Asymptotic behaviour of total generalised variation

Konstantinos Papafitsoros and Tuomo Valkonen

Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, United Kingdom

Abstract. The recently introduced second order total generalised variation functional $\mathrm{TGV}^2_{\beta,\alpha}$ has been a successful regulariser for image processing purposes. Its definition involves two positive parameters α and β whose values determine the amount and the quality of the regularisation. In this paper we report on the behaviour of $\mathrm{TGV}^2_{\beta,\alpha}$ in the cases where the parameters α, β as well as their ratio β/α becomes very large or very small. Among others, we prove that for sufficiently symmetric two dimensional data and large ratio β/α, $\mathrm{TGV}^2_{\beta,\alpha}$ regularisation coincides with total variation (TV) regularisation.

Keywords: Total Variation, Total Generalised Variation, Regularisation Parameters, Asymptotic Behaviour of Regularisers.

1 Introduction

Parameterisation of variational image processing models has not yet been solved to full satisfaction. Towards the better understanding of such models, we study the behaviour of their solutions as the parameters change. Within the constraints of these proceedings, we concentrate in particular on the asymptotic behaviour of total generalised variation [2].

In the variational image reconstruction approach, one typically tries to recover an improved version u of a corrupted image f as a solution of a minimisation problem of the type

$$\min_u \ \Phi(f, Tu) + \Psi(u), \tag{1.1}$$

where T is a linear operator that models the type of corruption. Here the term $\Phi(f, Tu)$ ensures the fidelity of the reconstruction to the initial data. The term $\Psi(u)$, the *regulariser*, imposes extra regularity on u and it is responsible for the overall quality of the reconstruction. The two terms are balanced by one or more parameters within Ψ. A typical example is $\Psi(u) = \alpha \mathrm{TV}(u)$, i.e., the total variation of u weighted by a positive parameter α [5,11]. While total variation regularisation leads to image reconstructions with sharp edges, it also promotes piecewise constant structures leading to the *staircasing effect*. The second order total generalised variation $\mathrm{TGV}^2_{\beta,\alpha}$ [2] resolves that issue by optimally balancing first and second order information in the image data. The $\mathrm{TGV}^2_{\beta,\alpha}$ functional

reads

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

where $\|\cdot\|_{\mathcal{M}}$ is the Radon norm, $\mathrm{BD}(\Omega)$ is the space of functions of bounded deformation in the domain Ω, $\mathcal{E}$ is the symmetrised gradient and $\alpha, \beta > 0$.

Since the values of α and β determine the amount and the quality of the reconstruction, it is important to understand their role in the regularisation process. In this paper we study the asymptotic behaviour of $\mathrm{TGV}^2_{\beta,\alpha}$ regularised solutions for the extremal cases, i.e., for large and small values of α, β and their ratio β/α. For simplicity we focus on the case where $\Phi(f, Tu) = \|f - u\|^2_{L^2(\Omega)}$ but in most cases, our results can be extended to more general fidelities.

Summary of our results: In Section 3.1 we show that as long as at least one of the parameters α, β is going to zero then the $\mathrm{TGV}^2_{\beta,\alpha}$ solutions converges to the data f. In one dimension we obtain a stronger result, showing in addition that for small values of β the solutions are continuous. In Section 3.2 we focus on the case when the ratio β/α is large, proving that in this regime $\mathrm{TGV}^2_{\beta,\alpha}$ is equivalent to TV modulo "an affine correction". In Section 3.3 we show that by setting the values of α and β large enough we obtain the linear regression of the data as a solution. In Section 3.4, we exploit the result of Section 3.2 and we show that for sufficiently symmetric data and large β/α, $\mathrm{TGV}^2_{\beta,\alpha}$ is equal to αTV. Our paper is furnished with some numerical experiments in Section 3.5, which verify our analytical results.

2   Preliminaries and Notation

In this section we briefly review the basic theory of functions of bounded variation, properties of TV and $\mathrm{TGV}^2_{\beta,\alpha}$ and we also fix our notation.

Let Ω be an open, bounded domain in $\mathbb{R}^d$. A function $u \in L^1(\Omega)$ is a *function of bounded variation* ($u \in \mathrm{BV}(\Omega)$) if its distributional derivative Du is represented by an $\mathbb{R}^d$–valued finite Radon measure. The total variation of u is defined as $\mathrm{TV}(u) = \|Du\|_{\mathcal{M}}$, where $\|\mathcal{T}\|_{\mathcal{M}}$ denotes the *Radon norm* of an $\mathbb{R}^\ell$–valued distribution $\mathcal{T}$ in Ω:

$$\|\mathcal{T}\|_{\mathcal{M}} := \sup\left\{\langle \mathcal{T}, v\rangle : \ v \in \mathcal{C}^\infty_c(\Omega; \mathbb{R}^\ell), \ \|v\|_\infty \le 1\right\}, \tag{2.1}$$

and it is equal to the total variation $|Du|(\Omega)$ of the measure Du when $u \in \mathrm{BV}(\Omega)$. The measure Du can be decomposed into the absolutely continuous and singular part with respect to the Lebesgue measure $\mathcal{L}^d$, $Du = D^a u + D^s u = \nabla u \mathcal{L}^d + D^s u$, where ∇u is the Radon-Nikodým derivative $D^a u/\mathcal{L}^d$. The space $\mathrm{BV}(\Omega)$ is a Banach space endowed with the norm $\|u\|_{\mathrm{BV}(\Omega)} = \|u\|_{L^1(\Omega)} + \|Du\|_{\mathcal{M}}$. We refer the reader to [1] for a complete account on the functions of bounded variation.

