Lieb, E. H., Loss, M.: Analysis, volume 14 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, second edition, (2001)
43. Maggi, F.: Sets of finite perimeter and geometric variational problems, volume 135 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, (2012)
44. Massari, U.: Frontiere orientate di curvatura media assegnata in L^p. Rend. Sem. Mat. Univ. Padova **53**, 37–52 (1975)
45. Oberlin, D.M., Stein, E.M.: Mapping properties of the Radon transform. Indiana Univ. Math. J. **31**(5), 641–650 (1982)
46. Pöschl, C., Scherzer, O.: Characterization of minimizers of convex regularization functionals. In: Frames and operator theory in analysis and signal processing, volume 451 of Contemp. Math., pp. 219–248. Amer. Math. Soc., Providence, RI, (2008)
47. Rockafellar, R.T., Wets, R.J.-B.: Variational analysis. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317. Springer-Verlag, Berlin (1998)
48. Rudin, L. I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D, 60(1-4):259–268, (1992). Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991)
49. Savaré, G.: On the regularity of the positive part of functions. Nonlinear Anal. **27**(9), 1055–1074 (1996)
50. Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., Lenzen, F.: Variational methods in imaging. Number 167 in Applied Mathematical Sciences. Springer, New York, (2009)
51. Schuster, T., Kaltenbacher, B., Hofmann, B., Kazimierski, K.S.: Regularization methods in Banach spaces. Radon Series on Computational and Applied Mathematics, vol. 10. Walter de Gruyter GmbH & Co. KG, Berlin (2012)
52. Strong, D. M., Chan, T. F.: Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. UCLA CAM Report 41-96, (1996)
53. Tamanini, I.: Regularity results for almost minimal oriented hypersurfaces in R^N. Quaderni del Dipartimento di Matematica dell'Università di Lecce, 1, (1984). available at http://siba-ese.unisalento.it/index.php/quadmat/issue/view/1073
54. Valkonen, T.: Regularisation, optimisation, subregularity. Inverse Prob. **37**(4), 045010 (2021)
55. Zălinescu, C.: Convex analysis in general vector spaces. World Scientific Publishing Co. Inc, River Edge, NJ (2002)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.