The jump set under geometric regularisation.
Part 2: Higher-order approaches

Tuomo Valkonen*

Abstract

In Part 1, we developed a new technique based on Lipschitz pushforwards for proving the jump set containment property $\mathcal{H}^{m-1}(J_u \setminus J_f) = 0$ of solutions u to total variation denoising. We demonstrated that the technique also applies to Huber-regularised TV. Now, in this Part 2, we extend the technique to higher-order regularisers. We are not quite able to prove the property for total generalised variation (TGV) based on the symmetrised gradient for the second-order term. We show that the property holds under three conditions: First, the solution u is locally bounded. Second, the second-order variable is of locally bounded variation, $w \in \mathrm{BV}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$, instead of just bounded deformation, $w \in \mathrm{BD}(\Omega)$. Third, w does not jump on J_u parallel to it. The second condition can be achieved for non-symmetric TGV. Both the second and third condition can be achieved if we change the Radon (or L^1) norm of the symmetrised gradient Ew into an L^p norm, $p > 1$, in which case Korn's inequality holds. We also consider the application of the technique to infimal convolution TV, and study the limiting behaviour of the singular part of Du, as the second parameter of TGV^2 goes to zero. Unsurprisingly, it vanishes, but in numerical discretisations the situation looks quite different. Finally, our work additionally includes a result on TGV-strict approximation in $\mathrm{BV}(\Omega)$.

Mathematics subject classification: 26B30, 49Q20, 65J20.

Keywords: bounded variation, higher-order, regularisation, jump set, TGV, ICTV.

1. Introduction

We introduced in Part 1 [36] the *double-Lipschitz comparability condition* of a regularisation functional R. Roughly

$$R(\overline{\gamma}_{\#}u) + R(\underline{\gamma}_{\#}u) - 2R(u) \leq T_{\overline{\gamma},\underline{\gamma}}|Du|(\mathrm{cl}\, U), \tag{1.1}$$

whenever $\overline{\gamma}, \underline{\gamma} : \Omega \to \Omega$ are bi-Lipschitz transformations reducing to the identity outside $U \subset \Omega$. Constructing specific Lipschitz shift transformations around a point $x \in J_u$, for which the constant $T_{\overline{\gamma},\underline{\gamma}} = O(\rho^2)$ for $\rho > 0$ the size of the shift, we were able to prove the jump set containment

$$\mathcal{H}^{m-1}(J_u \setminus J_f) = 0 \tag{J}$$

for $u \in \mathrm{BV}(\Omega)$ the solution of the denoising or regularisation problem

$$\min_{u \in \mathrm{BV}(\Omega)} \int_\Omega \phi(|f(x) - u(x)|)\, dx + R(u). \tag{P}$$

The admissible fidelities ϕ include here $\phi(t) = t^p$ for $1 < p < \infty$. For $p = 1$ we produced somewhat weaker results comparable to those for total variation (TV) in [21]. The admissible regularisers R included, obviously, total variation, for which the property was already proved previously by level set techniques [13]. We also showed the property for Huber-regularised total variation as a new contribution besides the technique. If non-convex total variation models and the Perona-Malik anisotropic diffusion were well-posed, we demonstrated that the technique would also apply to them.

The development of the new technique was motivated by higher-order regularisers, in particular by total generalised variation (TGV, [9]), for which the level set technique is not available due to the lack of a co-area formula. In this Part 2, we now aim to extend our Lipschitz pushforward technique to variants of TGV as well as infimal convolution TV (ICTV, [14]). In order to do this, we need to modify the double-Lipschitz

* Research Center on Mathematical Modeling, Escuela Politécnica Nacional de Quito, Ecuador.
E-mail: tuomo.valkonen@iki.fi.

comparability criterion (1.1) a little bit. Namely, we will in Section 3 introduce rigorously a *partial double-Lipschitz comparability condition* of the form

$$R(\overline{\gamma}_{\#}(u-v)+v)+R(\underline{\gamma}_{\#}(u-v)+v)-2R(u)\leq T_{\overline{\gamma},\underline{\gamma}}|D(u-v)|(\operatorname{cl}U)+\text{small terms}. \tag{1.2}$$

Here, in comparison to (1.1), we have subtracted v from u before the pushforward. The idea is the same as in the application the jump set containment result for TV to prove it for ICTV. Namely, as we may recall

$$\mathrm{ICTV}_{\vec{\alpha}}(u):=\min_{v\in W^{1,1}(\Omega),\nabla v\in\mathrm{BV}(\Omega;\mathbb{R}^m)}\alpha\|Du-\nabla v\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}+\beta\|D\nabla v\|_{F,\mathcal{M}(\Omega;\mathbb{R}^{m\times m})},$$

where $\vec{\alpha}=(\beta,\alpha)$. Now, if u solves (P) for $R=\mathrm{ICTV}_{\vec{\alpha}}$, then u solves

$$\min_{u\in\mathrm{BV}(\Omega)}\int_\Omega\phi(|f(x)-u(x)|)\,dx+\|Du-\nabla v\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}.$$

with v fixed. Otherwise written, $\bar{u}=u-v$ solves for $\bar{f}=f-v$ the total variation denoising problem

$$\min_{u\in\mathrm{BV}(\Omega)}\int_\Omega\phi(|\bar{f}(x)-\bar{u}(x)|)\,dx+\|D\bar{u}\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}.$$

Since $v\in W^{1,1}(\Omega)$ has no jumps, $J_{\bar{f}}=J_f$, the fact (J) that ICTV introduces no jumps follows from the corresponding result for TV.

The idea with v in (1.2) is roughly the same as this: to remove the second-order information from the problem, and reduce it into a first-order one. However, unlike in the case of ICTV, generally, we cannot reduce the problem to TV. Indeed, written in the differentiation cascade formulation [11], second-order TGV reads as

$$\mathrm{TGV}^2_{\vec{\alpha}}(u):=\min_{w\in\mathrm{BD}(\Omega)}\alpha\|Du-w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}+\beta\|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}. \tag{1.3}$$

Here $\mathrm{BD}(\Omega)$ is the space of vector fields of *bounded deformation* on Ω, and Ew the symmetrised gradient. Now, we do not generally have $w=\nabla v$ for any function v, which is the reason that the analysis is not as simple as that of ICTV. Standard TGV^2 is also significantly complicated by the symmetrised gradient Ew, and we cannot obtain as strong results for it, our results depending on assumptions on w. Namely, we need that w is "BV-differentiable", or, in practise that $w\in\mathrm{BV}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ instead of just $w\in\mathrm{BD}(\Omega)$, and that the projection $P^\perp_{z_\Gamma}(w^+(x)-w^-(x))=0$, on a Lipschitz graph Γ, representing J_u, parametrised on the plane orthogonal to $z_\Gamma\in\mathbb{R}^m$. These complications make us firstly consider the non-symmetric variant of TGV^2, nsTGV^2, where Ew in (1.3) is replaced by Dw and $w\in\mathrm{BV}(\Omega;\mathbb{R}^m)$. Secondly, we consider variants of TGV^2 employing for the second-order term L^q energies, $q>1$. These have the advantage that Korn's inequality holds. For all of these variants, and for ICTV, we obtain stronger results than for TGV^2 itself.

