Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Analytical aspects of spatially adapted total variation regularisation

Michael Hintermüller [1, 2], Konstantinos Papafitsoros [1], Carlos N. Rautenberg [2]

submitted: September 5, 2016

[1] Weierstrass Institute
    Mohrenstr. 39
    10117 Berlin
    Germany
    E-Mail: michael.hintermueller@wias-berlin.de
         kostas.papafitsoros@wias-berlin.de

[2] Department of Mathematics
    Humboldt-Universität zu Berlin
    Unter den Linden 6
    10099 Berlin. Germany
    E-Mail: carlos.rautenberg@math.hu-berlin.de

No. 2293

Berlin 2016

Abstract

In this paper we study the structure of solutions of the one dimensional weighted total variation regularisation problem, motivated by its application in signal recovery tasks. We study in depth the relationship between the weight function and the creation of new discontinuities in the solution. A partial semigroup property relating the weight function and the solution is shown and analytic solutions for simply data functions are computed. We prove that the weighted total variation minimisation problem is well-posed even in the case of vanishing weight function, despite the lack of coercivity. This is based on the fact that the total variation of the solution is bounded by the total variation of the data, a result that it also shown here. Finally the relationship to the corresponding weighted fidelity problem is explored, showing that the two problems can produce completely different solutions even for very simple data functions.

1 Introduction

A general task in mathematical signal reconstruction is to recover as best as possible a signal u_0, given a corrupted version f, which is generated by the following degradation process:

$$f = Tu_0 + \eta. \tag{1.1}$$

Here T denotes a bounded, linear operator and η is a random noise component. The mapping T might be related to blurring, downscaling, Fourier or wavelet transform, among several others. The problem aforementioned reconstruction problem (1.1) is typically ill-posed and variational regularisation methods are often employed for its solution. A specific, very successful regularisation model is given by total variation minimisation as introduced in the seminal work by Rudin, Osher and Fatemi [ROF92]. In that paper, the authors considered the case $T = id$, i.e., the denoising task, and they proposed to recover an approximation u of u_0 by solving the discrete version of the minimisation problem

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f - u)^2 dx + \alpha |Du|(\Omega), \tag{1.2}$$

where α is a positive scalar and $|Du|(\Omega)$ denotes the total variation of the function u. Here Ω represents a bounded, open domain with Lipschitz boundary. In image reconstruction tasks, Ω is typically a rectangle on which the image is defined.

Ever since, total variation minimisation has been employed for a variety of image restoration tasks mainly due its edge-preserving ability. This stems from the fact that the minimisation (1.2), is performed over the space of functions of bounded variation $\mathrm{BV}(\Omega)$. We note that an element of $\mathrm{BV}(\Omega)$ may exhibit jump discontinuities. One disadvantage of the model (1.2), on the other hand, is the promotion of piecewise constant structures in the

solution u, a phenomenon known as staircasing effect. To overcome this, higher order extensions of the total variation have been proposed in the literature. Here, we only mention the review paper [Ste15], as well as the references collected in the introduction of [Pap14].

Another drawback of (1.2) originates from the fact that the regularisation strength is uniform over the entire image domain, due to the regularisation parameter α being a scalar quantity only. This is particularly disadvantageous when the noise level or the amount of corruption in general, is not distributed uniformly throughout the image. Regularisation of uniform strength is also undesirable when both fine scales details, e.g. texture, and large homogeneous areas are present in an image. In that case, one ideally should strongly regularise in the smooth parts of the image and to a lesser degree in fine detailed areas in order for these details to be better preserved. Therefore, the introduction of spatially distributed weights in the minimising functional in (1.2) has been considered in the literature. This weight can be either introduced in the first term of (1.2), i.e., the so-called fidelity or data fitting term, or be incorporated into the total variation functional. For the denoising case, this leads to the following two models

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_{\Omega} w(f-u)^2 dx + |Du|(\Omega), \tag{1.3}$$

where $w \in L^{\infty}(\Omega)$ with $w \geq 0$, and

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_{\Omega} (f-u)^2 dx + \int_{\Omega} \alpha(x) d|Du|, \tag{1.4}$$

where $\alpha \in C(\overline{\Omega})$ with $\alpha > 0$. Here α and w are the two weight functions that determine locally the strength of the regularisation and the fidelity term respectively.

Versions of the weighted fidelity model (1.3) have been considered in [DHRC10a, DHRC10b] for Gaussian denoising and deblurring image restoration problems as well as in [HRC10] for reconstructing images that have been corrupted by impulse noise. In these works, the weight function w is selected based on local statistical estimators and the statistics of the extremes. An adaptation of this idea to TGV (total generalised variation) [BKP10], a higher order extension of the total variation, can be found in [BDH13]. A different statistical approach with variance estimators is considered in [ABCH07]. Variants of (1.3) are also studied in [FMM12, HMS^{+}12] using techniques based on a statistical multiresolution criterion. The model (1.3) is also considered in [BCRS] where a piecewise constant weight function is determined using a pre-segmentation of the image.

The weighted total variation model (1.4) has been considered recently for image restoration purposes in [HR16, HRWL16]. In these papers, the choice of the weight function α is done via a bilevel optimisation approach, see also [CCDIR^{+}15, CDIRS16]. Moreover, apart from the classical denoising and deblurring tasks, the fact that the fidelity term in (1.4) appears without weights, allows the authors of [HR16, HRWL16] to consider problems also in Fourier and wavelet domains, e.g., Fourier and wavelet impainting, something which highlights an advantage of the model (1.4) over (1.3). We also mention that recently, a weighted TV regularisation for vortex density models was studied in [AJNO15].

While the analysis of the regularisation properties of the scalar total variation regularisation (1.2) has received a considerable amount of attention in the literature [AV94, Rin00, Mey01, SC03, CE05, Gra07, CCN07, All08a, All08b, All09, DAG09, Val15, Jal15, CDPP16] this is not the case for the models (1.3) and (1.4). We note however two analytical contributions towards the weighted total variation model (1.4). Specifically, in [Jal14], the author

showed that the set of the jump discontinuities of the solution u of (1.4) is essentially contained in the set that consists of the jump discontinuity points of the data f and the jump discontinuity points of the gradient of the weight function α. This result shows that the solution u can potentially have jump discontinuities at points where the data function f is continuous. Hence, if the weight function is smooth, no new discontinuities are created. In the scalar parameter case, this was first shown in [CCN07] and subsequently in [Val15] using a different technique. In [AJNO15] the authors show, among others, that the maximal level set of the solution u is flat and has positive measure, as it is also the case for the scalar total variation regularisation [Jal15].

However, there are still several open questions regarding the models (1.3) and (1.4). For instance, concerning the weighted total variation model (1.4), one is interested to understand, under which specific conditions new discontinuities are created, and how these are related to the weight function α. It is also important to examine in what degree the structure of solutions of the weighted total variation minimisation resembles the one of the solutions of the standard scalar minimisation problem. Finally, it is of importance to understand the similarities and the differences of the two weighted models (1.3) and (1.4).

In view of this, the purpose of the present paper is to answer the questions raised above as well as related ones, thus filling in that knowledge gap in the literature. We do that by a extended fine scale analysis of the regularisation properties of the one dimensional versions of the problems (1.3) and (1.4). We should note, however, that the majority of our results concern the weighted total variation model (1.4), since, as we will see in the following sections, it is the one that exhibits a greater variety of interesting properties.

Summary of the results and organisation of the paper

For the reader's convenience we provide here a short summary of our results, stating as well the sections of the paper that each of these belongs to. The results are put into perspective with the literature in the corresponding sections.

Structure of solutions-creation of new discontinuities

After fixing the notation and recalling some preliminary facts in Section 2, we study in Sections 3.1–3.2 the weighted total variation problem (1.4) and the conditions under which new discontinuities are created in its solution u. We give a simple proof of a refined version of the result in [Jal14] in Proposition 3.3, showing that new jump discontinuities can potentially be created at the points where the weight function α is not differentiable. Note that in the case where $\alpha' \in \mathrm{BV}(\Omega)$ these are exactly the set of jump discontinuity points of α', i.e., the set of points $x \in \Omega$ such that $|D\alpha'|(\{x\}) > 0$. In fact, we show that in order for a new discontinuity to be created at x, there must hold $D\alpha'(\{x\}) > 0$, i.e., the derivative of α' must have a positive jump, Proposition 3.7. In contrast, if $D\alpha'(\{x\}) < 0$ then a plateau is created for the solution u around x. Furthermore, we show that in every point $x \in \Omega$, the following estimate holds

$$|Du|(\{x\}) \leq |Df|(\{x\}) + |D\alpha'|(\{x\}), \tag{1.5}$$

see Propositions 3.7, 3.9 as well as Corollary 3.8. Moreover, in the weighted case, the jump of u can have different direction from the one of the data f, something that does not occur in the scalar case. We show however that the jumps of f and u at a point, have the same direction when α is differentiable but are not necessarily *nested*, see Proposition 3.9

3

and the numerical examples of Figure 3. Finally it is shown that if α' is large enough in an area, then u is constant there, Proposition 3.10. Thus, it is not only high values of α that can produce flat areas as someone might expect, but also high values of α'.

A partial semigroup property

In Section 3.3 we show that, denoting by $S_\alpha(f)$ the solution of (1.4) with data f and weight α, it holds

$$S_{\alpha_1+\alpha_2}(f) = S_{\alpha_2}(S_{\alpha_1}(f)),$$

provided α_2 is a scalar. This is shown in Proposition 3.12 and we call this property *partial semigroup property*. On the other hand, one can easily construct counterexamples where this property fails even in the case where α_1 is scalar and α_2 is not, see Figures 4 and 5, i.e., unlike the full scalar case, this partial semigroup property is not commutative, something perhaps surprising.

Analytic solutions

In Section 3.4 we compute some analytic solutions for simple data and weight functions. In particular, we take as data a family of affine functions and as weight functions, a family of absolute value type functions. The formulae of the solutions are summarised in Proposition 3.13, also depicted in Figure 6. Note that this is the first example, where the creation of new discontinuities is computed analytically.

A bound on the total variation of the solution

In Section 4 we show that for the solution of the weighted total variation minimisation problem (1.4), the following estimate holds

$$|Du|(\Omega) \leq |Df|(\Omega). \tag{1.6}$$

Unlike the scalar case, the proof of (1.6) is quite involved and uses some fine scale analysis. We do that initially for differentiable weight α in Theorem 4.2 and then for continuous one in Theorem 5.4.

Vanishing weight function α

Provided that $f \in \mathrm{BV}(\Omega)$, we show in Section 5 the existence of solutions for (1.4) even when $\alpha \geq 0$, despite the lack of coercivity of the minimising functional, see Theorem 5.3. Letting α having zero values, can allow an exact recovery of piecewise constant functions, as we show in Proposition 5.5.

Relationship of the models (1.3) **and** (1.4)

In Section 6, we show that the structure of the solutions of the weighted fidelity problem (1.3) is simpler and resembles more the one of the scalar case. We prove that no new discontinuities are created, provided that $w > 0$, Proposition 6.2. Moreover, by considering the same family of simple affine data functions for which we computed analytic solutions for the problem (1.4), we see that the solutions here are much simpler, see Proposition 6.3. Interestingly, for these specific data functions, the sets of solutions of the problems (1.3)

and (1.4) are totally different, regardless of the choice of weight functions α and w. In fact, the only common solutions that they have are the ones that can be also obtained by the standard scalar total variation minimisation, see Proposition 6.4 and Figure 12. This shows how different can the models (1.3) and (1.4) be, even for very simple data functions.

2 Notation and Preliminaries

Functions of bounded variation play a central role in this paper. Standard references are the books [AFP00, ABM14, EG92, Giu84]. Here we follow the notation of [AFP00]. Let $\Omega \subseteq \mathbb{R}^d$ be a open set, $d \in \mathbb{N}$. Given a finite Radon measure $\mu \in \mathcal{M}(\Omega)$ we denote by $|\mu|$ its total variation measure and by $\mathrm{sgn}(\mu)$ the unique $L^1(\Omega, |\mu|)$ function such that $\mu = \mathrm{sgn}(u)|\mu|$. That is to say $\mathrm{sgn}(\mu)$ is the Radon-Nikodým derivative $\mathrm{sgn}(\mu) = \frac{d\mu}{d|\mu|}$, which is equal to 1 $|\mu|$–almost everywhere. A function $u \in L^1(\Omega)$ is said to be a *function of bounded variation* if its distributional derivative is represented by a $\mathbb{R}^d$-valued finite Radon measure, denoted by Du. Equivalently, u is a function of bounded variation if its *total variation* $\mathrm{TV}(u)$ is finite, where

$$
\mathrm{TV}(u) := \sup \left\{ \int_\Omega u \, \mathrm{div} v \, dx : \ v \in C_c^1(\Omega, \mathbb{R}^d), \ \|v\|_\infty \leq 1 \right\},
$$

and in that case it can be shown that $\mathrm{TV}(u) = |Du|(\Omega)$. The space of functions of bounded variation is denoted by $\mathrm{BV}(\Omega)$ and is a Banach space under the norm $\|u\|_{\mathrm{BV}(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega)$. The measure Du can be decomposed into the absolutely continuous and the singular part with respect to the Lebesgue measure $\mathcal{L}$, $D^a u$ and $D^s u$ respectively, i.e., $Du = D^a u + D^s u$.

In this paper, emphasis is given on the functions of bounded variation of one variable and in particular on the notion of *good representatives*. In order to define these, let $\Omega = (a, b)$ be a bounded open interval in $\mathbb{R}$. For a function $u : \Omega \to \mathbb{R}$, the *pointwise variation* of u in Ω is defined as

$$
\mathrm{pV}(u, \Omega) = \sup \left\{ \sum_{i=1}^{n-1} |u(x_{i+1}) - u(x_i)| : \ n \geq 2, \ a < x_1 < \cdots < x_n < b \right\},
$$

and the *essential variation* as

$$
\mathrm{eV}(u, \Omega) = \inf \left\{ \mathrm{pV}(v, \Omega) : v = u, \ \mathcal{L} - \text{a.e. in } \Omega \right\}.
$$

It turns out that when $u \in \mathrm{BV}(\Omega)$ then $|Du|(\Omega) = \mathrm{eV}(u, \Omega)$ and in fact the infimum in the definition of $\mathrm{eV}(u, \Omega)$ is attained. The functions in the equivalence class of u that attain this infimum are called good representatives of u. That is to say $\tilde{u}$ is a good representative of u if $\tilde{u} = u$ Lebesgue–almost everywhere and

$$
\mathrm{pV}(\tilde{u}) = \mathrm{eV}(u) = |Du|(\Omega).
$$

We denote by J_u the at most countable set of atoms of Du (jump set of u), i.e., $J_u = \{x \in \Omega : \ |Du|(\{x\}) \neq 0\}$. If $Du(\{x\}) > 0$ we say that u has a positive jump at x, whereas if $Du(\{x\}) < 0$ we say that u has a negative jump at x. It can be shown that there exists a unique $c \in \mathbb{R}$ such that the functions

$$
u^l(x) := c + Du((a, x)), \quad u^r(x) := c + Du((a, x]),
$$

are good representatives of u. Note that u^l and u^r are left and right continuous respectively. The following equalities also hold

$$u^l(x) = \lim_{\delta \downarrow 0} \int_{x-\delta}^x u(t)dt, \quad u^r(x) = \lim_{\delta \downarrow 0} \int_x^{x+\delta} u(t)dt, \quad \text{for all } x \in \Omega.$$

Any other function $\tilde{u} : \Omega \to \mathbb{R}$ is a good representative of u if and only if

$$\tilde{u}(x) \in \left\{ \theta u^l(x) + (1-\theta)u^r(x) : \ \theta \in [0,1] \right\}.$$

As a result, the following functions are also good representatives

$$\overline{u}(x) := \max(u^l(x), u^r(x)), \quad \underline{u}(x) := \min(u^l(x), u^r(x)).$$

The right and the left limits of any good representative $\tilde{u}$ exist at any point of $x \in \Omega$ and $\tilde{u}(x_+) = u^r(x), \tilde{u}(x_-) = u^l(x)$. Every good representative of u is continuous at the complement of the jump set of u i.e., in the set $\{x \in \Omega : \ |Du|(\{x\}) = 0\}$.

We denote by u' the density of $D^a u$ with respect to the Lebesgue measure, i.e.,

$$Du = u'\mathcal{L} + D^s u.$$

If $u \in W^{1,1}(\Omega)$ then u' is the standard weak derivative of u.

Recall some basic notions from convex analysis. If X, X^* are two vector spaces placed in duality and $F : X \to \mathbb{R} \cup \{+\infty\}$ then F^* denotes the convex conjugate of F

$$F^*(x^*) := \sup_{x \in X} \langle x^*, x \rangle - F(x).$$

The subdifferential of F is denoted as usual by ∂F. Given $A \subseteq X$ then $\mathcal{I}_A$ denotes the indicator function of A

$$\mathcal{I}(x) = \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \notin A. \end{cases}$$

We finally note that whenever we write *total variation regularisation* or *total variation minimisation* we always mean the total variation denoising problem with L^2 fidelity term.

3 Weighted total variation with strictly positive weight function α

The problem we are considering here is the one dimensional weighted total variation regularisation problem with L^2 fidelity term, i.e.,

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f-u)^2 dx + \int_\Omega \alpha(x)d|Du|, \tag{3.1}$$

where $\Omega = (a,b)$, $f \in \mathrm{BV}(\Omega)$ and $\alpha \in C(\overline{\Omega})$ with $\alpha > 0$. Thus, there exist constants $0 < c_\alpha \leq C_\alpha < \infty$ such that

$$0 < c_\alpha \leq \alpha(x) \leq C_\alpha < \infty, \quad \text{for all } x \in \Omega.$$

The well-posedness (existence and uniqueness) of (3.1) in all dimensions, i.e., when $\Omega \subseteq \mathbb{R}^d$, is proven via the direct method of calculus of variations taking advantage of the fact

that a strictly positive weight function α provides the necessary coercivity to the weighted TV functional, see [HR16] for details. Among others, the authors in [HR16] prove that for the anisotropic version of weighted TV it holds

$$
\int_{\Omega} \alpha(x)d|Du| = \sup \left\{ \int_{\Omega} u \operatorname{div} v \, dx : \ v \in H_0(\Omega, \operatorname{div}), \right.
$$
$$
\left. v_i(x)| \leq \alpha(x), \ \text{a.e. } i = 1, \ldots, d \right\}. \tag{3.2}
$$

Using also appropriate density arguments, the isotropic version of (3.2) reads

$$
\int_{\Omega} \alpha(x)d|Du| = \sup \left\{ \int_{\Omega} u \operatorname{div} v \, dx : \ v \in C_c^1(\Omega, \mathbb{R}^d), \right.
$$
$$
\left. |v(x)| \leq \alpha(x), \ \text{for every } x \in \Omega \right\}, \tag{3.3}
$$

where in the expression above $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$. It is then clear that the weighted TV is lower semicontinuous with respect to the strong convergence in L^1.

3.1 Optimality conditions

We now proceed to the derivation of the optimality conditions for the minimisation problem (3.1). This is done via the Fenchel-Rockafellar duality theory, see for instance [ET76].

