CrossMark

# Infimal Convolution Regularisation Functionals of BV and L$^p$ Spaces

## Part I: The Finite $p$ Case

**Martin Burger[1] · Konstantinos Papafitsoros[2] · Evangelos Papoutsellis[3] · Carola-Bibiane Schönlieb[3]**

**Abstract** We study a general class of infimal convolution type regularisation functionals suitable for applications in image processing. These functionals incorporate a combination of the total variation seminorm and L$^p$ norms. A unified well-posedness analysis is presented and a detailed study of the one-dimensional model is performed, by computing exact solutions for the corresponding denoising problem and the case $p = 2$. Furthermore, the dependency of the regularisation properties of this infimal convolution approach to the choice of $p$ is studied. It turns out that in the case $p = 2$ this regulariser is equivalent to the Huber-type variant of total variation regularisation. We provide numerical examples for image decomposition as well as for image denoising. We show that our model is capable of eliminating the staircasing effect, a well-known disadvantage of total variation regularisation. Moreover as $p$ increases we obtain almost piecewise affine reconstructions, leading also to a better preservation of hat-like structures.

## 1 Introduction

In this paper, we introduce a family of novel TV–L$^p$ infimal convolution type functionals with applications in image processing:

$$\mathrm{TVL}^p_{\alpha,\beta}(u) := \inf_{w \in \mathrm{L}^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{\mathrm{L}^p(\Omega)},$$
$$\alpha, \beta > 0 \quad \text{and} \quad p > 1. \tag{1.1}$$

Here $\| \cdot \|_{\mathcal{M}}$ denotes the Radon norm of a measure. The functional (1.1) is suitable to be used as a regulariser in the context of variational non-smooth regularisation in imaging applications. We study the properties of (1.1), its regularising mechanism for different values of $p$ and apply it successfully to image denoising.

### 1.1 Context

After the introduction of the total variation (TV) for image reconstruction purposes [38], the use of non-smooth regularisers has become increasingly popular during the last decades (cf. [7]). They are typically used in the context of variational regularisation, where the reconstructed image is obtained as a solution of a minimisation problem of the type:

$$\min_u \frac{1}{s} \|f - Tu\|^s_{\mathrm{L}^s(\Omega)} + \Psi(u). \tag{1.2}$$

The *regulariser* is denoted here by $\Psi$. We assume that the data $f$, defined on an open, bounded and connected domain $\Omega \subset \mathbb{R}^2$, have been corrupted through a bounded, linear

✉ Evangelos Papoutsellis
epapoutsellis@gmail.com; ep374@cam.ac.uk

Martin Burger
martin.burger@wwu.de

Konstantinos Papafitsoros
papafitsoros@hu-berlin.de

Carola-Bibiane Schönlieb
cbs31@cam.ac.uk

[1] Institute for Computational and Applied Mathematics, University of Münster, Münster, Germany

[2] Institute for Mathematics, Humboldt University of Berlin, Berlin, Germany

[3] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

operator $T$ and additive (random) noise. Different values of $s$ can be considered for the first term of (1.2), the *fidelity term*. For example, models incorporating an $L^2$ fidelity term (resp. $L^1$) have been shown to be efficient for the restoration of images corrupted by Gaussian noise (resp. impulse noise). Of course, other types of noise can also be considered and in those cases the form of the fidelity term is adjusted accordingly. Typically, one or more parameters within $\Psi$ balance the strength of regularisation against the fidelity term in the minimisation (1.2).

The advantage of using non-smooth regularisers is that the regularised images have sharp edges (discontinuities). For instance, it is a well-known fact that TV regularisation promotes piecewise constant reconstructions, thus preserving discontinuities. However, this also leads to blocky-like artefacts in the reconstructed image, an effect known as *staircasing*. Recall at this point that for two-dimensional images $u \in L^1(\Omega)$, the definition of the total variation functional reads

$$\mathrm{TV}(u) := \sup \left\{ \int_\Omega u \operatorname{div}\phi \, \mathrm{d}x : \phi \in C_c^\infty(\Omega, \mathbb{R}^2), \ \|\phi\|_\infty \leq 1 \right\}. \tag{1.3}$$

The total variation uses only first-order derivative information in the regularisation process. This can be seen from the fact that for $\mathrm{TV}(u) < \infty$ the distributional derivative $Du$ is a finite Radon measure and $\mathrm{TV}(u) = \|Du\|_{\mathcal{M}}$. Moreover if $u \in W^{1,1}(\Omega)$ then $\mathrm{TV}(u) = \int_\Omega |\nabla u| \, \mathrm{d}x$, i.e. the total variation is the $L^1$ norm of the gradient of $u$. Higher-order extensions of the total variation functional are widely explored in the literature e.g. [4,5,9,11,12,27,29,30,34]. The incorporation of second-order derivatives is shown to reduce or even eliminate the staircasing effect. The most successful regulariser of this kind is the second-order total generalised variation (TGV) introduced by Bredies et al. [5]. Its definition reads

$$\mathrm{TGV}^2_{\alpha,\beta}(u) := \min_{w \in BD(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}}. \tag{1.4}$$

Here $\alpha$, $\beta$ are positive parameters and $BD(\Omega)$ is the space of functions of bounded deformation, i.e. the space of all $L^1(\Omega)$ functions $w$, whose symmetrised distributional derivative $\mathcal{E}w$ is a finite Radon measure. This is a less regular space than the usual space of functions of bounded variation $BV(\Omega)$ for which the full gradient $Du$ is required to be a finite Radon measure. Note that if the variable $w$ in the definition (1.4) is forced to be the gradient of another function then we obtain the classical infimal convolution regulariser of Chambolle–Lions [9]. In that sense TGV can be seen as a particular instance of infimal convolution, optimally balancing first and second-order information.

In the discrete formulation of TGV (as well as for TV) the Radon norm is interpreted as an $L^1$ norm. The motivation for the current and the follow-up paper [8] is to explore the capabilities of $L^p$ norms within first-order regularisation functionals designed for image processing purposes. The use of $L^p$ norms for $p > 1$ has been exploited in different contexts—infinity and $p$-Laplacian (cf. e.g. [16] and [26] respectively).

## 1.2 Our Contribution

Comparing the definition (1.1) with the definition of TGV in (1.4), we see that the Radon norm of the symmetrised gradient of $w$ has been substituted by the $L^p$ norm of $w$, thus reducing the order of regularisation. Up to our knowledge, this is the first paper that provides a thorough analysis of TV–$L^p$ infimal convolution models (1.1) in this generality. We show that the minimisation in (1.1) is well-defined and that $\mathrm{TVL}^p_{\alpha,\beta}(u) < \infty$ if and only if $\mathrm{TV}(u) < \infty$. Hence $\mathrm{TVL}^p_{\alpha,\beta}$ regularised images belong to $BV(\Omega)$ as desired.

In order to get more insight in the regularising mechanism of the $\mathrm{TVL}^p_{\alpha,\beta}$ functional we provide a detailed and rigorous analysis of its one-dimensional version of the corresponding $L^2$ fidelity denoising problem

$$\min_{u \in BV(\Omega)} \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \mathrm{TVL}^p_{\alpha,\beta}(u). \tag{1.5}$$

For the denoising problem (1.5) with $p = 2$ we also compute exact solutions for simple one-dimensional data. We show that the obtained solutions are piecewise smooth, in contrast to TV (piecewise constant) and TGV (piecewise affine) solutions. Moreover, we show that for $p = 2$, the 2-homogeneous analogue of the functional (1.1)

$$\mathrm{TVL}^{2-hom}_{\alpha,\beta}(u) = \min_{w \in L^2(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \frac{\beta}{2}\|w\|^2_{L^2(\Omega)}, \tag{1.6}$$

is equivalent to a variant of *Huber* TV [24], with the functional (1.6) having a close connection to (1.1) itself. Huber total variation is a smooth approximation of total variation and even though it has been widely used in the imaging and inverse problems community, it has not been analysed adequately. Hence, as a by-product of our analysis, we compute exact solutions of the one-dimensional Huber TV denoising problem. An analogous connection of the $\mathrm{TVL}^p_{\alpha,\beta}$ functional with a generalised Huber TV regularisation is also established for general $p$.

We proceed with exhaustive numerical experiments focusing on (1.5). Our analysis is confirmed by the fact that the analytical results coincide with the numerical ones. Furthermore, we observe that even though a first-order regularisation

functional is used, we are capable of eliminating the staircasing effect, similar to Huber TV. By using the *Bregman iteration* version of our method [32], we are also able to enhance the contrast of the reconstructed images, obtaining results very similar in quality to the TGV ones. We observe numerically that high values of $p$ promote almost affine structures similar to second-order regularisation methods. We shed more light of this behaviour in the follow-up paper [8] where we study in depth the case $p = \infty$. Let us finally note that we also consider a modified version of the functional (1.1) where $w$ is restricted to be a gradient of another function leading to the more classical infimal convolution setting. Even though, this modified model is not so successful in staircasing reduction, it is effective in decomposing an image into piecewise constant and smooth parts.

### 1.3 Organisation of the Paper

After the introduction we proceed with the introduction of our model in Sect. 2. We prove the well-posedness of (1.1), we provide an equivalent definition and we prove its Lipschitz equivalence with the TV seminorm. We finish this section with a well-posedness result of the corresponding $\mathrm{TVL}_{\alpha,\beta}^p$ regularisation problem using standard tools.

In Sect. 3 we establish a link between the $\mathrm{TVL}_{\alpha,\beta}^p$ functional and its $p$-homogeneous analogue (using the $p$-th power of $\| \cdot \|_{\mathrm{L}^p(\Omega)}$). The $p$-homogeneous functional (for $p = 2$) is further shown to be equivalent to Huber total variation, while analogous results are obtained for $p \neq 2$.

We study the corresponding one-dimensional model in Sect. 4 focusing on the $\mathrm{L}^2$ fidelity denoising case. More specifically, after deriving the optimality conditions using Fenchel–Rockafellar duality in Sect. 4.1, we explore the structure of solutions in Sect. 4.2. In Sect. 4.3 we compute exact solutions for the case $p = 2$, considering a simple step function as data.

In Sect. 5 we present a variant of our model suitable for image decomposition purposes, i.e. geometric decomposition into piecewise constant and smooth structures.

Section 6 focuses on numerical experiments. Confirmation of the obtained one-dimensional analytical results is done in Sect. 6.2, while two-dimensional denoising experiments are performed in Sect. 6.3 using the split Bregman method. There, we show that our approach can lead to elimination of the staircasing effect and we also show that by using a Bregmanised version we can also enhance the contrast, achieving results very close to TGV, a method considered state of the art in the context of variational regularisation. We finish the section with some image decomposition examples and we summarise our results in Sect. 7.

In the appendix, we remind the reader of some basic facts from the theory of Radon measures and BV functions.

## 2 Basic Properties of the $\mathrm{TVL}_{\alpha,\beta}^p$ Functional

In this section, we introduce the $\mathrm{TVL}_{\alpha,\beta}^p$ functional (1.1) as well as some of its main properties. For $\alpha, \beta > 0$ and $1 < p \leq \infty$, we define $\mathrm{TVL}_{\alpha,\beta}^p : \mathrm{L}^1(\Omega) \to \overline{\mathbb{R}}$ (where $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$) as follows:

$$\mathrm{TVL}_{\alpha,\beta}^p(u) := \min_{w \in \mathrm{L}^p(\Omega)} \alpha \, \|Du - w\|_{\mathcal{M}} + \beta \, \|w\|_{\mathrm{L}^p(\Omega)}.$$

While in the present paper we mainly focus on the finite $p$ case, the results of this section are stated and proved for $p = \infty$ as well, since the proofs are similar.

The next proposition asserts that the minimisation in (1.1) is indeed well-defined. We omit the proof, which is based on standard coercivity and weak lower semicontinuity techniques:

**Proposition 2.1** *Let* $u \in \mathrm{BV}(\Omega)$ *with* $1 < p \leq \infty$ *and* $\alpha, \beta > 0$. *Then the minimum in the definition* (1.1) *is attained.*

Another useful formulation of the definition (1.1) is the *dual* formulation:

$$\mathrm{TVL}_{\alpha,\beta}^p(u) = \sup \left\{ \int_\Omega u \operatorname{div} \phi \, \mathrm{d}x : \phi \in C_c^1(\Omega), \right.$$
$$\left. \|\phi\|_\infty \leq \alpha, \ \|\phi\|_{\mathrm{L}^q(\Omega)} \leq \beta \right\}, \tag{2.1}$$

where $q$ denotes here the conjugate exponent of $p$, see (8.4). The following proposition shows that the two expressions coincide indeed. Recall first that for a functional $F : X \to \overline{\mathbb{R}}$ the effective domain is defined as $\mathrm{dom} F = \{x \in X : F(x) < \infty\}$, while the indicator and characteristic functions of $A \subseteq X$ are defined as

$$\mathbb{I}_A(x) = \begin{cases} 0, & \text{if } x \in A, \\ \infty, & \text{if } x \notin A, \end{cases} \quad \text{and} \quad \mathcal{X}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A, \end{cases}$$

respectively. As usual, we denote by $\langle \cdot, \cdot \rangle$ the duality product of $X$ and its dual $X^*$. Finally, recall that the convex conjugate $F^* : X^* \to \overline{\mathbb{R}}$ of $F$ is defined as $F^*(x^*) = \sup_{x \in X} \langle x^*, x \rangle - F(x)$.

**Proposition 2.2** *Let* $u \in \mathrm{BV}(\Omega)$ *and* $1 < p \leq \infty$ *then*

$$\min_{w \in \mathrm{L}^p(\Omega)} \alpha \, \|Du - w\|_{\mathcal{M}} + \beta \, \|w\|_{\mathrm{L}^p(\Omega)}$$
$$= \sup \left\{ \int_\Omega u \operatorname{div} \phi \, \mathrm{d}x : \phi \in C_c^1(\Omega), \right.$$
$$\left. \|\phi\|_\infty \leq \alpha, \ \|\phi\|_{\mathrm{L}^q(\Omega)} \leq \beta \right\}.$$

*Proof* First notice that in (2.1), we can replace $C_c^1(\Omega)$ by $C_0^1(\Omega)$, since $\overline{C_c^1(\Omega)} = C_0^1(\Omega)$ with the closure taken with respect to the $C^1$ norm, i.e. $\|\phi\| = \max\left\{\|\phi\|_\infty, \|\nabla\phi\|_\infty\right\}$. We define

$$X = C_0^1(\Omega),$$

$$F_1 : X \to \overline{\mathbb{R}}, \text{ with } F_1(\phi) = \mathbb{I}_{\{\|\cdot\|_{L^q(\Omega)} \leq \beta\}}(\phi),$$

$$F_2 : X \to \overline{\mathbb{R}}, \text{ with } F_2(\phi) = \mathbb{I}_{\{\|\cdot\|_\infty \leq \alpha\}}(\phi) - \int_\Omega u \operatorname{div}\phi \, dx.$$

Then, we can rewrite (2.1) as

$$\mathrm{TVL}_{\alpha,\beta}^p(u) = -\inf_{\substack{\phi \in X \\ \|\phi\|_\infty \leq \alpha \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \left\{ -\int_\Omega u \operatorname{div}\phi \, dx \right\}$$

$$= -\inf_{\phi \in X} \{F_1(\phi) + F_2(\phi)\}.$$

We can establish the following relation

$$\inf_{\phi \in X} \{F_1(\phi) + F_2(\phi)\} + \min_{w \in X^*} \left\{F_1^*(w) + F_2^*(w)\right\} = 0,$$

i.e. the absence of duality gap between the primal and the dual problems [15, Chapter III], provided that the set

$$\bigcup_{\lambda \geq 0} \lambda \left(\operatorname{dom} F_2 - \operatorname{dom} F_1\right),$$

is a closed vector space [2]. This is indeed true since on one hand we have

$$\bigcup_{\lambda \geq 0} \lambda \left(\operatorname{dom} F_2 - \operatorname{dom} F_1\right) \subset X,$$

and on the other hand, for every $\phi \in X$, we can write $\phi = \lambda(\lambda^{-1}\phi - 0)$ with $\left\|\lambda^{-1}\phi\right\|_\infty \leq \alpha$ and $0 \in \operatorname{dom} F_1$. Since $u \in \mathrm{BV}(\Omega)$, then $\mathrm{TVL}_{\alpha,\beta}^p$ is finite, see also Proposition 2.4. Hence, $F_1^*(w) < \infty$, $F_2^*(w) < \infty$ and

$$F_1^*(-w) = \sup_{\substack{\phi \in C_0^1(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \langle w, \phi \rangle$$

$$= \sup_{\substack{\phi \in L^q(\Omega) \\ \|\phi\|_{L^q(\Omega)} \leq \beta}} \langle w, \phi \rangle = \beta \|w\|_{L^p(\Omega)},$$

where we have used the fact that, with a density argument, the function $w : C_0^1 \to \mathbb{R}$ can be extended in the whole $L^q(\Omega)$ as a bounded, linear functional and using the Riesz representation theorem, we deduce that it is actually an $L^p$ function. Similarly, we have

$$F_2^*(w) = \sup_{\substack{\phi \in C_0^1(\Omega) \\ \|\phi\|_\infty \leq \alpha}} \langle w, \phi \rangle + \langle u, \operatorname{div}\phi \rangle = \alpha \|Du - w\|_{\mathcal{M}}.$$

Thus the desired equality is proven. $\qquad\square$

*Remark 2.3* Note that using the dual formulation of $\mathrm{TVL}_{\alpha,\beta}^p$ one can easily derive that the functional is lower semicontinuous with respect to the strong $L^1$ topology since it is a pointwise supremum of continuous functions.

The following lemma shows that the $\mathrm{TVL}_{\alpha,\beta}^p$ functional is Lipschitz equivalent to the total variation.

**Proposition 2.4** *Let $u \in L^1(\Omega)$ and $1 < p \leq \infty$. Then $\mathrm{TVL}_{\alpha,\beta}^p(u) < \infty$ if and only if $u \in \mathrm{BV}(\Omega)$. Moreover there exist constants $C_1 = \alpha$ and $C_2 = (C\tilde{C})^{-1}$, where $C = \max\left(1, |\Omega|^{\frac{1}{q}}\right)$ and $\tilde{C} = \max\left(\frac{1}{\alpha}, \frac{1}{\beta}\right)$ such that*

$$C_2 \|Du\|_{\mathcal{M}} \leq \mathrm{TVL}_{\alpha,\beta}^p(u) \leq C_1 \|Du\|_{\mathcal{M}}, \quad \text{for all } u \in \mathrm{BV}(\Omega). \tag{2.2}$$

*Finally in the special case where*

$$\frac{\beta}{\alpha} \geq |\Omega|^{\frac{1}{q}}, \tag{2.3}$$

*then*

$$\mathrm{TVL}_{\alpha,\beta}^p(u) = \alpha \|Du\|_{\mathcal{M}}, \quad \text{for all } u \in \mathrm{BV}(\Omega). \tag{2.4}$$

*Proof* Let $u \in \mathrm{BV}(\Omega)$. Using the definition (1.1) we have that

$$\mathrm{TVL}_{\alpha,\beta}^p(u) \leq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)},$$

for every $w \in L^p(\Omega)$. Setting $w = 0$ and $C_1 = \alpha$, we obtain

$$\mathrm{TVL}_{\alpha,\beta}^p(u) \leq C_1 \|Du\|_{\mathcal{M}}. \tag{2.5}$$

For the other direction, for any $w \in L^p(\Omega) \subset L^1(\Omega)$, by the triangle inequality we get

$$\|Du\|_{\mathcal{M}} \leq \|Du - w\|_{\mathcal{M}} + \|w\|_{L^1(\Omega)}$$

$$\leq \|Du - w\|_{\mathcal{M}} + |\Omega|^{\frac{1}{q}} \|w\|_{L^p(\Omega)}$$

$$\leq C(\|Du - w\|_{\mathcal{M}} + \|w\|_{L^p(\Omega)}), \tag{2.6}$$

with $C = \max\left(1, |\Omega|^{\frac{1}{q}}\right)$. By setting $C_2$ as in the statement of the lemma we obtain

$$C_2 \|Du\|_{\mathcal{M}} \leq \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)},$$

which, by minimising over $w$, yields the left-hand side inequality in (2.2).

Finally, observe simply that if (2.3) holds then from (2.6) we get

$$\alpha \, \|Du\|_{\mathcal{M}} \leq \alpha \, \|Du - w\|_{\mathcal{M}} + \beta \, \|w\|_{\mathrm{L}^p(\Omega)}$$

and thus minimising again over $w$ and combining (2.5) we get (2.4).

Notice that when (2.3) holds then the above proposition implies that $w = 0$ is an admissible solution to the definition of $\mathrm{TVL}^p_{\alpha,\beta}(u)$, i.e.

$$0 \in \underset{w \in \mathrm{L}^p(\Omega)}{\mathrm{argmin}} \; \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{\mathrm{L}^p(\Omega)},$$

for all $u \in \mathrm{BV}(\Omega)$.