Our analysis of the specific regularisation functionals is in Section 5 after we study local approximability of $w\in\mathrm{BD}(\Omega)$ and approximability in terms of TGV-strict convergence in Section 4. The analysis of the fidelity term $\int_\Omega\phi(|f(x)-u(x)|)\,dx$ is unchanged from Part 1 [36], and therefore the main lemma is only cited in Section 3, where we state our assumptions on R and ϕ, and prove (J) for (P) by combining the separate estimates for the fidelity and regularity terms. We concentrate on p-increasing fidelities for $1<p<\infty$. The case $p=1$ from Part 2 could also be extended, but we have chosen to concentrate on the case $p>1$ where stronger results exist. As an addendum to this qualitative study, we also study quantitatively in Section 6 the limiting behaviour of the singular part D^su of Du for TGV^2 as $\beta\searrow 0$. The behaviour is quite surprising, as on the discrete scale TGV^2 appears to preserve jumps in the limit, but analysis shows that the jumps disappear.

The class of problems (P) is of importance, in particular, for image denoising. We wish to know the structure of J_u in order to see that the denoising problem does not introduce undesirable artefacts, new edges, which in images model different materials and depth information. Higher-order geometric and other recently introduced image regularisers such a TGV [9], ICTV [14], Euler's elastica [16, 34], and many other variants [29, 12, 33, 15, 19, 20, 5] are, in fact, motivated by one serious artefact of the conventional total variation regulariser. This is the stair-casing effect. Further, non-convex total variation schemes and "lower-order fidelities" such as Meyer's G-norm and the Kantorovich-Rubinstein norm, have recently received increased attention in an attempt to, respectively, better model real image gradient statistics [27, 25, 26, 31, 24] or texture [30, 38, 28]. Very little is known about any of these analytically. For TGV^2 we primarily have the results on one-dimensional domains in [10, 32]. We hope that our work in this pair of papers provides an impetus and roots for a technique for the study of many of these and future approaches. We begin our study after going through the obligatory preliminaries in the following Section 2. We finish the study with a few final words in Section 7.

2. Notations and useful facts

We begin by introducing the tools necessary for our work. Much of this material is the same as in Part 1 [36]; we have however decided to make this manuscript to be mostly self-contained, legible without having to delve into the extensively detailed analysis of Part 1. We will also include additional information on tensor fields and functions of bounded deformation, $\mathrm{BD}(\Omega)$. These are crucial for the definition of TGV. First we introduce basic notations for sets, mappings, measures, and tensors. We then move on to tensor fields and Lipschitz mappings and graphs. Finally, we discuss distributional gradients of tensor fields, which allow us to define bounded variation and deformation in a unified way.

2.1. Basic notations

We denote by $\{e_1, \ldots, e_m\}$ the standard basis of $\mathbb{R}^m$. The boundary of a set A we denote by ∂A, and the closure by $\operatorname{cl} A$. The $\{0,1\}$-valued indicator function we write as χ_A. We denote the open ball of radius ρ centred at $x \in \mathbb{R}^m$ by $B(x,\rho)$. We denote by ω_m the volume of the unit ball $B(0,1)$ in $\mathbb{R}^m$.

For $z \in \mathbb{R}^m$, we denote by $z^\perp := \{x \in \mathbb{R}^m \mid \langle z, x\rangle = 0\}$ the hyperplane orthogonal to ν , whereas P_z denotes the projection operator onto the subspace spanned by z, and $P_z^\perp$ the projection onto $z^\perp$. If $A \subset z^\perp$, we denote by $\operatorname{ri} A$ the *relative interior* of A in $z^\perp$ as a subset of $\mathbb{R}^m$.

Let $\Omega \subset \mathbb{R}^m$ be an open set. We then denote the space of (signed) Radon measures on Ω by $\mathcal{M}(\Omega)$. If V is a vector space, then the space of Radon measures on Ω with values in V is denoted $\mathcal{M}(\Omega; V)$. The k-dimensional Hausdorff measure, on any given ambient space $\mathbb{R}^m$, ($k \le m$), is denoted by $\mathcal{H}^k$, while $\mathcal{L}^m$ denotes the Lebesgue measure on $\mathbb{R}^m$. The total variation (Radon) norm of a measure μ is denoted $\|\mu\|_{\mathcal{M}(\mathbb{R}^m)}$.

For vector-valued measures $\mu \in \mathcal{M}(\Omega; \mathbb{R}^k)$, we use the notation $\|\mu\|_{q,\mathcal{M}(\mathbb{R}^m)}$ to indicate that the finite-dimensional base norm is the q-norm. We use the same notation for vector fields $w \in L^p(\Omega; \mathbb{R}^k)$, namely

$$\|w\|_{q,L^p(\Omega)} := \left(\int_\Omega \|w(x)\|_q^p \, dx\right)^{1/p}.$$

For a measurable set A, we denote by $\mu \llcorner A$ the restricted measure defined by $(\mu \llcorner A)(B) := \mu(A \cap B)$. The notation $\mu \ll \nu$ means that μ is absolutely continuous with respect to the measure ν, and $\mu \perp \nu$ that μ and ν are mutually singular. The singular and absolutely continuous (with respect to the Lebesgue measure) part of μ are denoted μ^a and μ^s, respectively.

We denote the k-dimensional upper resp. lower density of μ by

$$\Theta_k^*(\mu; x) := \limsup_{\rho \searrow 0} \frac{\mu(B(x,\rho))}{\omega_k \rho^k}, \quad \text{resp.} \quad \Theta_{*,k}(\mu; x) := \liminf_{\rho \searrow 0} \frac{\mu(B(x,\rho))}{\omega_k \rho^k}.$$

The common value, if it exists, we denote by $\Theta_k(\mu; x)$.