We start with some useful definitions. For a finite Radon measure $\mu \in \mathcal{M}(\Omega)$ we define

$$
\operatorname{Sgn}(\mu) := \left\{ v \in L^{\infty}(\Omega) \cap L^{\infty}(\Omega, \mu) : \ \|v\|_{\infty} \leq 1, \ v = \operatorname{sgn}(\mu), \ |\mu| - \text{a.e.} \right\}, \tag{3.4}
$$

i.e., the set of all the functions v that are μ-almost everywhere equal to $\frac{d\mu}{d|\mu|}$ with the extra property that their absolute values is less than 1, Lebesgue–almost everywhere. The definition (3.4) originates from [BKV13]. For a function $\alpha \in C(\overline{\Omega})$, we also define

$$
\alpha(x)\operatorname{Sgn}(\mu) := \left\{ v \in L^{\infty}(\Omega) \cap L^{\infty}(\Omega, \mu) : \ v = \alpha \tilde{v} \text{ for some } \tilde{v} \in \operatorname{Sgn}(\mu) \right\}. \tag{3.5}
$$

Notice that we slightly abuse the notation in the definition (3.5) where we denote the set by "$\alpha(x)\operatorname{Sgn}(\mu)$" instead of "$\alpha\operatorname{Sgn}(\mu)$" in order to stress the fact that α is not necessarily a constant function.

The following proposition is an extension of [BKV13, Lemma 3.5] to the weighted case.

Lemma 3.1 (Subdifferential of the weighted Radon norm)**.** *Let* $\alpha \in C(\overline{\Omega})$*. Consider the map* $\|\cdot\|_{\mathcal{M},\alpha} : \mathcal{M}(\Omega) \to \mathbb{R}$ *where*

$$
\|\mu\|_{\mathcal{M},\alpha} = \int_{\Omega} \alpha(x)d|\mu|, \quad \mu \in \mathcal{M}(\Omega).
$$

Then for every $\mu \in \mathcal{M}(\Omega)$

$$
\partial \|\cdot\|_{\mathcal{M},\alpha}(\mu) \cap C_0(\Omega) = \alpha(x)\operatorname{Sgn}(\mu) \cap C_0(\Omega),
$$

Proof. Fix $\mu \in \mathcal{M}(\Omega)$ and let $v \in \partial \| \cdot \|_{\mathcal{M},\alpha}(\mu) \cap C_0(\Omega)$. Then

$$\int_\Omega \alpha(x)d|\mu| + \int_\Omega v(x)d(\nu - \mu) \leq \int_\Omega \alpha(x)d|\nu| \quad \text{for every } \nu \in \mathcal{M}(\Omega) \Rightarrow \quad (3.6)$$

$$\int_\Omega v(x)d(\nu - \mu) \leq \int_\Omega \alpha(x)d|\nu - \mu| \quad \text{for every } \nu \in \mathcal{M}(\Omega) \Rightarrow$$

$$\int_\Omega v(x)d\nu \leq \int_\Omega \alpha(x)d|\nu| \quad \text{for every } \nu \in \mathcal{M}(\Omega). \quad (3.7)$$

From the inequality (3.7) we deduce that

$$|v(x)| \leq \alpha(x) \quad \text{for every } x \in \Omega. \quad (3.8)$$

Observe that it also holds

$$\int_\Omega v(x)d\mu = \int_\Omega \alpha(x)d|\mu|. \quad (3.9)$$

Indeed, just consider (3.6) with $\nu = 0$ and $\nu = 2\mu$. One can readily check that if a function $v \in C_0(\Omega)$ satisfies (3.8)–(3.9) then $v \in \partial \| \cdot \|_{\mathcal{M},\alpha}(\mu) \cap C_0(\Omega)$. Then it just suffices to check that a function $v \in C_0(\Omega)$ satisfies (3.8)–(3.9) if and only if $v \in \alpha(x)\mathrm{Sgn}(\mu) \cap C_0(\Omega)$. The "if" implication is immediate from the definition of $\alpha(x)\mathrm{Sgn}(\mu)$. For the "only if" part, by considering the polar decomposition $\mu = \mathrm{sgn}(\mu)|\mu|$ we have

$$\int_\Omega v(x)d\mu = \int_\Omega \alpha(x)d|\mu| \Rightarrow \int_\Omega (v(x)\mathrm{sgn}(\mu)(x) - \alpha(x))d|\mu| = 0,$$

which, with the help of (3.8), implies that

$$v(x)\mathrm{sgn}(\mu)(x) = \alpha(x) \quad \text{for } |\mu|\text{-almost every } x \Rightarrow$$

$$v(x) = \mathrm{sgn}(\mu)(x)\alpha(x) \quad \text{for } |\mu|\text{-almost every } x.$$

Thus, $v \in \alpha(x)\mathrm{Sgn}(\mu) \cap C_0(\Omega)$ and the proof is complete. $\qquad \square$

We define now the predual problem of (3.1):

$$-\min\left\{ \frac{1}{2}\int_\Omega (v')^2 dx + \int_\Omega fv' dx : v \in H_0^1(\Omega), |v(x)| \leq \alpha(x), \text{ for every } x \in \Omega \right\}. \quad (3.10)$$

The fact that the minimum in (3.10) is attained by a unique H_0^1 function, can be shown easily using standard techniques. In order to be convinced that (3.10) is indeed the predual of (3.1) define

$$\Lambda : H_0^1(\Omega) \to L^2(\Omega) \quad \text{with} \quad \Lambda(v) = v',$$

$$G : L^2(\Omega) \to \mathbb{R} \quad \text{with} \quad G(\psi) = \frac{1}{2}\int_\Omega \psi^2 dx + \int_\Omega f\psi \, dx,$$

$$F : H_0^1(\Omega) \to \overline{\mathbb{R}} \quad \text{with} \quad F(v) = \mathcal{I}_{\{|\cdot(x)| \leq \alpha(x), \forall x \in \Omega\}}.$$

Then it is easy to verify that the problem (3.10) is equivalent to

$$- \min_{v \in H_0^1(\Omega)} F(v) + G(\Lambda v). \quad (3.11)$$

8

Now the dual problem of (3.11) is defined as [ET76]

$$\min_{u \in L^2(\Omega)^*} F^*(-\Lambda^* u) + G^*(u). \tag{3.12}$$

After a few computations the problem (3.12) can be shown to be equivalent with our main problem (3.1). The proof follows closely the analogue proofs in [Rin00], [BKV13] and [PB15] for the corresponding L^2–TV (scalar case), L^1–TGV and L^2–TGV minimisations and thus we omit it. We note here that the derivation of the predual problem of (3.1) in higher dimensions is more involved, see [HR16]. The solutions of the problems (3.11) and (3.12) are linked through the optimality conditions:

$$v \in \partial F^*(-\Lambda^* u),$$
$$\Lambda v \in \partial G^*(u),$$

which, after a few calculations, can be reformulated as

$$v' = f - u,$$
$$-v \in \alpha(x)\mathrm{Sgn}(Du).$$

Summarising, the following proposition holds.

Proposition 3.2 (Optimality conditions for weighted TV minimisation)**.** *Let $\Omega = (a,b)$, $f \in \mathrm{BV}(\Omega)$ and $\alpha \in C(\overline{\Omega})$ with $\alpha > 0$. A function $u \in \mathrm{BV}(\Omega)$ is the solution to the minimisation problem*

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2}\int_\Omega (f-u)^2 dx + \int_\Omega \alpha(x)d|Du|,$$

if and only if there exists a function $v \in H_0^1(\Omega)$ such that

$$v' = f - u, \tag{3.13}$$
$$-v \in \alpha(x)\mathrm{Sgn}(Du). \tag{3.14}$$

Observe here that Proposition 3.2 still holds when $f \in L^2(\Omega)$. This is useful in the context of image denoising, where f is a noisy, perhaps strongly oscillating function, modelled as an element outside $\mathrm{BV}(\Omega)$. Here, in contrast, we assume that $f \in \mathrm{BV}(\Omega)$, since in this study, we are more interested in the structural properties of weighted TV minimisation than addressing the entire reconstruction problem. Observe that since we are in dimension one, this also implies that we have more than H_0^1 regularity for the function v. Indeed, $v' \in \mathrm{BV}(\Omega) \subseteq L^\infty(\Omega)$ and in particular v is a Lipschitz function.

3.2 Structure of solutions – creation of new discontinuities

One can already notice a basic difference between the scalar and the weighted total variation regularisation. Indeed, when $\alpha(x) = \alpha \in \mathbb{R}$ for every $x \in \Omega$, the optimality conditions (3.13)–(3.14) imply that when $\overline{f} < \underline{u}$ (or $\underline{f} > \overline{u}$) then $Du = 0$ there. That is to say, the solution u is constant in the areas where it is not equal to the data f, a well-known characteristic of total variation minimisation [Rin00]. In the weighted case, however, the optimality conditions (3.13)–(3.14) do not enforce such a behaviour. In this section, using a series of propositions and numerical examples we highlight the differences between the

9

scalar and the weighted case as far as the structure of solutions is concerned. Particular emphasis is given on the discontinuities of the solution u. Recall here that one of the few analytical results concerning the weighted TV regularisation is that of Jalalzai [Jal14]. There, the author shows that given $\Omega \subseteq \mathbb{R}^d$ open, bounded with Lipschitz boundary, data $f \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$, and a bounded, Lipschitz continuous weight function α with the extra property that $\nabla \alpha \in \mathrm{BV}(\Omega)$, then

$$J_u \subseteq J_f \cup J_{\nabla\alpha}, \tag{3.15}$$

up to $\mathcal{H}^{d-1}$ negligible set. Here $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure. This result shows that new jump discontinuities can potentially appear in the solution u at points where the derivative of the weight function also has a jump. This is in strong contrast to the scalar TV minimisation where the discontinuities of the solution can only occur in points where the data f is discontinuous [CCN07, Val15]. Note that this also true in the weighted case when $\alpha \in C^1(\Omega)$ since then $J_{\nabla\alpha} = \emptyset$.

Here we investigate in detail, the creation of new discontinuities in the one dimensional regime. We will show with analytical and numerical results that at least in dimension one, the inclusion (3.15) is sharp. In order to develop an intuition for this phenomenon, we start with a simple proof of (3.15) in the one dimensional case. Note that we do not assume here that α is Lipschitz continuous with $\alpha' \in \mathrm{BV}(\Omega)$.

Proposition 3.3. *Let $u \in \mathrm{BV}(\Omega)$ be a solution to (3.1) and let $x \in \Omega$ such that α is differentiable at x and $|Df|(\{x\}) = 0$, i.e., $x \notin J_f$. Then $x \notin J_u$.*

Proof. Suppose, towards contradiction, that $x \in J_u$, i.e., $|Du|(\{x\}) > 0$. Without loss of generality we assume that $Du(\{x\}) > 0$ since the case $Du(\{x\}) < 0$ is treated analogously. Hence, we have

$$u^l(x) < u^r(x). \tag{3.16}$$

Since $|Df|(\{x\}) = 0$ we have that any good representative $\tilde{f}$ of f is continuous at x. Using (3.16), the continuity of $\tilde{f}$, the left and right continuity of $u^l(x)$ and $u^r(x)$, respectively, we have that there exist a small enough $\epsilon > 0$ and two constants $m < M$ such that

$$\sup_{t\in(x,x+\epsilon)} \tilde{f}(t) - u^r(t) \le m < M \le \inf_{t\in(x-\epsilon,x)} \tilde{f}(t) - u^l(t). \tag{3.17}$$

With the help of (3.13), the above inequalities are translated into

$$\operatorname*{ess\,sup}_{t\in(x,x+\epsilon)} v'(t) \le m < M \le \operatorname*{ess\,inf}_{t\in(x-\epsilon,x)} v'(t). \tag{3.18}$$

Since $Du(\{x\}) > 0$, condition (3.14) dictates that

$$v(x) = -\alpha(x).$$

Using now the fundamental theorem of calculus along with (3.18) we get that for every $t \in (x, x + \epsilon)$

$$
\begin{aligned}
v(t) &= -\alpha(x) + \int_x^t v'(t)dt \\
&\le -\alpha(x) + m(t-x),
\end{aligned}
$$

and for every $t \in (x - \epsilon, x)$

$$v(t) = -\alpha(x) + \int_t^x -v'(t)dt$$
$$\leq -\alpha(x) + M(t - x).$$

Using the fact that $-\alpha(t) \leq v(t)$ for every $t \in \Omega$ and condition (3.14), we further calculate

$$\lim_{t \to x-} \frac{\alpha(x) - \alpha(t)}{x - t} \leq \frac{\alpha(x) + v(t)}{x - t} \leq \frac{M(t - x)}{x - t} = -M \qquad (3.19)$$

and

$$\lim_{t \to x+} \frac{\alpha(t) - \alpha(x)}{t - x} \geq \frac{-v(t) - \alpha(x)}{t - x} \geq \frac{-m(t - x)}{t - x} = -m. \qquad (3.20)$$

The inequalities (3.19)–(3.20) contradict the differentiability of α at x and thus the proof is complete. $\qquad\square$

Even though it is now clear that non-differentiablity of α can potentially lead to the creation of new discontinuities, as the next proposition shows this is not always the case. In particular, we show in what follows that if α has an upward spike at a point x, then the solution u of (3.1) is constant in a neighbourhood of x; see Figure 1 for an illustration.

Proposition 3.4. *Let $\alpha \in C(\overline{\Omega})$ and $x \in \Omega$ such that $\alpha \in C(\overline{\Omega})$ is differentiable in a neighbourhood of x (but not at x) with*

$$\lim_{t \to x-} \alpha'(t) = +\infty \quad \text{and} \quad \lim_{t \to x+} \alpha'(t) = -\infty.$$

Then, if u is the solution of (3.1) *with weight function α and some given data $f \in \mathrm{BV}(\Omega)$, then there exists an $\epsilon > 0$ such that $|Du|((x - \epsilon, x + \epsilon)) = 0$, i.e., u is constant in $(x - \epsilon, x + \epsilon)$.*

Proof. We show first that there exists an $\epsilon > 0$ such that $|Du|((x, x + \epsilon)) = 0$. Indeed otherwise, using the condition (3.14), we can assume without loss of generality, that there exists a decreasing sequence $(t_n)_{n \in \mathbb{N}}$ such that $t_n \downarrow x$ with $x < t_n$ and

$$v(t_n) = -\alpha(t_n), \quad \text{for every } n \in \mathbb{N}.$$

But then, using the mean value theorem, we have for some $t_{n+1} < \xi_n < t_n$

$$\frac{|v(t_n) - v(t_{n+1})|}{|t_n - t_{n+1}|} = \frac{|\alpha(t_n) - \alpha(t_{n+1})|}{|t_n - t_{n+1}|}$$
$$= |\alpha'(\xi_n)|.$$

Since $|\alpha'(\xi_n)| \to \infty$, the equality above implies that v is not Lipschitz, a contradiction. Similarly we get $|Du|((x - \epsilon, x)) = 0$ for a small enough $\epsilon > 0$. Finally notice that it also holds that $|Du|(\{x\}) = 0$. Otherwise, again from condition (3.14), we would have that $v(x) = -\alpha(x)$ (or $v(x) = \alpha(x)$, with a similar proof) and using also the fact that $v \geq -\alpha$, we have for $t > x$

$$\frac{v(t) - v(x)}{t - x} \geq \frac{-\alpha(t) + \alpha(x)}{t - x} \to +\infty \quad \text{as } t \to x+, \qquad (3.21)$$

again contradicting the fact that v is Lipschitz. Hence, for a small enough $\epsilon > 0$ we have

$$|Du|((x - \epsilon, x + \epsilon)) = |Du|((x - \epsilon, x)) + |Du|(\{x\}) + |Du|((x, x + \epsilon)) = 0.$$
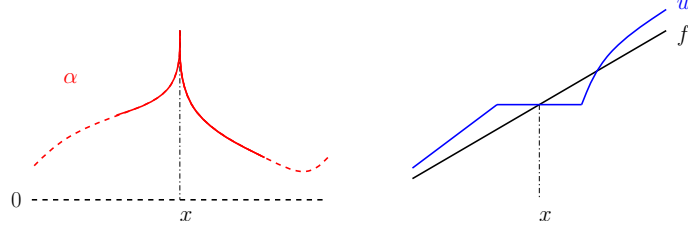
$\qquad\square$

Figure 1: Illustration of Proposition 3.4: When the weight function α has an upward spike at a point x (left plot) then the solution u of (3.1) is constant at an neighbourhood of x (right plot).

Observe that it is not essential to assume that α is differentiable at a set of the type $(x - \delta, x) \cup (x, x + \delta)$ for small enough $\delta > 0$. For example it would be enough to assume that α is concave at each of the intervals $(x - \delta, x)$ and $(x, x + \delta)$ and its graph does not satisfy the cone property at x.

After examining the case where α has an upward spike, it is natural to ask what happens if α exhibits a downward spike. The following proposition provides some intuition.

Proposition 3.5. *Let $f \in \mathrm{BV}(\Omega)$ such that f is continuous and strictly increasing. Suppose that $\alpha \in C(\overline{\Omega})$ is differentiable everywhere in Ω apart from a point x and*

$$\lim_{t \to x-} \alpha'(t) = -\infty \quad and \quad \lim_{t \to x+} \alpha'(t) = +\infty$$

with α attaining its minimum at x. Then, if u is the solution of (3.1) for the weight function α and data f, it has either a jump discontinuity at x or it is constant up to the boundary of Ω.

Proof. Similarly to the proof of Proposition 3.4 we can deduce that there exists an $\epsilon > 0$ such that $|Du|((x-\epsilon, x) \cup (x, x+\epsilon)) = 0$, i.e., u will be constant in each of the intervals $(x - \epsilon, x)$, $(x, x + \epsilon)$. Suppose now that u does not have a jump discontinuity at x, i.e., $Du(\{x\}) = 0$, and thus u is constant in $(x - \epsilon, x + \epsilon)$, say equal to c.

Case 1: $u(x) < f(x)$.

In this case we claim that u is constant, equal to c, in $[x - \epsilon, b)$. Suppose this is not true. Then note first that since f is strictly increasing, it is easily checked that u will be increasing as well. Recall from Proposition 3.3, that u will be continuous on (x, b) since α is differentiable there. Now choose $t_0 \in [x + \epsilon, b)$ such that

$$d := Du(x, t_0) \leq \frac{f(x) - u(x)}{2},$$

with d being strictly positive. Notice that this can be done since u is increasing in $[x+\epsilon, b)$ and not just equal to a constant. Define $\tilde{u}$ to be the following function:

$$\tilde{u}(t) = \begin{cases} u(t), & t \in (a, x), \\ u(x) + d, & t \in [x, t_0), \\ u(t), & t \in [t_0, b), \end{cases}$$

see also Figure 2 for an illustration. In other words, $\tilde{u}$ has all the variation of u in $(x+\epsilon, t_0)$ concentrated in x. Note that

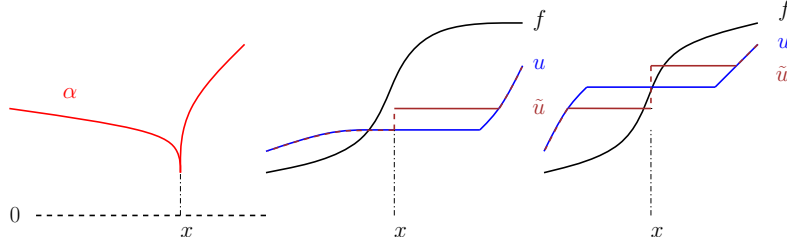$$\int_{\Omega} (f - \tilde{u})^2 dx < \int_{\Omega} (f - u)^2 dx,$$

12

Figure 2: The function $\tilde{u}$ from the proof of Proposition 3.5. Shifting the variation from areas which is costly into a single point where it is less costly. This favours the creation of a new discontinuity point.

and since α has a minimum at x we also have

$$\int_\Omega \alpha(x)d|D\tilde{u}| \leq \int_\Omega \alpha(x)d|Du|,$$

hence u is not optimal which is a contradiction.

Case 2: $u(x) > f(x)$.
This case is treated similarly to *Case 1*. If u does not have a jump discontinuity at x, then by similar arguments we conclude that u will be constant on an interval of the type $(a, x + \epsilon)$.