However, in general we cannot prove that this solution is unique.

Having shown the basic properties of the $\mathrm{TVL}^p_{\alpha,\beta}$ functional, we can use it as a regulariser for variational imaging problems of the type

$$\min_{u \in \mathrm{L}^s(\Omega) \cap \mathrm{BV}(\Omega)} \frac{1}{s} \|f - Tu\|_{\mathrm{L}^s(\Omega)}^s + \mathrm{TVL}^p_{\alpha,\beta}(u), \quad s \geq 1, \tag{2.7}$$

where $T : \mathrm{L}^s(\Omega) \to \mathrm{L}^s(\Omega)$ is a bounded, linear operator and $f \in \mathrm{L}^s(\Omega)$. We conclude our analysis with existence and uniqueness results for the minimisation problem (2.7).

**Theorem 2.5** *Let* $1 < p \leq \infty$ *and* $f \in \mathrm{L}^s(\Omega)$. *If* $T(\mathcal{X}_\Omega) \neq 0$ *then there exists a solution* $u \in \mathrm{L}^s(\Omega) \cap \mathrm{BV}(\Omega)$ *for the problem* (2.7). *If* $s > 1$ *and* $T$ *is injective then the solution is unique.*

*Proof* The proof is a straightforward application of the direct method of calculus of variations. We simply take advantage of the inequality (2.2) and the compactness theorem in $\mathrm{BV}(\Omega)$, see Appendix, along with the lower semicontinuity property of $\mathrm{TVL}^p_{\alpha,\beta}$. We also refer the reader to the corresponding proofs in [34,39]. □

Since we are mainly interested in studying the regularising properties of $\mathrm{TVL}^p_{\alpha,\beta}$, in what follows we focus on the case where $s = 2$ and $T$ is the identity function (denoising task) where rigorous analysis can be carried out. From now on, we also focus on the case where $p$ is finite, as the case $p = \infty$ is studied in the follow-up paper [8]. We thus define the following problem

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2} \|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \mathrm{TVL}^p_{\alpha,\beta}(u),$$

or equivalently

$$\min_{\substack{u \in \mathrm{BV}(\Omega) \\ w \in \mathrm{L}^p(\Omega)}} \frac{1}{2} \|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{\mathrm{L}^p(\Omega)}. \tag{$\mathcal{P}$}$$

## 3 The $p$-Homogeneous Analogue and Relation to Huber TV

Before we proceed to a detailed analysis of the one-dimensional version of ($\mathcal{P}$), in this section we consider its $p$-homogeneous analogue

$$\min_{\substack{u \in \mathrm{BV}(\Omega) \\ w \in \mathrm{L}^p(\Omega)}} \frac{1}{2} \|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{\mathrm{L}^p(\Omega)}^p, \quad 1 < p < \infty. \tag{$\mathcal{P}_{p-hom}$}$$

We show in Proposition 3.2 that there is a strong connection between the models ($\mathcal{P}$) and ($\mathcal{P}_{p-hom}$). The reason for the introduction of ($\mathcal{P}_{p-hom}$) is that, in certain cases, it is technically easier to derive exact solutions for ($\mathcal{P}_{p-hom}$) rather than for ($\mathcal{P}$) straightforwardly, see Sect. 4.3. Moreover, we can guarantee the uniqueness of the optimal $w^*$ in ($\mathcal{P}_{p-hom}$), since

$$w^* = \underset{w \in \mathrm{L}^p(\Omega)}{\mathrm{argmin}} \; \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta}{p} \|w\|_{\mathrm{L}^p(\Omega)}^p,$$

and thus $w^*$ is unique as a minimiser of a strictly convex functional. The next proposition states that, unless $f$ is a constant function then the optimal $w^*$ in ($\mathcal{P}_{p-hom}$) cannot be zero but nonetheless converges to zero as $\beta \to \infty$. In essence, this means that one cannot obtain TV type solutions with the $p$-homogeneous model.

**Proposition 3.1** *Let* $1 < p < \infty$, $f \in \mathrm{L}^2(\Omega)$ *and let* $(w^*, u^*)$ *be an optimal solution pair of the $p$-homogeneous problem* ($\mathcal{P}_{p-hom}$). *Then* $w^* = 0$ *if and only if* $f$ *is a constant function. For general data* $f$, *we have that* $w^* \to 0$ *in* $\mathrm{L}^p(\Omega)$ *when* $\beta \to \infty$.

*Proof* It follows immediately that if $f$ is constant then $(0, f)$ is the optimal pair for ($\mathcal{P}_{p-hom}$). Suppose that $(w^*, u^*)$ solve ($\mathcal{P}_{p-hom}$). Notice that in this case we also have

$$u^* = \underset{u \in \mathrm{BV}(\Omega)}{\mathrm{argmin}} \; \frac{1}{2} \|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|Du - w^*\|_{\mathcal{M}}. \tag{3.1}$$

Suppose now that $w^* = 0$. Then (3.1) becomes

$$u^* = \underset{u \in \mathrm{BV}(\Omega)}{\mathrm{argmin}} \; \frac{1}{2} \|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}}. \tag{3.2}$$

Furthermore, since $(0, u^*)$ solve ($\mathcal{P}_{p-hom}$), then for every $u \in C_c^\infty(\Omega)$ and $\epsilon > 0$, the pair $(\epsilon \nabla u, u^* + \epsilon u)$ is suboptimal for ($\mathcal{P}_{p-hom}$), i.e.

$$\frac{1}{2} \|f - u^*\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|Du^*\|_{\mathcal{M}} \leq \frac{1}{2} \|(f - u^*) - \epsilon u\|_{\mathrm{L}^2(\Omega)}^2$$
$$+ \alpha \|D(u^* + \epsilon u) - \epsilon \nabla u\|_{\mathcal{M}} + \frac{\beta}{p} \|\epsilon \nabla u\|_{\mathrm{L}^p(\Omega)}^p,$$

from which we take

$$
\frac{1}{2}\|f - u^*\|_{\mathrm{L}^2(\Omega)}^2 \le \frac{1}{2}\|(f - u^*) - \epsilon u\|_{\mathrm{L}^2(\Omega)}^2
$$
$$
+ \frac{\beta}{p}\|\epsilon \nabla u\|_{\mathrm{L}^p(\Omega)}^p \Leftrightarrow
$$
$$
0 \le \frac{\epsilon^2}{2}\|u\|_{\mathrm{L}^2(\Omega)}^2 - \epsilon \int_\Omega (f - u^*)u \, \mathrm{d}x
$$
$$
+ \frac{\beta \epsilon^p}{p}\|\nabla u\|_{\mathrm{L}^p(\Omega)}^p.
$$

By dividing the last inequality by $\epsilon$ and taking the limit $\epsilon \to 0$ we have that $\int_\Omega (f - u^*)u \, \mathrm{d}x \le 0$. By considering the analogous perturbations $u^* - \epsilon u$, we obtain similarly that $\int_\Omega (f - u^*)u \, \mathrm{d}x \ge 0$ and thus

$$
\int_\Omega (f - u^*)u \, \mathrm{d}x = 0, \quad \forall u \in C_c^\infty(\Omega).
$$

Hence $u^* = f$ and by taking the optimality condition of (3.2) we get that $0 \in \partial \|D(\cdot)\|_{\mathcal{M}}(f)$, which implies that $Df = 0$, i.e. $f$ is a constant function.

For the last part of the proposition, (supposing $f \ne 0$), simply observe that for every $u \in \mathrm{BV}(\Omega)$ and $w \in \mathrm{L}^p(\Omega)$ we have that

$$
\frac{1}{2}\|f - u^*\|_{\mathrm{L}^2(\Omega)}^2 + \alpha\|Du^* - w^*\|_{\mathcal{M}} + \frac{\beta}{p}\|w^*\|_{\mathrm{L}^p(\Omega)}^p
$$
$$
\le \frac{1}{2}\|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha\|Du - w\|_{\mathcal{M}} + \frac{\beta}{p}\|w\|_{\mathrm{L}^p(\Omega)}^p,
$$

and by setting $u = w = 0$, we obtain

$$
\frac{1}{p}\|w^*\|_{\mathrm{L}^p(\Omega)}^p \le \frac{1}{2\beta}\|f\|_{\mathrm{L}^2(\Omega)}^2,
$$

and thus $\|w^*\|_{\mathrm{L}^p(\Omega)}^p \to 0$ when $\beta \to \infty$. □

We further establish a connection between the 1-homogeneous $(\mathcal{P})$ and the $p$-homogeneous model $(\mathcal{P}_{p-hom})$:

**Proposition 3.2** *Let* $1 < p < \infty$ *and* $f \in \mathrm{L}^2(\Omega)$ *not a constant. A pair* $(w^*, u^*)$ *is a solution of* $(\mathcal{P}_{p-hom})$ *with parameters* $(\alpha, \beta_{p-hom})$ *if and only if it is also a solution of* $(\mathcal{P})$ *with parameters* $(\alpha, \beta_{1-hom})$ *where* $\beta_{1-hom} = \beta_{p-hom}\|w^*\|_{\mathrm{L}^p(\Omega)}^{p-1}$.

*Proof* Since $f$ is not a constant by the previous proposition we have that $w^* \ne 0$. Note that for an arbitrary function $u \in \mathrm{BV}(\Omega)$:

$$
w^* \in \operatorname*{argmin}_{w \in \mathrm{L}^p(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \frac{\beta_{p-hom}}{p}\|w\|_{\mathrm{L}^p(\Omega)}^p \Leftrightarrow
$$
$$
0 \in \alpha\partial\|Du - \cdot\|_{\mathcal{M}}(w^*) + \beta_{p-hom}|w^*|^{p-2}w^* \Leftrightarrow
$$
$$
0 \in \alpha\partial\|Du - \cdot\|_{\mathcal{M}}(w^*) + \frac{\beta_{1-hom}}{\|w^*\|_{\mathrm{L}^p(\Omega)}^{p-1}}|w^*|^{p-2}w^* \Leftrightarrow
$$
$$
w^* \in \operatorname*{argmin}_{w \in \mathrm{L}^p(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \beta_{1-hom}\|w\|_{\mathrm{L}^p(\Omega)}.
$$

This means that $w^*$ is an admissible solution for both problems $(\mathcal{P})$ and $(\mathcal{P}_{p-hom})$, with the corresponding set of parameters $(\alpha, \beta_{1-hom})$ and $(\alpha, \beta_{p-hom})$, respectively. The fact that the same holds for $u^*$ as well, comes from the fact that in both problems we have

$$
u^* \in \operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} \frac{1}{2}\|f - u\|_{\mathrm{L}^2(\Omega)}^2 + \alpha\|Du - w^*\|_{\mathcal{M}}.
$$

Finally, it turns out that for $p = 2$, problem $(\mathcal{P}_{p-hom})$ is essentially equivalent to the widely used Huber total variation regularisation, [24]. In fact we can show that for $1 < p < \infty$ $(\mathcal{P}_{p-hom})$ is equivalent to a generalised Huber total variation regularisation, see also [23]. This is proved in the next proposition.

**Proposition 3.3** *Let* $1 < p < \infty$ *and consider the functional* $\mathrm{TVL}_{\alpha,\beta}^{p-hom} : \mathrm{BV}(\Omega) \to \mathbb{R}$ *with*

$$
\mathrm{TVL}_{\alpha,\beta}^{p-hom}(u) = \min_{w \in \mathrm{L}^p(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \frac{\beta}{p}\|w\|_{\mathrm{L}^p(\Omega)}^p. \tag{3.3}
$$

*Then*

$$
\mathrm{TVL}_{\alpha,\beta}^{p-hom}(u) = \int_\Omega \varphi_p(\nabla u) \, \mathrm{d}x + \alpha|D^s u|(\Omega),
$$

*where* $\varphi_p : \mathbb{R}^d \to \mathbb{R}$ *with*

$$
\varphi_p(x) = \begin{cases} \alpha|x| - \left(1 - \frac{1}{p}\right)\frac{\alpha}{\lambda^{\frac{1}{p-1}}}, & |x| \ge \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ \frac{\beta}{p}|x|^p, & |x| \le \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases} \quad \text{where } \lambda := \frac{\beta}{\alpha},
$$

*and* $D^s u$ *denotes the singular part of the measure* $Du$*, cf. Appendix.*

*Proof* We have

$$
\mathrm{TVL}_{\alpha,\beta}^{p-hom}(u) = \min_{w \in \mathrm{L}^p(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \frac{\beta}{p}\|w\|_{\mathrm{L}^p(\Omega)}^p
$$
$$
= \alpha|D^s u|(\Omega) + \alpha \min_{w \in \mathrm{L}^p(\Omega)} \int_\Omega |\nabla u - w|
$$
$$
+ \frac{\beta}{p\alpha}|w|^p \, \mathrm{d}x.
$$

Thus we can focus on the minimisation problem

$$\min_{w \in L^p(\Omega)} \int_\Omega |\nabla u - w| + \frac{\beta}{p\alpha}|w|^p \, dx. \qquad (3.4)$$

Baring in mind that (as it can easily checked) for $c \in \mathbb{R}^d$ and $\lambda > 0$,

$$\operatorname*{argmin}_{y \in \mathbb{R}^d} |c - y| + \frac{\lambda}{p}|y|^p = \begin{cases} \frac{1}{\lambda^{\frac{1}{p-1}}}\frac{c}{|c|} & \text{if } |c| \ge \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ c & \text{if } |c| < \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases}$$

and

$$\min_{y \in \mathbb{R}^d} |c - y| + \frac{\lambda}{p}|y|^p = \begin{cases} |c| - \left(1 - \frac{1}{p}\right)\frac{1}{\lambda^{\frac{1}{p-1}}} & \text{if } |c| \ge \frac{1}{\lambda^{\frac{1}{p-1}}}, \\ \frac{\lambda}{p}|c|^p & \text{if } |c| < \frac{1}{\lambda^{\frac{1}{p-1}}}, \end{cases}$$

it is straightforwardly verified setting $\lambda = \beta/\alpha$ that the function

$$w^* = \lambda^{-\frac{1}{p-1}}\frac{\nabla u}{|\nabla u|}\mathcal{X}_{\left\{|\nabla u| \ge \lambda^{-\frac{1}{p-1}}\right\}} + \nabla u \mathcal{X}_{\left\{|\nabla u| < \lambda^{-\frac{1}{p-1}}\right\}},$$

belongs to $L^\infty(\Omega) \subset L^p(\Omega)$ and solves (3.4) with optimal value equal to $\frac{1}{\alpha}\int_\Omega \varphi_p(\nabla u) \, dx$.   □

Note that in the special case $p = 2$ we recover the classical Huber total variation regularisation since

$$\varphi_2(x) = \begin{cases} \alpha|x| - \frac{\alpha^2}{2\beta} & |x| \ge \frac{\alpha}{\beta}, \\ \frac{\beta}{2}|x|^2, & |x| \le \frac{\alpha}{\beta}, \end{cases}$$

i.e. in that case we have quadratic penalisation for small gradients ($p$–power penalisation for the general $1 < p < \infty$) and linear penalisation for large gradients.

For the reader's convenience, in Fig. 1 we have plotted some of the functions $\varphi_p$ in order to illustrate how their form changes when their parameters vary. Note for instance in Fig. 1a how $\phi_2$ is converging to an absolute type function when $\beta$ is getting large, i.e. approaching a total variation regularisation. This can also be seen from Proposition 3.1 where the optimal variable $w$ is converging to 0 when $\beta \to \infty$. On the other hand when $p$ is getting large, Fig. 1b, small gradients are essentially not penalised at all, allowing the gradient to be almost constant, equal to its maximum value, leading to piecewise affine structures. We refer to some of the numerical examples in Sect. 6.2 and also the second part of this paper [8] where the case $p = \infty$ is examined in detail.
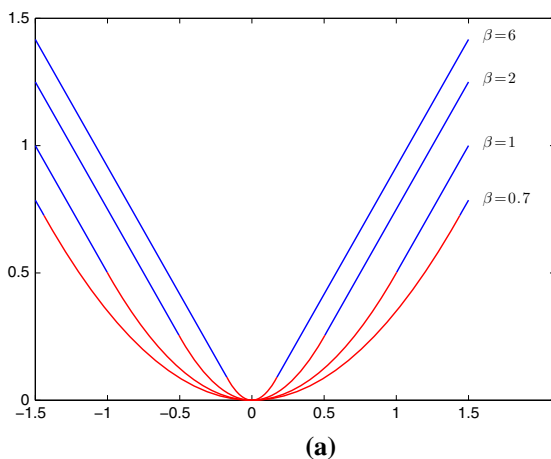
## 4 The One-Dimensional Case

In order to get more insights into the structure of solutions of the problem ($\mathcal{P}$), in this section we study its one-dimensional version. As above, we focus on the finite $p$ case, i.e. $1 < p < \infty$. The case $p = \infty$ leads to several additional complications and will be subject of a forthcoming paper [8]. For this section $\Omega \subset \mathbb{R}$ is an open and bounded interval, i.e. $\Omega = (a, b)$. Our analysis follows closely the ones in [6] and [33] where the one dimensional $L^1$–TGV and $L^2$–TGV problems are studied, respectively.

### 4.1 Optimality Conditions

In this section, we derive the optimality conditions for the one-dimensional problem ($\mathcal{P}$). We initially start our analysis by defining the predual problem ($\mathcal{P}^*$), proving existence and uniqueness for its solutions. We employ again the Fenchel–Rockafellar duality theory in order to find a connection between the solutions of the predual and primal problems.
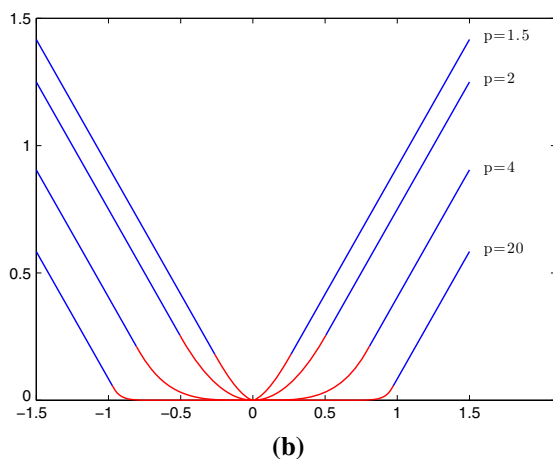


**(a)**      **(b)**

**Fig. 1** Illustration of the forms of the Huber-type functions $\varphi_p$ of Proposition 3.3. Their linear and $p$–power parts are plotted with *blue* and *red* colour, respectively. **a** Huber functions $\varphi_2$ with fixed $p = 2$, $\alpha = 1$ and varying $\beta$. **b** Generalised Huber functions $\varphi_p$ with fixed $\alpha = 1$, $\beta = 2$ and varying $p$ (Color figure online)

We define the predual problem $(\mathcal{P}^*)$ as

$$-\inf\left\{\int_\Omega f\phi'\mathrm{d}x + \frac{1}{2}\int_\Omega (\phi')^2\mathrm{d}x : \phi \in \mathrm{H}_0^1(\Omega), \|\phi\|_{\mathrm{L}^q(\Omega)} \le \beta, \|\phi\|_\infty \le \alpha\right\},$$
$$(\mathcal{P}^*)$$

where as always $q$ is the conjugate exponent of $p$. Existence and uniqueness for the solutions of $(\mathcal{P}^*)$ can be verified by standard arguments:

**Proposition 4.1** *For $f \in \mathrm{L}^2(\Omega)$, the predual problem $(\mathcal{P}^*)$ admits a unique solution in $\mathrm{H}_0^1(\Omega)$.*

The next proposition justifies the term predual for the problem $(\mathcal{P}^*)$.