Finally, we often denote by C, C', C''' arbitrary positive constants, and use the plus-minus notation $a^\pm = b^\pm$ in to mean that both $a^+ = b^+$ and $a^- = b^-$ hold.

2.2. Lipschitz and C^1 graphs

A set $\Gamma \subset \mathbb{R}^m$ is called a Lipschitz $(m-1)$-graph (of Lipschitz factor L), if there exist a unit vector z_Γ, an open set $V_\Gamma \subset z_\Gamma^\perp$, and a Lipschitz map $f_\Gamma : V_\Gamma \to \mathbb{R}$, of Lipschitz factor at most L, such that

$$\Gamma = \{v + f_\Gamma(v) z_\Gamma \mid v \in V_\Gamma\}.$$

If $f_\Gamma \in C^1(V_\Gamma)$, we cal Γ a C^1 $(m-1)$-graph. We also define $g_\Gamma : V_\Gamma \to \mathbb{R}^m$ by

$$g_\Gamma(v) = v + z_\Gamma f_\Gamma(v).$$

Then

$$\Gamma = g_\Gamma(V_\Gamma).$$

We denote the open domains "above" and "beneath" Γ, respectively, by

$$\Gamma^+ := \Gamma + (0,\infty) z_\Gamma, \quad \text{and} \quad \Gamma^- := \Gamma + (-\infty, 0) z_\Gamma.$$

We recall that by Kirszbraun's theorem, we may extend the domain of g_Γ from V_Γ to the whole space $z_\Gamma^\perp$ without altering the Lipschitz constant. Then Γ splits Ω into the two open halves $\Gamma^+ \cap \Omega$ and $\Gamma^- \cap \Omega$. We often use this fact.

2.3. Mappings from a subspace

We denote by $\mathcal{L}(V;W)$ the space of linear maps between the vector spaces V and W. If $L \in \mathcal{L}(V;\mathbb{R}^k)$, where $V \sim \mathbb{R}^n$, $(n \le k)$, is a finite-dimensional Hilbert space, Then $L^* \in \mathcal{L}(\mathbb{R}^k;V^*)$ denotes the adjoint, and the n-dimensional Jacobian is defined as [3]

$$\mathcal{J}_n[L] := \sqrt{\det(L^* \circ L)}.$$

With the gradient of a Lipschitz function $f : V \to \mathbb{R}^k$ defined in "components as columns order", $\nabla f(x) \in \mathcal{L}(\mathbb{R}^k;V)$, we extend this notation for brevity as

$$\mathcal{J}_n f(x) := \mathcal{J}_n[(\nabla f(x))^*].$$

If $\Omega \subset V$ is a measurable set, and $g \in L^1(\Omega)$, the *area formula* may then be stated

$$\int_{\mathbb{R}^k} \sum_{x \in \Omega \cap f^{-1}(y)} g(x)\, d\mathcal{H}^n(y) = \int_\Omega g(x)\mathcal{J}_n f(x)\, d\mathcal{H}^n(x). \tag{2.1}$$

That this indeed holds in our sitting of finite-dimensional Hilbert spaces $V \sim \mathbb{R}^n$ follows by a simple argument from the area formula for $f : \mathbb{R}^n \to \mathbb{R}^k$, stated in, e.g, [3]. We only use the cases $V = z^\perp$ for some $z \in \mathbb{R}^m$ $(n = m-1)$, or $V = \mathbb{R}^m$ $(n = m)$.

We also denote by

$$C^{2,\cap}(V) := \bigcap_{\lambda \in (0,1)} C^{2,\lambda}(V)$$

the class of functions that are twice differentiable (as defined above for tensor fields) with a Hölder continuous second differential for all exponents $\lambda \in (0,1)$.

The Lipschitz factor of a Lipschitz mapping f we denote by $\operatorname{lip} f$. We also recall that a Lipschitz transformation $\gamma : U \to V$ with $U, V \subset \mathbb{R}^m$ has the *Lusin N-property* if it maps $\mathcal{L}^m$-negligible sets to $\mathcal{L}^m$-negligible sets.

If $\gamma : \Omega \to \Omega$ is a 1-to-1 Lipschitz transformation, and $u : \Omega \to \Omega$ a Borel function, we define the pushforward $u_\gamma := \gamma_\# u := u \circ \gamma^{-1}$. Finally, we denote the identity transformation by $\iota(x) = x$.

2.4. Tensors and tensor fields

We now introduce tensors and tensor fields. We simplify the treatment from its full differential-geometric setting, as can be found in, e.g., [6], as we are working on finite-dimensional Hilbert spaces. These definitions and our approach to defining TGV^2 follow that in [37].

We let $V_1, \dots, V_k$ be finite-dimensional Hilbert spaces, $V_j \sim \mathbb{R}^{m_j}$ with corresponding bases $\{e_1^j, \dots, e_{m_j}^j\}$, $(j = 1, \dots, k)$. A k-tensor is then a k-linear mapping $A : V_1 \times \cdots \times V_k \to \mathbb{R}$. We denote $A \in \mathcal{T}(V_1, \dots, V_k)$. If $V_j = V$ for all $j = 1, \dots, k$, we write $\mathcal{T}^k(V) := \mathcal{T}(V_1, \dots, V_k)$. A symmetric tensor $A \in \operatorname{Sym}^k(V) \subset \mathcal{T}^k(V)$ satisfies for any permutation π of $\{1, \dots, k\}$ and any $c_1, \dots, c_k \in V$ that $A(c_{\pi 1}, \dots, c_{\pi k}) = A(c_1, \dots, c_k)$, For conciseness of notation, we often identify $V \sim \mathcal{T}^1(V)$ through the mapping $V(x) = \langle V, x \rangle$.

For a $A \in \mathcal{T}(V_1, \dots, V_k)$ and $B \in \mathcal{T}(V_{k+1}, \dots, V_{k+m})$ we define the $(m+k)$-tensor $A \otimes B \in \mathcal{T}(V_1, \dots, V_{k+m})$ by

$$(A \otimes B)(c_1, \dots, c_{k+m}) = A(c_1, \dots, c_k)B(c_{k+1}, \dots, c_{k+m}).$$

We define on $A, B \in \mathcal{T}(V_1, \dots, V_k)$ the inner product

$$\langle A, B \rangle := \sum_{p_1=1}^{m_1} \cdots \sum_{p_k=1}^{m_k} A(e_{p_1}^1, \dots, e_{p_k}^k) B(e_{p_1}^1, \dots, e_{p_k}^k),$$

and the Frobenius norm

$$\|A\|_F := \sqrt{\langle A, A \rangle}.$$

If $k = 1$,we simply denote $\|A\| := \|A\|_2 := \|A\|_F$, as the Frobenius norm agrees with the Euclidean norm.