Analogously we define the space of *functions of bounded deformation* $\mathrm{BD}(\Omega)$ as the set of all the $L^1(\Omega; \mathbb{R}^d)$ functions whose symmetrised distributional derivative $\mathcal{E}u$ is represented by an $\mathbb{R}^{d \times d}$–valued finite Radon measure [12]. Notationwise one can readily check that $\|\mathcal{E}u\|_{\mathcal{M}} = |\mathcal{E}u|(\Omega)$. The space $\mathrm{BV}(\Omega)$ is strictly contained in $\mathrm{BD}(\Omega)$ for $d > 1$ while $\mathrm{BD}(\Omega) = \mathrm{BV}(\Omega)$ for one dimensional domains Ω. We are not going to need much of the theory of BD functions apart

from the so-called *Sobolev–Korn inequality*. The latter says that if Ω has a Lipschitz boundary then there exists a constant $C_{\mathrm{BD}} > 0$ that depends only on Ω such that for every $w \in \mathrm{BD}(\Omega)$ there exists an element $r_w \in \mathrm{Ker}\mathcal{E}$ such that

$$\|w - r_w\|_{L^1(\Omega)} \leq C_{\mathrm{BD}}\|\mathcal{E}w\|_{\mathcal{M}}. \tag{2.2}$$

Here the kernel of $\mathcal{E}$ consists of all the functions of the form $r(x) = Ax + b$, where $b \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d \times d}$ is a skew symmetric matrix.

The *second order total generalised variation* $\mathrm{TGV}^2_{\beta,\alpha}(u)$ of a function $u \in L^1(\Omega)$ is defined as [2,3,4]

$$\mathrm{TGV}^2_{\beta,\alpha}(u) = \min_{w \in \mathrm{BD}(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}}, \tag{2.3}$$

for $\alpha, \beta > 0$. The above definition is usually referred to as the *differentiation cascade* definition of $\mathrm{TGV}^2_{\beta,\alpha}$, see [2] for the original formulation. It can be shown that $\mathrm{TGV}^2_{\beta,\alpha}$ is a seminorm and together with $\|\cdot\|_{L^1(\Omega)}$ form a norm equivalent to $\|\cdot\|_{\mathrm{BV}(\Omega)}$ [4], i.e., there exist constants $0 < c < C$ that depend only on Ω such that for every u with $\mathrm{TGV}^2_{\beta,\alpha}(u) < \infty$

$$c\|u\|_{\mathrm{BV}(\Omega)} \leq \|u\|_{L^1(\Omega)} + \mathrm{TGV}^2_{\beta,\alpha}(u) \leq C\|u\|_{\mathrm{BV}(\Omega)}. \tag{2.4}$$

Notice that the optimal w in (2.3) is not unique in general. In fact w is a solution of an L^1–$\|\mathcal{E}\|_{\mathcal{M}}$ problem (not strictly convex). Indeed since $\|Du\|_{\mathcal{M}} = \|D^a u\|_{\mathcal{M}} + \|D^s u\|_{\mathcal{M}}$, we have:

$$w \in \operatorname*{argmin}_{w \in \mathrm{BD}(\Omega)} \alpha\|Du - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}} \iff$$

$$w \in \operatorname*{argmin}_{w \in \mathrm{BD}(\Omega)} \int_\Omega |\nabla u - w| \, dx + \frac{\beta}{\alpha}\|\mathcal{E}u\|_{\mathcal{M}}. \tag{2.5}$$

In the following sections, we will take specific advantage of the fact that w solves (2.5), a problem which can be seen as an analogous one to L^1–TV minimisation.

Let us finally mention that properties of $\mathrm{TGV}^2_{\beta,\alpha}$ regularisation have been studied in the one dimensional case and when $\Phi(f, Tu) = \frac{1}{p}\|f - u\|^p_{L^p(\Omega)}$ for $p = 1$ or 2, in [3,9,10].

3 Asymptotic behaviour

3.1 $\beta \to 0$ while α is fixed and $\alpha \to 0$ while β is fixed

In this section we study the limiting behaviour of $\mathrm{TGV}^2_{\beta,\alpha}$ regularisation for small values of α, β. We first prove that by fixing α or β and sending β or α to zero respectively, then the regularised $\mathrm{TGV}^2_{\beta,\alpha}$ solution converges to the data f. For simplicity we work on the L^2–$\mathrm{TGV}^2_{\beta,\alpha}$ denoising problem, i.e., $T = Id$, but the next result can be extended in the more general case e.g. when the fidelity term reads $\frac{1}{p}\|f - Tu\|^p$, with $p \geq 1$ and T being a bounded, linear operator $T : L^p(\Omega) \to L^p(\Omega)$. For convenience we set

$$(u_{\beta,\alpha}, w_{\beta,\alpha}) = \operatorname*{argmin}_{\substack{u \in \mathrm{BV}(\Omega) \\ w \in \mathrm{BD}(\Omega)}} \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \alpha\|Du - w\|_{\mathcal{M}} + \beta\|\mathcal{E}w\|_{\mathcal{M}}. \tag{3.1}$$

Proposition 1 *Let $\Omega \subseteq \mathbb{R}^d$, open and bounded and $f \in L^2(\Omega) \cap \mathrm{BV}(\Omega)$. Then*

(i) Fixing $\alpha > 0$ we have that $\|f - u_{\beta,\alpha}\|^2_{L^2(\Omega)} \to 0$ as $\beta \to 0$.

(ii) Fixing $\beta > 0$ we have that $\|f - u_{\beta,\alpha}\|^2_{L^2(\Omega)} \to 0$ as $\alpha \to 0$.