Case 3: $u(x) = f(x)$.
The arguments are similar to the previous cases; see also the third graph in Figure 2. We just have to make sure that by choosing a small enough jump at x for the function $\tilde{u}$ we can achieve a better L^2 distance from f. This can be done, for instance, by choosing d smaller than $f(x + \frac{\epsilon}{3}) - u(x + \frac{\epsilon}{3})$.

$\square$

Remark 3.6. *Note that for the type of data (increasing) of the Proposition 3.5 the potential jump discontinuity at x can only be positive, i.e., $Du(\{x\}) > 0$. Indeed, it can be easily checked that if $Du(\{x\}) < 0$ then the function u would be not optimal.*

Summarising the findings so far, we can say that whenever the weight function α has a *spike* at a point x, no matter whether this *spike* is upward or downward, the solution will always be constant at each one of the intervals $(x - \epsilon, x)$ and $(x, x + \epsilon)$ for a small enough $\epsilon > 0$. If the spike is upward, then the solution u will be constant in the whole interval $(x - \epsilon, x + \epsilon)$. If the spike is downward then the solution u will be either constant in $(x - \epsilon, x + \epsilon)$ or piecewise constant with a jump discontinuity at x. In order to be convinced that the second alternative can indeed occur, think of the following corollary of Proposition 3.5. Suppose that f is a strictly increasing, continuous function with a graph which is symmetric with respect to $(\frac{b-a}{2}, f(\frac{b-a}{2}))$ and α is a similarly symmetric function with a downward spike at $\frac{b-a}{2}$, e.g., $\alpha(x) = \sqrt{\left|x - \frac{b-a}{2}\right|}$. Then unless u is a constant function, it will always have a jump discontinuity at $x = \frac{b-a}{2}$.

In fact new discontinuities can be created even with more regular weight function, i.e., when $\alpha' \in \mathrm{BV}(\Omega)$. While we will come back to this with specific examples in Section 3.4, the following proposition provides conditions on when this can indeed occur and it establishes a connection between the jump size of α' and the jump size of u. Note that for

13

such a function α we have

$$\{x \in \Omega : \ \alpha \text{ is not differentiable at } x\} = J_{\alpha'}. \tag{3.22}$$

Proposition 3.7. *Let $f \in \mathrm{BV}(\Omega)$ with f being continuous at a point $x \in \Omega$. Let $\alpha \in C(\overline{\Omega})$ be a weight function with $\alpha' \in \mathrm{BV}(\Omega)$ such that $|D\alpha'|(\{x\}) > 0$. Let u solve (3.1) with data f and weight function α. Then the following hold true:*

(i) If $D\alpha'(\{x\}) < 0$, then $|Du|((x - \epsilon, x + \epsilon)) = 0$, for a small enough $\epsilon > 0$.

(ii) If $D\alpha'(\{x\}) > 0$, then u has potentially a jump discontinuity at x with

$$|Du|(\{x\}) \le D\alpha'(\{x\}). \tag{3.23}$$

In the particular case where there exists an $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subseteq \mathrm{supp}(|Du|)$ then u has a jump discontinuity at x and

$$|Du|(\{x\}) = D\alpha'(\{x\}). \tag{3.24}$$

Proof. (i) We start with the first case. We show first that $|Du|(\{x\}) = 0$. Suppose towards contradiction that $|Du|(\{x\}) > 0$ and assume without loss of generality that $Du(\{x\}) > 0$. We claim that $Dv'(\{x\}) > 0$. Indeed we have

$$v(t) = v(x) + \int_x^t v'(s)ds, \quad -\alpha(t) = -\alpha(x) - \int_x^t \alpha'(s)ds, \quad \text{for all } x \le t,$$

$$v(t) = v(x) - \int_t^x v'(s)ds, \quad -\alpha(t) = -\alpha(x) + \int_t^x \alpha'(s)ds, \quad \text{for all } t \le x.$$

From condition (3.14) we have that $v(x) = -\alpha(x)$ and also $v(t) \ge -\alpha(t)$ for every $t \in \Omega$. Thus we can write

$$\int_x^t v'(s)ds \ge -\int_x^t \alpha'(s)ds, \quad \text{for all } x \le t, \tag{3.25}$$

$$-\int_t^x v'(s)ds \ge \int_t^x \alpha'(s)ds, \quad \text{for all } t \le x. \tag{3.26}$$

Since $-D\alpha'(\{x\}) > 0$ there exist a small enough $\epsilon > 0$ and two constants $m < M$ such that

$$\operatorname*{ess\,sup}_{s \in (x-\epsilon,x)} -\alpha'(s) < m < M < \operatorname*{ess\,inf}_{s \in (x,x+\epsilon)} -\alpha'(s). \tag{3.27}$$

In combination with (3.25)–(3.26), this implies that for all $\delta < \epsilon$

$$\frac{1}{\delta} \int_{x-\delta}^x v'(s)ds < m < M < \frac{1}{\delta} \int_x^{x+\delta} v'(s)ds. \tag{3.28}$$

Taking the limit in (3.28) as $\delta \to 0$ we get

$$(v')^l(x) \le m < M \le (v')^r(x).$$

This implies that $Dv'(\{x\}) > 0$. However from condition (3.13) we have that

$$Df(\{x\}) = Du(\{x\}) + Dv'(\{x\}) > 0,$$

14

which contradicts the continuity of f at x and hence $|Du|(\{x\}) = 0$. We now claim that not only $|Du|(\{x\}) = 0$ but there exists a small enough $\epsilon > 0$ such that $|Du|((x-\epsilon, x+\epsilon)) = 0$. If that was not the case, using condition (3.14), we can find a sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to x$ such that $|v(t_n)| = \alpha(t_n)$. Without loss of generality, we can assume that $v(t_n) = -\alpha(t_n)$ for all $n \in \mathbb{N}$. The proof is similar if we assume $v(t_n) = \alpha(t_n)$. From the continuity of v and α, this implies that $v(x) = -\alpha(x)$. Then by simply following again the steps above, we can derive again $Dv'(x) > 0$ and

$$Df(\{x\}) = Dv'(\{x\}) > 0,$$

which contradicts again the continuity of f at x.

(ii) Suppose now that $D\alpha'(\{x\}) > 0$. If u does not have a jump discontinuity at x then (3.23) holds trivially. Thus assume that $Du(\{x\}) > 0$. Working similarly to case (i), we arrive again at (3.25)–(3.26). Notice also that since we assumed that $Du(\{x\}) > 0$ we have from (3.13)

$$D(-v')(\{x\}) = Du(\{x\}) > 0,$$

which means that $(-v')^l(x) < (-v')^r(x)$ and thus for all $\delta > 0$ that are small enough we have

$$\frac{1}{\delta}\int_{x-\delta}^{x} -v'(s)ds < \frac{1}{\delta}\int_{x}^{x+\delta} -v'(s)ds. \qquad (3.29)$$

Inequality (3.29) together with (3.25)–(3.26) gives

$$\frac{1}{\delta}\int_{x-\delta}^{x}\alpha'(s)ds \leq \frac{1}{\delta}\int_{x-\delta}^{x} -v'(s)ds < \frac{1}{\delta}\int_{x}^{x+\delta} -v'(s)ds \leq \frac{1}{\delta}\int_{x}^{x+\delta}\alpha'(s)ds,$$

for all $\delta > 0$ small enough. Taking the limit $\delta \to 0$ in the expression above we end up with

$$(\alpha')^l(x) \leq (-v')^l(x) < (-v')^r(x) \leq (\alpha')^r(x),$$

and thus

$$Du(\{x\}) = D(-v')(\{x\}) \leq D\alpha'(\{x\}).$$

Assuming $Du(\{x\}) < 0$, by working similarly we derive

$$Du(\{x\}) \geq -D\alpha'(\{x\}),$$

and thus generally (3.23) holds.

For the second part of (ii), note first that since $(x - \epsilon, x + \epsilon) \subseteq \mathrm{supp}(|Du|)$ and v is continuous, from (3.14) it follows that there exists a sufficiently small $\delta > 0$ such that $v = 1$ everywhere in $(x - \delta, x + \delta)$ (or $v = -1$ everywhere in $(x - \delta, x + \delta)$). As a result, from condition (3.14) we get that $v = -\alpha$ or $v = \alpha$ in $(x - \delta, x + \delta)$ and condition (3.13) imposes there

$$-\alpha' = f - u \quad \text{or} \quad \alpha' = f - u.$$

Since $|Df|(\{x\}) = 0$ from the above we get that

$$Du(\{x\}) = D\alpha'(\{x\}) \quad \text{or} \quad Du(\{x\}) = -D\alpha'(\{x\}).$$

$\square$

By performing similar steps to the ones in the proof of Proposition 3.7, the following result can be shown.

Corollary 3.8. *Suppose that x is a jump discontinuity point for the data f and $D\alpha'(\{x\}) > 0$. Then, the following estimate holds:*

$$|Du|(\{x\}) \le |Df|(\{x\}) + D\alpha'(\{x\}). \tag{3.30}$$

Proof. We briefly sketch the proof. Suppose without loss of generality that $Df(\{x\}) > 0$.
• If $D(-v')(\{x\}) > 0$, then we follow the steps of the proof above starting from (3.29) and we derive $D(-v')(\{x\}) \le D\alpha'(x)$. Then from (3.13) we get

$$Du(\{x\}) = Df(\{x\}) + D(-v')(\{x\}) \le Df(\{x\}) + D\alpha'(\{x\}).$$

• If $-Df(\{x\}) \le D(-v')(\{x\}) < 0$, then obviously

$$Du(\{x\}) = Df(\{x\}) + D(-v')(\{x\}) < Df(\{x\}) \le Df(\{x\}) + D\alpha'(\{x\}).$$

• Lastly if $D(-v')(\{x\}) < -Df(\{x\})$ then it follows that $Du(\{x\}) < 0$. Then following exactly the steps of (ii) in the proof of Proposition 3.7 (only the signs are reversed) we end up to

$$Dv'(\{x\}) \le D\alpha'(\{x\}),$$

and thus in this case

$$0 > Du(\{x\}) = Df(\{x\}) + D(-v')(\{x\}) \ge Df(\{x\}) - D\alpha'(\{x\}).$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

We would like now to prove that if α is differentiable at a point x, then $|Du|(\{x\}) \le |Df|(\{x\})$. Notice that we cannot derive this straightforwardly from Corollary 3.8 as there we use the fact that $D\alpha'(\{x\}) > 0$. However, this can easily be shown independently as the next proposition shows.

Proposition 3.9. *Let u solve the weighted TV minimisation problem with data f and weight function $\alpha \in C(\overline{\Omega})$ with $\alpha' \in \mathrm{BV}(\Omega)$ and $\alpha > 0$. Then if $|D\alpha'|(\{x\}) = 0$, we have*

$$|Du|(\{x\}) \le |Df|(\{x\}). \tag{3.31}$$

Moreover, the jumps of u and f have the same direction.

Proof. If $|Df|(\{x\}) = 0$ we have nothing to prove since by Proposition 3.3 we have that $|Du|(\{x\}) = 0$ as well. Thus, suppose that $Df(\{x\}) > 0$. The case $Df(\{x\}) < 0$ is treated similarly. We first exclude the case $Du(\{x\}) < 0$. Suppose towards contradiction that this holds. From the left and right continuity properties of f and u and (3.13) we have that there exists an $\epsilon > 0$ and some real numbers $m < M$ such that

$$\operatorname*{ess\,sup}_{t \in (x-\epsilon, x)} v'(t) \le m < M \le \operatorname*{ess\,inf}_{t \in (x, x+\epsilon)} v'(t),$$

Bearing in mind that $v(x) = \alpha(x) > 0$ and the fact that

$$v(t) = v(x) + \int_x^t v'(s)ds, \quad x \le t,$$

$$v(t) = v(x) - \int_t^x v'(s)ds, \quad t \le x,$$

16

together with $v < \alpha$, we deduce that

$$\alpha(x) - \alpha(t) \geq M(x - t), \quad x \leq t,$$
$$\alpha(x) - \alpha(t) \leq m(x - t), \quad t \leq x,$$

which contradicts the fact that α is differentiable at x. Hence $Du(\{x\}) > 0$ and it now remains to prove (3.31). Notice first of all that by arguing similarly as above we can exclude the cases

$$\underline{u}(x) < \underline{f}(x) < \overline{f}(x) < \overline{u}(x), \quad \underline{u}(x) \leq \underline{f}(x) < \overline{f}(x) < \overline{u}(x) \quad \text{and} \quad \underline{u}(x) < \underline{f}(x) < \overline{f}(x) \leq \overline{u}(x).$$

We thus focus on the cases

$$\underline{f}(x) < \underline{u}(x) \leq \overline{f}(x) < \overline{u}(x), \tag{3.32}$$

$$\underline{f}(x) < \overline{f}(x) < \underline{u}(x) < \overline{u}(x), \tag{3.33}$$

$$\underline{u}(x) < \underline{f}(x) \leq \overline{u}(x) < \overline{f}(x), \tag{3.34}$$

$$\underline{u}(x) < \overline{u}(x) < \underline{f}(x) < \overline{f}(x). \tag{3.35}$$

and we will show that when these happen then (3.31) must hold.

We argue for (3.32) since (3.33), (3.34) and (3.35) can be treated similarly. Assume that (3.31) does not hold. This means that

$$\overline{f}(x) - \underline{f}(x) < \overline{u}(x) - \underline{u}(x).$$

Arguing in the same way as before, this implies that there exists an $\epsilon > 0$ and some real numbers $m < M$ such that

$$\operatorname*{ess\,sup}_{t \in (x, x+\epsilon)} v'(t) \leq m < M \leq \operatorname*{ess\,inf}_{t \in (x-\epsilon, x)} v'(t) < 0.$$

This, together with the fact that $v(x) = -\alpha(x)$ and $v \leq -\alpha$ contradicts again the differentiability of α at x. $\qquad \square$

Recall that in the standard scalar TV minimisation we always have at a jump point x of u,

$$\underline{f}(x) \leq \underline{u}(x) < \overline{u}(x) \leq \overline{f}(x). \tag{3.36}$$

Moreover, the jumps of u and f having the same directions, i.e.,

$$f^l(x) \leq u^l(x) < u^r(x) \leq f^r(x) \quad \text{or} \quad f^r(x) \leq u^r(x) < u^l(x) \leq f^l(x).$$

We now summarise our findings so far. Given $f \in \mathrm{BV}(\Omega)$ $\alpha \in C(\overline{\Omega})$ with $\alpha' \in \mathrm{BV}(\Omega)$, we have shown analytically the following:

(i) If $|D\alpha'|(\{x\}) = 0$ and $|Df|(\{x\}) = 0$ then $|Du|(\{x\}) = 0$; see Proposition 3.3 and (3.22).

(ii) If $D\alpha'(\{x\}) < 0$ then a *plateau* is created for u around x; see Proposition 3.7.

(iii) The estimate $|Du|(\{x\}) \leq |Df|(\{x\}) + |D\alpha'|(\{x\})$ holds in every point $x \in \Omega$.

(iv) If $|Df|(\{x\}) = 0$, $D\alpha'(\{x\}) > 0$ and $(x - \epsilon, x + \epsilon) \subseteq \operatorname{supp}(|Du|)$, then $|Du|(\{x\}) = D\alpha'(x)$; see Proposition 3.7.

17

Case	What is proved analytically	Is it possible...	Answer/Figure
$D\alpha'(\{x\}) > 0$	$|Du|(\{x\}) \leq |Df|(\{x\}) + |D\alpha'|(\{x\})$	for u to remain continuous?	Yes, Fig. 3a
$D\alpha'(\{x\}) > 0 \ \& \ Df(\{x\}) = 0$	$|Du|(\{x\}) \leq |D\alpha'|(\{x\})$	to have "$<$" ?	Yes, Fig. 3b
$D\alpha'(\{x\}) > 0 \ \& \ Df(\{x\}) = 0$	$|Du|(\{x\}) \leq |D\alpha'|(\{x\})$	to have "$=$" ?	Yes, Fig. 3c
$D\alpha'(\{x\}) = 0 \ \& \ Df(\{x\}) > 0$	$|Du|(\{x\}) \leq |Df|(\{x\})$	$f^l(x) < f^r(x) < u^l(x) < u^r(x)$?	Yes, Fig. 3d
$D\alpha'(\{x\}) > 0 \ \& \ Df(\{x\}) > 0$	$|Du|(\{x\}) \leq |Df|(\{x\}) + |D\alpha'|(\{x\})$	$u^l(x) < f^l(x) < f^r(x) < u^r(x)$?	Yes, Fig. 3e
$D\alpha'(\{x\}) > 0 \ \& \ Df(\{x\}) > 0$	$|Du|(\{x\}) \leq ||Df|(\{x\}) - D\alpha'(\{x\})|$ [1]	$u^l(x) < f^r(x) < f^l(x) < u^r(x)$?	Yes, Fig. 3f

Table 1: Summary of the questions that are answered with numerical examples in Figures 3a–3f

(v) If f and u jump at x in different directions then $|Du|(\{x\}) \leq ||Df|(\{x\}) - D\alpha'(\{x\})|$; see Corollary 3.8.

(vi) If $|D\alpha'|(\{x\}) = 0$ and u and f jump at x, then their jumps have the same direction; see Proposition 3.9.

Despite these first analytical results, several questions still need to be addressed. For instance, one wonders whether $D\alpha'(\{x\}) > 0$, always creates a jump discontinuity for u at x. Furthermore, we note that (3.36) is related to a loss of contrast in mathematical image processing. Here, one consequently is interested in understanding, whether such an effect still is possible in the weighted case, provided the weight is smooth. Table 1 summarises these and further questions. In Figures 3a–3f we provide numerical examples for all the cases discussed in Table 1.

In Figure 3a we have an example where both the data f and the solution u are continuous at a point x despite the fact that $D\alpha'(\{x\}) > 0$. Note also that a trivial example here would also be the case where α is so large that u would be a constant.

In Figure 3b we depict an example where a new discontinuity is created for the solution u at the point $x = 0$ where the data function f is continuous. Note that in this specific example, the creation of this discontinuity is guaranteed to happen since f is continuous and strictly increasing. This was shown in Proposition 3.5 for data functions with a downward spike, but it can be easily extended to an absolute value-type function as we have here. Note that the estimate $|Du|(\{x\}) \leq D\alpha'(\{x\})$ holds here with strict inequality. Observe that the jump of α' is very large at the point 0 in contrast to the jump of u there. In fact there is a further upper bound for $Du(\{x\})$ which is independent of α; see Theorem 4.2 of Section 4 and Theorem 5.4 of Section 5. As we mention in the introduction, in these theorems it is shown that $|Du|(\Omega) \leq |Df|(\Omega)$.

On the contrary, in Figure 3c we have an example where the estimate $|Du|(\{x\}) \leq D\alpha'(\{x\})$ holds with an equality. We use the same f as in Figure 3c and a similar weight function α with a small jump of α' at $x = 0$. Note that here it holds that $(-\epsilon, \epsilon) \subseteq \text{supp}(|Du|)$ for some small $\epsilon > 0$ and, as Proposition 3.7 predicts, we have $|Du|(\{x\}) = D\alpha'(\{x\})$.