Let then $u : \Omega \to \mathcal{T}(V_1, \ldots, V_k)$ be a Lebesgue-measurable function on the domain $\Omega \subset V_0$, where $V_0 \sim \mathbb{R}^m$ is also a finite-dimensional Hilbert space. We define the norms

$$\|u\|_{F,p} := \Big(\int_\Omega \|u(x)\|_F^p \, dx\Big)^{1/p}, \quad (p \in [1, \infty)), \quad \text{and} \quad \|u\|_{F,\infty} := \operatorname*{ess\,sup}_{x\in\Omega} \|u(x)\|_F,$$

and the spaces

$$L^p(\Omega; \mathcal{T}(V_1, \ldots, V_k)) = \{u : \Omega \to \mathcal{T}(V_1, \ldots, V_k) \mid u \text{ Borel}, \|u\|_{F,p} < \infty\}, \quad (p \in [1, \infty]).$$

The spaces $L^p(\Omega; \mathcal{T}^k(V))$ and $L^p(\Omega; \mathrm{Sym}^k(V))$ are defined analogously.

2.5. Distributional gradients and tensor-valued measures

For the definition of total generalised variation (TGV), we need to define the concept of a tensor-valued measure, as well as the distributional differential Du and the symmetrised distributional Eu on tensor fields. This is done now. If the reader is satisfied with a cursory understanding of TGV, this subsection may be skipped.

We start with tensor field divergences. Let $u \in C^1(\Omega; \mathcal{T}(V_1, \ldots, V_k))$, $(k \geq 0)$. The (Fréchet) differential $\mathrm{d}\, f(x) \in \mathcal{T}(V_0, V_1, \ldots, V_k)$ at $x \in \Omega$ is defined by the limit

$$\lim_{h\to 0} \frac{\|f(x + h) - f(x) - \mathrm{d}\, f(x)(h, \cdot, \ldots, \cdot)\|_F}{\|h\|_F} = 0.$$

If $k \geq 1$, if $V_0 = V_1$, we define the divergence, $\operatorname{div} u \in C(\Omega; \mathcal{T}(V_2, \ldots, V_k))$ by contraction as

$$[\operatorname{div} u(x)](c_2, \ldots, c_k) := \sum_{i=1}^{m_1} \mathrm{d}\, u(x)(\xi_i^1, \xi_i^1, c_2, \ldots, c_k).$$

Observe that if u is symmetric, then so is $\operatorname{div} u$. Moreover Green's identity

$$\int_\Omega \langle \mathrm{d}\, u(x), \phi(x)\rangle \, dx = \int_\Omega \langle u(x), -\operatorname{div} \phi(x)\rangle \, dx$$

holds for $u \in C^1(\Omega; \mathcal{T}(V_2, \ldots, V_k))$ and $\phi \in C_0^1(\Omega; \mathcal{T}(V_1, \ldots, V_k))$ with $\Omega \subset V_1 = V_0$.

Denoting by X^* the continuous linear functionals on the topological space X, we now define the distributional gradient

$$Du \in [C_c^\infty(\Omega; \mathcal{T}^{k+1}(\mathbb{R}^m))]^*$$

of $u \in L^1(\Omega; \mathcal{T}^k(\mathbb{R}^m))$ by

$$Du(\varphi) := -\int_\Omega \langle u(x), \operatorname{div} \varphi(x)\rangle \, dx, \quad (\varphi \in C_c^\infty(\Omega; \mathcal{T}^{k+1}(\mathbb{R}^m))).$$

Likewise we define the symmetrised distributional gradient

$$Eu \in [C_c^\infty(\Omega; \mathrm{Sym}^{k+1}(\mathbb{R}^m))]^*$$

of $u \in L^1(\Omega; \mathcal{T}^k(\mathbb{R}^m))$ by

$$Eu(\varphi) := -\int_\Omega \langle u(x), \operatorname{div} \varphi(x)\rangle \, dx, \quad (\varphi \in C_c^\infty(\Omega; \mathrm{Sym}^{k+1}(\mathbb{R}^m))).$$

We also define the "Frobenius unit ball"

$$V_{F,\mathrm{ns}}^k := \{\varphi \in C_c^\infty(\Omega; \mathcal{T}^k(\mathbb{R}^m)) \mid \|\varphi\|_{F,\infty} \leq 1\}.$$

and the "symmetric Frobenius unit ball"

$$V_{F,\mathrm{s}}^k := \{\varphi \in C_c^\infty(\Omega; \mathrm{Sym}^k(\mathbb{R}^m)) \mid \|\varphi\|_{F,\infty} \leq 1\}.$$

For our purposes it then suffices to define a tensor-valued measure $\mu \in \mathcal{M}(\Omega; \mathcal{T}^k(\mathbb{R}^m))$ as a linear functional $\mu \in [C_c^\infty(\Omega; \mathcal{T}^k(\mathbb{R}^m))]^*$ bounded in the sense that the total variation norm

$$\|\mu\|_{F,\mathcal{M}(\Omega;\mathcal{T}^k(\mathbb{R}^m))} := \sup\{\mu(\varphi) \mid \varphi \in V_{F,\mathrm{ns}}^k\} < \infty.$$

For a justification of this definition, we refer to [22]. The definition of a symmetric measure $\mu \in \mathcal{M}(\Omega; \mathrm{Sym}^k(\mathbb{R}^m))$ is analogous with $\mu \in [C_c^\infty(\Omega; \mathrm{Sym}^k(\mathbb{R}^m))]^*$ and

$$\|\mu\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^k(\mathbb{R}^m))} := \sup\{\mu(\varphi) \mid \varphi \in V^k_{F,\mathrm{s}}\} < \infty.$$

It follows that Du and Eu are measures when they are bounded on $V^k_{F,\mathrm{ns}}$ and $V^k_{F,\mathrm{s}}$, respectively. Observe that for $k = 0$, it holds $\mathcal{M}(\Omega; \mathcal{T}^0(\mathbb{R}^m)) = \mathcal{M}(\Omega; \mathrm{Sym}^0(\mathbb{R}^m)) = \mathcal{M}(\Omega)$, and for $k = 1$, it holds

$$\mathcal{M}(\Omega; \mathcal{T}^1(\mathbb{R}^m)) = \mathcal{M}(\Omega; \mathrm{Sym}^1(\mathbb{R}^m)) =: \mathcal{M}(\Omega; \mathbb{R}^m).$$

Remark 2.1. The choice of the Frobenius norm as the finite-dimensional norm in the above definitions, indicated by the subscript F, ensures isotropy and a degree of rotational invariance for tensor fields. Some alternative rotationally invariant norms, generalising the nuclear and the spectral norm for matrices, are discussed [37].