Proof. (i) Let $\epsilon > 0$ and $\{\rho_\delta\}_{\delta>0}$ be a standard family of mollifiers, i.e., $\rho_\delta(x) = \delta^{-d}\rho(x/\delta)$, where $\rho \in \mathcal{C}^\infty_c(\mathbb{R}^d)$, and set $f_\delta := \rho_\delta * f$. Because $(u_{\beta,\alpha}, w_{\beta,\alpha})$ is an optimal pair in (3.1) by setting $u = f_\delta$ and $w = \nabla f_\delta$ we have the following estimates, for some constant $C > 0$

$$\frac{1}{2}\|f - u_{\beta,\alpha}\|^2_{L^2(\Omega)} \le \frac{1}{2}\|f - u_{\beta,\alpha}\|^2_{L^2(\Omega)} + \alpha\|Du_{\beta,\alpha} - w_{\beta,\alpha}\|_{\mathcal{M}} + \beta\|\mathcal{E}w_{\beta,\alpha}\|_{\mathcal{M}}$$

$$\le \frac{1}{2}\|f - f_\delta\|^2_{L^2(\Omega)} + \beta\|\mathcal{E}(\nabla f_\delta)\|_{\mathcal{M}}$$

$$\le \frac{1}{2}\|f - f_\delta\|^2_{L^2(\Omega)} + \beta\|\mathcal{E}\rho_\delta * Df\|_{\mathcal{M}}$$

$$\le \frac{1}{2}\|f - f_\delta\|^2_{L^2(\Omega)} + \beta\frac{1}{\delta}\|Df\|_{\mathcal{M}}.$$

We set δ small enough such that $\|f - f_\delta\|^2_{L^2(\Omega)} \le \epsilon/2$. By choosing $\beta < \delta\epsilon/2\|Df\|_{\mathcal{M}}$, the result follows.

(ii) The proof is very similar to the (i) case, by setting $u = f_\delta$ and $w = 0$, instead. $\square$

Remark: Of course in both (i)–(ii) cases of Proposition 1, we also get $\|Du_{\beta,\alpha} - w_{\beta,\alpha}\|_{\mathcal{M}} \to 0$ as well as $\|\mathcal{E}w_{\beta,\alpha}\|_{\mathcal{M}} \to 0$ as $\beta \to 0$ or $\alpha \to 0$.

Another interesting behaviour occurs when $\beta \to 0$. In [13], it is proved that for an arbitrary dimension and a fixed $\alpha > 0$ we have

$$\|D^s u_{\beta,\alpha}\|_{\mathcal{M}} \to 0, \quad \text{as } \beta \to 0.$$

However it turns out that in dimension one we are able to prove something stronger, provided the data are bounded:

Proposition 2 *Let $\Omega = (a,b) \subseteq \mathbb{R}$, $f \in L^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ and $\alpha > 0$. Then there exists a threshold $\beta^* > 0$ such that for every $\beta < \beta^*$ we have that*

$$\|D^s u_{\beta,\alpha}\|_{\mathcal{M}} = 0 \quad \text{and} \quad w_{\beta,\alpha} = \nabla u_{\beta,\alpha}.$$

In particular this means that for $\beta < \beta^*$

$$u_{\beta,\alpha} = \operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \beta\|D^2 u\|_{\mathcal{M}}. \tag{3.2}$$

Proof. From the optimality conditions derived in [9], we have that $(u_{\beta,\alpha}, w_{\beta,\alpha})$ solve (3.1) if and only if there exists a dual variable $v \in H^2_0(\Omega)$ such that

$v'' = f - u_{\beta,\alpha}$ (C_f), $-v' \in \alpha\mathrm{Sgn}(Du_{\beta,\alpha} - w_{\beta,\alpha})$ (C_α), $v \in \beta\mathrm{Sgn}(Dw_{\beta,\alpha})$ (C_β),

where for a finite Radon measure μ we define

$$\mathrm{Sgn}(\mu) := \left\{ v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|) : \|v\|_\infty \le 1, \; v = \frac{\mu}{|\mu|}, \; |\mu| - a.e. \right\}.$$

Note also that there exists a constant C depending only on Ω such that the following interpolation inequality holds [8, Section 5.10, ex. 9]

$$\|Dv\|_{L^2(\Omega)} \le C\|v\|^{1/2}_{L^2(\Omega)}\|D^2v\|^{1/2}_{L^2(\Omega)}, \quad \text{for all } v \in H^2_0(\Omega). \tag{3.3}$$

Observe first that (denoting this dual function v by $v_{\beta,\alpha}$)

$$\|Dv_{\beta,\alpha}\|_{L^2(\Omega)} \to 0 \quad \text{as } \beta \to 0. \tag{3.4}$$

Indeed, from Proposition 1 and condition (C_f) we have that $\|D^2v_{\beta,\alpha}\|_{L^2(\Omega)} \to 0$ while from condition (C_β) we have that $\|v_{\beta,\alpha}\|_\infty \to 0$ and thus $\|v_{\beta,\alpha}\|_{L^2(\Omega)} \to 0$ as $\beta \to 0$. Then we just apply the estimate (3.3).