**Proposition 4.2** *The dual problem of $(\mathcal{P}^*)$ is equivalent to the problem $(\mathcal{P})$ in the sense that $(w, u)$ is a solution of the dual of $(\mathcal{P}^*)$ if and only if $(w, u) \in \mathrm{L}^p(\Omega) \times \mathrm{BV}(\Omega)$ and solves $(\mathcal{P})$.*

*Proof* Observe that we can also write down the predual problem $(\mathcal{P}^*)$ using the following equivalent formulation:

$$-\inf_{(\phi,\xi)\in X} F_1(\phi,\xi) + F_2(K(\phi,\xi)), \qquad (4.1)$$

where $X = \mathrm{H}_0^1(\Omega) \times \mathrm{H}_0^1(\Omega)$, $Y = \mathrm{H}_0^1(\Omega) \times \mathrm{L}^2(\Omega)$ and

$$K : X \to Y, K(\phi,\xi) = (\xi - \phi, \xi'),$$
$$F_1 : X \to \overline{\mathbb{R}}, \text{ with } F_1(\phi,\xi) = \mathbb{I}_{\{\|\cdot\|_{\mathrm{L}^q(\Omega)}\le\beta\}}(\phi)$$
$$\qquad + \mathbb{I}_{\{\|\cdot\|_\infty\le\alpha\}}(\xi),$$
$$F_2 : Y \to \overline{\mathbb{R}}, \text{ with } F_2(\phi,\psi) = \mathbb{I}_{\{0\}}(\phi) + \int_\Omega f\psi\,\mathrm{d}x$$
$$\qquad + \frac{1}{2}\int_\Omega \psi^2\mathrm{d}x. \qquad (4.2)$$

We denote the infimum in $(\mathcal{P}^*)$ by $\inf \mathcal{P}^*$. Then, it is immediate that

$$-\inf \mathcal{P}^* = -\inf_{(\phi,\xi)\in X} F_1(\phi,\xi) + F_2(K(\phi,\xi)).$$

The dual problem of (4.1), see [15], is defined as

$$\min_{(w,u)\in Y^*} F_1^*(-K^\star(w,u)) + F_2^*(w,u), \qquad (4.3)$$

where $K^\star$ here denotes the adjoint of $K$. Let $(\sigma, \tau)$ be elements of $X^* = \mathrm{H}_0^1(\Omega)^* \times \mathrm{H}_0^1(\Omega)^*$. The convex conjugate of $F_1$ can be written as

$$F_1^*(\sigma,\tau) = \sup_{\substack{(\phi,\xi)\in X \\ \|\phi\|_{\mathrm{L}^q(\Omega)}\le\beta\|\xi\|_\infty\le\alpha}} \langle\sigma,\phi\rangle + \langle\tau,\xi\rangle$$
$$= \beta \sup_{\substack{\phi\in\mathrm{H}_0^1(\Omega) \\ \|\phi\|_{\mathrm{L}^q(\Omega)}\le1}} \langle\sigma,\phi\rangle + \alpha \sup_{\substack{\xi\in\mathrm{H}_0^1(\Omega) \\ \|\xi\|_\infty\le1}} \langle\tau,\xi\rangle. \qquad (4.4)$$

By standard density arguments and using the Riesz representation theorem we have

$$F_1^*(\sigma,\tau) = \beta \sup_{\substack{\phi\in C_c^\infty(\Omega) \\ \|\phi\|_{\mathrm{L}^q(\Omega)}\le1}} \langle\sigma,\phi\rangle + \alpha \sup_{\substack{\xi\in C_c^\infty(\Omega) \\ \|\xi\|_\infty\le1}} \langle\tau,\xi\rangle$$
$$= \beta\|\sigma\|_{\mathrm{L}^p(\Omega)} + \alpha\|\tau\|_{\mathcal{M}}. \qquad (4.5)$$

Moreover, we have

$$\left\langle -K^\star(w,u), (\phi,\xi) \right\rangle$$
$$= -\langle(w,u), K(\phi,\xi)\rangle = -\left\langle(w,u),(\xi-\phi,\xi')\right\rangle$$
$$= -\langle w,\xi\rangle + \langle w,\phi\rangle - \left\langle u,\xi'\right\rangle = \langle Du - w,\xi\rangle + \langle w,\phi\rangle.$$

Since $F_1^*(-K^\star(w,u)) < \infty$, $F_2^*(w,u) < \infty$, we obtain that

$$F_1^*(-K^\star(w,u)) = \beta\|w\|_{\mathrm{L}^p(\Omega)} + \alpha\|Du - w\|_{\mathcal{M}}, \qquad (4.6)$$

and

$$F_2^*(w,u) = \sup_{\substack{(\phi,\psi)\in Y \\ \phi=0}} \langle w,\phi\rangle + \langle u,\psi\rangle - \langle f,\psi\rangle - \frac{1}{2}\int_\Omega \psi^2\mathrm{d}x$$
$$= \sup_{\psi\in\mathrm{L}^2(\Omega)} \langle u-f,\psi\rangle - \frac{1}{2}\int_\Omega \psi^2\mathrm{d}x$$
$$:= \left(\frac{1}{2}\|\cdot\|_{\mathrm{L}^2(\Omega)}^2\right)^* (u-f) = \frac{1}{2}\|u-f\|_{\mathrm{L}^2(\Omega)}^2. \qquad (4.7)$$

$\square$

We next verify that we have no duality gap between the two minimisation problems $(\mathcal{P})$ and $(\mathcal{P}^*)$. The proof of the following proposition follows the proof of the corresponding proposition in [6]. We slightly modify it for our case.

**Proposition 4.3** *Let $F_1, F_2, K, X, Y$ be defined as in (4.2). Then*

$$\bigcup_{\lambda\ge0} \lambda(\mathrm{dom}\,F_2 - K(\mathrm{dom}\,F_1)) = Y \qquad (4.8)$$

*and hence it is a closed vector space. Thus [2]*

$$\min_{(\phi,\xi)\in X} F_1(\phi,\xi) + F_2(K(\phi,\xi))$$
$$+ \min_{(w,u)\in Y^*} F_1^*(-K^\star(w,u)) + F_2^*(w,u) = 0. \qquad (4.9)$$

*Proof* Let $(v, \psi) \in Y$ and define $\psi_0(x) = c_1$, where

$$c_1 = \frac{1}{|\Omega|}\int_\Omega \psi(x)\,\mathrm{d}x. \qquad (4.10)$$

Now let $\xi(x) = \int_a^x (\psi_0 - \psi)(y)\,dy$. Since by construction, $\xi' = \psi_0 - \psi \in \mathrm{L}^2(\Omega)$ with $\xi(a) = \xi(b) = 0$, we have that

$\xi \in H_0^1(\Omega)$. Furthermore, let $\phi = -v + \xi \in H_0^1(\Omega)$ and $(\phi, \xi) \in X$ with

$$(v, \psi) = (\xi - \phi, \psi_0 - \xi') = (0, \psi_0) - (\xi - \phi, \xi')$$
$$= (0, \psi_0) - K(\phi, \xi).$$

Choosing appropriately $\lambda > 0$ such that $\left\| \lambda^{-1} \phi \right\|_{L^q(\Omega)} \leq \beta$, $\left\| \lambda^{-1} \xi \right\|_\infty \leq \alpha$, we can write

$$(v, \psi) = \lambda \left( (0, \lambda^{-1} \psi_0) - K(\lambda^{-1} \phi, \lambda^{-1} \xi) \right),$$

with $\mathrm{dom} F_2 = \{0\} \times L^2(\Omega)$ and $\mathrm{dom} F_1 = \{(\phi, \xi) : \|\phi\|_{L^q(\Omega)} \leq \beta, \|\xi\|_\infty \leq \alpha\}$. Since $(v, \psi) \in Y$ were chosen arbitrarily, (4.8) holds. $\qquad \square$

Since there is no duality gap, we can find a relationship between the solutions of $(\mathcal{P}^*)$ and $(\mathcal{P})$ via the following optimality conditions.

**Theorem 4.4** (Optimality conditions) *Let $1 < p \leq \infty$ and $f \in L^2(\Omega)$. A pair $(w, u) \in L^p(\Omega) \times BV(\Omega)$ is a solution of $(\mathcal{P})$ if and only if there exists a function $\phi \in H_0^1(\Omega)$ such that*

$$\phi' = u - f, \qquad (4.11)$$
$$\phi \in \alpha \mathrm{Sgn}(Du - w),$$

*and*

$$\begin{cases} \phi \in \{\tilde{\phi} \in L^q(\Omega) : \|\tilde{\phi}\|_{L^q(\Omega)} \leq \beta\} & \text{if } w = 0, \\ \phi = \beta \dfrac{|w|^{p-2} w}{\|w\|_{L^p(\Omega)}^{(p-1)}} & \text{if } w \neq 0. \end{cases} \qquad (4.12)$$

*Proof* Since there is no duality gap, the optimality conditions read [15, Prop. 4.1(III)]:

$$(\phi, \xi) \in \partial F_1^*(-K^\star(w, u)), \qquad (4.13)$$
$$K(\phi, \xi) \in \partial F_2^*(w, u), \qquad (4.14)$$

for every $(\phi, \xi)$ and $(w, u)$ that solve $(\mathcal{P}^*)$ and $(\mathcal{P})$, respectively. Hence, for every $(\sigma, \tau) \in X^*$, we have the following:

$$F_1^*(\sigma, \tau) \geq F_1^*(-K^\star(w, u)) + \langle (\sigma, \tau) + K^\star(w, u), (\phi, \xi) \rangle \Leftrightarrow$$

$$\alpha \|\tau\|_\mathcal{M} + \beta \|\sigma\|_{L^p(\Omega)} \geq \alpha \|Du - w\|_\mathcal{M} + \beta \|w\|_{L^p(\Omega)}$$
$$+ \langle (\sigma, \tau) - (w, Du - w), (\phi, \xi) \rangle \Leftrightarrow$$

$$\alpha \|\tau\|_\mathcal{M} + \beta \|\sigma\|_{L^p(\Omega)} \geq \alpha \|Du - w\|_\mathcal{M} + \beta \|w\|_{L^p(\Omega)}$$
$$+ \langle (\sigma - w), \phi \rangle + \langle \tau - (Du - w), \xi \rangle \Leftrightarrow$$

$$\begin{cases} \alpha \|\tau\|_\mathcal{M} \geq \alpha \|Du - w\|_\mathcal{M} + \langle \tau - (Du - w), \xi \rangle, \forall \tau \in H_0^1(\Omega)^*, \\ \beta \|\sigma\|_{L^p(\Omega)} \geq \beta \|w\|_{L^p(\Omega)} + \langle \sigma - w, \phi \rangle, \forall \sigma \in H_0^1(\Omega)^*, \end{cases} \Leftrightarrow$$

$$\begin{cases} \alpha \|\tau\|_\mathcal{M} \geq \alpha \|Du - w\|_\mathcal{M} + \langle \tau - (Du - w), \xi \rangle, \forall \tau \in \mathcal{M}(\Omega), \\ \beta \|\sigma\|_{L^p(\Omega)} \geq \beta \|w\|_{L^p(\Omega)} + \langle \sigma - w, \phi \rangle, \forall \sigma \in L^p(\Omega), \end{cases} \Leftrightarrow$$

$$\begin{cases} \xi \in \alpha \partial \|\cdot\|_\mathcal{M}(Du - w), \\ \phi \in \beta \partial \|\cdot\|_{L^p(\Omega)}(w). \end{cases} \qquad (4.15)$$

Since $\xi \in H_0^1(\Omega) \subset C_0(\Omega)$ in one dimension, we can make use of the fact that

$$\partial \|\cdot\|_\mathcal{M}(Du - w) \cap C_0(\Omega) = \mathrm{Sgn}(Du - w) \cap C_0(\Omega),$$

see (8.1), and write the expressions in (4.15) as

$$\xi \in \alpha \mathrm{Sgn}(Du - w), \qquad (4.16)$$

and

$$\begin{cases} \phi \in \{\tilde{\phi} \in L^q(\Omega) : \|\tilde{\phi}\|_{L^q(\Omega)} \leq \beta\} & \text{if } w = 0, \\ \phi = \beta \dfrac{|w|^{p-2} w}{\|w\|_{L^p(\Omega)}^{(p-1)}} & \text{if } w \neq 0. \end{cases} \qquad (4.17)$$

Indeed, the $L^p$ norm is an one-homogeneous functional and thus its subdifferential reads

$$\partial \|\cdot\|_{L^p(\Omega)}(w) = \left\{ z \in L^p(\Omega)^* : \langle z, w \rangle = \|w\|_{L^p(\Omega)}, \langle z, \sigma \rangle \right.$$
$$\left. \leq \|\sigma\|_{L^p(\Omega)}, \forall \sigma \in L^p(\Omega) \right\}.$$

Note that for $w = 0$, the above expression reduces to $\|\sigma\|_{L^p(\Omega)} \geq \langle z, \sigma \rangle$ for all $\sigma \in L^p(\Omega)$, which is valid for any $z \in L^q(\Omega)$ with $\|z\|_{L^q(\Omega)} \leq 1$, i.e. the unit ball of $L^q(\Omega)$. If $w \neq 0$ then the subdifferential is simply the Gâteaux derivative of the $L^p$ norm, i.e. $\partial \|\cdot\|_{L^p(\Omega)}(w) = \dfrac{|w|^{p-2} w}{\|w\|_{L^p(\Omega)}^{p-1}}$. Finally, from (4.14) we have for every $(\hat{w}, \hat{u}) \in Y^*$

$$F_2^*(\hat{w}, \hat{u}) \geq F_2^*(w, u) + \langle K(\phi, \xi), ((\hat{w}, \hat{u}) - (w, u) \rangle \Leftrightarrow$$

$$\frac{1}{2} \int_\Omega (f - \hat{u})^2 \, dx \geq \frac{1}{2} \int_\Omega (f - u)^2 \, dx + \langle (\xi - \phi, \xi'), (\hat{w} - w, \hat{u} - u) \rangle \Leftrightarrow$$

$$\frac{1}{2} \int_\Omega (f - \hat{u})^2 \, dx \geq \frac{1}{2} \int_\Omega (f - u)^2 \, dx + \langle (\xi - \phi, \hat{w} - w) + \langle \xi', \hat{u} - u \rangle \Leftrightarrow$$

$$\begin{cases} \langle \xi - \phi, \hat{w} - w \rangle \leq 0, \forall \hat{w} \in H_0^1(\Omega)^*, \\ \frac{1}{2} \int_\Omega (f - u)^2 dx + \langle \xi', \hat{u} - u \rangle \leq \frac{1}{2} \int_\Omega (f - \hat{u})^2 dx, \forall \hat{u} \in L^2(\Omega)^* \end{cases} \Leftrightarrow$$

$$\begin{cases} \xi = \phi, \\ \xi' \in \partial \left( \frac{1}{2} \|f - \cdot\|_{L^2(\Omega)}^2 \right)(u) = u - f. \end{cases}$$

Combining all the above results, we obtain the optimality conditions (4.11) and (4.12). $\qquad \square$

*Remark 4.5* We observe that if $w = 0$ then the conditions (4.11) coincide with the optimality conditions for the $L^2$–TV minimisation problem (ROF) with parameter $\alpha$, i.e.

$$\min_{u \in BV(\Omega)} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du\|_\mathcal{M}, \qquad (4.18)$$

see also [37]. On the other hand when $w \neq 0$, the additional condition (4.12) depends on the value of $p$ and as we will see later it allows a certain degree of smoothness in the final solution $u$.

## 4.2 Structure of the Solutions

The optimality conditions (4.11) and (4.12) can help us explore the structure of the solutions for the problem $(\mathcal{P})$ and how this structure is determined by the regularising parameters $\alpha$, $\beta$ and the value of $p$.

We initially discuss the cases where the solution $u$ of $(\mathcal{P})$ is a solution of a corresponding ROF minimisation problem i.e. $w = 0$. Note that the following proposition holds for $p = \infty$ as well.

**Proposition 4.6** (ROF solutions) *Let $q$ be the conjugate exponent of $p \in (1, \infty]$ as defined in (8.4). If*

$$\frac{\beta}{\alpha} \geq |\Omega|^{\frac{1}{q}},$$

*then $(0, u)$ is a solution pair for $(\mathcal{P})$ where $u$ solves the ROF minimisation problem (4.18).*

*Proof* The proof follows immediately from Proposition 2.4.

Proposition 4.6 is valid for any dimension $d \geq 1$. It provides a rough threshold for obtaining ROF type solutions in terms of the regularising parameters $\alpha$, $\beta$ and the image domain $\Omega$. However, the condition is not sharp in general since as we will see in the following sections we can obtain a sharper estimate for specific data $f$.

The following proposition in the spirit of [6,33] gives more insight into the structure of solutions of $(\mathcal{P})$.

**Proposition 4.7** *Let $f \in BV(\Omega)$ and suppose that $(w, u) \in L^p(\Omega) \times BV(\Omega)$ is a solution pair for $(\mathcal{P})$ with $p \in (1, \infty]$. Suppose that $u > f$ (or $u < f$) on an open interval $I \subset \Omega$ then $(Du - w) \lfloor I = 0$, i.e. $u' = w$ on $I$ and $|D^s u|(I) = 0$.*

The above proposition is formulated rigorously via the use of *good representatives* of BV functions, see [1], but for the sake of simplicity we rather not get into the details here. Instead we refer the reader to [6,33] where the analogue propositions are shown for the TGV regularised solutions and whose proofs are similar to the one of Proposition 4.7.

We now examine the case where the solution is constant in $\Omega$, which in fact coincides with the mean value $\tilde{f}$ of the data $f$:

$$\tilde{f} := \operatorname*{argmin}_{u\ \text{constant}} \frac{1}{2} \| f - u \|^2_{L^2(\Omega)} = \frac{1}{|\Omega|} \int_\Omega f\, dx. \quad (4.19)$$

**Proposition 4.8** (Mean value solution) *If the following conditions hold*

$$\begin{aligned} \alpha &\geq \| f - \tilde{f} \|_{L^1(\Omega)}, \\ \beta &\geq |\Omega|^{\frac{1}{q}} \| f - \tilde{f} \|_{L^1(\Omega)}, \end{aligned} \quad (4.20)$$

*then the solution of $(\mathcal{P})$ is constant and equal to $\tilde{f}$.*

*Proof* Clearly, if $u$ is a constant solution of $(\mathcal{P})$, then $Du = 0$ and from inequality (2.2) we get $\mathrm{TVL}^p_{\alpha,\beta}(u) = 0$. Hence, we have $u = \tilde{f}$.

In general, in order to have $u = \tilde{f}$, from the optimality conditions (4.11) and (4.12), it suffices to find a function $\phi \in \mathrm{H}^1_0(\Omega)$ such that

$$\phi' = f - \tilde{f}, \quad \|\phi\|_\infty \leq \alpha, \quad \|\phi\|_{L^q(\Omega)} \leq \beta.$$

Letting $\phi(x) = \int_a^x (f(s) - \tilde{f})\, ds$, then obviously $\phi \in \mathrm{H}^1_0(\Omega)$ since $\phi(a) = \phi(b) = 0$ and

$$|\phi(x)| \leq \int_a^x |f(s) - \tilde{f}|\, ds \leq \| f - \tilde{f} \|_{L^1(\Omega)} < \infty.$$

Therefore, $\|\phi\|_\infty \leq \| f - \tilde{f} \|_{L^1(\Omega)}$. Also, since $L^\infty(\Omega) \subset L^q(\Omega)$ we obtain

$$\|\phi\|_{L^q(\Omega)} \leq |\Omega|^{\frac{1}{q}} \|\phi\|_\infty \leq |\Omega|^{\frac{1}{q}} \| f - \tilde{f} \|_{L^1(\Omega)}.$$

Hence, it suffices to choose $\alpha$ and $\beta$ as in (4.20).

In Fig. 2, we summarise our results so far. There, we have partitioned the set $\{\alpha > 0, \beta > 0\}$ into different areas that correspond to different types of solutions of the problem $(\mathcal{P})$. The red area, arising from thresholds (4.20) corresponds to the choices of $\alpha$ and $\beta$ that produce constant solutions while the blue area corresponds to ROF type solutions, according to threshold (2.3). Therefore, we can determine the area where the non-trivial solutions are obtained, i.e. $w \neq 0$, see purple region. Note that since the conditions (2.3) and (4.20) are not sharp, the blue/red and the purple areas are potentially larger or smaller, respectively than it is shown in Fig. 2.
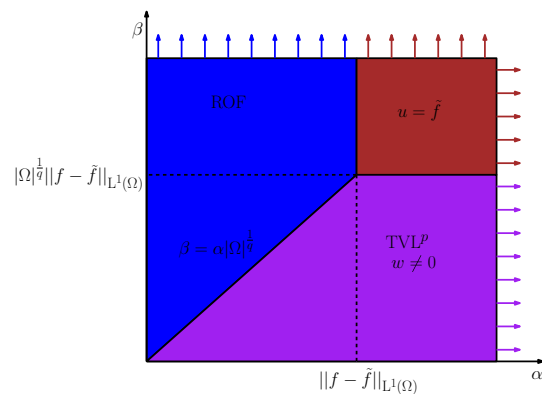


**Fig. 2** Characterisation of solutions of $(\mathcal{P})$ for any data $f$: The *blue/red* areas correspond to the ROF type solutions ($w = 0$) and the *purple* area corresponds to the TVL$^p$ solutions ($w \neq 0$) for $1 < p < \infty$. We note that the *blue/red* and *purple* areas are potentially larger or smaller, respectively as the conditions we have derived are not sharp (Color figure online)

Propositions 4.6 and 4.8 tell us how the solutions $u$ behave when $\beta/\alpha$ or one of the parameters $\alpha$ and $\beta$ is large. The other limiting case is also of interest, i.e. when the parameters are small. The analogous questions have been examined in [35] for the TGV case in arbitrary dimension. There it is shown that whenever $\beta \to 0$ while keeping $\alpha$ fixed or $\alpha \to 0$ while keeping $\beta$ fixed, the corresponding TGV solutions converge to the data $f$ strongly in $L^2$. The same result holds for the TVL$^p$ regularisation. The proof from [35] can be straightforwardly adapted to our case.