2.6. Functions of bounded variation

We say that a function $u : \Omega \to \mathbb{R}$ on a bounded open set $\Omega \subset \mathbb{R}^m$, is of *bounded variation* (see, e.g., [3] for a more thorough introduction), denoted $u \in \mathrm{BV}(\Omega)$, if $u \in L^1(\Omega)$, and the distributional gradient Du is a Radon measure. Given a sequence $\{u^i\}_{i=1}^\infty \subset \mathrm{BV}(\Omega)$, weak* convergence is defined as $u^i \to u$ strongly in $L^1(\Omega)$ along with $Du^i \overset{*}{\rightharpoonup} Du$ weakly* in $\mathcal{M}(\Omega)$. The sequence converges *strictly* if, in addition to this, $|Du^i|(\Omega) \to |Du|(\Omega)$.

We denote by S_u the approximate discontinuity set, i.e., the complement of the set where the Lebesgue limit $\widetilde{u}$ exists. The latter is defined by

$$\lim_{\rho \searrow 0} \frac{1}{\rho^m} \int_{B(x,\rho)} \|\widetilde{u}(x) - u(y)\|\, dy = 0.$$

The distributional gradient can be decomposed as $Du = \nabla u \mathcal{L}^m + D^j u + D^c u$, where the density ∇u of the *absolutely continuous part* of Du equals (a.e.) the approximate differential of u. We also define the *singular part* as $D^s u = D^j u + D^c u$. The *jump part* $D^j u$ may be represented as

$$D^j u = (u^+ - u^-) \otimes \nu_{J_u} \mathcal{H}^{m-1} \llcorner J_u,$$

where x is in the *jump set* $J_u \subset S_u$ of u if for some $\nu := \nu_{J_u}(x)$ there exist two *distinct* one-sided traces $u^+(x)$ and $u^-(x)$, defined as satisfying

$$\lim_{\rho \searrow 0} \frac{1}{\rho^m} \int_{B^\pm(x,\rho,\nu)} \|u^\pm(x) - u(y)\|\, dy = 0,$$

where $B^\pm(x,\rho,\nu) := \{y \in B(x,\rho) \mid \pm\langle y - x, \nu\rangle \geq 0\}$. It turns out that J_u is countably $\mathcal{H}^{m-1}$-rectifiable and ν is (a.e.) the normal to J_u. This former means that there exist Lipschitz $(m-1)$-graphs $\{\Gamma_i\}_{i=1}^\infty$ such that $\mathcal{H}^{m-1}(J_u \setminus \bigcup_{i=1}^\infty \Gamma_i) = 0$. Moreover, we have $\mathcal{H}^{m-1}(S_u \setminus J_u) = 0$. The remaining *Cantor part* $D^c u$ vanishes on any Borel set σ-finite with respect to $\mathcal{H}^{m-1}$.

We will depend on the following basic properties of densities of Du; for the proof, see, e.g., [3, Proposition 3.92].

Proposition 2.1. *Let $u \in \mathrm{BV}(\Omega)$ for an open domain $\Omega \subset \mathbb{R}^m$. Define*

$$\widetilde{S}_u := \{x \in \Omega \mid \Theta_{*,m}(|Du|; x) = \infty\}, \quad \text{and} \quad \widetilde{J}_u := \{x \in \Omega \mid \Theta_{*,m-1}(|Du|; x) > 0\}.$$

Then the following decomposition holds.

(i) $\nabla u = Du \llcorner (\Omega \setminus \widetilde{S}_u)$.
(ii) $D^j u = Du \llcorner \widetilde{J}_u$, *precisely* $\widetilde{J}_u \supset J_u$, *and* $\mathcal{H}^{m-1}(\widetilde{J}_u \setminus J_u) = 0$.
(iii) $D^c u = Du \llcorner (\widetilde{S}_u \setminus \widetilde{J}_u)$.

We will require the following property of the traces along a Lipschitz graph Γ.

Lemma 2.1 (Part 1)**.** *Let $u \in \mathrm{BV}(\Omega)$. Then there exists a Borel set Z_u with $\mathcal{H}^{m-1}(Z_u) = 0$ such that every $x \in J_u \setminus Z_u$ is a Lebesgue point of the one-sided traces $u^\pm$, and*

$$\Theta^*_{m-1}(|Du| \llcorner (\Gamma^x)^+; x) = 0, \text{ and } \Theta^*_{m-1}(|Du| \llcorner (\Gamma^x)^-; x) = 0$$

for a Lipschitz $(m-1)$-graph Γ^x, which satisfies the following. Firstly

$$V_{\Gamma^x} \supset B(P^\perp_{z_\Gamma} x, r(x))$$

for some $r(x) > 0$. Secondly the traces of u at x exist from both sides of Γ^x and agree with $u^\pm(x)$.

2.7. Functions of bounded deformation

Similarly to the definition of a function of bounded variation, a function $w \in L^1(\Omega;\mathbb{R}^m)$ for a domain $\Omega \subset \mathbb{R}^m$ is said to be of a *vector field (or function) of bounded deformation*, if the distributional symmetrised gradient $Ew \in \mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))$ [35]. We denote this space by $\mathrm{BD}(\Omega)$. The concept can also be generalised to tensor fields of higher orders [7], useful for the definition of TGV^k for $k > 2$.