From the fact that we are in dimension one and from (2.4) we have for a generic constant C

$$\begin{aligned}
\|u_{\beta,\alpha}\|_{L^\infty(\Omega)} \leq C\|u_{\beta,\alpha}\|_{\mathrm{BV}(\Omega)} &\leq C(\|u_{\beta,\alpha}\|_{L^2(\Omega)} + \mathrm{TGV}^2_{\beta,\alpha}(u_{\beta,\alpha})) \\
&\leq C(\|f - u_{\beta,\alpha}\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \mathrm{TGV}^2_{\beta,\alpha}(u_{\beta,\alpha})) \\
&\leq C(\|f\|_{L^2(\Omega)}, \mathrm{TGV}^2_{\beta,\alpha}(f)) := M,
\end{aligned}$$

which in combination with (C_f) and the fact that $f \in L^\infty(\Omega)$ implies that $\|D^2v_{\beta,\alpha}\|_\infty \leq M$. Thus from the Arzelà-Ascoli theorem we get the existence of a sequence $\beta_n \to 0$ and a continuous function $\tilde{v}$ such that $v_{\beta_n,\alpha} \to \tilde{v}$ uniformly. We immediately deduce using (3.4) that $v_{\beta,\alpha} \to 0$ uniformly as $\beta \to 0$. But then condition (C_α) implies that there must exist a β_0 such that for every $\beta < \beta_0$ we have $Du_{\beta,\alpha} = w_{\beta,\alpha}$, as measures, i.e., $D^su_{\beta,\alpha} = 0$ and $w_{\beta,\alpha} = \nabla u_{\beta,\alpha}$ since otherwise there would exist a point $x_{\beta_n,\alpha} \in (a,b)$ with $Dv_{\beta_n,\alpha}(x_{\beta_n,\alpha}) = \alpha$ for a sequence $(\beta_n)_{n\in\mathbb{N}}$ converging to 0, a contradiction. □

Remark: We believe that the above proof sets the basis for an analogue proof in higher dimensions even though admittedly this is a hard task. That would require an interpolation inequality for v, $\mathrm{div}v$ and div^2v analogous to (3.3), as well as a proof that the $\mathrm{TGV}^2_{\beta,\alpha}$ regularised solution remains bounded, for bounded data f.

3.2   Large ratio β/α

Recall from (2.5) that the optimal w is a solution to a L^1–$\|\mathcal{E}\|_\mathcal{M}$ type of problem. This motivates us to study some particular properties of the general form of such a problem:

$$\min_{w\in\mathrm{BD}(\Omega)} \|g - w\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w\|_\mathcal{M}, \quad g \in L^1(\Omega;\mathbb{R}^d),\ \lambda > 0. \tag{3.5}$$

The next theorem states that if the parameter λ is larger than a certain threshold (depending only on Ω) then a solution w of (3.5) will belong to Ker$\mathcal{E}$. This is analogous to the L^1–TV problem [6,7], where there for large enough value of the parameter λ, the solution is constant, i.e., belongs to the kernel of TV.

Proposition 3 *Let $\Omega \subseteq \mathbb{R}^d$ be an open, bounded set with Lipschitz boundary, $g \in L^1(\Omega;\mathbb{R}^d)$ and C_{BD} the constant that appears in the Sobolev–Korn inequality (2.2). Then if $\lambda > C_{\mathrm{BD}}$ and w_λ is a solution of (3.5) with parameter λ, then*

$$w_\lambda = m_\mathcal{E}(g) := \operatorname*{argmin}_{w\in\mathrm{Ker}\mathcal{E}} \|g - w\|_{L^1(\Omega;\mathbb{R}^d)}. \tag{3.6}$$

Proof. Since w_λ is a solution of (3.5), it is easy to check that if r_{w_λ} is the element of Ker$\mathcal{E}$ that corresponds to w_λ in the Sobolev-Korn inequality then,

$W_\lambda := w_\lambda - r_{w_\lambda}$ solves the following problem:

$$\min_{w \in \mathrm{BD}(\Omega)} \|(g - w_\lambda) - w\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w\|_{\mathcal{M}}. \qquad (3.7)$$

Indeed, we have for an arbitrary $w \in \mathrm{BD}(\Omega)$

$$\|(g - r_{w_\lambda}) - W_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}W_\lambda\|_{\mathcal{M}} \leq \|(g - r_{w_\lambda}) - w\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w\|_{\mathcal{M}},$$

$$\Longleftrightarrow$$

$$\|g - w_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w_\lambda\|_{\mathcal{M}} \leq \|g - (w + r_{w_\lambda})\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}(w + r_{w_\lambda})\|_{\mathcal{M}},$$

with the latter being true since

$$\|g - w_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w_\lambda\|_{\mathcal{M}} \leq \|g - w\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}w\|_{\mathcal{M}}, \quad \forall w \in \mathrm{BD}(\Omega).$$

Since W_λ solves (3.7), setting $G_\lambda := g - w_\lambda$ we have that

$$\|G_\lambda - W_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} + \lambda\|\mathcal{E}W_\lambda\|_{\mathcal{M}} \leq \|G_\lambda\|_{L^1(\Omega;\mathbb{R}^d)},$$

and using the Sobolev–Korn inequality $\|W_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} \leq C_{\mathrm{BD}}\|\mathcal{E}W_\lambda\|_{\mathcal{M}}$ we have

$$\|G_\lambda - W_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} + \frac{\lambda}{C_{\mathrm{BD}}}\|W_\lambda\|_{L^1(\Omega;\mathbb{R}^d)} \leq \|G_\lambda\|_{L^1(\Omega;\mathbb{R}^d)}. \qquad (3.8)$$

A simple application of the triangle inequality in (3.8) yields that if $\lambda > C_{\mathrm{BD}}$, then we must have $W_\lambda = 0$, i.e., $w_\lambda = r_{w_\lambda}$ from which (3.6) straightforwardly follows. $\qquad\square$

The notation $m_\mathcal{E}(g)$ can be interpreted as the median of g with respect to $\mathrm{Ker}\mathcal{E}$. If $d = 1$, then this is nothing else than the usual median since in that case $\mathrm{Ker}\mathcal{E}$ consists of all the constant functions. The following corollary follows immediately from (2.5) and Proposition (3). It says that for large β/α, $\mathrm{TGV}^2_{\beta,\alpha}$ is almost equivalent to TV up to an "affine correction".