The following proposition reveals more information about the structure of solutions in the case $w \neq 0$.

**Proposition 4.9** (TVL$^p$ solutions) *Let $f \in \mathrm{BV}(\Omega)$ and suppose that $(w, u) \in \mathrm{L}^p(\Omega) \times \mathrm{BV}(\Omega)$ is a solution pair for $(\mathcal{P})$ with $p \in (1, \infty)$ and $w \neq 0$. Suppose that $u > f$ (or $u < f$) on an open interval $I \subset \Omega$ then the solution $u$ of $(\mathcal{P})$ is obtained by*

$$-C(|u'(x)|^{p-2}u'(x))' + u(x) = f(x), \quad \forall x \in I$$
$$where \quad C = \frac{\beta}{\|w\|_{\mathrm{L}^p(\Omega)}^{p-1}}. \tag{4.21}$$

*Proof* Since $1 < p < \infty$, $w \neq 0$ using Proposition 4.7 and the second optimality condition of (4.12), we have that

$$\phi = \beta \frac{|u'|^{p-2}u'}{\|w\|_{\mathrm{L}^p(\Omega)}^{p-1}}.$$

Hence, using (4.11) we obtain (4.21) where $C = \frac{\beta}{\|w\|_{\mathrm{L}^p(\Omega)}^{p-1}}$. □

Let us make a few remarks regarding equation (4.21) which is in fact the $p$-Laplace equation. One cannot write down a priori the boundary conditions associated with this equation on an interval $I$ where $u > f$ (or $u < f$) as it depends on the data and the type of solution we are looking for. For instance see (4.31) for the kind of boundary conditions that might arise when we are seeking a particular exact solution. A general statement about the solvability of the equation cannot be made either. If the equation coupled with the boundary conditions (that arise when looking for a specific solution $u$) has a solution then indeed $u$ can possibly solve the minimisation problem. On the other hand, if the $p$-Laplace equation does not have a solution then the function $u$ that imposed the corresponding boundary conditions cannot be a minimiser. For more details on the $p$-Laplace equation and its solvability we refer the reader to [28] and the references therein.

### 4.3 Exact Solutions of $(\mathcal{P})$ for a Step Function

In what follows we compute explicit solutions of the TVL$^p$ denoising model $(\mathcal{P})$ for the case $p = 2$ for a simple data

function. We define the step function in $\Omega = (-L, L), L > 0$ as

$$f(x) = \begin{cases} 0 & \text{if } x \in (-L, 0], \\ h & \text{if } x \in (0, L). \end{cases} \tag{4.22}$$

We first investigate conditions under which we obtain ROF type solutions, that is $w = 0$.

#### 4.3.1 ROF Type Solutions

We note that in this case we can derive conditions for every $1 < p \leq \infty$. We are initially interested in solutions that respect the discontinuity at $x = 0$ and are piecewise constant. From the optimality conditions (4.11)–(4.12), it suffices to find a function $\phi \in \mathrm{H}_0^1(\Omega)$ which, apart from $\phi(-L) = \phi(L) = 0$, it also satisfies

$$\|\phi\|_\infty \leq \alpha, \quad \phi(0) = \alpha, \tag{4.23}$$

and it is piecewise affine. It is easy to see that by setting $\phi(x) = \frac{\alpha}{L}(L - |x|)$, the conditions (4.23) are satisfied and the solution $u$ is piecewise constant. The first condition of (4.12) implies that $\|\phi\|_{\mathrm{L}^q(\Omega)} \leq \beta \Leftrightarrow \frac{\beta}{\alpha} \geq (\frac{2L}{q+1})^{\frac{1}{q}}$ and provides a necessary and sufficient condition that needs to be fulfilled in order for $u$ to be piecewise constant, that is to say

$$u(x) = \begin{cases} \frac{\alpha}{L} & \text{if } x \in (-L, 0], \\ h - \frac{\alpha}{L} & \text{if } x \in (0, L) \end{cases} \Leftrightarrow \frac{\beta}{\alpha} \geq \left(\frac{2L}{q+1}\right)^{\frac{1}{q}}. \tag{4.24}$$

A special case of the ROF type solution is when $u$ is constant, i.e. when $u = \tilde{f}$, the mean value of $f$. We define $\phi(x) = \frac{h}{2}(L - |x|)$ and in that case we have that $\|\phi\|_\infty \leq \alpha \Leftrightarrow \alpha \geq \frac{hL}{2}$ and $\|\phi\|_{\mathrm{L}^q(\Omega)} \leq \beta \Leftrightarrow \beta \geq \frac{h}{2}(\frac{2L^{q+1}}{q+1})^{\frac{1}{q}}$. This implies that

$$u = \tilde{f} = \frac{h}{2} \quad \Leftrightarrow \quad \alpha \geq \frac{hL}{2} \quad \text{and} \quad \beta \geq \frac{h}{2}\left(\frac{2L^{q+1}}{q+1}\right)^{\frac{1}{q}}. \tag{4.25}$$

Using now (4.24)–(4.25) we can draw the exact regions in the quadrant of $\{\alpha > 0, \beta > 0\}$ that correspond to these two types of solutions, see the left graph in Fig. 4 for the special case $p = 2$. Notice that in these regions $w = 0$.

#### 4.3.2 TVL$^2$ Type Solutions

For simplicity reasons, we examine here only the case $p = 2$ with $w \neq 0$ in $\Omega$. However, we refer the reader to Sect. 6.2 where we compute numerically solutions for $p \neq 2$. Using

Proposition 4.9, we observe that the solution is given by the following second-order differential equation:

$$-Cu''(x) - u(x) = f(x), \quad \text{subject to} \quad C = \frac{\beta}{\|w\|_{L^2(\Omega)}}.$$

(4.26)

Even though we can tell that the solution of (4.26) has an exponential form, the fact that the constraint on $C$ depends on the solution $w$, creates a difficult computation in order to recover $u$ analytically. In order to overcome this obstacle, we consider the one-dimensional version of the 2-homogeneous analogue of $(\mathcal{P})$ that was introduced in Sect. 3:

$$\min_{\substack{u \in BV(\Omega) \\ w \in L^2(\Omega)}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \frac{\beta_{2-hom}}{2} \|w\|_{L^2(\Omega)}^2.$$

(4.27)

One can derive the optimality conditions for (4.27) similarly to Sect. 4.1. A pair $(w, u)$ is a solution of (4.27) if and only if there exists a function $\phi \in H_0^1(\Omega)$ such that

$$\phi' = u - f,$$
$$\phi \in \alpha \operatorname{Sgn}(Du - w),$$
$$\phi = \beta_{2-hom} w.$$

(4.28)

In view of Proposition 3.2, in order to recover analytically the solutions of $(\mathcal{P})$ for $p = 2$ and determine exactly the purple region in Fig. 2, it suffices to solve the equivalent model (4.27) in which $w \neq 0$ holds always. We may restrict our computations only on $(-L, 0] \subset \Omega$ and due to symmetry the solution in $(0, L)$ is given by $u(x) = h - u(-x)$. The optimality condition (4.28) results to

$$-u''(x) + ku(x) = 0, \quad \text{where} \quad k^2 = \frac{1}{\beta} \quad \text{and} \quad x \in (-L, 0].$$

(4.29)

Then, we get $u(x) = c_1 e^{kx} + c_2 e^{-kx}$ with $\phi(x) = \frac{c_1}{k} e^{kx} - \frac{c_2}{k} e^{-kx} + c_3$ for all $x \in (-L, 0]$. Firstly, we examine solutions that are continuous which due to symmetry must satisfy $u(0) = \frac{h}{2}$. Since $\phi \in H_0^1(-L, L)$, we have $\phi(-L) = 0$ and also $u'(-L) = 0$. Finally, we require that $\phi(0) < \alpha$. After some computations, we conclude that

$$u(x) = \begin{cases} c_1 e^{kx} + c_2 e^{-kx} & \text{if } x \in (-L, 0], \\ h - c_1 e^{-kx} - c_2 e^{kx} & \text{if } x \in (0, L) \end{cases} \Leftrightarrow \frac{\tanh(kL)}{k} < \frac{2\alpha}{h},$$

(4.30)

where $c_1 = c_2 e^{2kL}$, $c_2 = \frac{h}{2(e^{2kL}+1)}$ and $k = \frac{1}{\sqrt{\beta}}$.

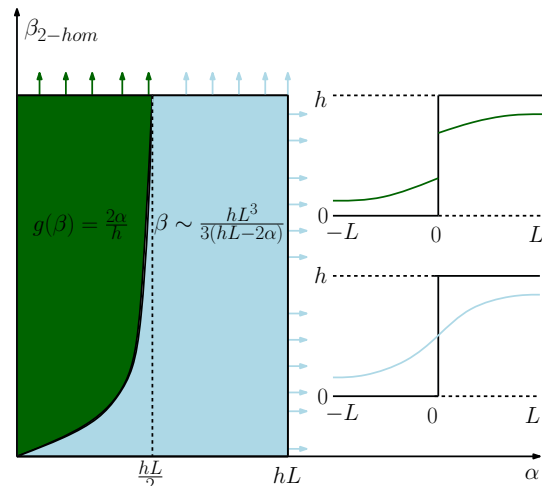On the other hand, in order to get solutions that preserve the discontinuity at $x = 0$, we require the following:



**Fig. 3** Characterisation of solutions of (4.27) for data $f$ being a step function. The *green region* corresponds to solutions that preserve the discontinuity at $x = 0$, (4.32), while the *blue region* corresponds to continuous solutions, (4.30), both having an exponential form (Color figure online)

$$\phi(-L) = 0, \quad u'(-L) = 0,$$
$$u(0) < \frac{h}{2}, \quad \phi(0) = \alpha.$$

(4.31)

Then we get

$$u(x) = \begin{cases} c_1 e^{kx} + c_2 e^{-kx} & \text{if } x \in (-L, 0], \\ h - c_1 e^{-kx} - c_2 e^{kx} & \text{if } x \in (0, L) \end{cases} \Leftrightarrow \frac{\tanh(kL)}{k} > \frac{2\alpha}{h},$$

(4.32)

where $c_1 = c_2 e^{2kL}$, $c_2 = \frac{\alpha k}{e^{2kL}-1}$ and $k = \frac{1}{\sqrt{\beta}}$. Notice that the conditions for $\alpha$ and $\beta$ in (4.30) and (4.32) are supplementary and thus only these type of solutions can occur, see the quadrant of $\{\alpha > 0, \beta > 0\}$ as it presented in Fig. 3. Letting $g(\beta) = \sqrt{\beta} \tanh\left(\frac{L}{\sqrt{\beta}}\right)$, if $g(\beta) < \frac{2\alpha}{h}$ then the solution is of the form (4.30), see the blue region in Fig. 3. On the other hand in the complementary green region we obtain the solution (4.32). For extreme cases where $\beta \to \infty$, i.e. $k \to 0$ we obtain $\frac{\tanh(kL)}{k} \to L$, which means that there is an asymptote of $g$ at $\alpha = \frac{hL}{2}$. Although, we know the form of the inverse function of the hyperbolic tangent, we cannot compute analytically the inverse $f^{-1}$. However, we can obtain an approximation using a Taylor expansion which leads to

$$\sqrt{\beta} \tanh\left(\frac{L}{\sqrt{\beta}}\right) = L - \frac{L^3}{3\beta} + \mathcal{O}\left(\frac{1}{\beta^2}\right) = \frac{2\alpha}{h} \Leftrightarrow$$
$$\beta = \frac{hL^3}{3(hL - 2\alpha)},$$

(4.33)

where $\alpha > 0$ and $\alpha \neq \frac{hL}{2}$.

Finally, we would like to describe the solution on the limiting case $\beta \to \infty$. Letting $\beta \to \infty$ in (4.30), we have that $c_1, c_2 \to \frac{h}{2}$ and $u(x) \to \frac{h}{2}$ for every $x \in \Omega$, which in fact is the mean value solution of $(\mathcal{P})$. For the discontinuous solutions, we have that $c_1, c_2 \to \frac{\alpha}{2L}$ and

$$u(x) \to \begin{cases} \frac{\alpha}{L} & \text{if } x \in (-L, 0], \\ h - \frac{\alpha}{L} & \text{if } x \in (0, L), \end{cases}$$

i.e. the limiting solution is (4.24). We also get that

$$w(x) = kc_2 \begin{cases} e^{2kL+kx} - e^{-kx} & \text{if } x \in (-L, 0], \\ e^{2kL-kx} - e^{kx} & \text{if } x \in (0, L], \end{cases} \tag{4.34}$$

with $\|w\|_{L^2(\Omega)} = c_2 k \sqrt{2} e^{kL} (\sinh(2kL) - 2kL)^{\frac{1}{2}}$ and $c_2$ is given either from (4.30) or (4.32). Then, in both cases we have $w \to 0$ as $k \to 0$. Observe that the product of $\beta_{2-hom} \|w\|_{L^2(\Omega)}$ is bounded as $\beta_{2-hom} \to \infty$ for both types of solutions and in fact corresponds to the bounds found in (4.24) and (4.25). Indeed, since

$$\frac{(\sinh(2kL) - 2kL)^{\frac{1}{2}}}{k^{\frac{3}{2}}} \to 2\sqrt{\frac{L^3}{3}}, \quad \text{as } k \to 0,$$

if $\alpha > \frac{hL}{2}$ then

$$\beta_{2-hom} \|w\|_{L^2(\Omega)} \to \frac{h}{2}\sqrt{\frac{2L^3}{3}}, \quad \text{as } \beta_{2-hom} \to \infty,$$

while if $\alpha \le \frac{hL}{2}$

$$\beta_{2-hom} \|w\|_{L^2(\Omega)} \to \alpha\sqrt{\frac{2L}{3}}, \quad \text{as } \beta_{2-hom} \to \infty.$$

The last result is yet another verification of Proposition 3.2 and it shows that there is an one to one correspondence between $\beta_{2-hom} \|w\|_{L^2(\Omega)}$ and $\beta_{1-hom}$. The solutions that belong to the purple region of Fig. 4 are the same to the solutions that are shown in Fig. 3.

In the next proposition we summarise the type of solutions for $(\mathcal{P})$ for the step function:

**Proposition 4.10** *There are four different types of solutions for the problem* $(\mathcal{P})$ *with* $p = 2$ *taking the step function* (4.22) *as data:*

(1) *A piecewise constant solution given in* (4.24) *(blue region in Fig. 4).*
(2) *A constant solution, equal to the mean value of the data, given in* (4.25) *(brown region in Fig. 4).*
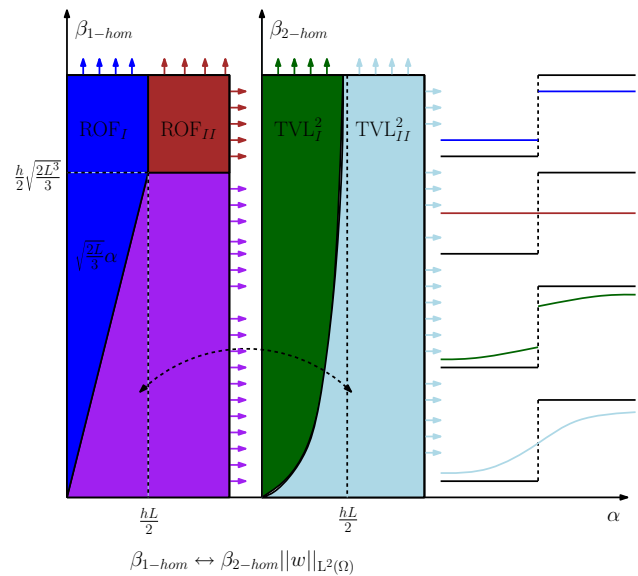(3) *A continuous exponential solution given in* (4.30) *(light-blue region in Fig. 4).*



**Fig. 4** Characterisation of solutions of $(\mathcal{P})$ for $p = 2$ for data $f$ being a step function. The type of solutions in the *purple region* of the *left graph* are exactly the solutions obtained for the 2-homogeneous problem (4.27), on the *right graph* (Color figure online)

(4) *A discontinuous piecewise exponential solution given in* (4.32) *(green region in Fig. 4).*

*Furthermore, there is an one to one correspondence between the purple and the green/light blue regions in Fig. 4.*

## 5 An Image Decomposition Approach

In this section, we present another formulation of the problem $(\mathcal{P})$, where we decompose an image into a BV part (piecewise constant) and a part that belongs to $W^{1,p}(\Omega)$ (smooth). Let $1 < p \le \infty$ and $\Omega \subset \mathbb{R}^d$ and consider the following minimisation problem:

$$\min_{\substack{u \in \mathrm{BV}(\Omega) \\ v \in \mathrm{W}^{1,p}(\Omega)}} L(u, v) := \frac{1}{2} \|f - u - v\|_{L^2(\Omega)}^2 + \alpha \|Du\|_{\mathcal{M}}$$
$$+ \beta \|\nabla v\|_{L^p(\Omega)}. \tag{5.1}$$

In this way, we can decompose our image into two components of different structure. The second term captures the piecewise constant structures in the image, whereas the third term captures the smoothness that depends on the value of $p$. In the one-dimensional setting, we can prove that the problems $(\mathcal{P})$ and (5.1) are equivalent.

**Proposition 5.1** *Let* $\Omega = (a, b) \subset \mathbb{R}$ *and* $1 < p \le \infty$. *Then a pair* $(v^*, u^*) \in \mathrm{W}^{1,p}(\Omega) \times \mathrm{BV}(\Omega)$ *is a solution of* (5.1) *if and only if* $(\nabla v^*, u^* + v^*) \in L^p(\Omega) \times \mathrm{BV}(\Omega)$ *is a solution of* $(\mathcal{P})$.

*Proof* Let $\overline{u} = u + v$ then, we have the following

$$(v^*, u^*) \in \operatorname*{argmin}_{\substack{u \in \mathrm{BV}(\Omega) \\ v \in \mathrm{W}^{1,p}(\Omega)}} \frac{1}{2} \|f - u - v\|_{\mathrm{L}^2(\Omega)}^2$$
$$+ \alpha \|Du\|_{\mathcal{M}} + \beta \|\nabla v\|_{\mathrm{L}^p(\Omega)} \Leftrightarrow$$

$$(v^*, u^*) \in \operatorname*{argmin}_{\substack{u \in \mathrm{BV}(\Omega) \\ v \in \mathrm{W}^{1,p}(\Omega)}} \frac{1}{2} \|f - u - v\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_\infty \leq 1}} \{\langle u, \phi' \rangle\}$$
$$+ \beta \|\nabla v\|_{\mathrm{L}^p(\Omega)} \Leftrightarrow$$

$$(v^*, \overline{u^*}) \in \operatorname*{argmin}_{\substack{\overline{u} \in \mathrm{BV}(\Omega) \\ v \in \mathrm{W}^{1,p}(\Omega)}} \frac{1}{2} \|f - \overline{u}\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \sup_{\substack{\phi \in C_c^\infty(\Omega) \\ \|\phi\|_\infty \leq 1}} \{\langle \overline{u}, \phi' \rangle + \langle \nabla v, \phi \rangle\}$$
$$+ \beta \|\nabla v\|_{\mathrm{L}^p(\Omega)}$$

$$(w^*, \overline{u^*}) \in \operatorname*{argmin}_{\substack{\overline{u} \in \mathrm{BV}(\Omega) \\ w = \nabla v \\ v \in \mathrm{W}^{1,p}(\Omega)}} \frac{1}{2} \|f - \overline{u}\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|D\overline{u} - w\|_{\mathcal{M}}$$
$$+ \beta \|w\|_{\mathrm{L}^p(\Omega)}.$$

However, the constraints $w = \nabla v$, $v \in \mathrm{W}^{1,p}(\Omega)$ can simply be substituted by $w \in \mathrm{L}^p(\Omega)$ since

$$\left\{ w \in \mathrm{L}^p(\Omega) : \exists v \in \mathrm{W}^{1,p}(\Omega), \ w = \nabla v \right\} = \mathrm{L}^p(\Omega). \quad (5.2)$$

Indeed, let $w \in \mathrm{L}^p(\Omega) \subset \mathrm{L}^1(\Omega)$ for $p \in (1, \infty)$ and define $v(x) = \int_a^x w(s)\,ds$ for $x \in \Omega \subset \mathbb{R}$. Clearly, $v' = w$ a.e. and by Hölder's inequality we have for every $x \in (a, b)$

$$|v(x)|^p = \left| \int_a^x w(s)\,ds \right|^p \leq (x - a)^{p-1}$$
$$\int_a^x |w(s)|^p\,ds < C < \infty.$$

Thus $v \in \mathrm{W}^{1,p}(\Omega)$ for $p \in (1, \infty)$. If $p = \infty$, suppose again $w \in \mathrm{L}^\infty(\Omega)$ and let $C > 0$ be a constant such that $|w(x)| \leq C$ a.e. on $\Omega$. In that case we have $|v(x)| \leq \int_a^x |w(s)|\,ds \leq C|\Omega| < \infty$, i.e. $v \in \mathrm{L}^\infty(\Omega)$ and hence $v \in \mathrm{W}^{1,\infty}(\Omega)$ since $v' = w$. Therefore,

$$(w^*, \overline{u^*}) \in \operatorname*{argmin}_{\substack{\overline{u} \in \mathrm{BV}(\Omega) \\ w \in \mathrm{L}^p(\Omega)}} \frac{1}{2} \|f - \overline{u}\|_{\mathrm{L}^2(\Omega)}^2 + \alpha \|D\overline{u} - w\|_{\mathcal{M}}$$
$$+ \beta \|w\|_{\mathrm{L}^p(\Omega)},$$

where $\overline{u^*} = u^* + v^*$ and $w^* = \nabla v^*$.