Similar to BV, we have the decomposition [1]

$$Ew = \mathcal{E}w\mathcal{L}^m + E^j w + E^c w,$$

where $\mathcal{E}w$ is the absolutely continuous part. For smooth functions

$$\mathcal{E}w(x) = \frac{1}{2}\big(\nabla w(x) + [\nabla w(x)]^T\big).$$

Generally this expression holds at points of *approximate differentiability* of w, at $\mathcal{L}^m$-a.e. $x \in \Omega$ [1, 23]. The jump part satisfies

$$E^j w = \frac{1}{2}\big(\nu_{J_u} \otimes (w^+ - w^-) + (w^+ - w^-) \otimes \nu_{J_w}\big)\mathcal{H}^{m-1} \llcorner J_w,$$

where the one-sided traces $w^\pm$, the jump set J_w and its approximate normal ν_{J_w} are as in the case of functions bounded variation. Likewise, the Cantor part vanishes on any Borel set σ-finite with respect to $\mathcal{H}^{m-1}$. Similarly to Proposition 2.1, defining

$$\widetilde{J}_u := \{x \in \Omega \mid \Theta_{*,m-1}(|Eu|;x) > 0\},$$

we have

$$\widetilde{J}_u \supset J_u, \quad \text{and} \quad \mathcal{H}^{m-1}(\widetilde{J}_u \setminus J_u) = 0. \tag{2.2}$$

Many other results are however not as strong in $\mathrm{BD}(\Omega)$ as in $\mathrm{BV}(\Omega)$. For one, we only have [1] $|Ew|(S_w \setminus J_w) = 0$ instead of the stronger result $\mathcal{H}^{m-1}(S_w \setminus J_w) = 0$, which were to hold if $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. In fact, this result can be made a little stronger. Namely, $|Ew|(S_w \setminus J_v) = 0$ for $v, w \in \mathrm{BD}(\Omega)$.

Instead of Poincaré's inequality in $\mathrm{BV}(\Omega;\mathbb{R}^n)$, which says that on Lipschitz domains we can approximate zero-mean $\|u - \bar{u}\|$ for $\bar{u} = ⨍_\Omega u\,dy$ by $C_\Omega|Du|(\Omega)$, in $\mathrm{BD}(\Omega)$ we have the Sobolev-Korn inequality. This says that there exists a constant $C_\Omega > 0$ and for each $w \in \mathrm{BD}(\Omega)$ an element $\bar{w} \in \ker E$ such that

$$\|w - \bar{w}\|_{2,L^1(\Omega;\mathbb{R}^m)} \leq C_\Omega \|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}.$$

The kernel of E consists of affine maps $\bar{w}(x) = Ax + c$ for A a skew-symmetric matrix. The Sobolev-Korn inequality can also be extended to symmetric tensor fields of higher-order than the present $k = 1$, in which case the kernel is also a higher-order polynomial [7].

We will not really need these latter properties. The point is that $\mathrm{BD}(\Omega)$ has significantly weaker regularity than $\mathrm{BV}(\Omega;\mathbb{R}^m)$. This will have implications to our work. What we will use is Korn's inequality, which holds for $1 < p < \infty$ but *notoriously* not for $p = 1$. The form most suitable for our purposes, easily obtainable from the versions in [1, 18, 17], states the existence of a constant $C_{\Omega,q} > 0$ such that

$$\int_\Omega \|\nabla w(x)\|_F^q \,dx \leq C_{\Omega,q} \int_\Omega \|\mathcal{E}w(x)\|_F^q \,dx, \tag{2.3}$$

for bounded domains Ω, and vector fields $w \in W_0^{1,q}(\Omega;\mathbb{R}^m)$. Our reason for the zero boundary condition, as opposed to $\Omega = \mathbb{R}^m$, a Sobolev-Korn type $\|\nabla(w - \bar{w})(x)\|_F^q$ on the left, or extra $\|w\|_{2,L^q(\Omega;\mathbb{R}^m)}$ on the right, is that in our application, we do not want to *directly* enforce $w \in L^q(\Omega;\mathbb{R}^m)$. This will typically however follow a posteriori from (2.3) and the Gagliardo-Nirenberg-Sobolev inequality.

3. Problem statement

Before stating our main results rigorously, we introduce our assumptions on regularisation functionals and fidelities. The definition of an admissible regularisation functional, and our assumptions on the fidelity ϕ are unchanged from Part 1, but we replace the double-Lipschitz comparability by a notion of partial double-Lipschitz comparability, and limit the set of admissible Lipschitz transformations to one that operates along a specific direction.

3.1. Admissible regularisation functionals and fidelities

We begin by stating our assumptions on R, which are formulated in Definition 3.1 and Definition 3.4.

Definition 3.1. We call R an *admissible regularisation functional* on $L^1(\Omega)$, where the domain $\Omega \subset \mathbb{R}^m$, if it is convex, lower semi-continuous with respect to weak* convergence in $\mathrm{BV}(\Omega)$, and there exist $C, c > 0$ such that

$$\|Du\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \leq C\big(1 + \|u\|_{L^1(\Omega)} + R(u)\big), \quad (u \in L^1(\Omega)). \tag{3.1}$$

The next two technical definitions will be required by Definition 3.4.

Definition 3.2. We denote by $\mathcal{F}(\Omega)$ the set of one-to-one Lipschitz transformations $\gamma : \Omega \to \Omega$ with γ^{-1} also Lipschitz and both satisfying the Lusin N-property. With $U \subset \Omega$ an open set, and $z \in \mathbb{R}^m$ a unit vector, we set

$$\begin{aligned} \mathcal{F}(\Omega, U) &:= \{\gamma \in \mathcal{F}(\Omega) \mid \gamma(x) = x \text{ for } x \notin U\}, \quad \text{and} \\ \mathcal{F}(\Omega, U, z) &:= \{\gamma \in \mathcal{F}(\Omega, U) \mid P_z^\perp \gamma(y) = P_z^\perp y \text{ for all } y \in \Omega\}. \end{aligned}$$

With $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega)$, we then define the *basic double-Lipschitz comparison constants*

$$G_{\overline{\gamma},\underline{\gamma}} := \sup_{x\in\Omega, \|v\|=1} \|A_{\overline{\gamma}}(x)v\| + \|A_{\underline{\gamma}}(x)v\| - 2\|v\|.$$

and

$$J_{\overline{\gamma},\underline{\gamma}} := \sup_{x\in\Omega} |\mathcal{J}_m\overline{\gamma}(x) + \mathcal{J}_m\underline{\gamma}(x) - 2|.$$