Corollary 4 *Let $\Omega \subseteq \mathbb{R}^d$ be an open, bounded set with Lipschitz boundary and let $\alpha, \beta > 0$ such that $\beta/\alpha > C_{\mathrm{BD}}$. Then for every $u \in \mathrm{BV}(\Omega)$*

$$\mathrm{TGV}^2_{\beta,\alpha}(u) = \alpha\|Du - m_\mathcal{E}(\nabla u)\|_{\mathcal{M}}.$$

3.3   Thresholds for regression

In this section we show that there exist some thresholds for α and β above which the solution to the L^2–$\mathrm{TGV}^2_{\beta,\alpha}$ regularisation problem is the L^2-linear regression of the data f, denoted by $f^\star$:

$$f^\star := \operatorname*{argmin}_{\phi \text{ affine}} \|f - \phi\|^2_{L^2(\Omega)}.$$

We are going to need the following proposition proved in [4]:

Proposition 5 ([4, Proposition 4.1]) *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded, open set with Lipschitz boundary. Then for every $1 \leq p \leq d/(d-1)$, there exists a constant $C_{\mathrm{BGV}}(\beta/\alpha) > 0$, that depends only on Ω, p and the ratio β/α such that*

$$\|u - u^\star\|_{L^p(\Omega)} \leq C_{\mathrm{BGV}}(\beta/\alpha)\mathrm{TGV}^2_{\beta/\alpha,1}(u). \qquad (3.9)$$

In the next proposition we show the existence of these regression thresholds for $d = 2$ and also for $d > 2$ under the extra assumption that the L^p norm of the data f controls the L^p norm of the solution for some $p \in [d, \infty]$.

Proposition 6 *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded, open set with Lipschitz boundary. Suppose that either*

* *(i) $d = 2$ and $f \in \mathrm{BV}(\Omega)$ or*
* *(ii) $d > 2$, $f \in L^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ and there exists a constant $C > 0$ depending only on the domain and $p \in [d, \infty]$ such that $\|u\|_{L^p(\Omega)} \le C\|f\|_{L^p(\Omega)}$ for u solution to (3.1),*

then there exist $\alpha^\star, \beta^\star > 0$ such that whenever $\alpha > \alpha^\star$, $\beta > \beta^\star$ then the solution to the L^2–$\mathrm{TGV}^2_{\beta,\alpha}$ regularisation problem is equal to $f^\star$.

Proof. Suppose initially that $d = 2$ and $f \in \mathrm{BV}(\Omega)$. Then using the Hölder inequality along with (3.9) and the fact that any function $u \in \mathrm{BV}(\Omega)$ that solves the L^2–$\mathrm{TGV}^2_{\beta,\alpha}$ problem has a L^2 norm bounded by a constant C depending only on f and not on α, β

$$
\frac{1}{2}\|f - f^\star\|^2_{L^2(\Omega)} = \min_{\phi \text{ affine}} \frac{1}{2}\|f - \phi\|^2_{L^2(\Omega)} \le \frac{1}{2}\|f - u^\star\|^2_{L^2(\Omega)}
$$

$$
= \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \frac{1}{2}\|u - u^\star\|^2_{L^2(\Omega)} + \int_\Omega (f - u)(u - u^\star)dx
$$

$$
\le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + C(f)\|u - u^\star\|_{L^2(\Omega)} \qquad (3.10)
$$

$$
\le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + C(f)C_{\mathrm{BGV}}(\beta/\alpha)\mathrm{TGV}^2_{\beta/\alpha,1}(u).
$$

Setting $\alpha^\star = C(f)C_{\mathrm{BGV}}(1)$ and $\beta^\star = \alpha^\star$ we have that if $\alpha > \alpha^\star$ and $\beta > \beta^\star$ we can further estimate

$$
\frac{1}{2}\|f - f^\star\|^2_{L^2(\Omega)} \le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + C(f)C_{\mathrm{BGV}}(1)\mathrm{TGV}^2_{\beta^\star/\alpha^\star,1}(u)
$$

$$
\le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \alpha^\star\mathrm{TGV}^2_{\beta^\star/\alpha^\star,1}(u)
$$

$$
\le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \mathrm{TGV}^2_{\beta^\star,\alpha^\star}(u)
$$

$$
\le \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + \mathrm{TGV}^2_{\beta,\alpha}(u).
$$

The proof goes through for the case (*ii*) as well, where the only difference is that Hölder inequality in (3.10) gives two terms $\|u - u^\star\|_{L^p(\Omega)}\|u - u^\star\|_{L^{p^*}(\Omega)}$ and $\|f - u\|_{L^p(\Omega)}\|u - u^\star\|_{L^{p^*}(\Omega)}$, where $p^* = p/(p-1)$ and $\infty^* := 1$. These terms can be further bounded using inequality (3.9) (note that $p^* \le d/(d-1)$) and the fact that $\|u\|_{L^p(\Omega)} \le C\|f\|_{L^p(\Omega)}$. □

More explicit regression thresholds have been given in [9] both for general and specific one dimensional data f. Let us point out that the condition $\|u\|_{L^p(\Omega)} \le C\|f\|_{L^p(\Omega)}$ and in particular $\|u\|_\infty \le C\|f\|_\infty$ (which can be derived easily for TV regularisation with $C = 1$), as natural as it may seems, it cannot be shown easily. However, if proved, it will also have positive implications as far as the inclusion of the jump set of the solution to the jump set of the data is concerned, see [13].