In Figure 3d we encounter another, perhaps unexpected situation. Even though we are using a smooth weight function α and, as Proposition 3.3 states, the jumps of u should occur at the same points where f has jumps, condition (3.36) is violated. Indeed, here we have that the whole jump of u is above that of f, i.e., $f^l(x) < f^r(x) < u^l(x) < u^r(x)$. Nevertheless, there still holds $|Du|(\{x\}) \leq |Df|(\{x\})$ in accordance with Proposition
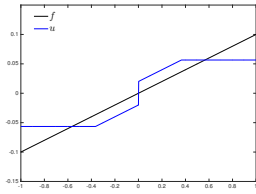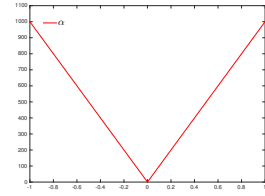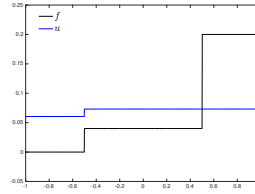

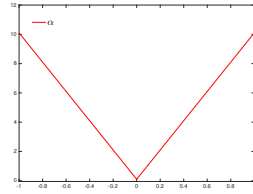[1] when the jumps of u and f have opposite directions
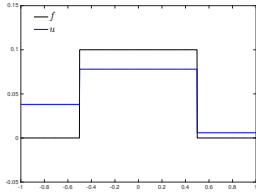
(a) Case where even though α' has a positive jump, the solution u remains continuous at this point
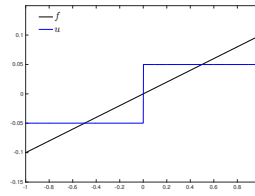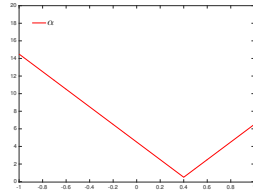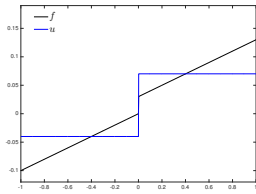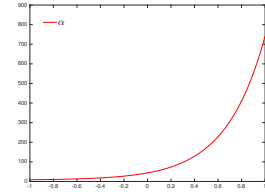
(b) Case where $|Du|(\{x\}) < D\alpha'(\{x\})$. Note that $x = 0$ is not an interior point of $\operatorname{supp}(|Du|)$

(c) Case where $|Du|(\{x\}) = D\alpha'(\{x\})$. Note that this is predicted by Proposition 3.7, since $(-\epsilon, \epsilon) \subseteq \operatorname{supp}(|Du|)$

(d) Case where even though the weight function α is smooth, the jump of u is above the jump of f, i.e., $f^l(x) < f^r(x) < u^l(x) < u^r(x)$

(e) Case where the jump of u is larger than the jump of f and has the same direction, i.e., $u^l(x) < f^l(x) < f^r(x) < u^r(x)$
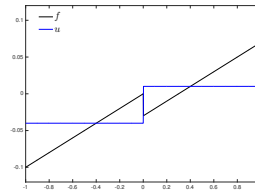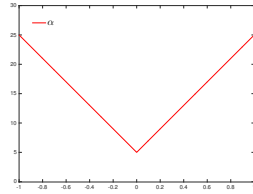
(f) Case where the jump of u is larger than the jump of f and has the opposite direction, i.e., $u^l(x) < f^r(x) < f^l(x) < u^r(x)$

Figure 3: Series of numerical examples that confirm that certain cases, as these are described in Table 1, can indeed occur using weighted TV regularisation

3.9.

In Figure 3e we have an example where the size of the jump of u is larger than the one of f since we have that α' also jumps there. Note that in this example the jump directions are the same.

On the other hand, in Figure 3f we have a similar situation, i.e., the jump of u is larger than the one of f, but also the orientation of the jumps is different as $Du(\{0\}) > 0$, while $Df(\{0\}) < 0$.

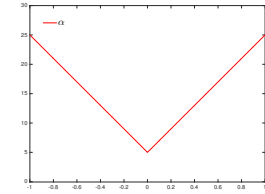Let us note here that situations where condition (3.36) is violated, like the examples in Figures 3d and 3f in which the jump is moving upwards or changing orientations, shows that the weighted TV has fundamentally different regularising properties when compared to the scalar TV. Moreover, this is also true in comparison with other popular regularisers like Huber-TV or TGV. Indeed, for (the scalar parameter versions of) one dimensional Huber-TV and TGV, condition (3.36) always holds [BKV13, PB15, BPPS16].

Finally we note that even though in Figure 3 we present only numerical results, one can easily calculate analytically the explicit forms of the solutions depicted there. Indeed, in Section 3.4 we do that exhaustively for the examples in Figures 3b–3c.

Before we finish this section, we mention two last results. The first one says that the solution u is constant in areas where the weight function α has high gradient. This is essentially a corollary of the fact that the solution u is bounded by the data f with respect to the uniform norm.

Proposition 3.10. *Let $f \in \mathrm{BV}(\Omega)$ and α be a weight function, differentiable in an open interval $I \subseteq \Omega$, such that*

$$|\alpha'(x)| > 2\|f\|_\infty, \quad \text{for all } x \in I. \tag{3.37}$$

Let u be the solution of (3.1) with data f and weight α. Then u is constant in I.

Proof. One can easily observe that for the solution u the following maximum principle holds true:

$$\operatorname*{ess\,inf}_{x \in \Omega} f(x) \le \operatorname*{ess\,inf}_{x \in \Omega} u(x) \le \operatorname*{ess\,sup}_{x \in \Omega} u(x) \le \operatorname*{ess\,sup}_{x \in \Omega} f(x). \tag{3.38}$$

This stems from the fact that the function

$$u_{\inf f}^{\sup f}(x) := \max(\min(\operatorname*{ess\,sup}_{t \in \Omega} f(t), \tilde{u}(x)), \operatorname*{ess\,inf}_{t \in \Omega} f(t)),$$

has always equal or less energy than u, where here $\tilde{u}$ is any good representative of u. Using a simple translation argument, we can assume that

$$\operatorname*{ess\,sup}_{x \in \Omega} f(x) = \|f\|_\infty,$$
$$\operatorname*{ess\,inf}_{x \in \Omega} f(x) = -\|f\|_\infty.$$

This together with (3.38) implies that $\|f - u\|_\infty \le 2\|f\|_\infty$ and, thus, for the dual variable v it must hold

$$\|v'\|_\infty \le 2\|f\|_\infty. \tag{3.39}$$

Suppose now that u is not constant in I. Then from condition (3.14) we have that there exists a point $x \in I$ such that $|v(x)| = \alpha(x)$ and, without loss of generality, we assume

that $v(x) = \alpha(x)$. We can also assume without loss of generality that $\alpha' < -2\|f\|_\infty$ in I. Using the fundamental theorem of calculus, the mean value theorem, the fact that $v(x) = \alpha(x)$ and $v \leq \alpha$, we get for every $t > x$ with $t \in I$

$$\frac{1}{t-x} \int_x^t v'(s)ds \leq \frac{\alpha(x) - \alpha(t)}{x-t} < -2\|f\|_\infty. \tag{3.40}$$

We claim now that for every $\epsilon > 0$ small enough there exists a set $A \subseteq (x, x+\epsilon)$ of positive Lebesgue measure such that $v' < -2\|f\|_\infty$ in A, which of course contradicts (3.39). Indeed if there was a small enough $\epsilon > 0$ such that $v' \geq -2\|f\|_\infty$ almost everywhere on $(x, x+\epsilon) \subseteq I$ then we would get for every $x < t < x+\epsilon$

$$-2\|f\|_\infty = \frac{1}{t-x} \int_x^t -2\|f\|_\infty \, ds \leq \frac{1}{t-x} \int_x^t v'(s)ds,$$

which contradicts (3.40). $\qquad\square$

The last result we state is that, as in the scalar case, solutions of the weighted TV minimisation are constant near the boundary of Ω. The proof is very simple and stems from the facts that α is bounded away from zero as well as $v \in H_0^1(\Omega)$.

Proposition 3.11. *Let $\Omega = (a, b)$, $f \in \mathrm{BV}(\Omega)$ and $\alpha \in C(\overline{\Omega})$ with $\alpha > 0$. Then there exist $a < x_1 < x_2 < b$ such that the solution u of (3.1) with data f and weight α, is constant in (a, x_1) and (x_2, b).*

Proof. Suppose that there does not exist a $x_1 \in (a, b)$ such that u is constant in (a, x_1). Then from condition (3.14) it follows that there exists a sequence $(a_n)_{n\in\mathbb{N}} \in (a, b)$, converging to a and

$$|v(a_n)| = \alpha(a_n), \quad \text{for every } n \in \mathbb{N}.$$

From the continuity of v and the fact that it belongs to $H_0^1(\Omega)$ we have

$$0 = \lim_{n\to\infty} |v(a_n)| = \lim_{n\to\infty} \alpha(a_n) \geq \min_{x\in(a,b)} \alpha(x) > 0,$$

which is a contradiction. Similarly we deal with the right part of the boundary of Ω. $\qquad\square$

3.3 A partial semigroup property

For this section it is convenient to introduce the following notation

$$S_{\alpha(x)}(f) := \operatorname*{argmin}_{u\in\mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f-u)^2 dx + \int_\Omega \alpha(x)d|Du|,$$

i.e., $S_{\alpha(x)}(f)$ denotes the solution of (3.1) with data $f \in \mathrm{BV}(\Omega)$ and weight function $\alpha \in C(\overline{\Omega})$. Again we slightly abuse the notation by writing $S_{\alpha(x)}(f)$ when α is not a constant function and $S_\alpha(f)$ when it is. It is a well-known fact that the following *semigroup* property holds for the one dimensional scalar total variation problem [SGG$^+$09]

$$S_{\alpha_1+\alpha_2}(f) = S_{\alpha_2}\left(S_{\alpha_1}(f)\right) = S_{\alpha_1}\left(S_{\alpha_2}(f)\right), \quad \alpha_1, \alpha_2 > 0. \tag{3.41}$$

In other words, one can obtain the solution of TV regularisation with scalar parameter $\alpha_1 + \alpha_2$ and data f by applying TV regularisation with parameter α_2 to the result which

is obtained by applying TV regularisation with parameter α_1 to the data f. The result remains the same if we apply TV regularisation first with parameter α_2 and then with α_1. Here we examine whether this property holds true in the weighted TV regularisation or not. The next proposition states that this is indeed the case when the second regularisation parameter is a scalar.

Proposition 3.12 (Partial semigroup property). *Let $f \in \mathrm{BV}(\Omega)$, $\alpha_1 \in C(\overline{\Omega})$ with $\alpha_1 > 0$, and $\alpha_2 > 0$ be a scalar. Then*

$$S_{\alpha_1(x)+\alpha_2}(f) = S_{\alpha_2}\left(S_{\alpha_1(x)}(f)\right). \tag{3.42}$$

Proof. Let $u_1 := S_{\alpha_1(x)}(f)$ and $u_2 := S_{\alpha_2}\left(S_{\alpha_1(x)}(f)\right)$. The optimality conditions for the corresponding minimisation problems read

$$
\begin{aligned}
v_1' &= f - u_1, & v_2' &= u_1 - u_2, \\
-v_1 &\in \alpha_1(x)\mathrm{Sgn}(Du_1), & -v_2 &\in \alpha_2\mathrm{Sgn}(Du_2),
\end{aligned}
$$

where both $v_1, v_2 \in H_0^1(\Omega)$. Defining $v_{12} = v_1 + v_2$, we have $v_{12} \in H_0^1(\Omega)$ and

$$v_{12}' = f - u_2.$$

Thus, for (3.42) it suffices to show that

$$- v_1 - v_2 \in (\alpha_1(x) + \alpha_2)\mathrm{sgn}(Du_2). \tag{3.43}$$

Since $|v_1(x)| \leq \alpha_1(x)$ for Lebesgue–almost every $x \in \Omega$ and $\|v_2\|_\infty \leq \alpha_2$, it obviously holds that $|v_{12}(x)| \leq \alpha_1(x) + \alpha_2$ for Lebesgue–almost every $x \in \Omega$. Consequently, it suffices to check that

$$- v_{12} = (\alpha_1(x) + \alpha_2)\mathrm{sgn}(Du_2), \quad |Du_2| - \text{a.e.} \tag{3.44}$$

In what follows we work with the continuous representatives of $\mathrm{sgn}(Du_1)$ and $\mathrm{sgn}(Du_2)$. Let $x \in \Omega$ with $x \in \mathrm{supp}(|Du_2|)$ and $\mathrm{sgn}(Du_2)(x) = 1$, then we will show that $\mathrm{sgn}(Du_1)(x) = 1$ as well (similarly we proceed with -1). Note that in this case (any good representative of) u_2 is increasing near x in the sense that there exists a small enough $\epsilon > 0$ such that

$$u_2(s) \leq u_2(s') < u_2(t) \leq u_2(t'), \quad \text{for every } x - \epsilon < s < s' < x < t < t' < x + \epsilon.$$

Indeed this comes from the fact that

$$u_2^r(s) - u_2^r(t) = \int_{(s,t]} \mathrm{sgn}(Du_2)(y)d|Du_2|, \quad s < x < t,$$

combined with the continuity of $\mathrm{sgn}(Du_2)$ and the fact that $\mathrm{sgn}(Du_2)(x) = 1$. We claim then, that $x \in \mathrm{supp}(|Du_1|)$ as well. Indeed, if this is not true then $|Du_1|((x - \epsilon, x + \epsilon)) = 0$ for a small $\epsilon > 0$, and thus the same would hold for u_2 (which follows directly when using the optimality conditions). Since $x \in \mathrm{supp}(|Du_1|)$, then either $\mathrm{sgn}(Du_1)(x) = 1$ or $\mathrm{sgn}(Du_1)(x) = -1$. However, the latter case can be excluded and thus $\mathrm{sgn}(Du_1)(x) = 1$. Indeed, from the optimality conditions, and by considering different cases depending on whether x it is a jump point or not, it follows that the data u_1 and the solution u_2 cannot be decreasing and increasing respectively near x. As a result,
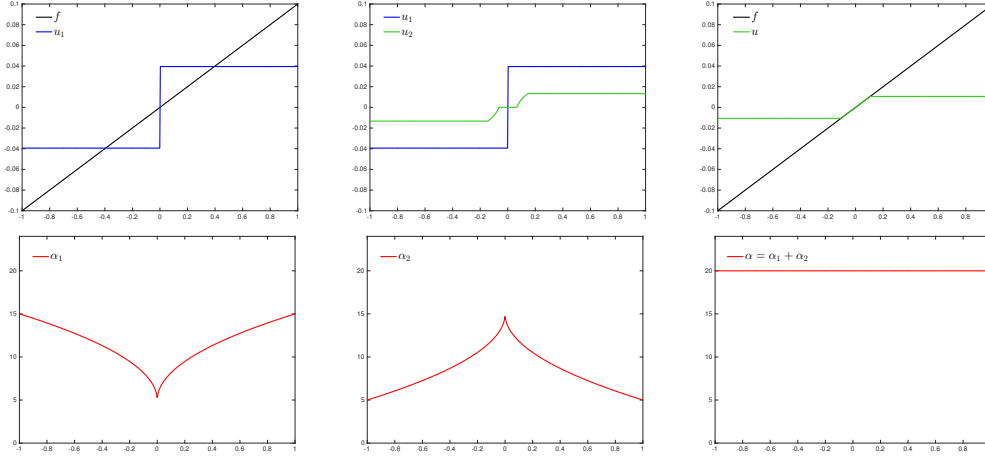
(a) Regularisation of the initial data with a downward spike weight function α_1. Creation of a new discontinuity

(b) Regularisation of the result in Figure 4a with an upward spike weight function $\alpha_2 = 20 - \alpha_1$. Creation of a plateau

(c) Regularisation of the initial data with the sum of the previous weight functions $\alpha = \alpha_1 + \alpha_2$. The result is different than the one obtained by applying the regularisation in two steps

Figure 4: Failure of the semigroup property for the weighted total variation model in the case where both weights α_1 and α_2 are not constant

if $\operatorname{sgn}(Du_2)(x) = 1$ it follows that $\operatorname{sgn}(Du_1)(x) = 1$. From the optimality conditions we obtain

$$-v_2(x) = \alpha_2 \quad \text{and} \quad -v_1(x) = \alpha_1(x),$$

respectively. Hence $-v_{12}(x) = \alpha_1(x) + \alpha_2$. Similarly if $\operatorname{sgn}(Du_2)(x) = 1$ then $-v_{12}(x) = -\alpha_1(x) - \alpha_2$ and thus (3.44) holds. □

It is natural to ask whether (3.42) holds when both α_1, α_2 are non-scalar continuous weight functions. Propositions 3.4 and 3.5 indicate that this is not necessarily true. We illustrate this fact with a numerical example in Figure 4. We start with an affine function f and first apply weighted TV regularisation with a square root type weight function α_1. Note the discontinuity that it is created in the solution u_1 exactly at the point where α_1 has the downward spike; see Figure 4a. By further applying weighted TV regularisation to the result with weight function $\alpha_2 = 20 - \alpha_1$, a plateau is created at this point; compare Figure 4b. The final result is different to the solution of the TV regularisation of the initial data f with weight $\alpha_1 + \alpha_2 = 20$, which is shown in Figure 4c.

However, what is perhaps even more surprising is that property (3.42) of Proposition 3.12 is non-commutative, i.e., it might fail even in the case where we first regularise with a constant weight function and then with a non constant one; hence the term *partial semigroup property* in Proposition 3.12, i.e.,

$$S_{\alpha_1(x) + \alpha_2}(f) = S_{\alpha_2}\left(S_{\alpha_1(x)}(f)\right) \neq S_{\alpha_1(x)}\left(S_{\alpha_2}(f)\right). \tag{3.45}$$
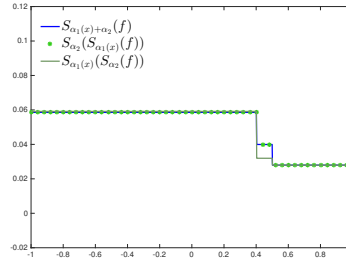
23
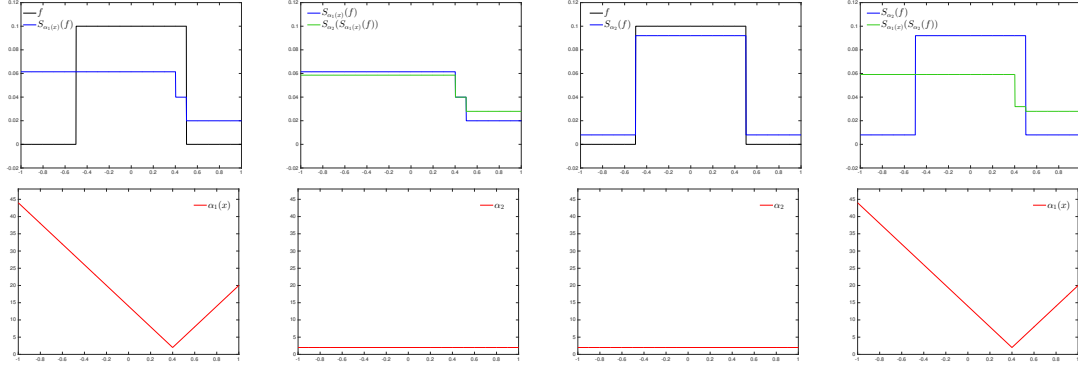
(a) Regularisation of the initial data with a non-constant weight $\alpha_1(x)$

(b) Regularisation of the result in Figure 5a with a constant weight α_2

(c) Regularisation of the initial data with a constant weight α_2

(d) Regularisation of the result in Figure 5c with a non-constant weight $\alpha_1(x)$

(e) The result $S_{\alpha_1(x)+\alpha_2}(f)$ obtained by solving (3.1) using as weight the sum of a non-constant $\alpha_1(x)$ and a constant function α_2 is not also obtained by successively solve (3.1) using the constant and then with the non-constant weight, $S_{\alpha_1(x)}(S_{\alpha_2}(f))$ but can be obtained by doing so, first with the non-constant and then with the constant weight $S_{\alpha_2}(S_{\alpha_1(x)}(f))$, as Proposition 3.12 dictates

Figure 5: Non-commutativity of the semigroup property even when one of the weights is constant

In order to get an intuition for (3.45), observe that the optimality conditions read

$$v_1' = f - u_1, \qquad\qquad v_2' = u_1 - u_2,$$
$$-v_1 \in \alpha_2 \mathrm{Sgn}(Du_1), \qquad -v_2 \in \alpha_1(x)\mathrm{Sgn}(Du_2).$$

Recalling $(v_1 + v_2)' = f - u_2$, for u_2 to be the result obtained after regularising f with $\alpha_1 + \alpha_2(x)$ requires that

$$-v_1 - v_2 = (\alpha_1(x) + \alpha_2)\mathrm{Sgn}(Du_2), \quad |Du_2| - \text{a.e.} \tag{3.46}$$

However, condition (3.46) could fail if for instance u_2 has a new discontinuity created at a point x around which u_1 is constant and below (or above) f. That would enforce $|v_2(x)| = \alpha_1(x)$ and $|v_1(x)| < \alpha_2$ and thus (3.46) would not hold.