Here the norm is the operator norm, I the identity mapping on $\mathbb{R}^m$, and

$$A_\gamma(x) := \nabla\gamma^{-1}(\gamma(x))\mathcal{J}_m\gamma(x).$$

We also define the *distances-to-identity*

$$D_\gamma := \sup_{x\in\Omega} \|\nabla\gamma^{-1}(\gamma(x)) - I\|, \quad \text{and} \quad J_\gamma := \sup_{x\in\Omega} |\mathcal{J}_m\gamma(x) - 1|,$$

as well as the *normalised transformation distance*

$$\overline{M}_\gamma := \sup_{\substack{U:\gamma\in\mathcal{F}(\Omega,U)\\ u\in\mathrm{BV}(\Omega)}} \int_\Omega \frac{\|\gamma_\# u(y) - u(y)\|}{\operatorname{diam}(U)|Du|(U)}\, dy. \tag{3.2}$$

Finally we combine these all into the overall *double-Lipschitz comparison constant*

$$T_{\overline{\gamma},\underline{\gamma}} := G_{\overline{\gamma},\underline{\gamma}} + J_{\overline{\gamma},\underline{\gamma}} + D_{\overline{\gamma}}^2 + D_{\underline{\gamma}}^2 + J_{\overline{\gamma}}^2 + J_{\underline{\gamma}}^2 + \overline{M}_{\overline{\gamma}}^2 + \overline{M}_{\underline{\gamma}}^2.$$

Observe that by Poincaré's inequality, if $\operatorname{supp}(\gamma - \iota)$ has Lipschitz boundary, then $\overline{M}_\gamma < \infty$ for $\mathcal{L}^m$-a.e. $x \in \Omega$, small enough $r > 0$, and $\gamma \in \mathcal{F}(\Omega, U)$ for $U \subset B(0, r)$.

Definition 3.3. Given $u, v \in L^1(\Omega)$, and $\gamma \in \mathcal{F}(\Omega)$, we define the *partial pushforward*

$$\gamma_\#[\![u, v]\!] := \gamma_\#(u - v) + v.$$

Finally, we may state rigorously our most central concept.

Definition 3.4. Let $x_0 \in \Omega$ and $u \in \mathrm{BV}(\Omega)$. We say that R is *partially double-Lipschitz comparable for u at x_0*, if there exists a constant $R^a > 0$ and a function $v \in W^{1,1}(\Omega)$, $x_0 \notin S_v$, satisfying the following: for every $\epsilon > 0$, for some $r_0 > 0$, if $U \subset B(x_0, r)$, $0 < r < r_0$ and $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega, U)$ with $T_{\overline{\gamma},\underline{\gamma}} < 1$, then

$$R(\overline{\gamma}_\#[\![u, v]\!]) + R(\underline{\gamma}_\#[\![u, v]\!]) - 2R(u) \leq R^a T_{\overline{\gamma},\underline{\gamma}} |D(u - v)|(\operatorname{cl} U) + (T_{\overline{\gamma},\underline{\gamma}}^{1/2} + r)\epsilon r^m. \tag{3.3}$$

We also say that R is *partially double-Lipschitz comparable at x_0 for u in the direction z* for some unit vector $z \in \mathbb{R}^m$, if (3.3) holds with the change that $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega, U, z)$.

Remark 3.1. Usually R^a will be a universal constant for R, but we do not need this in this work. The function v will depend on both u and x_0. The bound $T_{\overline{\gamma},\underline{\gamma}} < 1$ is mostly about aesthetics. We could instead allow $T_{\overline{\gamma},\underline{\gamma}} < \delta$ for arbitrary $\delta > 0$; we however cannot allow δ to be determined by $\epsilon > 0$ for the proof of our main result Theorem 3.2. It can only depend on u and x_0 similarly to v. The only purpose of the bound is to allow the use of the single constant $T_{\overline{\gamma},\underline{\gamma}}$ in front of both of the terms on the right hand side of (3.3), replacing any second-order terms that we might get in front of the remainder term ϵr^m by first-order terms, which suffice there; compare the proof of Proposition 5.1. For this the fixed bound suffices: $T_{\overline{\gamma},\underline{\gamma}} \leq T_{\overline{\gamma},\underline{\gamma}}^{1/2}$. Instead of this, we could also replace $T_{\overline{\gamma},\underline{\gamma}}$ by two arbitrary polynomials of the square roots of the variables in its definition, the one in front of $|D(u-v)|(\operatorname{cl} U)$ being of lowest order 2, and the one in front of ϵr^m of lowest order 1. Then we would not have to bound $T_{\overline{\gamma},\underline{\gamma}} < 1$. The reason for introducing the *normalised* transformation distance is likewise aesthetical.

We will strive to prove the following property of the regularisation functionals that we study. We will only use the more involved case (ii) in this work.

Assumption 3.1. We assume that R is an admissible regularisation functional on $L^1(\Omega)$ that satisfies the following for every $u \in \mathrm{BV}(\Omega)$ and every Lipschitz $(m-1)$-graph $\Gamma \subset \Omega$.

(i) R is partially double-Lipschitz comparable for u at $\mathcal{L}^m$-a.e. $x \in \Omega$.
(ii) R is partially double-Lipschitz comparable for u in the direction z_Γ at $\mathcal{H}^{m-1}$-a.e. $x \in \Gamma$.

In order to show the existence of solutions to (P), we require the following property from ϕ.

Definition 3.5. Let the domain $\Omega \subset \mathbb{R}^m$. We call $\phi : [0,\infty) \to [0,\infty]$ an *admissible fidelity function on* Ω if it is convex, lower semi-continuous, $\phi(0) = 0$, and satisfies for some $C > 0$ the coercivity condition

$$\|u\|_{L^1(\Omega)} \leq C \left(\int_\Omega \phi(|u(x)|)\, dx + 1 \right), \quad (u \in L^1(\Omega)). \tag{3.4}$$

Throughout this paper, we extend the domain of ϕ to $\mathbb{R}$ by defining

$$\phi(t) := \phi(-t), \quad (t < 0).$$

This is in order to simplify the notation $\phi(|u(x)|)$ to $\phi(u(x))$.

For the study of the jump set J_u of solutions to (P), we require additionally the following increase criterion to be satisfied by ϕ.

Definition 3.6. We say that ϕ is *p-increasing* for $p \geq 1$, if there exists a constant $C_\phi > 0$ for which

$$\phi(x) - \phi(y) \leq C_\phi (x-y)|x|^{p-1}, \quad (x, y \geq 0).$$

As we have seen in Part 1, the functions $\phi(t) = t^p$, $(p \geq 1)$, in particular are p-increasing and admissible. Moreover, the problem (P) is well-posed under the above assumptions.