Even though for $d = 1$ it is true that every $\mathrm{L}^p$ function can be written as a gradient, this is not true in higher dimensions. In fact, as we show in the following sections, this constraint is quite restrictive and for example the staircasing effect cannot be always eliminated in the denoising process, see for instance Fig. 20.

The existence of minimisers of (5.1) is shown following again the same techniques as in Theorem 2.5. Moreover, due

to the strict convexity of the fidelity term in (5.1), one can prove that the sum $u + v \in \mathrm{BV}(\Omega)$ is unique for a solution $(u, v) \in \mathrm{W}^{1,p}(\Omega) \times \mathrm{BV}(\Omega)$. This result reflects the uniqueness for the problem $(\mathcal{P})$ for $\overline{u}$. However one cannot show that the solutions $(u, v)$ are unique in general. Yet, one can say something more about this issue. if $(u_1, v_1)$, $(u_2, v_2)$ are two minimisers of (5.1), then from the convexity of $L(u, v)$ we have for $0 \leq \lambda \leq 1$

$$L(\lambda(u_1, v_1) + (1 - \lambda)(u_2, v_2)) \leq \lambda L(u_1, v_1)$$
$$+ (1 - \lambda) L(u_2, v_2).$$

Since $(u_1, v_1)$, $(u_2, v_2)$ are both minimisers, the above inequality is in fact an equality. Since $u_1 + v_1 = u_2 + v_2$, we obtain

$$\alpha \|D(\lambda u_1 + (1-\lambda)u_2)\|_{\mathcal{M}} + \beta \|\nabla(\lambda v_1 + (1-\lambda)v_2)\|_{\mathrm{L}^p(\Omega)}$$
$$= \alpha(\lambda \|Du_1\|_{\mathcal{M}} + (1-\lambda)\|Du_2\|_{\mathcal{M}}) + \beta(\lambda \|\nabla v_1\|_{\mathrm{L}^p(\Omega)}$$
$$+ (1 - \lambda) \|\nabla v_2\|_{\mathrm{L}^p(\Omega)}). \quad (5.3)$$

If we assume that

$$\|\nabla(\lambda v_1 + (1-\lambda)v_2)\|_{\mathrm{L}^p(\Omega)} < \lambda \|\nabla v_1\|_{\mathrm{L}^p(\Omega)}$$
$$+ (1 - \lambda) \|\nabla v_2\|_{\mathrm{L}^p(\Omega)},$$

then we contradict the equality on (5.3). Hence, the Minkowski inequality becomes an equality, i.e.

$$\|\nabla(\lambda v_1 + (1-\lambda)v_2)\|_{\mathrm{L}^p(\Omega)} = \lambda \|\nabla v_1\|_{\mathrm{L}^p(\Omega)}$$
$$+ (1 - \lambda) \|\nabla v_2\|_{\mathrm{L}^p(\Omega)},$$

which is equivalent to the existence of $\mu \geq 0$ such that $\nabla v_2 = \mu \nabla v_1$. In other words, we have proved the following proposition which was also shown in [25] in a similar context:

**Proposition 5.2** *Let $(u_1, v_1)$, $(u_2, v_2)$ be two minimisers of* (5.1)*. Then*

$$u_1 + v_1 = u_2 + v_2, \quad and \quad (5.4)$$

*there exists a $\mu \geq 0$ such that $\nabla v_2 = \mu \nabla v_1$.* $\quad (5.5)$

## 6 Numerical Experiments

In this section we present our numerical simulations for the problem $(\mathcal{P})$. We begin with the one-dimensional case where we verify numerically the analytical solutions obtained in Sect. 4.3. Through examples we also investigate the type of structures that are promoted for different values of $p$. Finally, we proceed to the two-dimensional case where we focus on

image denoising tasks and in particular on the elimination of the staircasing effect.

We start by defining the discretised version of problem $(\mathcal{P})$

$$\min_{u \in \mathbb{R}^{n \times m}} \frac{1}{2} \| f - u \|_2^2 + \mathrm{TVL}_{\alpha,\beta}^p(u). \qquad (6.1)$$

Here $\mathrm{TVL}_{\alpha,\beta}^p : \mathbb{R}^{n \times m} \to \mathbb{R}$ is defined as

$$\mathrm{TVL}_{\alpha,\beta}^p(u) = \operatorname*{argmin}_{w \in (\mathbb{R}^{n \times m})^2} \alpha \| \nabla u - w \|_1 + \beta \| w \|_p, \qquad (6.2)$$

where for $x \in \mathbb{R}^{n \times m}$, we set $\| x \|_p = (\sum_{i,j=1}^{n,m} |x(i,j)|^p)^{\frac{1}{p}}$ and for $x = (x_1, x_2) \in (\mathbb{R}^{n \times m})^2$ we define

$$\| x \|_p = \left( \sum_{i,j=1}^{n,m} \left( \sqrt{(x_1(i,j))^2 + (x_2(i,j))^2} \right)^p \right)^{\frac{1}{p}}. \qquad (6.3)$$

We denote by $\nabla = (\nabla_1, \nabla_2)$ the discretised gradient with forward differences and zero Neumann boundary conditions defined as

$$(\nabla_1 u)_{i,j} = \begin{cases} u(i+1,j) - u(i,j) & \text{if } 1 \le i < n, \ 1 \le j \le m, \\ 0 & \text{if } i = n, \ 1 \le j \le m, \end{cases}$$

$$(\nabla_2 u)_{i,j} = \begin{cases} u(i,j+1) - u(i,j) & \text{if } 1 \le i \le n, \ 1 \le j < m, \\ 0 & \text{if } 1 \le i \le n, \ j = m. \end{cases}$$

The discrete version of the divergence operator is defined as the adjoint of discrete gradient. That is, for every $w = (w_1, w_2) \in (\mathbb{R}^{n \times m})^2$ and $u \in \mathbb{R}^{n \times m}$, we have $\langle -\mathrm{div} w, u \rangle = \langle w, \nabla u \rangle$ with

$$(\mathrm{div} w)_{i,j} = \begin{cases} w_1(i,j) - w_1(i,j-1) & \text{if } 1 < j < m, \ 1 \le i \le n, \\ w_1(i,j) & \text{if } j = 1, \ 1 \le i \le n, \\ -w_1(i,j-1) & \text{if } j = m, \ 1 \le i \le n, \end{cases}$$

$$+ \begin{cases} w_2(i,j) - w_2(i-1,j) & \text{if } 1 < i < n, \ 1 \le j \le m, \\ w_2(i,j) & \text{if } i = 1, \ 1 \le j \le m, \\ -w_2(i-1,j) & \text{if } i = m, \ 1 \le j \le m. \end{cases} \qquad (6.4)$$

We solve the minimisation problem (6.1) in two ways. The first one is by using the CVX optimisation package with MOSEK solver (interior point methods) [19]. This method is efficient for small–medium scale optimisation problems and thus it is a suitable choice in order to replicate one-dimensional solutions. On the other hand, we prefer to solve large scale two-dimensional versions of (6.1) with the split Bregman method [18] which has been widely used for the fast solution of non-smooth minimisation problems.

## 6.1 Split Bregman for $\mathrm{L}^2$–$\mathrm{TVL}^p$

In this section we describe how we adapt the split Bregman algorithm to our discrete model (6.1). We first transform the unconstrained problem (6.1) into a constrained one by setting $z = \nabla u - w$:

$$\min_{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2 \\ z \in (\mathbb{R}^{n \times m})^2}} \frac{1}{2} \| f - u \|_2^2 + \alpha \| z \|_1 + \beta \| w \|_p,$$
$$\text{such that} \quad z = \nabla u - w. \qquad (6.5)$$

Replacing the constraint, using a Lagrange multiplier $\lambda$, we obtain the following unconstrained formulation:

$$\min_{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2 \\ z \in (\mathbb{R}^{n \times m})^2}} \frac{1}{2} \| f - u \|_2^2 + \alpha \| z \|_1 + \beta \| w \|_p + \frac{\lambda}{2} \| z - \nabla u + w \|_2^2. \qquad (6.6)$$

The Bregman iteration [32], that corresponds to the minimisation (6.6) leads to the following two-step algorithm:

$$(u^{k+1}, z^{k+1}, w^{k+1}) = \operatorname*{argmin}_{u,z,w} \frac{1}{2} \| f - u \|_2^2 + \alpha \| z \|_1 + \beta \| w \|_p$$
$$+ \frac{\lambda}{2} \| b^k - z + \nabla u - w \|_2^2, \qquad (6.7)$$
$$b^{k+1} = b^k + z^{k+1} - \nabla u^{k+1} - w^{k+1}. \qquad (6.8)$$

Since solving (6.7) at once is a difficult task, we employ a splitting technique and minimise alternately for $u$, $z$ and $w$. This yields the split Bregman iteration for our method:

$$u^{k+1} = \operatorname*{argmin}_{u \in \mathbb{R}^{n \times m}} \frac{1}{2} \| f - u \|_2^2 + \frac{\lambda}{2} \| b^k + z^k - \nabla u + w^k \|_2^2, \qquad (6.9)$$

$$z^{k+1} = \operatorname*{argmin}_{z \in (\mathbb{R}^{n \times m})^2} \alpha \| z \|_1 + \frac{\lambda}{2} \| b^k + z - \nabla u^{k+1} + w^k \|_2^2, \qquad (6.10)$$

$$w^{k+1} = \operatorname*{argmin}_{w \in (\mathbb{R}^{n \times m})^2} \beta \| w \|_p + \frac{\lambda}{2} \| b^k + z^{k+1} - \nabla u^{k+1} + w \|_2^2, \qquad (6.11)$$

$$b^{k+1} = b^k + z^{k+1} - \nabla u^{k+1} - w^{k+1}. \qquad (6.12)$$

Next, we discuss how we solve each of the subproblems (6.9)–(6.11). The first-order optimality condition of (6.9) results into the following linear system:

$$\underbrace{(I - \lambda \Delta)}_{A} u = \underbrace{f - \lambda \mathrm{div}(b^k + z^k - w^k)}_{c}. \qquad (6.13)$$

Here $A$ is a sparse, symmetric, positive definite and strictly diagonal dominant matrix, thus we can easily solve (6.13) with an iterative solver such as conjugate gradients or Gauss–Seidel. However, due to the zero Neumann boundary conditions, the matrix $A$ can be efficiently diagonalised by the two-dimensional discrete cosine transform,

$$A = W_{nm}^{\mathsf{T}} D W_{nm}, \tag{6.14}$$

where here $W_{nm}$ is the discrete cosine matrix and $D = diag(\mu_1, \cdots, \mu_{nm})$ is the diagonal matrix of the eigenvalues of $A$. In that case, $A$ has a particular structure of a block symmetric *Toeplitz-plus-Hankel* matrix with *Toeplitz-plus-Hankel* blocks and one can obtain the solution of (6.9) by three operations involving the two-dimensional discrete cosine transform [20] as follows: Firstly, we calculate the eigenvalues of $A$ by multiplying (6.14) with $e_1 = (1, 0, \cdots, 0)^{\mathsf{T}}$ from both sides and using the fact that $W_{nm}^{\mathsf{T}} W_{nm} = W_{nm} W_{nm}^{\mathsf{T}} = I_{nm}$, we get

$$D_{i,i} = \frac{[W_{nm} A e_1]_i}{[W_{nm} e_1]_i}, \ i = 1, 2, \cdots, nm. \tag{6.15}$$

Then, the solution of (6.9) is computed exactly by

$$u = W_{nm}^{\mathsf{T}} D^{-1} W_{nm} c. \tag{6.16}$$

The solution of the subproblem (6.10) is obtained in a closed form via the following shrinkage operator:

$$
\begin{aligned}
z_i^{k+1} &= \text{shrink}_{\frac{\alpha}{\lambda}}(\underbrace{b_i^k - \nabla_i u^{k+1} + w_i^k}_{g_i}) \\
&:= \max\left(\|g\|_2 - \frac{\alpha}{\lambda}\right) \frac{g_i}{\|g\|_2}.
\end{aligned} \tag{6.17}
$$

Finally, we discuss the solution of the subproblem (6.11). In the spirit of [40], we solve (6.11) by a fixed-point iteration scheme. Letting $\kappa = \frac{\beta}{\lambda}$ and $\eta = -b^k - z^{k+1} + \nabla u^{k+1}$, the first-order optimality condition of (6.11) becomes

$$\kappa \frac{|w|^{p-2} w}{\|w\|_p^{p-1}} + w - \eta = 0. \tag{6.18}$$

For given $w^k$, we obtain $w^{k+1}$ by the following fixed-point iteration

$$w_i^{k+1} = \frac{\eta_i \|w^k\|_p^{p-1}}{\kappa |w^k|^{p-2} + \|w^k\|_p^{p-1}}, \tag{6.19}$$

under the convention that $0/0 = 0$. We can also consider solving the $p$-homogenous analogue $(\mathcal{P}_{p-hom})$, where for certain values of $p$, e.g. $p = 2$, we can solve exactly the corresponding version of (6.11), since in that case $w_i^{k+1} = $

$\frac{\eta_i}{\kappa+1}$. However, we have observed empirically that there is no significant computational difference between these two methods.

Since we do not solve all the subproblems (6.9)–(6.11) exactly in every iteration, we cannot guarantee convergence for our version of the split Bregman iteration. Moreover, convergence of the split Bregman algorithm when more than two splittings are performed have not been yet fully established, even though this has been an active field of research lately see for instance [13,17,31]. Let us note that the three subproblems in the split Bregman algorithm can be modified into two subproblems (inexact linearised ADMM) with a small cost in the speed of convergence, see for instance [14,21,22]. However in practice, the algorithm converges to the right solutions. This claim is supported by the study presented in Fig. 5 where the solutions of the split Bregman iteration are compared to the corresponding solutions obtained with the CVX package for which we have convergence guarantees. There, we have solved the TVL$^p$ minimisations that correspond to Figs. 13d–f and 14d, i.e. for $p = \frac{3}{2}, 2, 3, 7$ using both split Bregman and CVX. We plot the relative differences of the split Bregman iterates $u^k$ and the CVX solution $u_{\text{CVX}}$ until they are sufficiently close to each other, i.e.

$$\frac{\|u^k - u_{\text{CVX}}\|_2}{\|u_{\text{CVX}}\|_2} < \text{tol}. \tag{6.20}$$

In all the plots of Fig. 5, we observe that the split Bregman iterates in practice converge to the CVX solution. Their relative difference becomes of the order $10^{-5}$ in around 100 iterations except for $p = 7$, Fig. 5d, where the error tolerance is still reached but only after approximately 10000 iterations.

In Table 1, we compare the computational times of the split Bregman algorithm until (6.20) is satisfied and the computational times of CVX for the same examples as in Fig. 5. The implementations were done in MATLAB (2013) using 2.4 GHz Intel Core 2 Duo and 2 GB of memory. Notice that unless $p = 2$, second line in Table 1, CVX needs more than an hour to converge, in contrast to split Bregman where only a few seconds are required for small values of $p$. Note that the split Bregman algorithm is significantly slower for large values of $p$, e.g. $p = 7$, see fourth line in Table 1, mainly due to the fixed-point iteration in the subproblem (6.19). We would like to point out that the computational speed can be significantly reduced in the $p = \infty$ case, since the corresponding subproblem is solved exactly, see [8,36] and in the same time we can obtain similar results to the ones obtained for high values of $p$.

### 6.2 One-Dimensional Results

In this section, we present some numerical results in dimension one, i.e. $m = 1$, $u \in \mathbb{R}^{n \times 1}$ and $w \in \mathbb{R}^{n \times 1}$. Initially, we
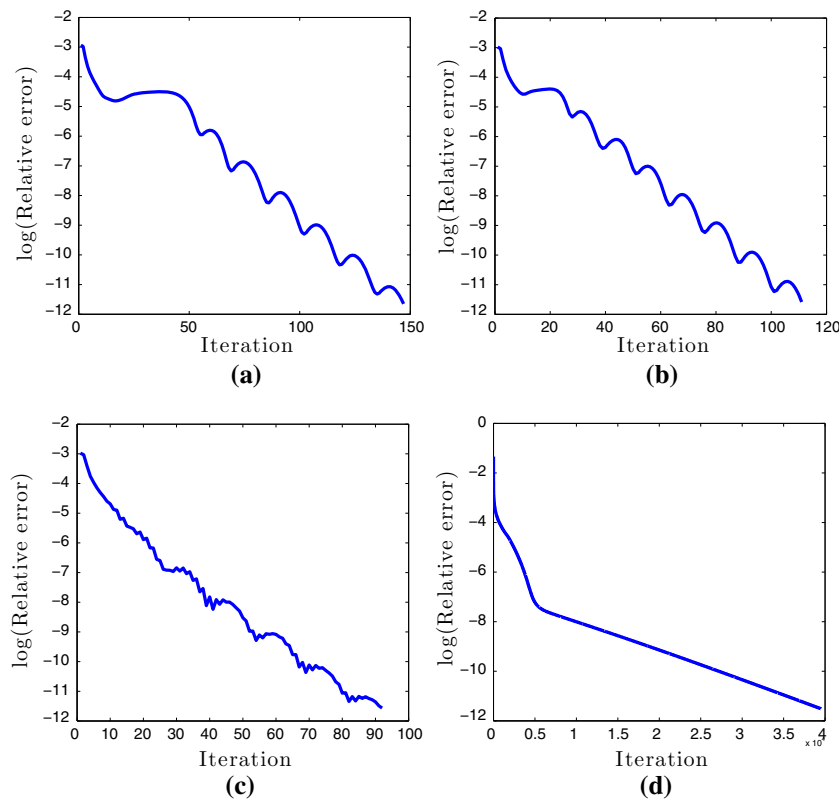
**Fig. 5** Plots of the relative differences between the split Bregman iterates and the CVX solutions until (6.20) is true with tol = $10^{-5}$, for the examples in Figs. 13d–f and 14d. In all cases the split Bregman algorithm converges to the solutions given by CVX. **a** Relative residual error between the split Bregman iterates and the CVX solution for the example in Fig. 13d. **b** Relative residual error between the split Bregman iterates and the CVX solution for the example in Fig. 13e. **c** Relative residual error between the split Bregman iterates and the CVX solution for the example in Fig. 13f. **d** Relative residual error between the split Bregman iterates and the CVX solution for the example in Fig. 14d

**Table 1** Computational times of the split Bregman algorithm until (6.20) is true with tol = $10^{-5}$ and comparison to the computational times of CVX for the same examples as in Fig. 5

|  | Split Bregman (s) | Iterations | Relative error | CVX (s) |
|---|---|---|---|---|
| $p = 1.5$ | 3.58 | 147 | $8.72 \times 10^{-6}$ | 3433 |
| $p = 2$ | 1.99 | 111 | $9.31 \times 10^{-6}$ | 193.14 |
| $p = 3$ | 1.58 | 92 | $9.49 \times 10^{-6}$ | 3418 |
| $p = 7$ | 2266 | 39518 | $9.68 \times 10^{-6}$ | 3532 |

compare our numerical solutions with the analytical ones, obtained in Sect. 4.3 for the step function. We set $p = 2$, $h = 100$, $L = 1$ and $\Omega = (-1, 1)$ which is discretised into 2000 points. We first examine the cases where ROF solutions are obtained, i.e. the parameters $\alpha$ and $\beta$ are selected according to the conditions (4.24) and (4.25). We have done that in Fig. 6 where we see that the analytical solutions coincide with the numerical ones.

We proceed by computing the non-ROF solutions. The numerical solutions are solved using the 2-homogeneous analogue (4.27), since we have proved that the 1-homogeneous and $p$-homogeneous problems are equivalent modulo an appropriate rescaling of the parameter $\beta$, see Proposition 3.2. Indeed, as it is described in Fig. 4, in order to obtain solu-

tions from the purple region, it suffices to seek for solutions of the 2-homogeneous (4.27). Recall also that these solutions are exactly the solutions obtained solving a Huber TV problem, see Proposition 3.3. The analytical solutions are given in (4.30) and (4.32) and are compared to the numerical ones in Fig. 7, where we observe that they coincide. We also verify the equivalence between the 1-homogeneous and 2-homogeneous problems where $\alpha$ is fixed and $\beta$ is obtained from Proposition 3.2, see Fig. 7c.