We provide a numerical example in Figure 5. In Figures 5a–5b, we display the result which is obtained when we first regularise with a non-constant weight $\alpha_1(x)$ and then with a constant α_2. Figures 5c–5d, on the other hand, show the result when first regularising with α_2 and then with $\alpha_1(x)$. Figure 5e confirms that (3.45) holds for this example.

24

3.4 Analytic solutions for simple data and weight functions

Next we compute several analytic solutions for the problem (3.1). We get a further intuition about the structure of solutions and how it changes with respect to different weight functions.

Suppose $\Omega = (-L, L)$, where $L > 0$, and

$$f(x) = \lambda x, \quad x \in (-L, L), \quad \lambda > 0. \tag{3.47}$$

We also consider an absolute value type function as weight, i.e.,

$$\alpha(x) = \mu|x| + c, \quad x \in (-L, L), \quad \mu, c > 0. \tag{3.48}$$

Note that the symmetry of f and α imply that the solution u is also symmetric. Thus it suffices to describe the solution in the interval $[0, L)$. According to Proposition 3.11, u will be constant at the right boundary of $[0, L)$, say at the interval (x_0, L). Assuming that the solution is strictly increasing on $(0, x_0]$, the optimality conditions (3.13)–(3.14) yield $-\alpha'(x) = f - u$. Note that according to Proposition 3.7, in this case, u will have a jump discontinuity at $x = 0$. Thus, we study the ansatz

$$u(x) = \begin{cases} \lambda x + \mu, & \text{if } x \in (0, x_0], \\ \lambda x_0 + \mu, & \text{if } x \in (x_0, L). \end{cases}$$

From the optimality conditions (3.13)–(3.14) it follows that v has the form

$$v(x) = \begin{cases} -\alpha(x), & \text{if } x \in (0, x_0], \\ \frac{1}{2}\lambda x^2 - (\lambda x_0 + \mu)x + d, & \text{if } x \in (x_0, L), \end{cases}$$

for some constant d. From the continuity of v at x_0 and $v \in H_0^1(\Omega)$, we infer

$$\lim_{x \to x_0+} v(x) = -\alpha(x_0), \tag{3.49}$$

$$\lim_{x \to L} v(x) = 0. \tag{3.50}$$

Using conditions (3.49)–(3.50), one computes

$$x_0 = L - \frac{\sqrt{2\lambda\mu L + 2\lambda c}}{\lambda}. \tag{3.51}$$

Notice that $|v(x)| \leq \alpha(x)$ is satisfied for all $x \in (0, L)$. Finally, we need $0 < x_0$. This holds if and only if

$$\mu L + c < \frac{\lambda L^2}{2}. \tag{3.52}$$

Next we examine the case $x_0 = 0$, that is the solution is constant in $(0, L)$ but nonetheless has a jump discontinuity at $x = 0$. Hence, our ansatz here is

$$u(x) = M, \quad x \in (0, L), \quad M > 0.$$

Now the variable v is of the form

$$v(x) = \frac{1}{2}\lambda x^2 - Mx + d, \quad x \in (0, L).$$

From the fact that u has a jump discontinuity at $x = 0$ (note that the jump must be positive) and $v \in H_0^1(\Omega)$, we infer

$$v(0) = -\alpha(0), \quad \lim_{x \to L} v(x) = 0,$$

which in turn give $M = \frac{\lambda L}{2} - \frac{c}{L}$. Since we require $M > 0$ we must have

$$c < \frac{\lambda L^2}{2}. \tag{3.53}$$

Finally, in order to guarantee that $|v(x)| \leq \alpha(x)$ for all x, it suffices to enforce $v'(0) \geq -\mu$, i.e.,

$$\mu L + c \geq \frac{\lambda L^2}{2}. \tag{3.54}$$

Note that the inequalities (3.53)–(3.54) form necessary and sufficient conditions for the occurrence of this kind of solution.

The last alternative for the solution is to be constant, equal to the mean value of the data, i.e., $u = 0$. Working similarly as in the previous cases we deduce that this solution occurs if and only if

$$c \geq \frac{\lambda L^2}{2}. \tag{3.55}$$

Note that the conditions (3.52), (3.53)–(3.54) and (3.55) define a partition of the quadrant $\{\mu \geq 0, \ c \geq 0\}$. We summarise our findings in the following proposition.

Proposition 3.13. *Let $\Omega = (-L, L)$, $f(x) = \lambda x$ and $\alpha(x) = \mu|x| + c$ with $L, \lambda, \mu, c > 0$. Then the solution u of the problem*

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f - u)^2 dx + \int_\Omega \alpha(x) d|Du|,$$

is given by the following formulae:

If $\mu L + c < \dfrac{\lambda L^2}{2}$,
$$u(x) = \begin{cases} -\lambda x_{\mu,c} - \mu, & \text{if} \quad x \in (-L, -x_{\mu,c}], \\ -\lambda x - \mu, & \text{if} \quad x \in (-x_{\mu,c}, 0), \\ \lambda x + \mu, & \text{if} \quad x \in [0, x_{\mu,c}), \\ \lambda x_{\mu,c} + \mu, & \text{if} \quad x \in [x_{\mu,c}, L), \end{cases}$$

where $x_{\mu,c} = L - \dfrac{\sqrt{2\lambda\mu L + 2\lambda c}}{\lambda}$.

If $\mu L + c \geq \dfrac{\lambda L^2}{2}$ & $c < \dfrac{\lambda L^2}{2}$,
$$u(x) = \begin{cases} -\frac{\lambda L}{2} + \frac{c}{L}, & \text{if} \quad x \in (-L, 0), \\ \frac{\lambda L}{2} - \frac{c}{L}, & \text{if} \quad x \in [0, L). \end{cases}$$

If $c \geq \dfrac{\lambda L^2}{2}$,
$$u(x) = 0, \quad x \in (-L, L).$$

In Figure 6 we depict the different areas of the parameter space of the absolute type weight function, $\{\mu \geq 0, \ c \geq 0\}$ that correspond to the different types of solutions.

In Figure 7 we summarise numerical examples to better understand the solution dependence on the weight parameters c and μ. In Figure 7a, keeping μ fixed, we vary the parameter c by adding constants. Among others, one can observe here the numerical verification of the partial semigroup property of Proposition 3.12. The first weight (light blue
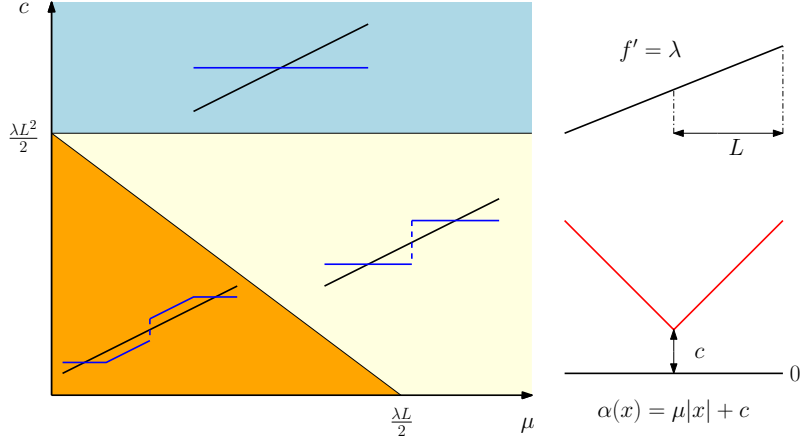
Figure 6: Different types of solutions of the weighted TV minimisation problem (3.1) with data $f(x) = \lambda x$ in $(-L, L)$ and weight function $\alpha(x) = \mu|x| + c$

(a) Varying c by keeping μ fixed

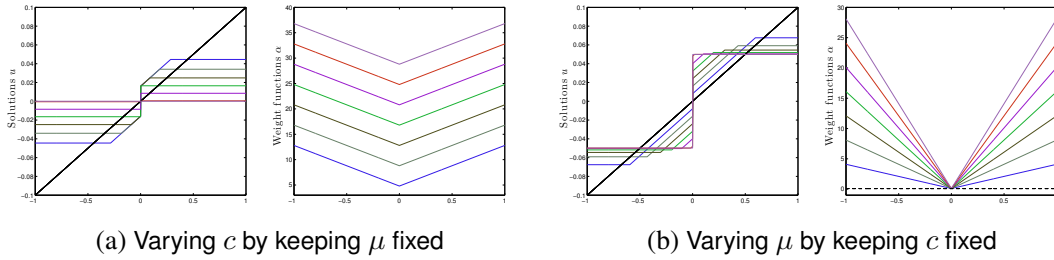(b) Varying μ by keeping c fixed

Figure 7: Numerical solutions of the weighted TV minimisation (3.1) with data $f(x) = \lambda x$ in $(-L, L)$ and weight functions of the type $\alpha(x) = \mu|x| + c$

line) produces a result, say u_1, that has a new discontinuity at the origin. As we increase the weight function by a constant the solutions we obtain are the corresponding solutions of the scalar TV problem with data u_1 and the parameter being the very constant.

In Figure 7b, by keeping c fixed, we vary the parameter μ by increasing the slope of the weight function. Notice that as the slope increases, i.e., $D\alpha'(\{0\})$ increases, so does the jump discontinuity of the solution u. This is in accordance with Proposition 3.7, i.e., we have $Du(\{x\}) = D\alpha'(\{0\})$, since at least when μ is not too large, a whole neighbourhood of $x = 0$ belongs to $\mathrm{supp}(|Du|)$. This behaviour stops when μ, and thus α', becomes too large (light purple line). Then the jump size of the solutions stops growing, no matter the size of μ and it depends only on the value of c, cf. Propositions 3.10 and 3.13.

4 Bound of the total variation of the solution by the total variation of the data

It is a standard result in the scalar total variation minimisation that the total variation of the solution is bounded by the total variation of data, i.e., if α is a positive constant and u

solves

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f - u)^2 dx + \alpha |Du|(\Omega),$$

then

$$|Du|(\Omega) \leq |Df|(\Omega). \tag{4.1}$$

This is simply proven by comparing the energy of the minimiser u and the data f, that is

$$\alpha |Du|(\Omega) \leq \frac{1}{2} \int_\Omega (f - u)^2 dx + \alpha |Du|(\Omega) \leq \alpha |Df|(\Omega).$$

Note that this holds for any dimension. A similar argument for the weighted case would only give

$$\int_\Omega \alpha(x) d|Du| \leq \int_\Omega \alpha(x) d|Df| \quad \Rightarrow \quad |Du|(\Omega) \leq \frac{\max_{x \in \Omega} \alpha(x)}{\min_{x \in \Omega} \alpha(x)} |Df|(\Omega). \tag{4.2}$$

However, the estimate in (4.2) is not satisfactory since the constant in the right-hand side depends on the weight α and in fact blows up as α tends to zero. In this section, we show that the estimate (4.1) holds in the weighted case as well. At first glance, (4.1) is counterintuitive as weighted TV may create new discontinuities, thus, increasing the variation locally. Our proof here uses fine scale analysis of the structure of solutions of the weighted TV problem and thus we are only able to show this result in dimension one.

We start by showing that weighted TV does not introduce oscillations in the solution u.

Proposition 4.1. *Let $u \in \mathrm{BV}(\Omega)$ solve the weighted TV problem (3.1) with data $f \in \mathrm{BV}(\Omega)$ and a differentiable weight function $\alpha \in C(\overline{\Omega})$ with $\alpha > 0$. Then $\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}$ is an open set and hence it is the countable union of disjoint intervals I_n. In each one of these intervals I_n, the solution u has at most one initial decreasing part, followed by at most one increasing part. The analogous result holds for the set $\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}$. There the solution u has at most one initial increasing part, followed by an at most one decreasing part.*

Proof. Observe first that

$$\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\} = \{\underline{u} > \overline{f}\} \cup \{\underline{f} < \underline{u} \leq \overline{f} < \overline{u}\}. \tag{4.3}$$

The fact that the set $\{\underline{u} > \overline{f}\}$ is open is shown in [BKV13]. Let now $x \in \{\underline{f} < \underline{u} \leq \overline{f} < \overline{u}\}$ and suppose without loss of generality that $Df(\{x\}) > 0$. Then, $Du(\{x\}) > 0$ as well, since α is differentiable, cf. Proposition 3.9. This means that

$$f^l(x) < u^l(x) \leq f^r(x) < u^r(x).$$

Bearing in mind that $f^l(t) = c_1 + Df((a, t))$, $u^l(t) = c_2 + Du((a, t))$, from the left continuity of f^l and u^l it follows that there exists an $\epsilon > 0$ and two real numbers $m < M$ such that

$$c_1 + Df((a, t)) < m < M < c_2 + Du((a, t)), \quad \text{for all } t \in (x - \epsilon, x). \tag{4.4}$$

But since $Df((\alpha, t_0]) = \lim_{t \to t_0^+} Df((\alpha, t))$, and analogously for u, we have that

$$c_1 + Df((a, t]) \leq m < M \leq c_2 + Du((a, t]), \quad \text{for all } t \in (x - \epsilon, x), \tag{4.5}$$

28

also holds and hence every t in $(x - \epsilon, x)$ also belongs to $\{\underline{u} > \overline{f}\}$. Similarly we show that there exists an $\epsilon > 0$ such that every t in $(x, x + \epsilon)$ also belongs to $\{\underline{u} > \overline{f}\}$ and hence from (4.3) the set $\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}$ is open. Thus it can be written as a countable union of disjoint intervals I_n.

We focus on a single interval I_n. From the optimality conditions (3.13)–(3.14) we have that $v' < 0$ almost everywhere in I_n and thus v is strictly decreasing there. Suppose now that u has an increasing part in I_n followed by a decreasing part. Then there would exist points x_1 and x_2 in I_n with $x_1 < x_2$ and

$$\mathrm{sgn}(Du)(x_1) = 1 \quad \text{and} \quad \mathrm{sgn}(Du)(x_2) = -1.$$

However, again from the optimality conditions one obtains

$$v(x_1) = -\alpha(x_1) < 0 \quad \text{and} \quad v(x_2) = \alpha(x_2) > 0,$$

which is a contradiction since v is decreasing in $(x_1, x_2) \subseteq I_n$.

$\square$

We are now ready to prove the main theorem of the section. Note that we will assume for the time being that the weight function α is differentiable. In Section 5 we will extend this result to continuous weights α via a Γ-convergence argument.

Theorem 4.2 ($|Du|(\Omega) \le |Df|(\Omega)$, for differentiable α). *Let $\Omega = (a, b) \subseteq \mathbb{R}$, $\alpha \in C(\overline{\Omega})$ differentiable with $\alpha > 0$ and $f \in \mathrm{BV}(\Omega)$. If*

$$u = \operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega (f - u)^2 dx + \int_\Omega \alpha(x) d|Du|,$$

then

$$|Du|(\Omega) \le |Df|(\Omega).$$

Proof. Since the weight function α is differentiable we have that, according to Proposition 3.3, the jump discontinuities of the solution u are contained in the ones of the data f. Moreover, $\{\underline{u} \le \underline{f} < \overline{f} < \overline{u}\} \cup \{\underline{u} < \underline{f} < \overline{f} \le \overline{u}\}$ is empty; cf. Proposition 3.9. As J_f is at most countable we can write

$$|Du|(\Omega) = |Du|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) + |Du|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\})$$
$$+ |Du|(\{\underline{f} \le \underline{u} < \overline{u} \le \overline{f}\}) + |Du|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}). \qquad (4.6)$$

Notice that the sets that appear in (4.6) are disjoint. We focus first on the set $\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}$. According to Proposition 4.1, this can be written as a countable union of disjoint open intervals. Let (x_1, x_2) be one of these intervals and assume for the moment that $a < x_1 < x_2 < b$. Then we have that (x_1, x_2) is maximal in the sense that in the endpoints x_1 and x_2, at least one of the conditions $\underline{u}(x) > \underline{f}(x)$, $\overline{u}(x) > \overline{f}(x)$ does not hold. Recall from Proposition 4.1 that in this interval, u starts with a decreasing part followed by an increasing one (both not necessarily strict). In what follows, we will infer bounds on $|Du|((x_1, x_2))$ by considering different alternative cases.

(i) Consider first that f, and hence u, is continuous at both endpoints x_1 and x_2. From the maximality of the associated interval we have

$$\underline{u}(x_1) = \overline{u}(x_1) = \underline{f}(x_1) = \overline{f}(x_1) \quad \text{and} \quad \underline{u}(x_2) = \overline{u}(x_2) = \underline{f}(x_2) = \overline{f}(x_2).$$

29

Then from the monotonicity structure of u and the fact that in this interval $\operatorname{ess\,inf}_{t \in (x_1, x_2)} f(t) < \operatorname{ess\,inf}_{t \in (x_1, x_2)} u(t)$ we have that

$$|Du|((x_1, x_2)) \leq |Df|((x_1, x_2)).$$

(ii) The second case is that f is continuous at x_1 and has a jump discontinuity at x_2. Note that this jump must be a positive one as otherwise the condition $\underline{u} > \overline{f}$ would be violated inside the interval, unless we have $\overline{u}(x_2) = \overline{f}(x_2)$ in which case the estimate $|Du|((x_1, x_2)) \leq |Df|((x_1, x_2))$ holds. Notice also that due to the maximality of x_2 we cannot have $\underline{f}(x) < \underline{u}(x) \leq \overline{f}(x) < \overline{u}(x)$. So we must have

$$\underline{u}(x_1) = \overline{u}(x_1) = \underline{f}(x_1) = \overline{f}(x_1) \quad \text{and} \quad \underline{f}(x_2) \leq \underline{u}(x_2) < \overline{u}(x_2) \leq \overline{f}(x_2), \quad (4.7)$$

if u has a jump in x_2, or

$$\underline{u}(x_1) = \overline{u}(x_1) = \underline{f}(x_1) = \overline{f}(x_1) \quad \text{and} \quad \underline{f}(x_2) \leq \underline{u}(x_2) = \overline{u}(x_2) \leq \overline{f}(x_2), \quad (4.8)$$

if u does not have a jump in x_2. Suppose we have (4.7), then arguing similarly as before we have that

$$|Du|((x_1, x_2)) \leq |Df|((x_1, x_2)) + \underline{u}(x_2) - \underline{f}(x_2),$$

i.e., in that case we need to add to the right-hand side also a part of the jump of f at x_2, which is however below the jump of u. If we have (4.8) we estimate again

$$|Du|((x_1, x_2)) \leq |Df|((x_1, x_2)) + u(x_2) - \underline{f}(x_2).$$

(iii) The third case is that f has a jump discontinuity at x_1 and it is continuous at x_2. This is treated analogously to the second case.