3.4   Equivalence to TV for large ratio β/α and sufficiently symmetric data

In Corollary 4 we obtained a more precise characterisation of $\mathrm{TGV}^2_{\beta,\alpha}$ for large values of the ratio β/α. In this section we show that at least for symmetric enough data f, $\mathrm{TGV}^2_{\beta,\alpha}$ regularisation is actually equivalent to $\alpha\mathrm{TV}$ regularisation. For the sake of the simplicity of the analysis we assume here that Ω is a two dimensional domain, i.e., $\Omega \subseteq \mathbb{R}^2$. We will also need some symmetry for Ω, for the time being let Ω be a square centered at the origin. We shall prove the following theorem.

Theorem 7 *Suppose that $\Omega \subseteq \mathbb{R}^2$ is a bounded, open square, centred at the origin and let $f \in \mathrm{BV}(\Omega)$ satisfy the following symmetry properties:*

(i) f is symmetric with respect to both axes, i.e.,
$$f(x_1, x_2) = f(-x_1, x_2), \quad f(x_1, x_2) = f(x_1, -x_2), \quad \text{for a.e. } (x_1, x_2) \in \Omega.$$
(ii) f is invariant under $\pi/2$ rotations, i.e., $f(O_{\pi/2}x) = f(x)$, where $O_{\pi/2}$ denotes counterclockwise rotation by $\pi/2$ degrees.

Then if $\beta/\alpha > C_{\mathrm{BD}}$, the problems
$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{p} \int_\Omega |f - u|^p dx + \mathrm{TGV}^2_{\beta,\alpha}(u) \quad \text{and} \quad \min_{u \in \mathrm{BV}(\Omega)} \frac{1}{p} \int_\Omega |f - u|^p dx + \alpha\mathrm{TV}(u)$$
for $p \geq 1$ are equivalent.

Remark 8 The proof of Theorem 7 is essentially based on the fact that the symmetry of the data f is inherited to the solution u and thus to ∇u. In that case we can show that $m_{\mathcal{E}}(\nabla u) = 0$ something that shows the equivalence of $\mathrm{TGV}^2_{\beta,\alpha}$ and $\alpha\mathrm{TV}$. Other symmetric domains, e.g. circles, rectangles, together with appropriate symmetry conditions for f can also guarantee that ∇u has the desired symmetry properties as well. The same holds for any fidelities $\Phi(f, Tu)$ that ensure that the symmetry of f is passed to u.

Let us also mention that abusing the notation a bit, by $m_{\mathcal{E}}(\nabla u) = 0$ we mean that zero is a solution of the problem (3.6) with $g = \nabla u$.

Proof (of Theorem 7). Since $\beta/\alpha > C_{\mathrm{BD}}$, from Corollary 4 we have that the $\mathrm{TGV}^2_{\beta,\alpha}$ regularisation problem is equivalent to
$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{p} \int_\Omega |f - u|^p dx + \alpha\|Du - m_{\mathcal{E}}(\nabla u)\|_{\mathcal{M}}. \tag{3.11}$$
Thus it suffices to show that $m_{\mathcal{E}}(\nabla u) = 0$. Since f satisfies the symmetry properties (i)–(ii), from the rotational invariance of $\mathrm{TGV}^2_{\beta,\alpha}$ [2] we have that the same conditions hold for the $\mathrm{TGV}^2_{\beta,\alpha}$ regularised solution u. This also means that $\nabla u = (\nabla_1 u, \nabla_2 u)$ has the following properties for almost all $x = (x_1, x_2) \in \Omega$:
$$\nabla_1 u(x_1, x_2) = \nabla_1 u(x_1, -x_2), \qquad \nabla_2 u(x_1, x_2) = \nabla_2 u(-x_1, x_2), \tag{3.12}$$
$$\nabla u(x) = -\nabla u(-x), \qquad \nabla_1 u(O_{\pi/2}x) = \nabla_2 u(x). \tag{3.13}$$

Recalling that

$$m_{\mathcal{E}}(\nabla u) = \underset{w \in \mathrm{Ker}\mathcal{E}}{\mathrm{argmin}} \|\nabla u - w\|_{L^1(\Omega;\mathbb{R}^d)},$$

and that $m_{\mathcal{E}}(\nabla u)$ has the form $Ax + b$ it is easy to check, see the following lemma, that $m_{\mathcal{E}}(\nabla u) = 0$. $\square$

Lemma 9 *Let Ω be a square centred at the origin and suppose that $g = (g_1, g_2) \in L^1(\Omega; \mathbb{R}^2)$ satisfies the symmetry properties*

$$g(x) = -g(-x), \quad g_1(x_1, x_2) = g(x_1, -x_2), \quad g_2(x_1, x_2) = g_2(-x_1, x_2), \quad (3.14)$$

$$g_1(O_{\pi/2}x) = g_2(x), \quad\quad (3.15)$$

for almost every $x = (x_1, x_2) \in \Omega$. Then the minimisation problem

$$\min_{w \in \mathrm{Ker}\mathcal{E}} \|g - w\|_{L^1(\Omega;\mathbb{R}^2)}, \quad\quad (3.16)$$

admits $w = 0$ as a solution.