We continue our experiments for general values of $p$ focusing on the structure of the solutions as $p$ increases. In order to compare the solutions for different values $p \in (1, \infty)$, we fix the parameter $\alpha$ and choose appropriate values of $\beta$. Since we are mainly interested in non-ROF
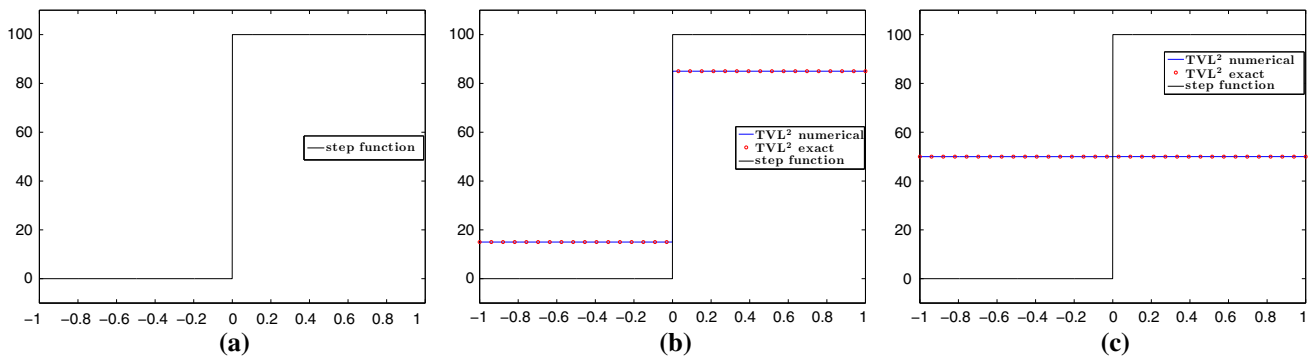
**Fig. 6** Step function: comparison between numerical solutions of $(\mathcal{P})$ and the corresponding analytical solutions obtained in Sect. 4.3. The parameters $\alpha$ and $\beta$ are chosen so that conditions (4.24) and (4.25) are satisfied which result in ROF solutions. **a** Original data. **b** TVL$^2$: $\alpha = 15$, $\beta = 500$. **c** TVL$^2$: $\alpha = 60$, $\beta = 1300$ (Color figure online)
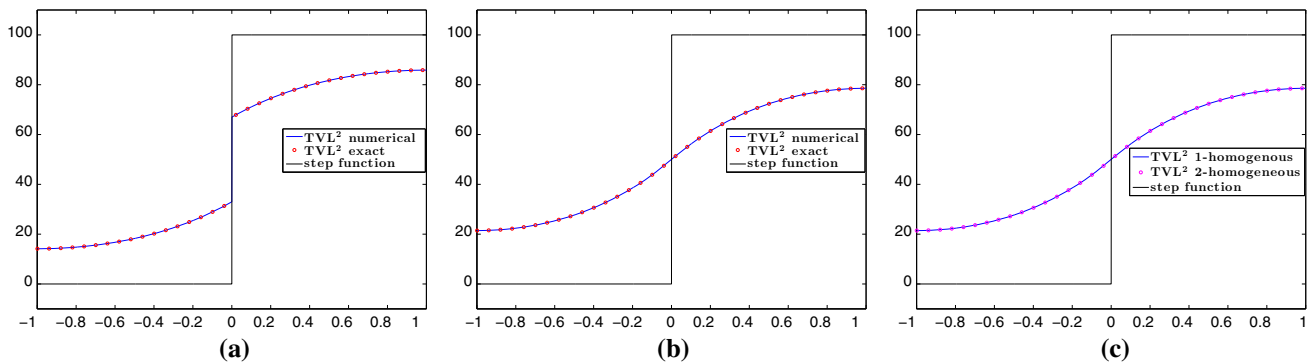
**Fig. 7** Step function: comparison between numerical and analytical solutions obtained in Sect. 4.3, by solving the 2-homogeneous problem (4.27). The parameters $\alpha$ and $\beta$ are chosen so that conditions (4.30) and (4.32) are satisfied which result in non-ROF solutions. The last plot shows the equivalence between the 1-homogeneous problem $(\mathcal{P})$ and 2-homogeneous (4.27). **a** TVL$^2$ : $\alpha = 20$, $\beta_{2-hom} = 450$. **b** TVL$^2$ : $\alpha = 60$, $\beta_{2-hom} = 450$. **c** Equivalence of 1 and 2-homogeneous models: $\alpha = 15$, $\beta_{2-hom} = 450$, $\beta_{1-hom} = \beta_{2-hom} \|w\|_2$ (Color figure online)
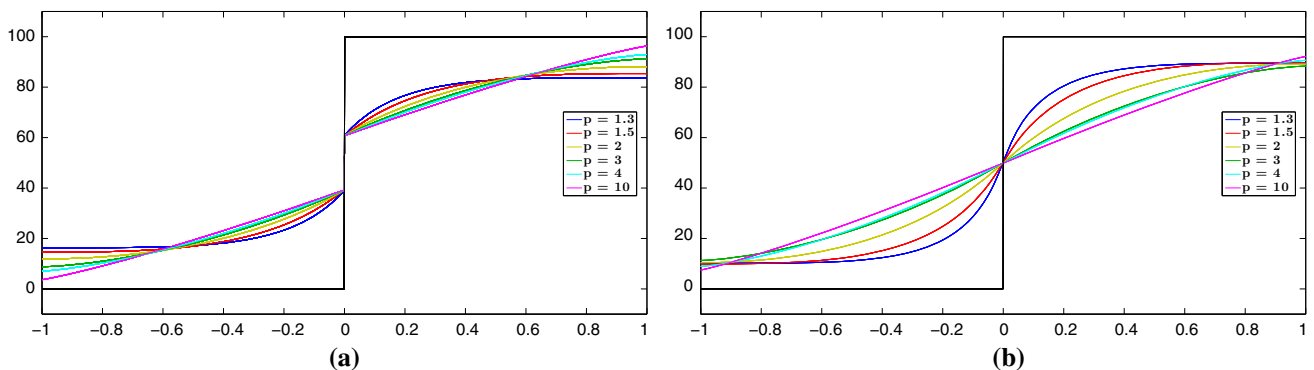
**Fig. 8** Step function: the types of solutions for the problem $(\mathcal{P})$ for different values of $p$. **a** TVL$^p$ discontinuous solutions for $p = \{\frac{4}{3}, \frac{3}{2}, 2, 3, 4, 10\}$. **b** TVL$^p$ continuous solutions for $p = \{\frac{4}{3}, \frac{3}{2}, 2, 3, 4, 10\}$ (Color figure online)

solutions, we choose $\alpha$ and $\beta$ so that they belong to the purple region of Fig. 4, i.e. $\beta < (\frac{2L}{q+1})^{\frac{1}{q}}\alpha$ and $\beta < \frac{h}{2}(\frac{2L^{q+1}}{q+1})^{\frac{1}{q}}$. We set $p = \{\frac{4}{3}, \frac{3}{2}, 2, 3, 4, 10\}$ and in order to get solutions that preserve the discontinuity we set $\beta = \{72, 140, 430, 1350, 2400, 6800\}$ with fixed $\alpha = 20$, see

Fig. 8a. In order to obtain continuous solutions, we set $\alpha = 60$ and $\beta = \{50, 110, 430, 1700, 3000, 9500\}$, see Fig. 8b. We observe that for $p = \frac{4}{3}$, the solution has a similar behaviour to $p = 2$, but with a steeper gradient at the discontinuity point and the solution becomes almost constant near the boundary of $\Omega$. On the other hand, as we increase $p$, the slope
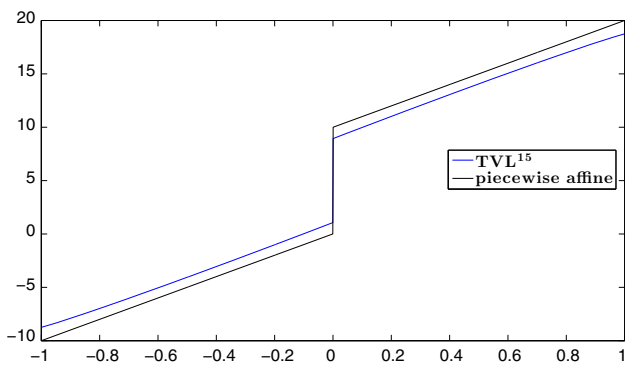
**Fig. 9** Piecewise affine data: TVL$^{15}$ solution with $\alpha = 1$, $\beta = 620$ (Color figure online)

of the solution near the discontinuity point reduces and it becomes almost linear with a relative small constant part near the boundary.

The linear structure of the solutions that appears for large $p$ motivates us to examine the case of piecewise affine data $f$ defined as

$$f(x) = \begin{cases} \lambda x & \text{if } x \in (-L, 0], \\ \lambda x + h & \text{if } x \in (0, L], \end{cases} \tag{6.21}$$

see Fig. 9. We set again $\Omega = (-1, 1)$ and $\lambda = \frac{1}{10}$. As we observe, the reconstruction for $p = 15$ behaves almost linearly everywhere in $\Omega$ except near the boundary. In the follow-up paper [8], where the case $p = \infty$ is examined in detail and exact solutions are computed for the data (6.21), the occurrence of this linear structure is justified also analytically.

In the last part of this section, we present some numerical examples of the image decomposition approach presented in Sect. 5. We use as data a more complicated one-dimensional noiseless signal with piecewise constant, affine and quadratic components and solve the discretised version of (5.1) using CVX. In Fig. 10, we verify numerically the equivalence between (5.1) and ($\mathcal{P}$) for $p = 2$, i.e. we show that $(\nabla v, u + v)$ corresponds to $(w, \overline{u})$ where $(v, u)$ and $(w, \overline{u})$ are the solutions of (5.1) and ($\mathcal{P}$), respectively. We also compare the decomposed parts $u$, $v$ for $p = \frac{4}{3}$ and $p = 10$. In order to have a reasonable comparison on the correspond-
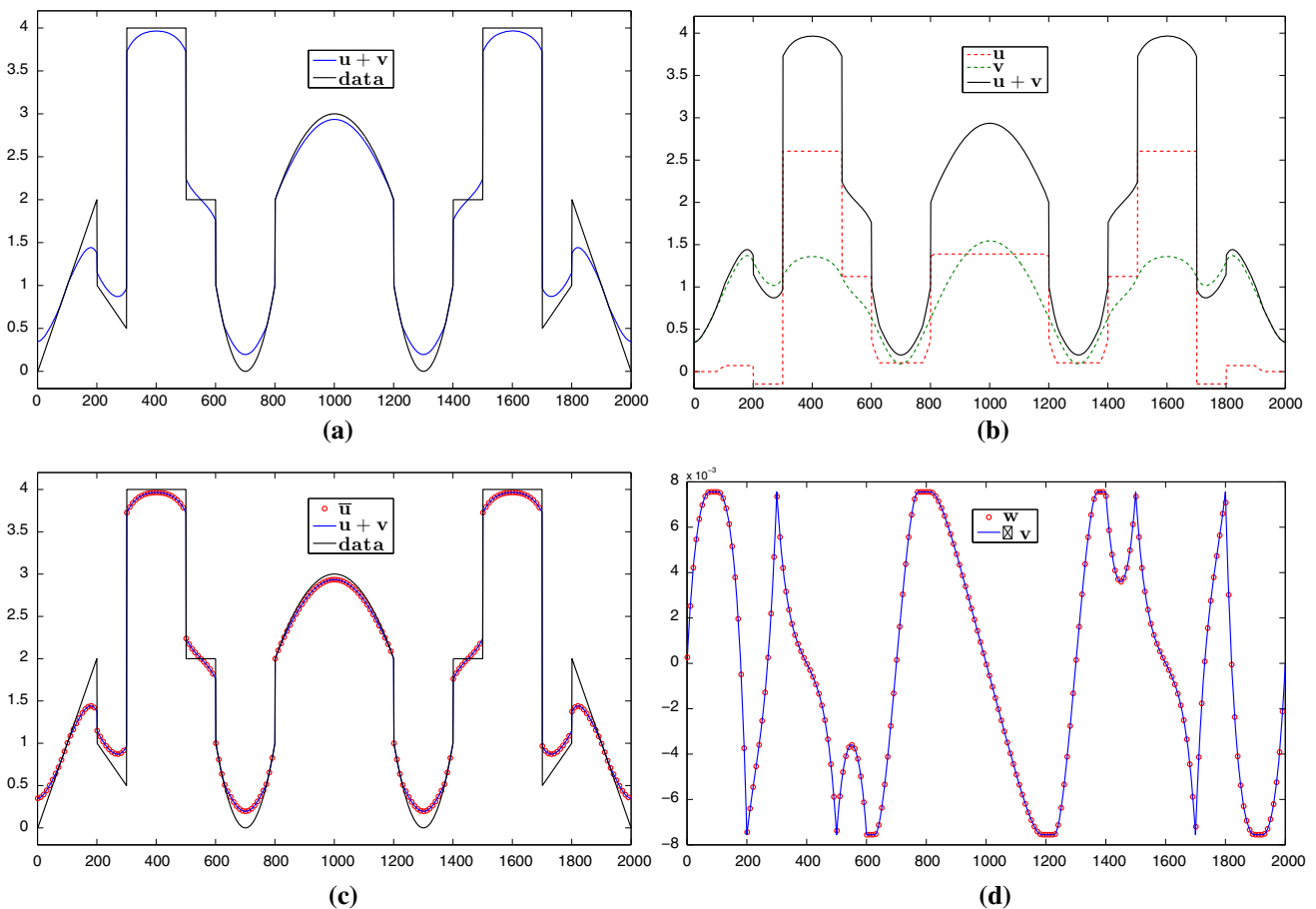


**Fig. 10** Numerical examples of the image decomposition approach (5.1) for $p = 2$. **a** Solution $u + v$ of (5.1). **b** Decomposition into $u$, $v$ parts. **c** Equivalence of ($\mathcal{P}$) and (5.1): $\overline{u} = u + v$. **d** Equivalence of ($\mathcal{P}$) and (5.1): $w = \nabla v$ (Color figure online)
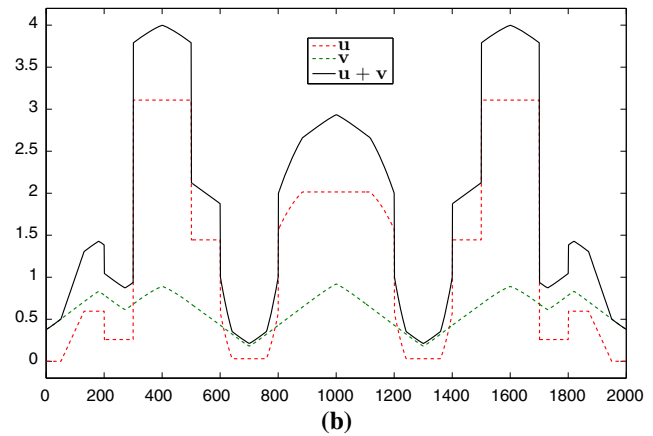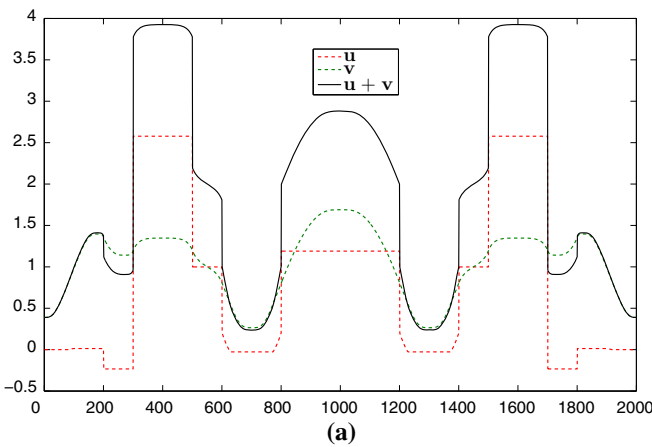
**Fig. 11** Decomposition of the data in Fig. 10a into $u$, $v$ parts for $p = \frac{4}{3}$ and $p = 10$. The value $p = \frac{4}{3}$ produces a $v$ component with flat structures while $p = 10$ produces a component with almost affine structures.

In both cases we have $\|f - u - v\|_2 = 6.667$. **a** Decomposition of the data in Fig. 10a for $p = \frac{4}{3}$. **b** Decomposition of the data in Fig. 10a for $p = 10$ (Color figure online)
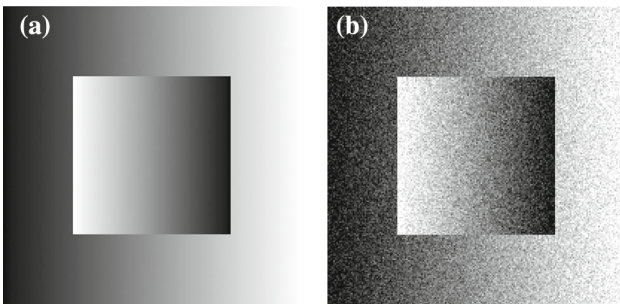


**Fig. 12** Square with piecewise affine structures and its noisy version with $\sigma = 0.01$. **a** Square. **b** Noisy square: PSNR = 20.66 and SSIM = 0.1791

a [0, 1] intensity range. The noisy image, Fig. 12b, is a corrupted version of the original image, Fig. 12a, with Gaussian noise of zero mean and standard deviation $\sigma = 0.01$.

In Fig. 13, we present the best reconstructions results in terms of two quality measures, the classical *Peak Signal to Noise Ratio* (PSNR) and the *Structural Similarity Index* (SSIM), see [41] for the definition of the latter. In each case, the values of $\alpha$ and $\beta$ are selected appropriately for optimal PSNR and SSIM. We use here the split Bregman algorithm as this is described in Sect. 6.1. Our stopping criterion is the relative residual error becoming less than $10^{-6}$, i.e.

$$\frac{\left\|u^{k+1} - u^k\right\|_2}{\left\|u^{k+1}\right\|_2} \leq 10^{-6}. \tag{6.22}$$

For computational efficiency, we set $\lambda = 10\alpha$ when $1 < p < 4$ and $\lambda = 1000\alpha$ when $p \geq 4$ (empirical rule).

Observe that the best reconstructions in terms of the PSNR have no visual difference for $p = \frac{3}{2}$, 2 and 3 and staircasing is present, Fig. 13a–c. This is one more indication that the PSNR—which is based on the squares of the pointwise differences between the ground truth and the reconstruction—does not correspond to the optimal visual results. The best reconstructions in terms of SSIM are visually better. They exhibit significantly reduced staircasing for $p = \frac{3}{2}$ and $p = 3$ and is essentially absent in the case of $p = 2$, see Fig. 13d–f.

We can also get a total staircasing elimination by setting higher values for the parameters $\alpha$ and $\beta$, as we show in Fig. 14. There, one observes that on one hand as we increase $p$, almost affine structures are promoted—see the middle row profiles in Fig. 14—and on the other hand these choices of $\alpha$, $\beta$ produce a serious loss of contrast that however can be easily treated via the *Bregman iteration* that we briefly discuss next.

ing solutions, the parameters $\alpha$, $\beta$ are selected such that the residual $\|f - u - v\|_2$ is the same for both values of $p$. As we observe, the $v$ decomposition with $p = \frac{4}{3}$ exhibits some *flatness* compared to $p = 2$, compare Figs. 10b and 11a. On the other hand for $p = 10$, the $v$ component consists again of almost affine structures, Fig. 11b. Notice, that in both cases the $v$ components are continuous. This is expected since in dimension one, we have $W^{1,p}(\Omega) \subset C(\overline{\Omega})$ for every $1 < p < \infty$.

### 6.3 Two-Dimensional Results

In this section we consider the two-dimensional case where $u \in \mathbb{R}^{n \times m}$, $w \in (\mathbb{R}^{n \times m})^2$ with $m > 1$ and $\Omega$ is a rectangular image domain. We focus on image denoising tasks and on eliminating the staircasing effect for different values of $p$.