(iv) The fourth case is that f has jump discontinuities at both points x_1 and x_2. The only case the jump of f at x_1 is positive is when $\overline{u}(x_1) = \overline{f}(x_1)$ and in that case we have

$$|Du|((x_1, x_2)) \leq |Df|((x_1, x_2)) \quad \text{or} \quad |Du|((x_1, x_2)) \leq |Df|((x_1, x_2)) + \underline{u}(x_2) - \underline{f}(x_2),$$

depending on whether the jump of f at x_2 is negative or positive. So we are left with the case where the jumps of f at x_1 and x_2 are negative and positive, respectively. We then have

$$\underline{f}(x_1) \leq \underline{u}(x_1) \leq \overline{u}(x_1) \leq \overline{f}(x_1) \quad \text{and} \quad \underline{f}(x_2) \leq \underline{u}(x_2) \leq \overline{u}(x_2) \leq \overline{f}(x_2), \quad (4.9)$$

where we estimate

$$|Du|((x_1, x_2)) \leq |Df|((x_1, x_2)) + (\underline{u}(x_1) - \underline{f}(x_1)) + (\underline{u}(x_2) - \underline{f}(x_2)).$$

Note that analogous estimates hold if $x_1 = a$ or $x_2 = b$. By summing over the countable set of disjoint intervals related to $\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}$, the following inequality holds

$$|Du|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) \leq |Df|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\underline{u}(x) - \underline{f}(x)).$$

$$(4.10)$$

This is also due to the fact that any term of the form $\underline{u}(x) - \underline{f}(x))$ appears in only one of the corresponding estimates for the intervals I_n. Similarly, for $\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}$ we obtain

$$|Du|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}) \leq |Df|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}) + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\overline{f}(x) - \overline{u}(x)).$$

(4.11)

The final estimate then follows:

$$
\begin{aligned}
|Du|(\Omega) &= |Du|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) + |Du|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}) \\
&\quad + |Du|(\{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}) + |Du|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}) \\
&\leq |Df|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) + |Df|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}) \\
&\quad + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\underline{u}(x) - \underline{f}(x)) + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\overline{f}(x) - \overline{u}(x)) \\
&\quad + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\overline{u}(x) - \underline{u}(x)) + |Df|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}) \\
&\leq |Df|(\{\underline{u} > \underline{f}\} \cap \{\overline{u} > \overline{f}\}) + |Df|(\{\underline{u} < \underline{f}\} \cap \{\overline{u} < \overline{f}\}) \\
&\quad + \sum_{x \in \{\underline{f} \leq \underline{u} < \overline{u} \leq \overline{f}\}} (\overline{f}(x) - \underline{f}(x)) \\
&\quad + |Df|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}) \\
&\leq |Df|(\Omega).
\end{aligned}
$$

Here we also use Lemma 4.3 to infer

$$|Du|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}) = |Df|(\{\underline{u} = \overline{u} = \underline{f} = \overline{f}\}).$$

$\square$

It remains to prove the following lemma that was used in the final estimate above.

Lemma 4.3. *Let* $\Omega = (a,b)$, $u, v \in \mathrm{BV}(\Omega)$ *and let*

$$A = \{x \in \Omega : \ u, v \text{ are continuous at } x \text{ and } u(x) = v(x)\}.$$

Then

$$|Du|(A) = |Dv|(A).$$

Proof. Note first that A can also be written as

$$A = \{x \in \Omega : \ \underline{u}(x) = \overline{u}(x) = \underline{v}(x) = \overline{v}(x)\}.$$

We claim that A is a G_δ set, i.e., a countable intersection of open sets. Indeed recall first, that the set of continuity points of a function is a G_δ; see for instance [Olm56]. Therefore, the sets of continuity points of $\underline{u}$ and $\underline{v}$, i.e., J_u^c and J_v^c, respectively, are G_δ sets. Note that A can also be written as

$$A = J_u^c \cap J_v^c \cap \{x \in \Omega : \ \underline{u}(x) = \underline{v}(x)\}. \tag{4.12}$$

Observe that $\underline{u}$, $\underline{v}$ as well as their difference $w = \underline{u} - \underline{v}$ are lower semicontinuous functions. In particular, for any $c \in \mathbb{R}$ the sets $w^{-1}((c, \infty))$ and $w^{-1}((-\infty, c])$ are open and closed, respectively. Thus we have

$$\{x \in \Omega : \underline{u}(x) = \underline{v}(x)\} = w^{-1}(\{0\})$$
$$= \underbrace{w^{-1}((-\infty, 0])}_{\text{closed}} \cap \underbrace{\bigcap_{n \in \mathbb{N}} w^{-1}\left(\left(-\frac{1}{n}, \infty\right)\right)}_{G_\delta},$$

and hence the set $\{x \in \Omega : \underline{u}(x) = \underline{v}(x)\}$ is G_δ as an intersection of G_δ sets (recall here that in any metric space every closed set F is G_δ since $F = \bigcap_{n=1}^{\infty}\{x : \mathrm{dist}(x, F) < \frac{1}{n}\}$). Hence from (4.12) we have that A is G_δ as well. Likewise, if $I \subseteq \Omega$ is an open interval then $I \cap A$ is also G_δ. It is now convenient to consider $I \cap A$ without its potential isolated points [2] $\{a_n\}_{n \in \mathbb{N}}$. Thus we define

$$\tilde{A} = (I \cap A) \cap (\{a_n\}_{n \in \mathbb{N}})^c,$$

which remains a G_δ set as an intersection of two G_δ sets. Hence, we can write

$$\tilde{A} = \bigcap_{n \in \mathbb{N}} A_n, \quad A_n = \bigcup_{k \in \mathbb{N}} I_n^k,$$

where the sequence $(A_n)_{n \in \mathbb{N}}$ can be chosen to be decreasing and with each $\{I_n^k\}_{k \in \mathbb{N}}$ being a disjoint family of open intervals.

We are now ready to proceed with the main part of the proof. For every $x_1, x_2 \in \tilde{A}$ with $x_1 < x_2$ we have the following equalities

$$Du((x_1, x_2)) = u(x_2) - u(x_1) = v(x_2) - v(x_1) = Dv((x_1, x_2)).$$

We fix an interval I_n^k and claim that $Du(I_n^k) = Dv(I_n^k)$. Since we have removed all the isolated points from A (left and right ones), we can assume, by potentially making I_n^k smaller, that its endpoints can be approximated by points of $\tilde{A}$, i.e.,

$$I_n^k = \bigcup_{i \in \mathbb{N}}(x_1^i, x_2^i), \quad x_1^i, x_2^i \in \tilde{A}, \quad \text{with} \quad (x_1^i, x_2^i) \subseteq (x_1^{i+1}, x_2^{i+1}), \ \ i \in \mathbb{N}.$$

This implies that

$$Du(I_n^k) = Du\left(\bigcup_{i \in \mathbb{N}}(x_1^i, x_2^i)\right) = \lim_{i \to \infty} Du((x_1^i, x_2^i))$$
$$= \lim_{i \to \infty} Dv((x_1^i, x_2^i)) = Dv\left(\bigcup_{i \in \mathbb{N}}(x_1^i, x_2^i)\right) = Dv(I_n^k).$$

It then follows that for every $n \in \mathbb{N}$

$$Du(A_n) = \sum_{k \in \mathbb{N}} Du(I_n^k) = \sum_{k \in \mathbb{N}} Dv(I_n^k) = Dv(A_n),$$

---

[2] Here by isolated points, we mean that there exists $\epsilon > 0$ such that $(\alpha_n, \alpha_n + \epsilon)$ or $(\alpha_n - \epsilon, \alpha_n)$ does not intersect $I \cap A$.

and from continuity of the measures Du and Dv, since $(A_n)_{n\in\mathbb{N}}$ is decreasing, we have

$$Du(\tilde{A}) = \lim_{n\to\infty} Du(A_n) = \lim_{n\to\infty} Dv(A_n) = Dv(\tilde{A}).$$

Finally from the fact that the isolated points of $I \cup A$, $\{a_n\}_{n\in\mathbb{N}}$ are at most countable and u, v are both continuous there, we have that

$$Du\left(\bigcup_{n\in\mathbb{N}} \{a_n\}\right) = Dv\left(\bigcup_{n\in\mathbb{N}} \{a_n\}\right) = 0,$$

and hence

$$Du(I \cap A) = Du(\tilde{A}) = Dv(\tilde{A}) = Dv(I \cap A),$$

with this equality being true for any open interval $I \subseteq \Omega$. This implies that $|Du|(A) = |Dv|(A)$. $\qquad\square$

5 Weighted total variation with vanishing weight function α

In this section we are considering the problem (3.1) in the case where the weight function $\alpha \in C(\overline{\Omega})$ can also have zero values, i.e., $\alpha \geq 0$. This is motivated by the fact that a vanishing weight function imposes locally no regularisation. In terms of image processing this is useful when one wants the reconstructed image to be equal to the data in some areas. Moreover, as we analyse in Section 5.2, a vanishing weight function can result to a better preservation of contrast and exact reconstruction of piecewise constant data.

5.1 The well-posedness

Concerning existence of solutions to (3.1) with $\alpha \geq 0$, note that a simple application of the direct method of calculus of variations fails here, due to the absence of coercivity of the objective functional. We overcome this complication by employing a Γ-convergence argument. First we focus on the lower semicontinuity of the weighted TV functional.

Proposition 5.1. *Let $\alpha \in C(\overline{\Omega})$ with $\alpha \geq 0$. Then the map $u \mapsto \int_\Omega \alpha(x)d|Du|$ defined on $\mathrm{BV}(\Omega)$ is lower semicontinuous with respect to the weak* convergence in $\mathrm{BV}(\Omega)$.*

Proof. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence converging to u weakly* in $\mathrm{BV}(\Omega)$. In particular, $Du_n \to Du$ weakly* in the space of Radon measures, i.e.,

$$\lim_{n\to\infty} \int_\Omega v \, dDu_n \to \int_\Omega v \, dDu, \quad \text{for every } v \in C_0(\Omega).$$

This implies that

$$\lim_{n\to\infty} \int_\Omega v\alpha \, dDu_n \to \int_\Omega v\alpha \, dDu, \quad \text{for every } v \in C_0(\Omega),$$

since $v\alpha \in C_0(\Omega)$ for every $v \in C_0(\Omega)$. Thus, we also have $\alpha Du_n \to \alpha Du$ weakly* in the space of Radon measures. From the lower semicontinuity of the total variation map with respect to the weak* convergence in the space of Radon measures we get

$$\int_\Omega \alpha(x)d|Du| = |\alpha Du|(\Omega) \leq \liminf_{n\to\infty} |\alpha Du_n|(\Omega) = \liminf_{n\to\infty} \int_\Omega \alpha(x)d|Du_n|.$$

$\qquad\square$

We now proceed with the following Γ-convergence result.

Proposition 5.2. *Let $\alpha \in C(\overline{\Omega})$ with $\alpha \geq 0$ and $(\alpha_n)_{n \in \mathbb{N}} \subseteq C(\overline{\Omega})$ being a decreasing sequence with $\alpha_n \geq 0$ for every $n \in \mathbb{N}$ and $\alpha_n \to \alpha$, uniformly. Defining $F, F_n : \mathrm{BV}(\Omega) \to \mathbb{R}$, by*

$$F(u) := \int_\Omega \alpha(x) d|Du|, \quad F_n(u) := \int_\Omega \alpha_n(x) d|Du|,$$

then, the sequence $(F_n)_{n \in \mathbb{N}}$ Γ-converges to F with the underlying topology being the topology of weak* convergence in $\mathrm{BV}(\Omega)$.

Proof. It is straightforward to show that $(F_n)_{n \in \mathbb{N}}$ is a decreasing sequence which converges to F pointwise since for every $u \in \mathrm{BV}(\Omega)$ we have

$$|F_n(u) - F(u)| \leq \int_\Omega |\alpha_n(x) - \alpha(x)| d|Du| \leq \|\alpha_n - \alpha\|_\infty |Du|(\Omega) \to 0 \quad \text{as } n \to \infty.$$

Then the conclusion follows from [DM93, Prop. 5.7] and the fact that F is lower semicontinuous with respect to the weak* convergence in $\mathrm{BV}(\Omega)$. $\qquad\square$

Note that the above results hold in arbitrary dimension. However, the proof of the well-posedness of (3.1) with vanishing weight α, strongly relies on Theorem 4.2 and, thus, it only holds in dimension one.

Theorem 5.3. *Let $\Omega = (a, b)$, $f \in \mathrm{BV}(\Omega)$ and $\alpha \in C(\overline{\Omega})$ with $\alpha \geq 0$. Then the minimisation problem*

$$\min_{u \in \mathrm{BV}(\Omega)} J(u) := \frac{1}{2} \int_\Omega (f - u)^2 dx + \int_\Omega \alpha(x) d|Du|,$$

has a unique solution $u^* \in \mathrm{BV}(\Omega)$. Moreover if

$$J_n(u) := \frac{1}{2} \int_\Omega (f - u)^2 dx + \int_\Omega \alpha_n(x) d|Du|,$$

where $(\alpha_n)_{n \in \mathbb{N}} \subseteq C^\infty(\overline{\Omega})$ is a decreasing sequence, uniformly convergent to α, with $\alpha_n > 0$ for every $n \in \mathbb{N}$, then

$$J(u^*) = \lim_{n \to \infty} \min_{u \in \mathrm{BV}(\Omega)} J_n(u).$$

Proof. Note first that since α is continuous, using a standard mollification argument, we can indeed construct a decreasing sequence $(\alpha_n)_{n \in \mathbb{N}}$ in $C^\infty(\overline{\Omega})$, uniformly convergent to α, and $\alpha_n > 0$ for every $n \in \mathbb{N}$. Using Proposition 5.2 and the fact that the map $u \mapsto \frac{1}{2} \int_\Omega (f - u)^2 dx$ is lower semicontinuous with respect to the weak* convergence in $\mathrm{BV}(\Omega)$ we have [DM93, Prop. 6.25] that the sequence $(J_n)_{n \in \mathbb{N}}$ Γ-converges to J. Let now

$$u_n = \operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} J_n(u).$$

We then have

$$\frac{1}{2} \int_\Omega (f - u_n)^2 dx \leq \frac{1}{2} \int_\Omega (f - u_n)^2 dx + \int_\Omega \alpha_n(x) d|Du_n| \leq \frac{1}{2} \int_\Omega f^2 dx.$$

34

Thus the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$ and hence also in $L^1(\Omega)$. Using now Theorem 4.2 we have that $(u_n)_{n\in\mathbb{N}}$ is also bounded in $\mathrm{BV}(\Omega)$, i.e., there exists a positive constant C such that

$$u_n \in K := \{u \in \mathrm{BV}(\Omega): \ \|u\|_{\mathrm{BV}(\Omega)} \leq C\}, \quad \text{for all } n \in \mathbb{N}.$$

This means that for every $n \in \mathbb{N}$

$$\min_{u\in\mathrm{BV}(\Omega)} J_n(u) = \min_{u\in K} J_n(u),$$

witk K being a weakly* sequentially compact set. Then the conclusions of the theorem follow straightforwardly from [DM93, Thm. 7.4] and the fact that the functional J is strictly convex.

$\square$

We would like to finish this section with an extension of Theorem 4.2 for continuous weight functions α that are also not necessarily bounded away from zero.

Theorem 5.4 ($|Du|(\Omega) \leq |Df|(\Omega)$, for continuous α). *Let $\Omega = (a,b) \subseteq \mathbb{R}$, $\alpha \in C(\overline{\Omega})$ with $\alpha \geq 0$ and $f \in \mathrm{BV}(\Omega)$. If*

$$u = \operatorname*{argmin}_{u\in\mathrm{BV}(\Omega)} \frac{1}{2}\int_\Omega (f-u)^2 dx + \int_\Omega \alpha(x)d|Du|,$$

then

$$|Du|(\Omega) \leq |Df|(\Omega).$$

Proof. Define the weights $(\alpha_n)_{n\in\mathbb{N}} \in C^\infty(\overline{\Omega})$ and the functionals J, $(J_n)_{n\in\mathbb{N}}$ as in Theorem 5.3. Let also u and $(u_n)_{n\in\mathbb{N}}$ be the corresponding minimisers. Note that we have already shown in Theorem 4.2 that

$$|Du_n|(\Omega) \leq |Df|(\Omega), \quad \text{for all } n \in \mathbb{N}. \tag{5.1}$$

Since the sequence $(J_n)_{n\in\mathbb{N}}$ Γ-converges to J, we have that any weak* cluster point of $(u_n)_{n\in\mathbb{N}}$ is equal to u [DM93, Cor. 7.20]. This implies that the whole sequence $(u_n)_{n\in\mathbb{N}}$ converges to u weakly* in $\mathrm{BV}(\Omega)$. Hence from the lower semicontinuity of total variation with respect to L^1 convergence and (5.1) we have

$$|Du|(\Omega) \leq \liminf_{n\to\infty} |Du_n|(\Omega) \leq |Df|(\Omega).$$

$\square$

Notice that in view of Theorem 5.4, Theorem 5.3 holds also in the case where $(\alpha_n)_{n\in\mathbb{N}} \subseteq C(\overline{\Omega})$.

5.2 Application: Exact reconstruction of piecewise constant noisy data

It is well-known that scalar total variation regularisation is very efficient in recovering noisy piecewise constant functions. Nevertheless the reconstruction typically suffers from a loss of contrast. Techniques like Bregman iteration [OBG$^+$05] have been used to reduce this

effect. In this section we show that by choosing a suitable weight function α, it is possible to avoid this loss of contrast and, under some mild assumptions on the noise, to exactly recover the piecewise constant function.

In what follows, let f_0 be a piecewise constant function, i.e.,

$$f_0(x) = \sum_{i=1}^{N} f_i \mathcal{X}_{I_i}(x), \quad x \in \Omega, \tag{5.2}$$

where $(I_i)_{i=1}^{N}$ are disjoint intervals with $\bigcup_{i=1}^{N} I_i = \Omega$ and $f_i \in \mathbb{R}$ for every $i = 1, \ldots, N$ with $f_i \neq f_j$ for $i \neq j$. Moreover let η denote an oscillatory function that belongs to $\mathrm{BV}(\Omega)$ and satisfies

$$\int_{I_i} \eta \, dx = 0 \quad \text{for all } i = 1, \ldots, N. \tag{5.3}$$

Proposition 5.5. *Let $f \in \mathrm{BV}(\Omega)$ with*

$$f = f_0 + \eta,$$

where f_0 and η are as described above. Define α to be a continuous piecewise affine function such that

(i) $\alpha(x_i) = 0$, where $x_i = \sup I_i$, $i = 1, \ldots, N - 1$.

(ii) $\alpha \geq 0$ and it is of the form $-\mu_i |x - c_i| + d_i$ in every interval I_i, with $\mu_i > 2\|f\|_\infty$ for every $i = 1, \ldots, N$.