Proof. Recalling that $\mathrm{Ker}\mathcal{E}$ consists of all the functions of the form $r(x) = Ax + b$ with A being a skew symmetric function, we have that the minimisation (3.16) is equivalent to

$$\min_{A,b} \int_{\Omega} |g(x) - Ax - b| dx, \quad\quad (3.17)$$

with corresponding optimality conditions

$$\int_{\Omega} \left\langle \frac{g(x) - AO_{\pi/2}x - b}{|g(x) - AO_{\pi/2}x - b|}, O_{\pi/2}x \right\rangle dx = 0, \quad \text{with } O_{\pi/2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Using the equalities $g_2(x_1, x_2) = g_1(-x_2, x_1)$ and $g_1(-x_2, x_1) = g_1(-x_2, -x_1) = -g_1(x_1, x_2)$ we have that $A = 0$, $b = 0$ solve (3.17) if

$$\int_{\Omega} \left\langle \frac{O_{\pi/2}g(x)}{|O_{\pi/2}g(x)|}, x \right\rangle dx = 0 \iff \int \frac{x_2 g_1(x_1, x_2) - x_1 g_2(x_1, x_2)}{|g(x)|} dx = 0 \iff$$

$$\int_{\Omega} \frac{x_2 g_1(x_1, x_2) - x_1 g_1(-x_2, x_1)}{\sqrt{g_1(x_1, x_2)^2 + g_1(-x_2, x_1)^2}} dx = 0 \iff \int_{\Omega} (x_1 + x_2) \frac{g_1(x_1, x_2)}{|g_1(x_1, x_2)|} dx = 0,$$

with last equality being true since $-g(-x) = g(x)$. $\square$

3.5   Numerical experiments

In this section we verify some of our results using numerical experiments. In Figure 1 we confirm Theorem 7. There, we apply αTV and $\mathrm{TGV}^2_{\beta,\alpha}$ denoising with L^2 fidelity, to a characteristic function of a disk centred at the middle of the domain, Figure 1(a) and away from it, Figure 1(e). Notice that the symmetry properties of Theorem 7 are satisfied for the first case. There, we observe that by choosing the ratio β/α large enough, $\mathrm{TGV}^2_{\beta,\alpha}$ and αTV solutions coincide, Figures 1(b) and 1(c) . However, they do not coincide for small ratio β/α, Figure 1(d), see also the middle row slices in Figure 1(i). In this case $\mathrm{TGV}^2_{\beta,\alpha}$ produces a piecewise smooth result in comparison to the piecewise constant one of αTV. Note that when the symmetry is broken, αTV and $\mathrm{TGV}^2_{\beta,\alpha}$ solutions do not coincide even for large ratio β/α, Figures 1(g), 1(h) and 1(j).
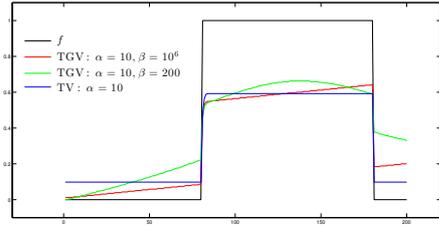
(a) Symmetric data

(b) TV solution, $\alpha = 10$

(c) TGV solution, $\alpha = 10$, $\beta = 10^6$

(d) TGV solution, $\alpha = 10$, $\beta = 200$
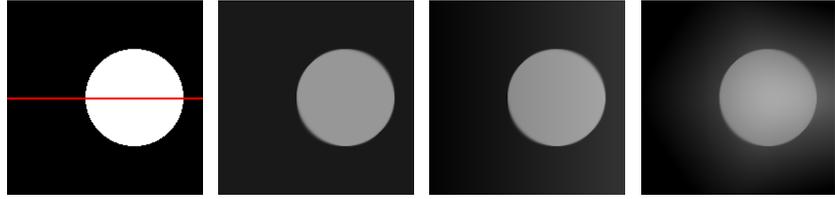
(e) Non-symmetric data

(f) TV solution, $\alpha = 10$

(g) TGV solution, $\alpha = 10$, $\beta = 10^6$

(h) TGV solution, $\alpha = 10$, $\beta = 10^6$
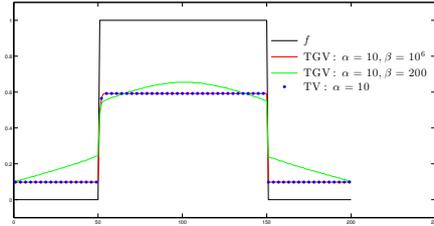
(i) Corresponding middle row slices for symmetric data

(j) Corresponding middle row slices for non-symmetric data

Fig. 1: Illustration of the two dimensional αTV and TGV$^2_{\beta,\alpha}$ equivalence for symmetric data when β/α is large enough. Notice that the equivalence does not hold once the symmetry is broken.

Figure 2 depicts another example of an image that satisfies the symmetry properties of Theorem 7. The αTV solution coincides with the TGV$^2_{\beta,\alpha}$ one for large ratio β/α, Figures 2(b) and 2(c), but not for small ratio, Figure 2(d).