We start with the image in Fig. 12, i.e. a square with piecewise affine structures. The image size is $200 \times 200$ pixels at
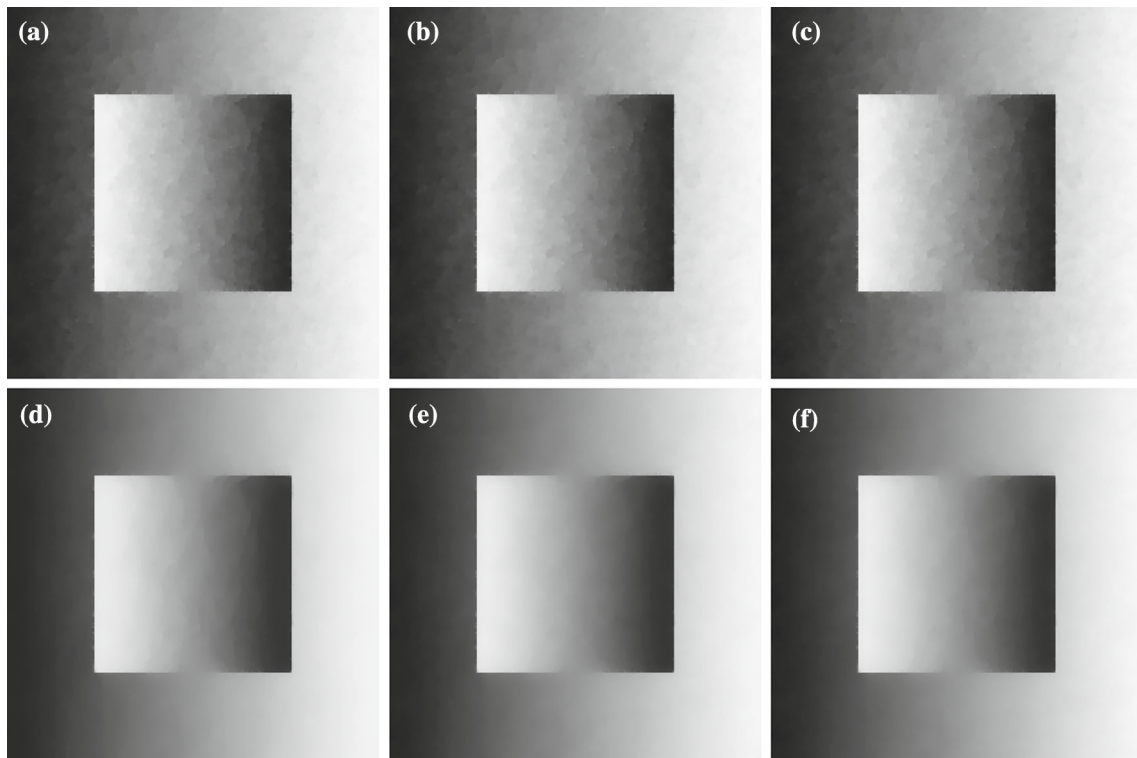
**Fig. 13** Best reconstructions in terms of PSNR and SSIM for $p = \frac{3}{2}$, 2, 3. **a** TVL$^{\frac{3}{2}}$: $\alpha = 0.1$, $\beta = 2.5$, PSNR = 33.63. **b** TVL$^2$: $\alpha = 0.1$, $\beta = 13.5$, PSNR = 33.68. **c** TVL$^3$: $\alpha = 0.1$, $\beta = 76$, PSNR = 33.70. **d** TVL$^{\frac{3}{2}}$: $\alpha = 0.3$, $\beta = 7.7$, SSIM = 0.9669. **e** TVL$^2$: $\alpha = 0.3$, $\beta = 34$, SSIM = 0.9706. **f** TVL$^3$: $\alpha = 0.3$, $\beta = 182$, SSIM = 0.9709
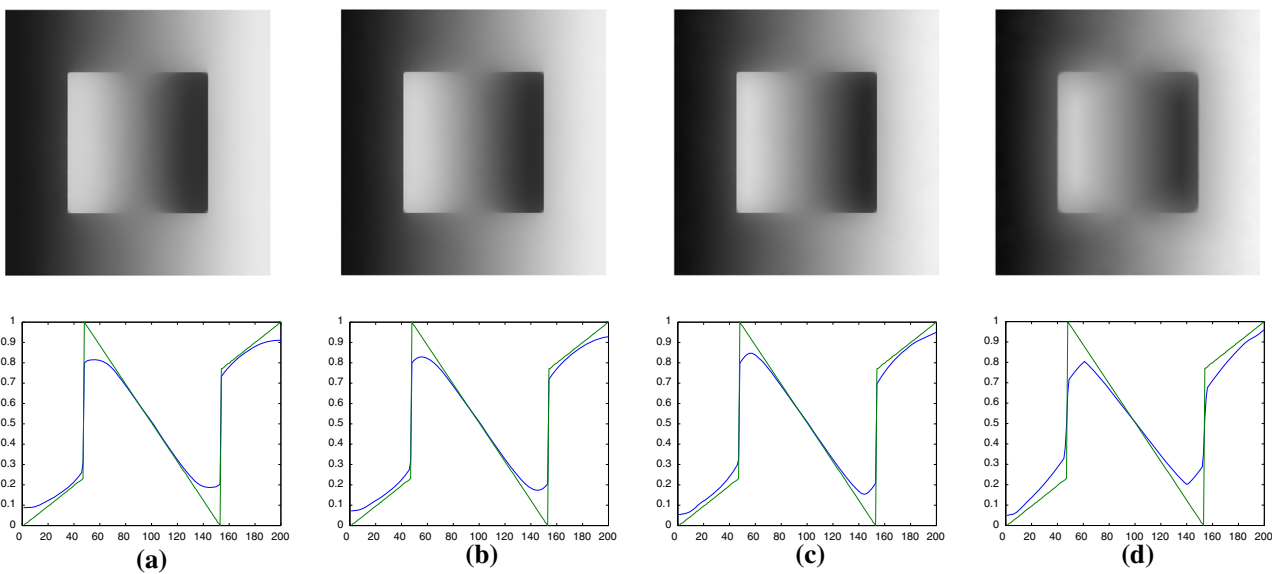


**Fig. 14** Staircasing elimination for $p = \frac{3}{2}$, 2, 3 and 7. High values of $p$ promote almost affine structures. **a** TVL$^{\frac{3}{2}}$: $\alpha = 1$, $\beta = 25$, SSIM = 0.9391. **b** TVL$^2$: $\alpha = 1$, $\beta = 116$, SSIM = 0.9433. **c** TVL$^3$: $\alpha = 1$, $\beta = 438$, SSIM = 0.9430. **d** TVL$^7$: $\alpha = 2$, $\beta = 5000$, SSIM = 0.9001
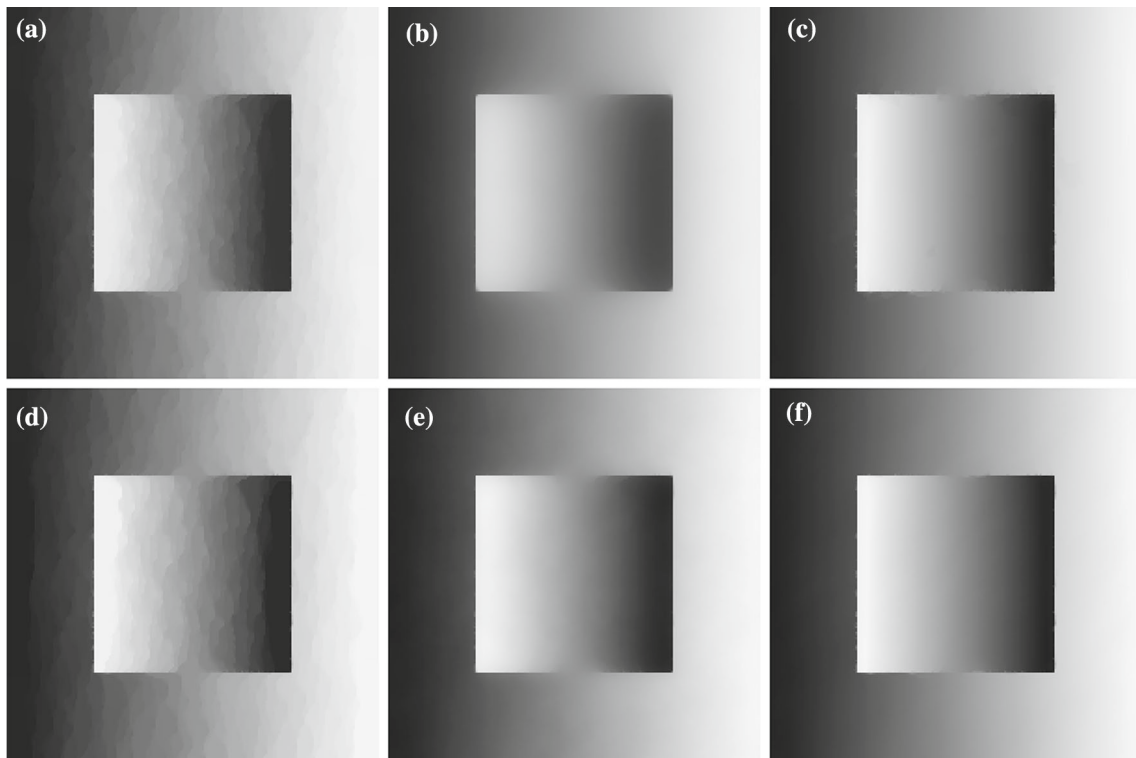
**Fig. 15** *First row* Best reconstructions in terms of SSIM for TV, TVL$^2$ and TGV. *Second row* Best reconstructions in terms of SSIM for the Bregman iteration versions of TV, TVL$^2$ and TGV. **a** TV: $\alpha = 0.2$, SSIM = 0.9387. **b** TVL$^2$: $\alpha = 1$, $\beta = 116$, SSIM = 0.9433. **c** TGV: $\alpha = 0.12$, $\beta = 0.55$, SSIM = 0.9861. **d** TV Bregman iteration: $\alpha = 1$, SSIM = 0.9401, 4th iteration. **e** TVL$^2$ Bregman iteration: $\alpha = 2$, $\beta = 220$, SSIM = 0.9778, 4th iteration. **f** TGV Bregman iteration: $\alpha = 2$, $\beta = 10$, SSIM = 0.9889, 8th iteration

Contrast enhancement via Bregman iteration was introduced in [32], see also [3] for an application to higher-order models. It involves solving a modified version of the minimisation problem. Setting $u^0 = f$, one solves for $k = 1, 2, \ldots$:

$$
u^{k+1} = \underset{\substack{u \in \mathbb{R}^{n \times m} \\ w \in (\mathbb{R}^{n \times m})^2}}{\operatorname{argmin}} \ \frac{1}{2} \left\| f + \tilde{v}^k - u \right\|_2^2 + \alpha \left\| \nabla u - w \right\|_1 + \beta \left\| w \right\|_p,
$$

$$
\tilde{v}^{k+1} = \tilde{v}^k + f - u^{k+1}. \tag{6.23}
$$

Instead of solving (6.1) once for fixed $\alpha$ and $\beta$, we solve a sequence of similar problems adding back a noisy residual in each iteration. For stopping criteria regarding the Bregman iteration we refer to [32].

In Fig. 15 we present our best Bregman iteration results for $p = 2$ in terms of SSIM along with the corresponding TV and TGV results for which the Bregman iteration has also been employed for the sake of fair comparison. We also show the best SSIM results where no Bregman iteration is used. We solve the TGV minimisation using the Chambolle–Pock primal-dual method [10]. We notice that the Bregman iteration version of TVL$^2$ leads to a significant contrast improvement, compare for instance Fig. 15b, e. In
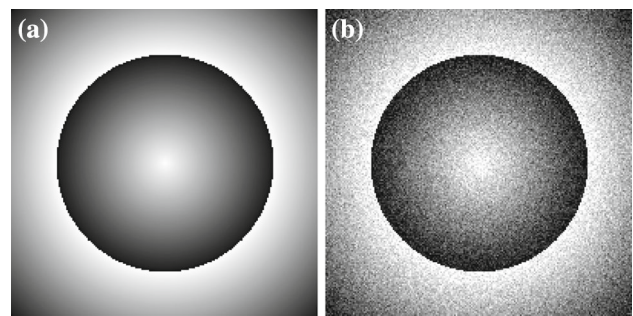


**Fig. 16** Image with symmetric radial structures and its noisy version with $\sigma = 0.01$. **a** Circle. **b** Noisy circle: SSIM = 0.2457

fact, it can achieve a reconstruction which is visually close to the Bregman iteration version of TGV, compare Fig. 15e, f.

We continue our demonstration with a radially symmetric image, see Fig. 16. As in the previous example, we can achieve staircasing-free reconstructions using TVL$^p$ regularisation for different values of $p$, see Fig. 17. In fact, as we increase $p$, we obtain results that preserve the spike in the centre of the circle, see the corresponding middle row slice in Fig. 17d. This provides us with another motivation to examine the $p = \infty$ case in [8]. The loss of contrast can
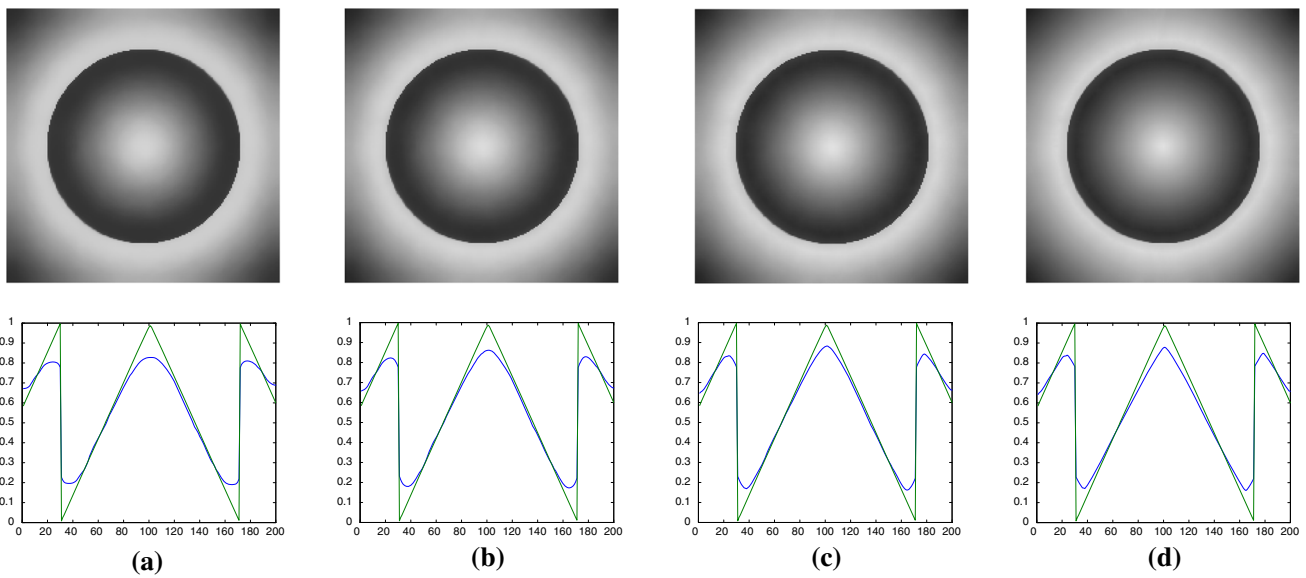
**Fig. 17** Better preservation of spike-like structures for large values of $p$. **a** $\text{TVL}^{\frac{3}{2}}$: $\alpha = 0.8$, $\beta = 17$, SSIM = 0.8909. **b** $\text{TVL}^2$: $\alpha = 0.8$, $\beta = 79$, SSIM = 0.8998. **c** $\text{TVL}^3$: $\alpha = 0.8$, $\beta = 405$, SSIM = 0.9019. **d** $\text{TVL}^7$: $\alpha = 0.8$, $\beta = 3700$, SSIM = 0.9024
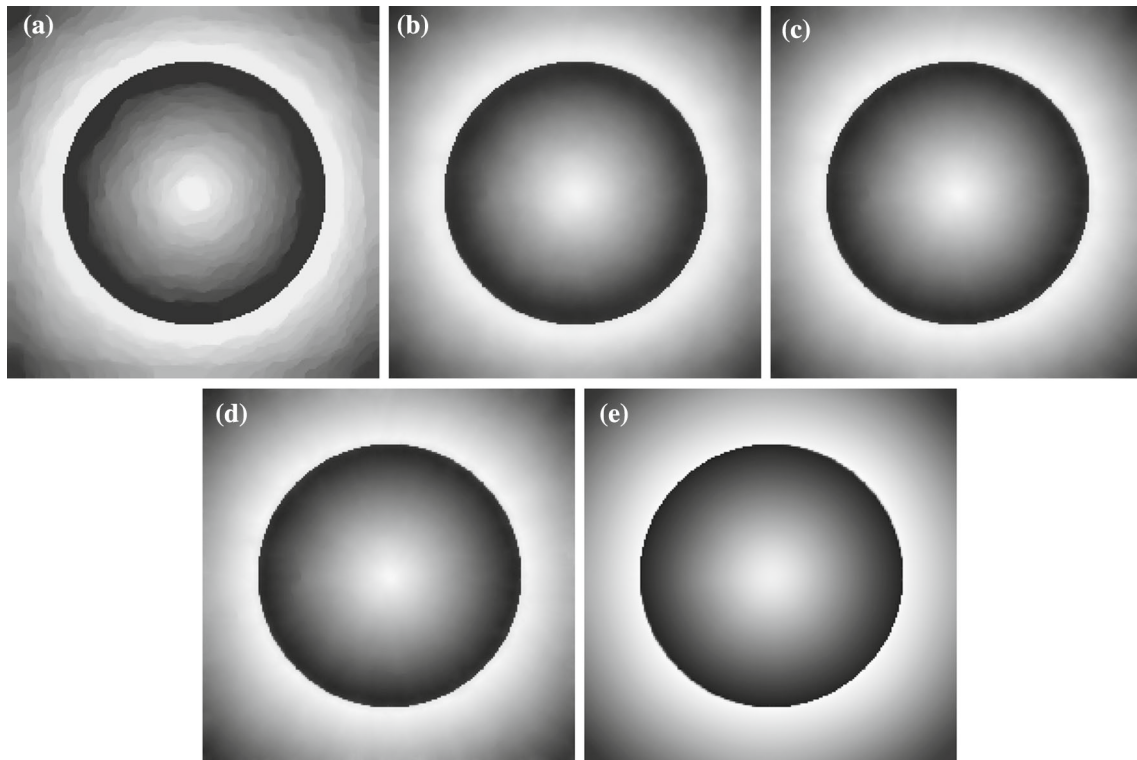


**Fig. 18** Best reconstruction in terms of SSIM for the Bregman iteration versions of $\text{TVL}^2$, $\text{TVL}^4$, $\text{TVL}^7$ and TGV. **a** TV Bregman iteration: $\alpha = 2$, SSIM = 0.8912, 6th iteration. **b** $\text{TVL}^2$ Bregman iteration: $\alpha = 5$, $\beta = 625$, SSIM = 0.9718, 12th iteration. **c** $\text{TVL}^4$ Bregman iteration: $\alpha = 5$, $\beta = 8000$, SSIM=0.9802, 13th iteration. **d** $\text{TVL}^7$ Bregman iteration: $\alpha = 3$, $\beta = 15000$, SSIM = 0.9807, 9th iteration. **e** TGV Bregman iteration: $\alpha = 2$, $\beta = 10$, SSIM = 0.9913, 8th iteration

be again treated using the Bregman iteration (6.23). The best results of the latter in terms of SSIM are presented in Fig. 18, for $p = 2$, 4 and 7 and they are also compared to the corresponding Bregman iteration version of TGV. We observe

that we can obtain reconstructions that are visually close to the TGV ones and in fact notice that for $p = 7$, the spike on the centre of the circle is better reconstructed compared to TGV, see also the surface plots in Fig. 19.
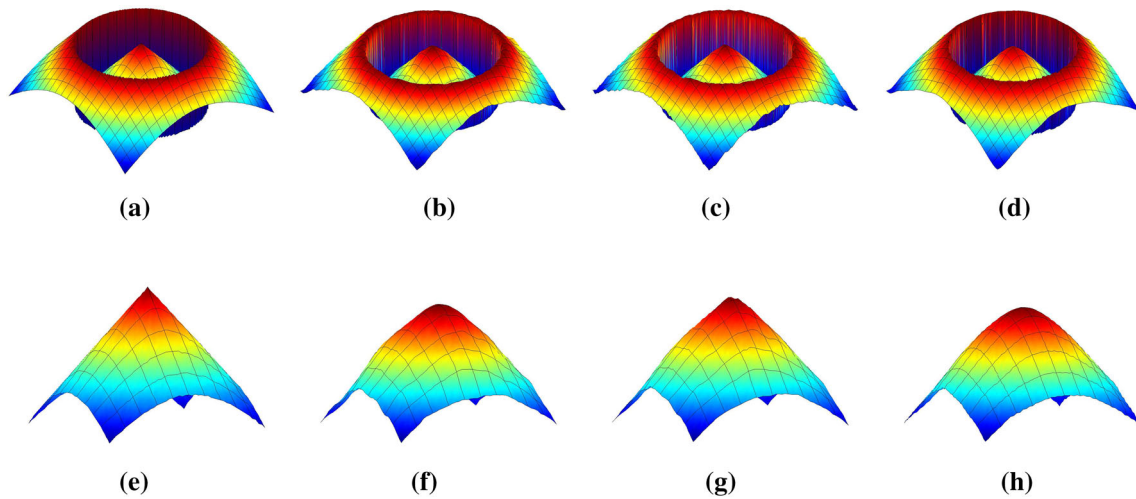
**Fig. 19** Surface plots of the images in Fig. 18. Notice how high values of $p$ can better preserve the sharp spike in the middle of the image. **a** Original. **b** TVL$^2$ Bregman iteration. **c** TVL$^7$ Bregman iteration. **d** TGV Bregman iteration. (e) Original: central part zoom. **f** TVL$^2$ Bregman iteration: central part zoom. **g** TVL$^7$ Bregman iteration: central part zoom. **h** TGV Bregman iteration: central part zoom (Color figure online)
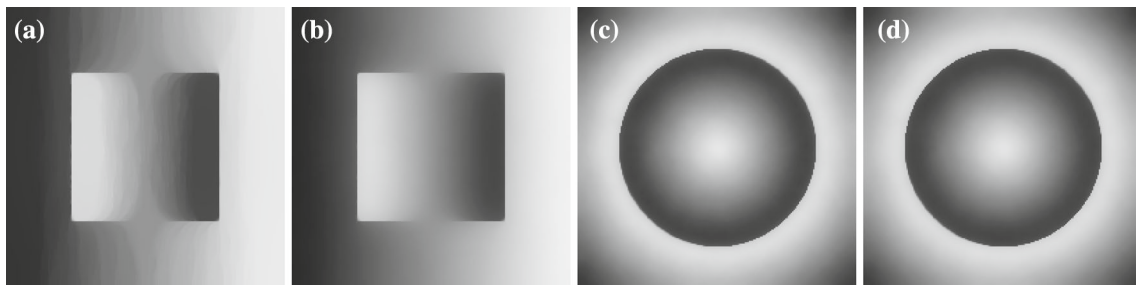