Then we have that

$$f_0 = \underset{u \in \mathrm{BV}(\Omega)}{\mathrm{argmin}} \frac{1}{2} \int_\Omega (f - u)^2 dx + \int_\Omega \alpha(x) d|Du|. \tag{5.4}$$

Proof. Consider first the corresponding minimisation problems with weights $\alpha_n := \alpha + 1/n$ and solutions u_n. Let u be the solution of the weighted total variation minimisation with weight α. We have to show that $u = f_0$. Since $|\alpha'_n(x)| > 2\|f\|_\infty$ in every interval I_i, from Proposition 3.10 in combination with Proposition 3.7 we have that u_n will be constant in the interior of every I_i. Moreover, the Γ-convergence argument from the proof of Theorem 5.4 implies that $u_n \to u$ weakly* in $\mathrm{BV}(\Omega)$ and thus u will be also constant in the interior of every I_i. From the fact that $\alpha(x_i) = 0$ we have that

$$\int_\Omega \alpha(x) d|Du| = 0,$$

i.e.,

$$u = \underset{\phi \in \mathrm{Ker} \int \alpha(x) d|D\cdot|}{\mathrm{argmin}} \frac{1}{2} \int_\Omega (f - \phi)^2 dx, \tag{5.5}$$

where

$$\mathrm{Ker} \int \alpha(x) d|D \cdot | = \left\{ u \in \mathrm{BV}(\Omega) : \; u = \sum_{i=1}^{N} u_i \mathcal{X}_{I_i}(x), \; u_i \in \mathbb{R}, \; i = 1, \ldots, N \right\}.$$

It is then easy to see that the minimisation in (5.5) can be separated into N minimisation problems each of which corresponds to an interval I_i. Consequently, the solution u is of the form

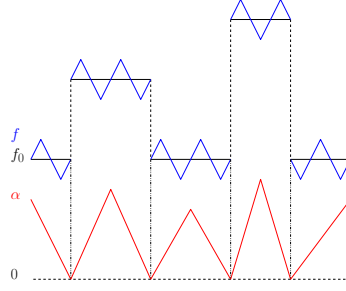$$u = \sum_{i=1}^{N} u_i \mathcal{X}_{I_i}(x),$$

Figure 8: Schematic explanation of Proposition 5.5. A noisy piecewise constant function can be recovered exactly using weighted TV regularisation under a suitable weight function α which vanishes exactly at the jump points of f_0 and has large gradient everywhere else
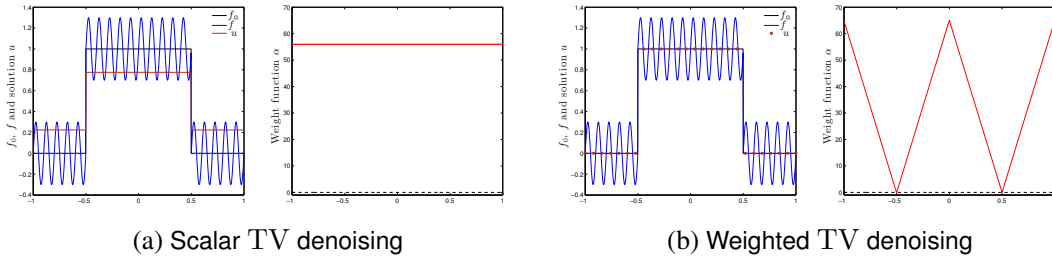

(a) Scalar TV denoising

(b) Weighted TV denoising

Figure 9: Numerical examples: Unlike the scalar case, weighted TV can recover exactly noisy piecewise constant functions

where

$$
u_i = \operatorname*{argmin}_{c \in \mathbb{R}} \frac{1}{2} \int_{I_i} (f - c)^2 dx = \frac{1}{|I_i|} \int_{I_i} f\, dx = \frac{1}{|I_i|} \int_{I_i} f_0 + \eta\, dx
$$
$$
\stackrel{(5.3)}{=} \frac{1}{|I_i|} \int_{I_i} f_0\, dx = f_i,
$$

and thus $u = f_0$. □

We should remark here that this exact recovery of piecewise constant functions can be achieved for a whole family of weight functions α that are not necessarily of the form $-\mu_i|x - c_i| + d_i$ in every interval I_i. For instance any weight function α which is initially increasing and then decreasing in every interval I_i with the property $|\alpha'(x)| > \|f\|_\infty$ (except at its maximum point in I_i) will also lead to an exact recovery of f_0.

Figure 8 depicts the result of Proposition 5.5. Note that α must be zero at the jump points of f so that these jumps are not penalised at all. On the other hand, α should be steep enough so the remaining parts of the solution u are constant according to Proposition 3.10. We also provide some numerical examples in Figure 9. There, the data f is a perturbation of a piecewise constant function f_0, under a $\sin$ type of noisy function that satisfies the mean value property (5.3). For the scalar TV denoising, we have chosen the smallest α such that we get a piecewise constant solution. As expected, this results in a significant loss of contrast, Figure 9a. However, by applying weighted TV regularisation with a weight as it is described in Proposition 5.5, an exact recovery is achieved, Figure

37

9b. One might expect that since the values of α remain small close to the jump points, then the noise would be still present in that area. However, this is not true, since constant parts do not only result due to the large magnitude of α but also due to high values of α', Proposition 3.10.

6 Total variation regularisation with weighted fidelity term

As we have already mentioned in the introduction, one has two options when it comes to spatially adapted regularisation, i.e., introducing a weight function either in the regulariser or in the fidelity term. In this final section of the paper we would like to briefly examine the second case and in particular study the total variation regularisation problem with weighted fidelity term

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_{\Omega} w(f-u)^2 dx + |Du|(\Omega), \tag{6.1}$$

where $\Omega = (a,b)$, $f \in L^2(\Omega)$ and $w \in L^\infty(\Omega)$ with $w \geq 0$.

Existence of solutions for the problem (6.1) follows straightforwardly using standard methods. Note that the possibility of w vanishing in some areas does not pose extra difficulties as in the weighted total variation case. It is clear, however, that the solution of (6.1) is not always unique. The solution will be unique if the operator $T_w : L^2(\Omega) \to L^2(\Omega)$ with $T_w u(x) = \sqrt{w(x)} u(x)$, is injective, i.e., when the set $\{w = 0\}$ is of zero Lebesgue measure.

First order optimality conditions for (6.1) are stated next, see, e.g., [Rin00] for a proof.

Proposition 6.1 ([Rin00])**.** *Let $\Omega = (a,b)$, $f \in L^2(\Omega)$ and $w \in L^\infty(\Omega)$ with $w \geq 0$. A function $u \in \mathrm{BV}(\Omega)$ is a solution of* (6.1) *if and only if there exists a function $v \in H_0^1(\Omega)$ such that*

$$v' = w(f-u), \tag{6.2}$$

$$-v \in \mathrm{Sgn}(Du). \tag{6.3}$$

As it can be readily seen from (6.2)–(6.3) and also shown in [Rin00] we have that $Du = 0$ when $f \neq u$, or more rigorously on the open sets $\{\underline{f} > \overline{u}\}$ and $\{\overline{f} < \underline{u}\}$, provided $w > 0$ there. This is in strong contrast to the weighted TV case where this is not true, in general. Furthermore, unlike weighted TV, no new discontinuities can be created with the one dimensional version of (6.1). This is shown in the following proposition whose proof is similar to the one dimensional scalar TV case.

Proposition 6.2. *Let $\Omega = (a,b)$, $f \in \mathrm{BV}(\Omega)$ and $w \in L^\infty(\Omega)$ with $w \geq 0$. If u is a solution to* (6.1) *then*

$$J_u \cap \mathrm{supp}_{l,r} w \subseteq J_f,$$

where

$$\mathrm{supp}_{l,r} w = \big\{ x \in \Omega : \text{ for every } \epsilon > 0, \ w \text{ is not zero a.e. in each of the sets}$$
$$(x - \epsilon, x) \text{ and } (x, x + \epsilon) \big\}.$$

In particular if $w > 0$ a.e. then

$$J_u \subseteq J_f.$$

Proof. Let $x \in J_u \cap \operatorname{supp}_{l,r} w$ and suppose that $x \notin J_f$. Without loss of generality, we can assume that $Du(\{x\}) > 0$. Then from condition (6.3) we have $v(x) = -1$, where v is the corresponding $H_0^1(\Omega)$ variable of Proposition 6.1. Observe that since u has a positive jump at x and f is continuous there we have that there exists $m > 0$ and a sufficiently small $\epsilon > 0$ such that

$$\text{either} \quad 0 < m < \operatorname*{ess\,inf}_{t \in (x-\epsilon, x)} f(t) - u(t) \quad \text{or} \quad \operatorname*{ess\,sup}_{t \in (x, x+\epsilon)} f(t) - u(t) < -m < 0.$$

It is also possible that both inequalities above hold. If the first inequality holds, then, in view of (6.2) and the fact that $x \in \operatorname{supp}_{l,r} w$, we have that

$$v(x) - v(x - \delta) = \int_{x-\delta}^{\delta} v'(s) ds > 0, \quad \text{for all } \delta \in (0, \epsilon).$$

This implies that $v(x - \delta) < v(x) = -1$. If the second inequality holds then we have

$$v(x + \delta) - v(x) = \int_{x}^{x+\delta} v'(s) ds < 0, \quad \text{for all } \delta \in (0, \epsilon),$$

which implies that $v(x + \delta) < v(x) = -1$. Thus, in both cases we end up in a contradiction since we must have $|v| \leq 1$ in Ω according to (6.3). □

Next we study relations between (3.1) and (6.1) by means of an explicit example, which offers a good insight into different structural properties of the associated solutions.

Proposition 6.3. *Let $\Omega = (-L, L)$, $f(x) = \lambda x$ with $\lambda > 0$. Moreover let $w \in L^\infty(\Omega)$ with $w \geq 0$. Then, a solution to the problem*

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_{\Omega} w(f - u)^2 dx + |Du|(\Omega),$$

will be of the following form:

$$u(x) = \begin{cases} \lambda x_1, & \text{if} \quad x \in (-L, x_1), \\ \lambda x, & \text{if} \quad x \in [x_1, x_2], \\ \lambda x_2, & \text{if} \quad x \in (x_2, L), \end{cases} \quad \text{where} - L < x_1 \leq x_2 < L. \quad (6.4)$$

Proof. We will show that for every weight function $w \in L^\infty(\Omega)$, a function u of the type (6.4), will always satisfy the optimality conditions (6.2)–(6.3) together with an appropriate function $v \in H_0^1(\Omega)$. Since $v \in H_0^1(\Omega)$ and it satisfies (6.2) we have

$$v(x_1) = \lambda \int_{-L}^{x_1} w(s)(s - x_1) ds, \quad v(x_2) = -\lambda \int_{x_2}^{L} w(s)(s - x_2) ds.$$

Notice that if we set

$$\phi_1(x) := \lambda \int_{-L}^{x} w(s)(s - x) ds, \quad \phi_2(x) := -\lambda \int_{x}^{L} w(s)(s - x) ds, \quad x \in [-L, L],$$

then we have that $\phi_1(-L) = \phi_2(L) = 0$, ϕ_1 and ϕ_2 are respectively decreasing and increasing continuous functions, and hence there exists a point $\xi \in (-L, L)$ such that
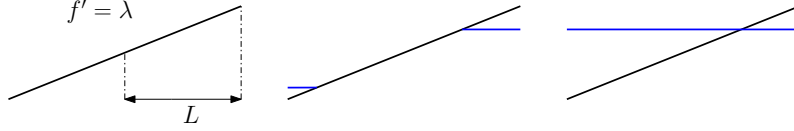
Figure 10: Different types of solutions of the total variation minimisation problem with weighted fidelity term (6.1) with data $f(x) = \lambda x$ in $(-L, L)$ and for any weight function w

$\phi_1(\xi) = \phi_2(\xi) = k < 0$.

Case I: $k \leq -1$. In that case we set

$$x_1 = \min\{x \in \Omega : \phi_1(x) = -1\} \quad \text{and} \quad x_2 = \max\{x \in \Omega : \phi_2(x) = -1\},$$

and we have $-L < x_1 < x_2 < L$. Then it is easy to check that

$$v(x) = \begin{cases} \phi_1(x), & \text{if} \quad x \in (-L, x_1), \\ -1, & \text{if} \quad x \in [x_1, x_2], \\ \phi_2(x), & \text{if} \quad x \in (x_2, L), \end{cases}$$

belongs to $H_0^1(\Omega)$ and satisfies the optimality conditions (6.2)–(6.3) together with the function u defined in (6.4).

Case II: $k > -1$. In this case, we simply set

$$v(x) = \begin{cases} \phi_1(x), & \text{if} \quad x \in (-L, \xi], \\ \phi_2(x), & \text{if} \quad x \in (\xi, L), \end{cases}$$

and the optimality conditions are satisfied with this function v and the constant function $u = \lambda \xi$. □

As the proof above shows, if the weight function is small enough in the sense that either

$$\phi_1(L) \leq -1 \quad \text{or} \quad \phi_2(-L) \leq -1, \tag{6.5}$$

then the solution u of (6.1) will be a constant.

In Figure 10 we depict all these possible solutions. These should be compared to the ones of the weighted TV problem in Figure 6. Note that when $x_1 = -x_2$ (which can be achieved for instance using a symmetric weight function w), then the resulting solution also belongs to the set of solutions of the scalar total variation minimisation problem.

We have also performed some numerical simulations for the problem (6.1) with data $f(x) = \lambda x$; see Figure 11. Note that, as predicted by the Proposition 6.3, all the solutions are of the form (6.4). In order to avoid the non-uniqueness issues, we have set the minimum value of each weight function w to be a small positive constant. We briefly now comment on these results:

(a) In Figure 11a we have chosen the weight functions to have a small value away from the boundary of the domain, while increasing their values close to the boundary. Observe that the solutions u converge to f as the values of the weight functions near the boundary are increasing. Thus, for this example, in order to recover f almost perfectly, it suffices to choose w to have large values only near the boundary.

(a) Weight functions w with large values near the boundary

(b) Linear weight functions w

(c) Weight functions w of small mass

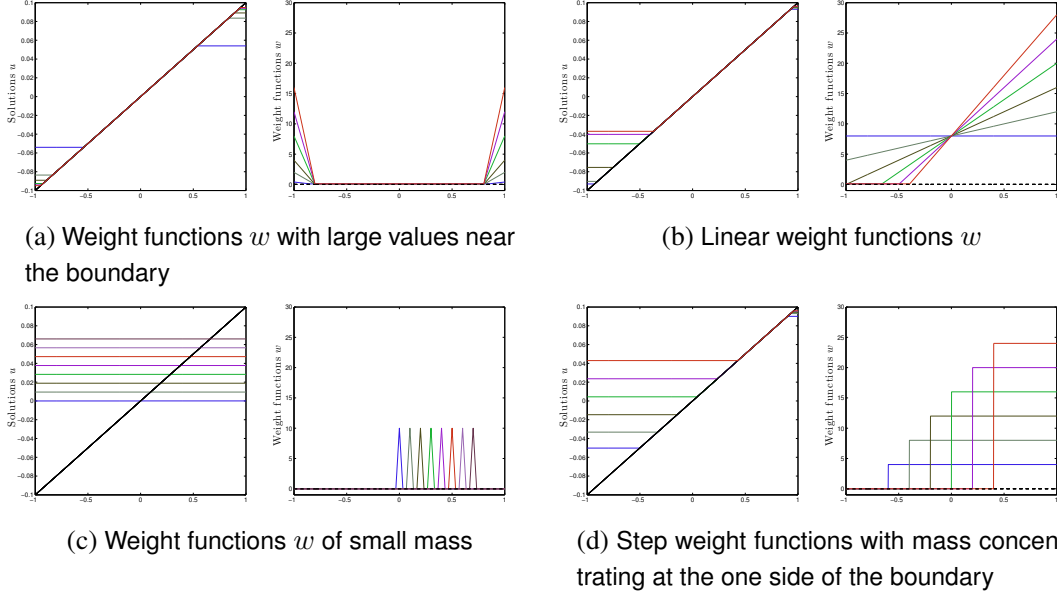(d) Step weight functions with mass concentrating at the one side of the boundary

Figure 11: Numerical solutions of the problem (6.1) with data $f(x) = \lambda x$ in $(-L, L)$ for different types of weight functions w. In every case the solutions are of the form (6.4)

(b) In Figure 11b, the weight functions have been chosen to be linear with increasing gradient. As expected, as the gradient of w increases, resulting in large values close to the point L, the solution u approximates f better, while the approximation becomes worse close to $-L$ where w has small values.

(c) In Figure 11c, all the weight functions w have been chosen such that their mass is small and is shifted towards the boundary. In all the cases, the conditions (6.5) are satisfied and hence constant solutions are obtained.

(d) Lastly, in Figure 11d, we have chosen weight functions w whose mass concentrates at the right side of the boundary of Ω, observing a similar behaviour as in Figure 11b.

The following proposition states that at least for the affine data considered above, the weighted TV problem (3.1) cannot produce those solutions of (6.1) that are not symmetric with respect to the origin, i.e., the ones that cannot be obtained by scalar total variation regularisation. We note that this is regardless of the choice of the weight function α.

Proposition 6.4. *Let* $\Omega = (-L, L)$, $\alpha \in C(\overline{\Omega})$, $\alpha > 0$, $f(x) = \lambda x$, *for every* $x \in \Omega$, $\lambda > 0$. *Then the weighted total variation problem* (3.1) *with data* f *and weight* α *cannot have any solution* u *of the type* (6.4) *unless* $x_1 = -x_2$, *that is, unless the solution is symmetric with respect to the origin.*

Proof. Suppose without loss of generality that there exists a weight function $\alpha \in C(\overline{\Omega})$ such that the solution of (3.1) is a function of the type (6.4), where $x_1 < -x_2$. From conditions (3.13)–(3.14) we have that this can only happen if $\alpha(x) = \alpha_0$ for every $x \in (x_1, x_2)$. Moreover, the dual function v would have to be quadratic in the intervals $(0, x_1]$ and $[x_2, L)$ satisfying the conditions $v(0) = 0$, $v(x_1) = -\alpha_0$, $v'(x_1) = 0$, and $v(L) = 0$, $v(x_2) = -\alpha_0$, $v'(x_2) = 0$. One can easily see that such a v cannot exist but note

41

also that in that case we would have $-\alpha_0 \leq v(x) \leq 0$ for every $x \in \Omega$, i.e., u and v would also satisfy the optimality conditions of standard total variation minimisation with scalar parameter α_0. Of course this is a contradiction since we know that any solution to the scalar parameter problem is not of that form.

$\square$

In contrast to the previous result, one can easily see that for the specific data function of Proposition 6.4, there are choices of non-constant weight functions $\alpha \in C(\overline{\Omega})$ and $w \in L^\infty(\Omega)$ such that the corresponding solutions of the problems (3.1) and (6.1) are solutions of scalar total variation problems. These are exactly the solutions of the type (6.4) with $x_1 = -x_2$. Figure 12 visualises the relation between (3.1) and (6.1).
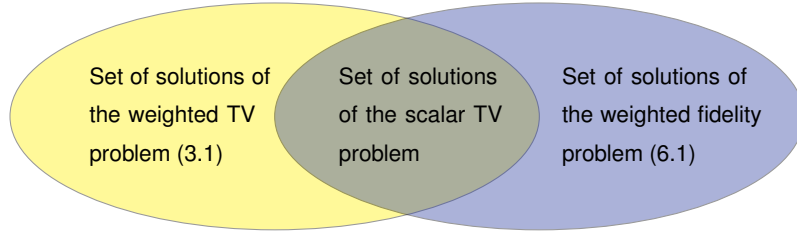

Figure 12: Data function $f(x) = \lambda x$: The solutions that can be obtained by both problems (3.1) and (6.1) are exactly those that can be obtained by the classical scalar total variation minimisation.

In Section 5.2 we examined the capability of the weighted TV model to recover exactly piecewise constant data, provided the weight function α vanishes at the jump points of the data. On the contrary, the weighted fidelity model (6.1) cannot recover piecewise constant functions even in the absence of noise, regardless of the weight function w. Indeed, it can be easily checked using the optimality conditions (6.2)–(6.3) that only constant functions can be recovered exactly using (6.1). However, as the next proposition shows, a piecewise constant function can be recovered as a limit of a sequence of solutions of (6.1) using suitable weight functions, which have mass concentrating at the jump points of the data.