Finally in Figure 3, we solve the L^2–TGV$^2_{\beta,\alpha}$ regularisation problem in a noisy image. We observe that for very small values of β or α, essentially we have no regularisation at all, see Figures 3(c) and 3(d) respectively, verifying Proposition 1. In Figure 3(e), we choose a large ratio β/α, obtaining a TV–like result which is nevertheless quite different than the αTV result, Figure 3(f), having staircasing only inside the ellipse. This is due to the "affine" correction predicted by Corollary 4, see also the corresponding diagonal slices in Figure 3(i). Figure 3(g) depicts a typical TGV solution with no staircasing while in
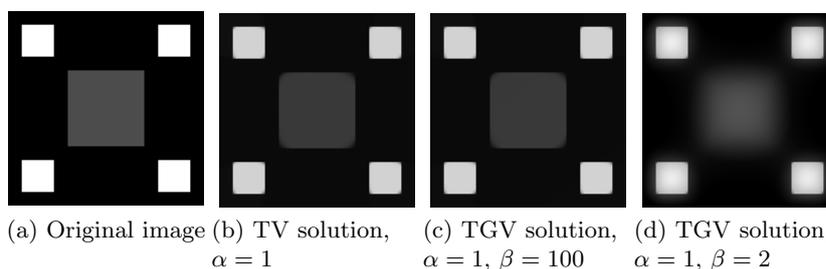
(a) Original image (b) TV solution, (c) TGV solution, (d) TGV solution,
                   $\alpha = 1$ $\alpha = 1$, $\beta = 100$ $\alpha = 1$, $\beta = 2$

Fig. 2: Illustration of the two dimensional αTV and $\text{TGV}^2_{\beta,\alpha}$ equivalence for symmetric data when β/α is large enough.

(a) Original image   (b) Noisy image   (c) TGV solution,   (d) TGV solution,
                                        $\alpha = 0.1$, $\beta = 10^{-4}$ $\alpha = 10^{-4}$, $\beta = 0.15$

(e) TGV solution,   (f) TV solution,   (g) TGV solution,   (h) TGV solution,
$\alpha = 0.1$, $\beta = 100$ $\alpha = 0.1$ $\alpha = 0.1$, $\beta = 0.15$ $\alpha = 100$, $b = 1000$

(i) Corresponding diagonal slices
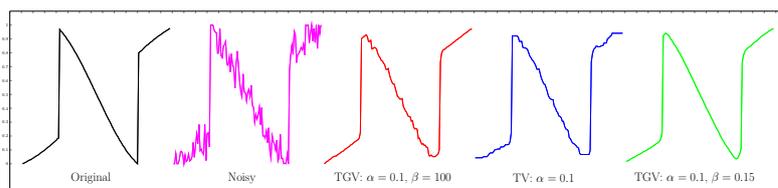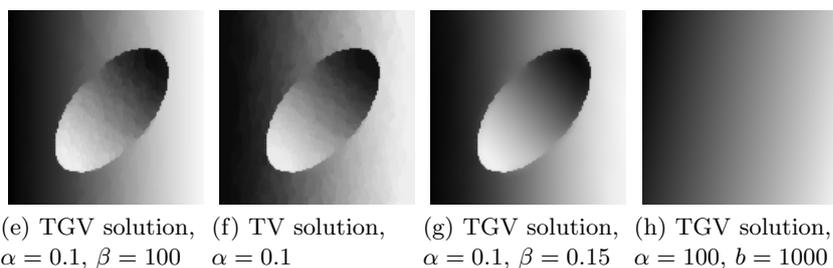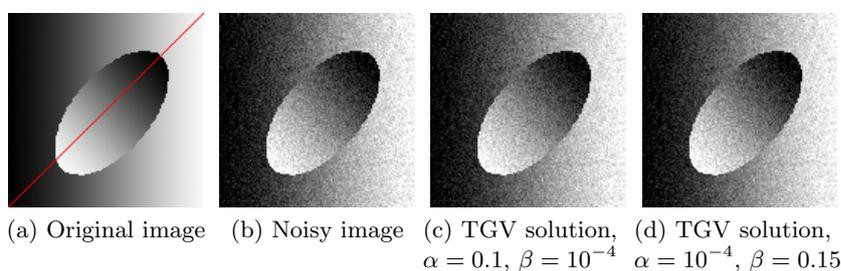
Fig. 3: L^2–$\text{TGV}^2_{\beta,\alpha}$ denoising for extremal values of α and β.

Figure 3(h) we set α and β large enough and we obtain the linear regression of the data, as expected from Proposition 6.

References

1. L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford University Press, USA, 2000.
2. K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.
3. K. Bredies, K. Kunisch, and T. Valkonen. Properties of L^1-TGV^2 : The one-dimensional case. *Journal of Mathematical Analysis and Applications*, 398(1):438 – 454, 2013.
4. K. Bredies and T. Valkonen. Inverse problems with second-order total generalized variation constraints. In *Proceedings of SampTA 2011 - 9th International Conference on Sampling Theory and Applications, Singapore*, 2011.
5. A. Chambolle and P. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188.
6. T. Chan and S. Esedoglu. Aspects of total variation regularized L^1 function approximation. *SIAM Journal on Applied Mathematics*, pages 1817–1837, 2005.
7. V. Duval, J. Aujol, and Y. Gousseau. The TVL1 model: a geometric point of view. *SIAM Journal on Multiscale Modeling and Simulation*, 8(1):154–189.
8. L. Evans. *Partial Differential Equations, volume 19 of Graduate Studies in Mathematics, Second Edition.* American Mathematical Society, 2010.
9. K. Papafitsoros and K. Bredies. A study of the one dimensional total generalised variation regularisation problem. *Inverse Problems and Imaging*, 9(2), 2015.
10. C. Pöschl and O. Scherzer. Exact solutions of one-dimensional total generalized variation. *Communications in Mathematical Sciences*, 13(1):171–202, 2015.
11. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268.
12. R. Temam and G. Strang. Functions of bounded deformation. *Archive for Rational Mechanics and Analysis*, 75(1):7–21, 1980.
13. T. Valkonen. The jump set under geometric regularisation. Part 2: Higher-order approaches. *arXiv preprint 1407.2334.*