**Fig. 20** Comparison between the model (5.1) for $p = 2$ and TVL$^2$: Staircasing cannot be always eliminated by using the decomposition approach (5.1). **a** Solution $u + v$ of (5.1): $p = 2$, $\alpha = 0.8$, $\beta = 120$, SSIM $= 0.9268$. **b** TVL$^2$: $\alpha = 1$, $\beta = 116$, SSIM $= 0.9433$. **c** Solution $u + v$ of (5.1): $p = 2$, $\alpha = 0.8$, $\beta = 70$, SSIM $= 0.8994$. **d** TVL$^2$: $\alpha = 0.8$, $\beta = 79$, SSIM $= 0.8998$
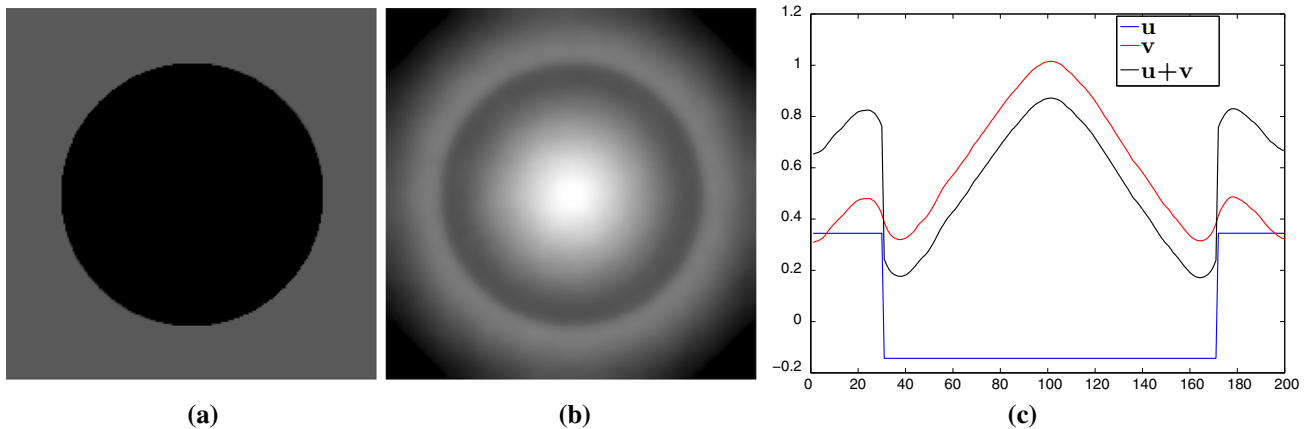


**Fig. 21** Efficient geometric decomposition of the image in Fig. 20c into a piecewise constant and smooth component, $u$ and $v$, respectively, by solving (5.1). **a** Piecewise constant component $u$. **b** Smooth component $v$, **c** Middle row profiles (Color figure online)

We conclude with numerical results for the image decomposition approach of Sect. 5 which we solve again using the split Bregman algorithm. Recall that in dimension two, the solutions of (5.1) are not necessarily the same with the ones of ($\mathcal{P}$). In fact, we observe that (5.1) cannot always eliminate the staircasing, see for instance Figure 20. Even though,

as we have already seen, we can eliminate the staircasing both in the square and in the circle by applying TVL$^p$ regularisation, Fig. 20b, d, we cannot obtain equally satisfactory results by solving (5.1). While using the latter we can get rid of the staircasing in the circle, Fig. 20c, this is not possible for the square, Fig. 20a, where we observe—after extensive experimentation—that no values of $\alpha$ and $\beta$ lead to a staircasing elimination. This is analogous to the difference between TGV and the TV–TV$^2$ infimal convolution of Chambolle–Lions [9].

However, the strength of the formulation (5.1) lies on its ability to efficiently decompose an image into piecewise constant and smooth parts. We depict that in Fig. 21, where we show the components $u$ and $v$ of the result in Fig. 20c.

## 7 Conclusion

We have introduced a novel first-order, one-homogeneous TV–L$^p$ infimal convolution type functional suitable for variational image regularisation. The TVL$^p$ functional constitutes a very general class of regularisation functionals exhibiting diverse smoothing properties for different choices of $p$. In the case $p = 2$ the well-known Huber TV regulariser is recovered.

We studied the corresponding one-dimensional denoising problem focusing on the structure of its solutions. We computed exact solutions of this problem for the case $p = 2$ for simple one-dimensional data. Hence, as an additional novelty in our paper we presented exact solutions of the one-dimensional Huber TV denoising problem.

Numerical experiments for several values of $p$ indicate that our model leads to an elimination of the staircasing effect. We show that we can further enhance our results by increasing the contrast via a Bregman iteration scheme and thus obtaining results of similar quality to those of TGV. Furthermore, as $p$ increases the structure of the solutions changes from piecewise smooth to piecewise linear and the model, in contrast to TGV, is capable of preserving sharp spikes in the reconstruction. This observation motivates a more detailed study of the TVL$^p$ functional for large $p$ and in particular for the case $p = \infty$.

This concludes the first part of the study of the TVL$^p$ model for $p < \infty$. The second part [8], is devoted to the $p = \infty$ case. There we explore further, both in an analytical and an experimental level, the capability of the TVL$^\infty$ model to promote affine and spike-like structures in the reconstructed image and we discuss several applications.

## Appendix: Radon Measures and Functions of Bounded Variation

In what follows $\Omega \subset \mathbb{R}^d$ is an open, bounded set with Lipschitz boundary whose Lebesgue measure is denoted by $|\Omega|$. We denote by $\mathcal{M}(\Omega, \mathbb{R}^d)$ (and $\mathcal{M}(\Omega)$ if $d = 1$) the space of finite Radon measures on $\Omega$. The total variation measure of $\mu \in \mathcal{M}(\Omega, \mathbb{R}^d)$ is denoted by $|\mu|$, while we denote the polar decomposition of $\mu$ by $\mu = \text{sgn}(\mu)|\mu|$, where $\text{sgn}(\mu) = 1$ $|\mu|$-almost everywhere.

Recall that the *Radon norm* of a $\mathbb{R}^d$-valued distribution $\mathcal{T}$ on $\Omega$ is defined as

$$\|\mathcal{T}\|_{\mathcal{M}} := \sup \left\{ \langle u, \phi \rangle : \ \phi \in C_c^\infty(\Omega, \mathbb{R}^d), \ \|\phi\|_\infty \leq 1 \right\}.$$

It can be shown that $\|\mathcal{T}\|_{\mathcal{M}} < \infty$ if and only if $\mathcal{T}$ can be represented by a measure $\mu \in \mathcal{M}(\Omega, \mathbb{R}^d)$ and in that case $\|\mu\|_{\mathcal{M}} = |\mu|(\Omega)$. Regarding the subdifferential of the Radon norm we have that it can be characterised, at least for $C_0$ functions, as follows [6]

$$\partial \|\cdot\|_{\mathcal{M}} (\mu) \cap C_0(\Omega) = \text{Sgn}(\mu) \cap C_0(\Omega), \qquad (8.1)$$

where here $\text{Sgn}(\mu)$ denotes the set-valued sign

$$\text{Sgn}(\mu) = \big\{ v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|) : \|v\|_\infty \leq 1, \\ v = \text{sgn}(\mu), |\mu| - a.e. \big\}. \qquad (8.2)$$

A function $u \in L^1(\Omega)$ is a function of bounded variation if its distributional derivative $Du$ is representable by a finite Radon measure. We denote by $\text{BV}(\Omega)$, the space of functions of bounded variation which is a Banach space under the norm

$$\|u\|_{\text{BV}(\Omega)} := \|u\|_{L^1(\Omega)} + \|Du\|_{\mathcal{M}}.$$

The term

$$\|Du\|_{\mathcal{M}} = \sup \left\{ \int_\Omega u \, \text{div}\phi \, dx : \phi \in C_c^\infty(\Omega, \mathbb{R}^d), \ \|\phi\|_\infty \leq 1 \right\},$$

is called the total variation of $u$, also commonly denoted by TV$(u)$. From the Radon–Nikodym theorem, the measure $\mu$ can be decomposed into an absolutely continuous and a singular part with respect to the Lebesgue measure $\mathcal{L}^d$, that is $Du = D^a u + D^s u$. Here, $D^a u = \nabla u \mathcal{L}^d$, i.e. $\nabla u$ denotes the Radon–Nikodym derivative of $D^a u$ with respect to $\mathcal{L}^d$. Note that we also have $\|Du\|_{\mathcal{M}} = \|D^a u\|_{\mathcal{M}} + \|D^s u\|_{\mathcal{M}}$. When $d = 1$, $\nabla u$ is simply denoted by $u'$. Finally, we recall that BV$(\Omega) \hookrightarrow L^r(\Omega)$ for $1 \le r \le \frac{d}{d-1}$ (resp. $1 \le r < \frac{d}{d-1}$) with continuous embedding (resp. compact embedding) .

Recall also the following basic inequality regarding inclusions of L$^p$ spaces

$$\|h\|_{L^{p_1}(\Omega)} \le |\Omega|^{\frac{1}{p_1} - \frac{1}{p_2}} \|h\|_{L^{p_2}(\Omega)}, \quad 1 \le p_1 < p_2 \le \infty. \tag{8.3}$$

Unless otherwise stated $q$ denotes the Hölder conjugate of the exponent $p$, i.e.

$$q = \begin{cases} \frac{p}{p-1} & \text{if } p \in (1, \infty), \\ 1 & \text{if } p = \infty. \end{cases} \tag{8.4}$$

## References

1. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Science Publications, Oxford (2000)
2. Attouch, H., Brezis, H.: Duality for the sum of convex functions in general Banach spaces. North-Holland Mathematical Library **34**, 125–133 (1986)
3. Benning, M., Brune, C., Burger, M., Müller, J.: Higher-order TV methods—enhancement via Bregman iteration. J. Sci. Comput. **54**(2–3), 269–310 (2013). doi:10.1007/s10915-012-9650-3
4. Bergounioux, M., Piffet, L.: A second-order model for image denoising. Set-Valued Var. Anal. **18**(3–4), 277–306 (2010). doi:10.1007/s11228-010-0156-6
5. Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM J. Imaging Sci. **3**(3), 492–526 (2010). doi:10.1137/090769521
6. Bredies, K., Kunisch, K., Valkonen, T.: Properties of $L^1$-TGV$^2$: the one-dimensional case. J. Math. Anal. Appl. **398**(1), 438–454 (2013). doi:10.1016/j.jmaa.2012.08.053
7. Burger, M., Osher, S.: A guide to the TV zoo. Level Set and PDE Based Reconstruction Methods in Imaging, pp. 1–70. Springer, Berlin (2013)
8. Burger, M., Papafitsoros, K., Papoutsellis, E., Schönlieb, C.-B.: Infimal convolution regularisation functionals of BV and L$^p$ spaces. The case $p = \infty$. Submitted (2015)
9. Chambolle, A., Lions, P.L.: Image recovery via total variation minimization and related problems. Numerische Mathematik **76**, 167–188 (1997). doi:10.1007/s002110050258
10. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011). doi:10.1007/s10851-010-0251-1
11. Chan, T., Marquina, A., Mulet, P.: High-order total variation-based image restoration. SIAM J. Sci. Comput. **22**(2), 503–516 (2001). doi:10.1137/S1064827598344169
12. Chan, T.F., Esedoglu, S., Park, F.E.: Image decomposition combining staircase reduction and texture extraction. J. Vis. Commun. Image Represent. **18**(6), 464–486 (2007). doi:10.1016/j.jvcir.2006.12.004
13. Chen, C., He, B., Ye, Y., Yuan, X.: The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent, Math. Program. 1–23 (2014)
14. Yuan, X., Dai, Y., Han, D., Zhang, W.: A sequential updating scheme of Lagrange multiplier for separable convex programming, Math. Comput., to appear (2013)
15. Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. SIAM, Philadelphia (1999)
16. Elion, C., Vese, L.: An image decomposition model using the total variation and the infinity Laplacian. Proc. SPIE **6498**, 64980W–64980W-10 (2007). doi:10.1117/12.716079
17. Esser, E.: Applications of Lagrangian-based alternating direction methods and connections to split Bregman. CAM Rep. **9**, 31 (2009)
18. Goldstein, T., Osher, S.: The split Bregman algorithm method for $L_1$-regularized problems. SIAM J. Imaging Sci. **2**, 323–343 (2009). doi:10.1137/080725891
19. Grant, M., Boyd, S.: CVX: Matlab software for disciplined convex programming, version 2.1, (2014)
20. Hansen, P.: Discrete inverse problems. Soc. Ind. Appl. Math. (2010). doi:10.1137/1.9780898718836
21. He, B., Hou, L., Yuan, X.: On full jacobian decomposition of the augmented lagrangian method for separable convex programming, preprint (2013)
22. He, B., Tao, M., Yuan, X.: Convergence rate and iteration complexity on the alternating direction method of multipliers with a substitution procedure for separable convex programming, Math. Oper. Res., under revision 2 (2012)
23. Hintermüller, M., Stadler, G.: An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. SIAM J. Sci. Comput. **28**(1), 1–23 (2006). doi:10.1137/040613263
24. Huber, P.J.: Robust regression: asymptotics, conjectures and monte carlo. Ann. Stat. **1**, 799–821 (1973). http://www.jstor.org/stable/2958283
25. Kim, Y., Vese, L.: Image recovery using functions of bounded variation and Sobolev spaces of negative differentiability. Inverse Probl Imaging **3**, 43–68 (2009). doi:10.3934/ipi.2009.3.43
26. Kuijper, A.: P-Laplacian driven image processing. IEEE Int. Conf. Image Process. **5**, 257–260 (2007). doi:10.1109/ICIP.2007.4379814
27. Lefkimmiatis, S., Bourquard, A., Unser, M.: Hessian-based norm regularization for image restoration with biomedical applications. IEEE Trans. Image Process. **21**, 983–995 (2012). doi:10.1109/TIP.2011.2168232
28. Lindqvist, P.: Notes on the $p$-Laplace equation, University of Jyvaskyla (2006)
29. Lysaker, M., Lundervold, A., Tai, X.C.: Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. IEEE Trans. Image Process. **12**(12), 1579–1590 (2003). doi:10.1109/TIP.2003.819229
30. Lysaker, M., Tai, X.C.: Iterative image restoration combining total variation minimization and a second-order functional. Int. J. Comput. Vis. **66**(1), 5–18 (2006). doi:10.1007/s11263-005-3219-7
31. Nien, H., Fessler, J.A.: A convergence proof of the split Bregman method for regularized least-squares problems, arXiv:1402.4371, preprint (2014)
32. Osher, S., Burger, M., Goldfarb, D., Xu, J., Yin, W.: An iterative regularization method for total variation based image restoration. SIAM Multisc. Model. Simul. **4**, 460–489 (2005). doi:10.1137/040605412

33. Papafitsoros, K., Bredies, K.: A study of the one dimensional total generalised variation regularisation problem. Inverse Probl. Imaging **9**, 511–550 (2015). doi:10.3934/ipi.2015.9.511

34. Papafitsoros, K., Schönlieb, C.B.: A combined first and second order variational approach for image reconstruction. J. Math. Imaging Vis. **48**(2), 308–338 (2014). doi:10.1007/s10851-013-0445-4

35. Papafitsoros, K., Valkonen, T.: Asymptotic behaviour of total generalised variation. Scale Space and Variational Methods in Computer Vision **9087**, 702–714 (2015)

36. Papoutsellis, E.: First-order gradient regularisation methods for image restoration. reconstruction of tomographic images with thin structures and denoising piecewise affine images, Ph.D. thesis, University of Cambridge, Cambridge (2015)

37. Ring, W.: Structural properties of solutions to total variation regularisation problems. ESAIM Math. Model. Numer. Anal. **34**, 799–810 (2000). doi:10.1051/m2an:2000104

38. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D: Nonlinear Phenom. **60**, 259–268 (1992). doi:10.1016/0167-2789(92)90242-F

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

40. Vogel, C., Oman, M.: Iterative methods for total variation denoising. SIAM J. Sci. Comput. **17**(1), 227–238 (1996). doi:10.1137/0917016

41. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**, 600–612 (2004). doi:10.1109/TIP.2003.819861

**Konstantinos Papafitsoros** received a 5-year diploma degree from the School of Applied Mathematics and Physical Sciences, National Technical University of Athens, in 2007. He holds two M.Sc. degrees, one in Pure Mathematics from the University of Athens in 2009 and one in Mathematical Modelling and Scientific Computing from the University of Oxford in 2010. He received his Ph.D. degree in 2014 from the University of Cambridge where he also stayed as an EPSRC Doctoral Prize fellow at the Department of Applied Mathematics and Theoretical Physics until April 2015. He is currently a Humboldt foundation postdoctoral fellow at the Institute for Mathematics, Humboldt University of Berlin. His research interests lie in the area of variational methods and higher order regularisation in imaging.
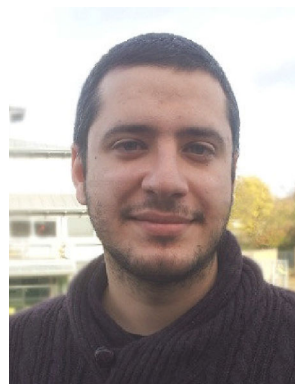


**Evangelos Papoutsellis** completed his undergraduate studies at the School of Applied Mathematics and Physical Sciences in National Technical University of Athens, obtaining a 5 year diploma in June 2010. He continued his studies at a graduate level in University of Cambridge, where he received a Master of Advanced Study in Mathematics (Part III of the Mathematical Tripos), in July 2011 and a Ph.D. degree, in October 2015. His research interests lie in the area of mathematical image processing and medical imaging as well as functional and convex analysis. He is a member of the Department of Applied Mathematics and Theoretical Physics and the Cambridge Image Analysis group at the University of Cambridge.



**Martin Burger** studied Mathematics at Johannes Kepler University Linz, and University of Milano, he received his Ph.D. in Linz in 2000. After periods as postdoc and assistant professor at the University of California, Los Angeles and the Johannes Kepler University he moved to the Westflische Wilhelms-Universitt Mnster in 2006, where he holds the position of a full professor in applied mathematics and leads the Mathematical Imaging Group since then. His research interests are inverse problems and mathematical imaging as well as mathematical models in life and social sciences. He received numerous awards and recognitions for his work, including the Calderon Prize for Inverse Problems 2009 and an ERC Consolidator Grant 2014. Since 2012 he is PI and Board Member of the Cluster of Excellence Cells in Motion in Mnster.



**Carola-Bibiane Schönlieb** graduated from the Institute for Mathematics, University of Salzburg (Austria) in 2004. From 2004 to 2005 she held a teaching position in Salzburg. She received her Ph.D. degree from the University of Cambridge (UK) in 2009. After one year of postdoctoral activity at the University of Gttingen (Germany), she became a Lecturer in Applied and Computational Analysis at the Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge (UK) in 2010, promoted to Reader in 2015. There, she is head of the Cambridge Image Analysis group and since 2011 a fellow of Jesus College Cambridge. Her current research interests range from nonlinear partial differential equations to computational- and convex analysis, with applications in digital image processing and inverse problems.