Proposition 6.5. *Let $f_0 \in \mathrm{BV}(\Omega)$ be a piecewise constant function as defined in* (5.2) *and let*

$$f = f_0 + \eta,$$

where η is a continuous function which vanishes at the jump points of f_0 (not necessarily satisfying (5.3)*). Define the sequence of weight functions*

$$w_n(x) = \sum_{i=1}^{N-1} n^2 \mathcal{X}_{B(x_i, \frac{1}{n})},$$

where x_i, $i = 1, \ldots, N-1$ are the jump points of f_0. Then for any sequence $(u_n)_{n \in \mathbb{N}}$ such that

$$u_n \in \operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \int_\Omega w_n (f - u)^2 dx + |Du|(\Omega),$$

we have

$$\|f_0 - u_n\|_\infty \to 0 \quad \text{as } n \to \infty.$$

Proof. We can assume that n is large enough so that the balls $B(x_i, \frac{1}{n})$ are all inside Ω. Note as well that since each u_n is optimal, we have for every $n \in \mathbb{N}$

$$\frac{1}{2} \int_\Omega w_n (f - u_n)^2 dx \leq \frac{1}{2} \int_\Omega w_n (f - u_n)^2 dx + |Du_n|(\Omega) \leq |Df|(\Omega).$$

Using the above estimate, we find

$$\frac{1}{|B(x_i, \frac{1}{n})|} \int_{B(x_i, \frac{1}{n})} (f - u_n)^2 dx = \frac{1}{n^2 |B(x_i, \frac{1}{n})|} \int_{B(x_i, \frac{1}{n})} w_n (f - u_n)^2 dx$$
$$\leq \frac{2n}{n^2} |Df|(\Omega),$$

and hence

$$\frac{1}{|B(x_i, \frac{1}{n})|} \int_{B(x_i, \frac{1}{n})} (f - u_n)^2 dx \to 0, \quad \text{as} \quad n \to \infty, \tag{6.6}$$

for every x_i, $i = 1, \ldots, N - 1$. Note that (6.6) in fact implies

$$\sup_{x \in (x_i, x_i + \frac{1}{n})} (f(x) - u_n(x))^2 \to 0 \quad \text{and} \quad \sup_{x \in (x_i - \frac{1}{n}, x_i)} (f(x) - u_n(x))^2 \to 0, \tag{6.7}$$

as $n \to \infty$, for every $i = 1, \ldots, N - 1$. Indeed, suppose that (6.7) does not hold, say that the first limit fails for some x_i. Then there exists an $\epsilon > 0$ such that for every $k \in \mathbb{N}$ there exists an $n_k \in \mathbb{N}$, $n_k > k$, and a $x_{n_k} \in (x_i, x_i + \frac{1}{n_k})$ such that

$$(f(x_{n_k}) - u_{n_k}(x_{n_k}))^2 \geq \epsilon. \tag{6.8}$$

Since f is right continuous in x_i, we can pick a large enough $k \in \mathbb{N}$ such that $|f(x_{i+}) - f(x)| < \sqrt{\epsilon}/4$ for every $x \in (x_i, x_i + \frac{1}{n_k})$. Then from (6.8) and Proposition 6.1 it follows that u_{n_k} is constant in $(x_i, x_i + \frac{1}{n_k})$, i.e., $u_{n_k} = c_k$ and $|f(x) - c_k| \geq \sqrt{\epsilon}/2$ for every $x \in (x_i, x_i + \frac{1}{n_k})$. However, this violates (6.6) as for large k we would have

$$\frac{1}{|B(x_i, \frac{1}{n_k})|} \int_{B(x_i, \frac{1}{n_k})} (f - u_{n_k})^2 dx \geq \frac{2n_k}{n_k} \inf_{x \in (x_i, x_i + \frac{1}{n_k})} (f(x) - u_{n_k}(x))^2 \geq \epsilon.$$

Thus (6.7) holds. Notice also that it holds as well if we replace the squares with absolute values. Note now that from Proposition 6.2, u_n is continuous in the intervals $(x_i, x_i + \frac{1}{n})$ and $(x_i - \frac{1}{n}, x_i)$. Then it can be readily checked that (any solution) u_n will be just a monotonic interpolation between $u_n^l(x_i + \frac{1}{n})$ and $u_n^r(x_{i+1} - \frac{1}{n})$, for $i = 1, \ldots, N - 1$ and constant in $(a, x_1 - \frac{1}{n})$ and $(x_{N-1} + \frac{1}{n}, b)$ taking the values $u_n^r(x_1 - \frac{1}{n})$ and $u_n^l(x_{N-1} + \frac{1}{n})$ respectively in those intervals. Since $f_0 = f + \eta$ with η being continuous and vanishing at every x_i, using (6.7) we have that

$$\sup_{x \in (x_i, x_i + \frac{1}{n})} |f_0(x) - u_n(x)| \to 0 \quad \text{and} \quad \sup_{x \in (x_i - \frac{1}{n}, x_i)} |f_0(x) - u_n(x)| \to 0, \tag{6.9}$$

as $n \to \infty$. It is then straightforward that u_n converges to f_0 uniformly. $\qquad\square$

While the result of Proposition 6.5 could be considered of theoretical value only, note however that using similar techniques as in Section 3.2 one can show that for the solution u of (6.1) it holds

$$\underline{f}(x) \leq \underline{u}(x) \leq \overline{u}(x) \leq \overline{f}(x), \quad \text{for every } x \in J_f.$$
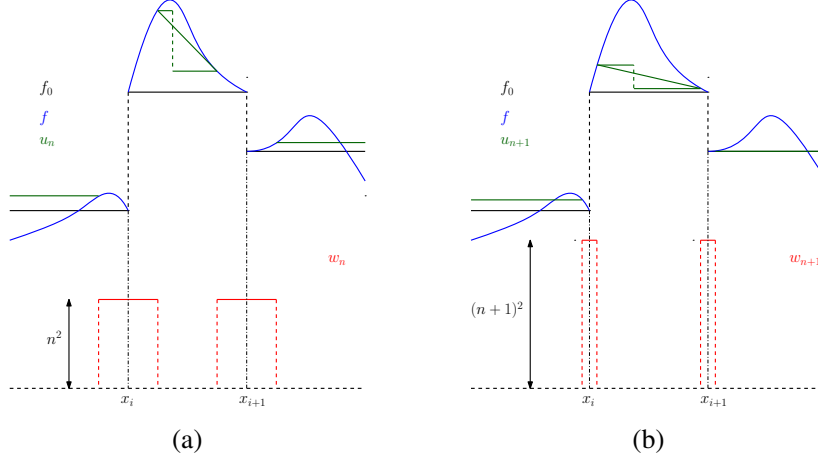
43

Figure 13: Illustration of the result of Proposition 6.5. While a piecewise constant function cannot be recovered exactly using the weighted fidelity total variation regularisation (6.1), this can be achieved in the limit, as the mass of the weight functions w_n is concentrated with a high rate at the jump points of the data f

Thus, if the noise function η changes the values of f_0 at its jump points one cannot expect to recover f_0 using the limiting process of Proposition 6.5. It might happen for instance that $f = f_0 + \eta$ satisfies $\underline{f_0}(x) < \underline{f}(x) < \overline{f}(x) < \overline{f_0}(x)$ at a jump point x and thus any solution u of (6.1) would satify

$$\underline{f_0}(x) < \underline{f}(x) \leq \underline{u}(x) \leq \overline{u}(x) \leq \overline{f}(x) < \overline{f_0}(x).$$

Let us remark also here that it is essential that the values of w_n grow at a high rate around x_i. For instance, if $\|w_n\|_{L^1(\Omega)} \to 0$, we have that the solutions u_n tend to a constant, up to a subsequence. Indeed, in that case we have

$$|Du_n|(\Omega) \leq \frac{1}{2} \int_\Omega w_n(f - u_n)^2 dx + |Du_n|(\Omega) \leq \frac{1}{2} \int_\Omega w_n f^2\, dx$$

$$\leq \frac{1}{2}\|f^2\|_\infty \|w_n\|_{L^1(\Omega)} \to 0, \quad \text{as } n \to \infty.$$

Denoting by u_Ω the mean value of u in Ω, from the Poincaré inequality we get

$$|(u_\Omega)_n| \leq C(\|u_n\|_{L^1(\Omega)} + \|u_n - (u_\Omega)_n\|_{L^1(\Omega)}) \leq C(1 + |Du_n|(\Omega)) < C,$$

for a generic constant C. Thus, there exists a subsequence $(u_\Omega)_{n_k} \to c$ as $k \to \infty$. Then

$$\|u_{n_k} - c\|_{L^1(\Omega)} \leq \|(u_\Omega)_{n_k} - c\|_{L^1(\Omega)} + \|u_{n_k} - (u_\Omega)_{n_k}\|_{L^1(\Omega)}$$

$$\leq |\Omega||(u_\Omega)_{n_k} - c| + C|Du_{n_k}|(\Omega) \to 0, \quad \text{as } k \to \infty.$$

Figure 13 describes the limiting process of Proposition 6.5. As n tends to infinity and the mass of the weight w is concentrated at the jump points x_i, the solution u_n of (6.1) approaches f (and hence f_0) near these jump points, while in the areas where w_n is zero, u_n is a monotonic interpolation. Note that this interpolation is not unique as depicted in the middle part of the Figures 13a–13b.

Finally for the reader's convenience, we summarise in Table 2 the differences between the weighted total variation (3.1) and the weighted fidelity (6.1) models.

44

	Weighted TV model	**Weighted fidelity model**
Creation of new discontinuities	Potentially	No
Solution u is constant when $u \neq f$	Not necessarily	Yes
Exact recovery of piecewise constant functions	Yes, setting $\alpha = 0$ at jump points	Only in the limit $w \to \infty$ at jump points

Table 2: Summary of the differences between the weighted total variation (3.1) and the weighted fidelity (6.1) models

7 Conclusions

We have performed a thorough analysis of the weighted total variation regularisation problem in dimension one. We studied conditions under which discontinuities are created in the solution, at points where the data function is continuous. A variety of analytical and numerical results was provided. Moreover, in contrast to the standard scalar total variation regularisation only a partial version of the parameter semigroup property holds in the weighted case. We further computed some analytic solutions for simple data and weight functions. We were able to infer a bound of the total variation of the solution by the total variation of the data, an estimate which was used to show the well-posedness of the weighted total variation minimisation problem even in the case of vanishing weight function. It was shown, that by using vanishing weights one can recover exactly piecewise constant data. Finally, the total variation problem with weighted fidelity term was considered and it was shown that even for very simple data, its solutions can be very different to those of the weighted total variation problem.

Even though a few of our results hold in arbitrary dimension, the main statements of this paper are in dimension one and their proofs rely on a fine scale one dimensional analysis. Extension of these results in higher dimensions is a subject of future research. We note that a higher dimension study poses some extra challenges, regarding for example the derivation of the predual problem, a characterisation of the subdifferential of the weighted TV, as well as deriving a bound on the total variation of the solution as we have done in Section 4.

References

[ABCH07] A. Almansa, C. Ballester, V. Caselles, and G. Haro, *A TV based restoration model with local constraints*, Journal of Scientific Computing **34** (2007), no. 3, 209–236, `http://dx.doi.org/10.1007/s10915-007-9160-x`.

[ABM14] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces: Applications to PDEs and optimization*, vol. 17, SIAM, 2014.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, USA, 2000.

[AJNO15] P. Athavale, R.L. Jerrard, M. Novaga, and G. Orlandi, *Weighted TV minimization and application to vortex density models*, preprint (2015), `http://cvgmt.sns.it/media/doc/paper/2790/AJNO.pdf`.

[All08a]   W. Allard, *Total variation regularization for image denoising, I. Geometric theory*, SIAM Journal on Mathematical Analysis **39** (2008), no. 4, 1150–1190, `http://dx.doi.org/10.1137/060662617`.

[All08b]   _____, *Total variation regularization for image denoising, II. Examples*, SIAM Journal on Imaging Sciences **1** (2008), no. 4, 400–417, `http://dx.doi.org/10.1137/070698749`.

[All09]   _____, *Total variation regularization for image denoising, III. Examples.*, SIAM Journal on Imaging Sciences **2** (2009), no. 2, 532–568, `http://dx.doi.org/10.1137/070711128`.

[AV94]   R. Acar and C.R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems **10** (1994), no. 6, 1217–1229, `http://dx.doi.org/10.1088/0266-5611/10/6/003`.

[BCRS]   M. Bertalmio, V. Caselles, B. Rougé, and A. Solé, *TV based image restoration with local constraints*, Journal of Scientific Computing **19**, no. 1, 95–122, `http://dx.doi.org/10.1023/A:1025391506181`.

[BDH13]   K. Bredies, Y. Dong, and M. Hintermüller, *Spatially dependent regularization parameter selection in total generalized variation models for image restoration*, International Journal of Computer Mathematics **90** (2013), no. 1, 109–123, `http://dx.doi.org/10.1080/00207160.2012.700400`.

[BKP10]   K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM Journal on Imaging Sciences **3** (2010), no. 3, 492–526, `http://dx.doi.org/10.1137/090769521`.

[BKV13]   K. Bredies, K. Kunisch, and T. Valkonen, *Properties of L^1-TGV^2 : The one-dimensional case*, Journal of Mathematical Analysis and Applications **398** (2013), no. 1, $438 - 454$, `http://dx.doi.org/10.1016/j.jmaa.2012.08.053`.

[BPPS16]   M. Burger, K. Papafitsoros, E. Papoutsellis, and C.B. Schönlieb, *Infimal convolution regularisation functionals of BV and L^p spaces. Part I: The finite p case*, Journal of Mathematical Imaging and Vision **55** (2016), no. 3, 343–369, `http://dx.doi.org/10.1007/s10851-015-0624-6`.

[CCDIR$^+$15] L. Calatroni, V.C. Cao, J.C. De los Reyes, C.B. Schönlieb, and T. Valkonen, *Bilevel approaches for learning of variational imaging models*, arXiv preprint 1505.02120 (2015), `https://arxiv.org/abs/1505.02120`.

[CCN07]   V. Caselles, A. Chambolle, and M. Novaga, *The discontinuity set of solutions of the TV denoising problem and some extensions*, Multiscale Modeling & Simulation **6** (2007), no. 3, 879–894, `http://dx.doi.org/10.1137/070683003`.

[CDIRS16]   V.C. Cao, J.C. De los Reyes, and C.B. Schönlieb, *Learning optimal spatially-dependent regularization parameters in total variation image restoration*, arXiv preprint 1603.09155 (2016), `http://arxiv.org/abs/1603.09155`.

[CDPP16]   A. Chambolle, V. Duval, G. Peyré, and C. Poon, *Geometric properties of solutions to the total variation denoising problem*, arXiv preprint 1602.00087 (2016), `http://arxiv.org/abs/1602.00087`.

[CE05]   T.F. Chan and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, SIAM Journal on Applied Mathematics (2005), 1817–1837, http://dx.doi.org/10.1137/040604297.

[DAG09]   V. Duval, J.F. Aujol, and Y. Gousseau, *The $TVL1$ model: a geometric point of view*, SIAM Journal on Multiscale Modeling & Simulation **8** (2009), no. 1, 154–189, http://dx.doi.org/10.1137/090757083.

[DHRC10a] Y. Dong, M. Hintermüller, and M. M. Rincon-Camacho, *Automated regularization parameter selection in multi-scale total variation models for image restoration*, Journal of Mathematical Imaging and Vision **40** (2010), no. 1, 82–104, http://dx.doi.org/10.1007/s10851-010-0248-9.

[DHRC10b] _____, *A multi-scale vectorial L^T-TV framework for color image restoration*, International Journal of Computer Vision **92** (2010), no. 3, 296–307, http://dx.doi.org/10.1007/s11263-010-0359-1.

[DM93]   G. Dal Maso, *Introduction to Γ-convergence*, Birkhäuser, 1993.

[EG92]   L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992.

[ET76]   I. Ekeland and R. Temam, *Convex analysis and variational problems*, vol. 1, North Holland, 1976.

[FMM12]   K. Frick, P. Marnitz, and A. Munk, *Statistical multiresolution dantzig estimation in imaging: Fundamental concepts and algorithmic framework*, Electronic Journal of Statistics **6** (2012), 231–268, http://dx.doi.org/10.1214/12-EJS671.

[Giu84]   E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser, 1984.

[Gra07]   M. Grasmair, *The equivalence of the taut string algorithm and BV-regularization*, Journal of Mathematical Imaging and Vision **27** (2007), no. 1, 59–66, http://dx.doi.org/10.1007/s10851-006-9796-4.

[HMS$^+$12] Y. Hotz, P. Marnitz, R. Stichtenoth, L. Davies, Z. Kabluchko, and A. Munk, *Locally adaptive image denoising by a statistical multiresolution criterion*, Computational Statistics & Data Analysis **56** (2012), no. 3, 543 – 558, http://dx.doi.org/10.1016/j.csda.2011.08.018.

[HR16]   M. Hintermüller and C.N. Rautenberg, *Optimal selection of the regularization function in a generalized total variation model. Part I: Modelling and theory*, WIAS Preprint No. 2235 (2016), http://www.wias-berlin.de/preprint/2235/wias_preprints_2235.pdf.

[HRC10]   M. Hintermüller and M.M. Rincon-Camacho, *Expected absolute value estimators for a spatially adapted regularization parameter choice rule in L^1-TV-based image restoration*, Inverse Problems **26** (2010), no. 8, 085005, http://stacks.iop.org/0266-5611/26/i=8/a=085005.

[HRWL16]   M. Hintermüller, C.N. Rautenberg, T. Wu, and A. Langer, *Optimal selection of the regularization function in a generalized total variation*

model. Part II: Algorithm, its analysis and numerical tests, WIAS Preprint No. 2236 (2016), `http://www.wias-berlin.de/preprint/2236/wias_preprints_2236.pdf`.

[Jal14]   K. Jalalzai, *Discontinuities of the minimizers of the weighted or anisotropic total variation for image reconstruction*, arXiv preprint 1402.0026 (2014), `http://arxiv.org/abs/1402.0026`.

[Jal15]   _____, *Some remarks on the staircasing phenomenon in total variation-based image denoising*, Journal of Mathematical Imaging and Vision **54** (2015), no. 2, 256–268, `http://dx.doi.org/10.1007/s10851-015-0600-1`.

[Mey01]   Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*, vol. 22, American Mathematical Society, 2001.

[OBG$^+$05] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Modeling & Simulation **4** (2005), no. 2, 460–489, `http://dx.doi.org/10.1137/040605412`.

[Olm56]   J.M.H. Olmsted, *Real variables*, Appleton Century-Crofts, Inc., New York, 1956.

[Pap14]   K. Papafitsoros, *Novel higher order regularisation methods for image reconstruction*, Ph.D. thesis, University of Cambridge, 2014, `https://www.repository.cam.ac.uk/handle/1810/246692`.

[PB15]   K. Papafitsoros and K. Bredies, *A study of the one dimensional total generalised variation regularisation problem*, Inverse Problems and Imaging **9** (2015), no. 2, 511–550, `http://dx.doi.org/10.3934/ipi.2015.9.511`.

[Rin00]   W. Ring, *Structural properties of solutions to total variation regularization problems*, ESAIM: Mathematical Modelling and Numerical Analysis **34** (2000), no. 4, 799–810, `http://dx.doi.org/10.1051/m2an:2000104`.

[ROF92]   L.I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena **60** (1992), no. 1-4, 259–268, `http://dx.doi.org/10.1016/0167-2789(92)90242-F`.

[SC03]   D. Strong and T. Chan, *Edge-preserving and scale-dependent properties of total variation regularization*, Inverse Problems **19** (2003), no. 6, S165, `http://dx.doi.org/10.1088/0266-5611/19/6/059`.

[SGG$^+$09] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*, Springer, New York, 2009.

[Ste15]   G. Steidl, *Combined first and second order variational approaches for image processing*, Jahresbericht der Deutschen Mathematiker-Vereinigung 2015 **117** (2015), no. 2, 133–160, `http://dx.doi.org/10.1365/s13291-015-0113-2`.

[Val15]   T. Valkonen, *The jump set under geometric regularization. Part 1: Basic technique and first-order denoising*, SIAM Journal on Mathematical Analysis **47** (2015), no. 4, 2587–2629, `http://dx.doi.org/10.1137/140976248`.