Theorem 3.1 (Part 1 & standard)**.** *Let $f \in L^1(\Omega)$ satisfy $\int_\Omega \phi(f(x))\, dx < \infty$. Suppose that R is an admissible regularisation functional on $L^1(\Omega)$, and ϕ an admissible fidelity function for Ω. Then there exists a solution $u \in L^1(\Omega)$ to (P), and any solution satisfies $u \in \mathrm{BV}(\Omega)$.*

3.2. Jump set containment

Our main result in this paper is the following theorem combined with the corresponding partial double-Lipschitz comparability estimates for higher-order regularisers in Section 5.

Theorem 3.2. *Let the domain $\Omega \subset \mathbb{R}^m$ be bounded with Lipschitz boundary, and $\phi : [0,\infty) \to [0,\infty)$ be an admissible p-increasing fidelity function for some $1 < p < \infty$. Let $f \in \mathrm{BV}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$, and suppose $u \in \mathrm{BV}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$ solves (P). If R satisfies Assumption 3.1(ii), then*

$$\mathcal{H}^{m-1}(J_u \setminus J_f) = 0.$$

Remark 3.2. Observe that we require u to be locally bounded. This does not necessarily hold, and needs to be proved separately. In imaging applications we are however not usually interested in unbounded data , and nearly always $\|f\|_{L^\infty(\Omega)} \leq M$ for some known dynamic range M. So one would think that it suffices to add the constraint $\|u\|_{L^\infty(\Omega)} \leq M$ to the problem (P). This would even work under the simpler double-Lipschitz comparability (1.1) of Part 1, as the constraint is invariant under pushforwards $\gamma_\# u$.

It is, however, not generally invariant under the partial pushforward $\gamma_\# [\![u, v]\!]$, which might not satisfy the constraint if $|\widetilde{u}(x_0)| = M$. If $|\widetilde{u}(x_0)| < M$, and the radius $r_0 > 0$ is small enough, the constraint will still be satisfied for otherwise well-behaved u and typical constructions of v. What this says is that if (well-behaved) u jumps outside J_f, then it will jump to activate the constraint. Whether in practise the v prescribed by the partial double-Lipschitz property of any particular regulariser satisfies $\|\gamma_\# [\![u, v]\!]\|_{L^\infty(\Omega)} \leq M$, is as interesting an open question as boundedness itself.

The proof of Theorem 3.2 is based on combining the double-Lipschitz estimate for the regulariser with a separate estimate for the fidelity, for specific "shift" transformations $\gamma_{\rho,r}$. In Part 1, we proved the following lemmas about these.

Lemma 3.1 (Part 1). *Suppose $\gamma \in \mathcal{F}(\Omega, U, z)$ for some $z \in \mathbb{R}^m$ and $U \subset \mathbb{R}^m$. Let $u \in \mathrm{BV}(\Omega)$. Then*

$$\int_U |u(\gamma(x)) - u(x)| \, dx \leq M_\gamma |Du|(U),$$

where

$$M_\gamma := \sup_{x \in \Omega} \|\gamma(x) - x\|.$$

Proof. This is proved in Part 1 for specific transformations, but everything in the proof only depends on $\gamma \in \mathcal{F}(\Omega, U, z)$. □

Lemma 3.2 (Part 1). *Let $\Omega \subset \mathbb{R}^m$, and $\Gamma \subset \Omega$ be a Lipschitz $(m-1)$-graph, $x_0 \in \Gamma$. There exist $r_0 > 0$ and Lipschitz transformations $\gamma_{\rho,r} \in \mathcal{F}(\Omega, U, z_\Gamma)$, $(-1 < \rho < 1,\ 0 < r < r_0)$, with*

$$U_r := x_0 + z_\Gamma^\perp \cap B(0, r) + (3 + \operatorname{lip} f_\Gamma)(-r, r) z_\Gamma.$$

Moreover, there exists a constant $C > 0$ such that

$$T_{\gamma_{\rho,r}, \gamma_{-\rho,r}} \leq C\rho^2.$$

Proof. Only the facts $J_{\gamma_{\rho,r}} \leq C|\rho|$ and $\overline{M}_{\gamma_{\rho,r}} \leq C|\rho|$, which are required for the bound $T_{\gamma_{\rho,r}, \gamma_{-\rho,r}}$, are not directly proved in Part 1. The former follows immediately from the expression calculated for $\mathcal{J}_m \gamma_{\rho,r}$ in Part 1. Regarding $\overline{M}_{\gamma_{\rho,r}}$, it follows from Lemma 3.1 that

$$\overline{M}_\gamma \leq \sup_{U' : \gamma \in \mathcal{F}(\Omega, U', z)} \frac{M_\gamma}{\operatorname{diam}(U')}, \quad (\gamma \in \mathcal{F}(\Omega, U, z)).$$

This is why we call $\overline{M}_\gamma$ the *normalised* transformation distance. In Part 1, we proved that $M_{\gamma_{\rho,r}} = |\rho| r$. Therefore, there exists a constant $C > 0$ satisfying

$$\overline{M}_{\gamma_{\rho,r}} \leq \frac{|\rho| r}{\operatorname{diam}(U_r)} \leq C|\rho|.$$ □

Lemma 3.3 (Part 1). *Suppose ϕ is admissible and p-increasing with $1 < p < \infty$, and both $u, f \in \mathrm{BV}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$. Let $x_0 \in J_u \setminus (S_f \cup Z_u)$. Then there exist $\theta \in (0, 1)$, $r_0 > 0$, independent of ρ, and a constant $C = C(\phi, u^\pm(x_0), \widetilde{f}(x_0)) > 0$, such that whenever $0 < r < r_0$ and $0 < \rho < 1$, the functions*

$$\bar{u}_{\rho,r}(x) = \theta u(x) + (1 - \theta) \gamma_{\rho,r\#} u(x), \tag{3.5}$$

satisfy

$$\int_\Omega \phi(\bar{u}_{\rho,r}(x) - f(x)) \, dx + \int_\Omega \phi(\bar{u}_{-\rho,r}(x) - f(x)) \, dx - 2 \int_\Omega \phi(u(x) - f(x)) \, dx \leq -C\rho r^m. \tag{3.6}$$

With these, we may without much difficulty prove Theorem 3.2.

Proof of Theorem 3.2. Since J_u is rectifiable, there exists a family $\{\Gamma_i\}_{i=1}^\infty$ of Lipschitz graphs with $\mathcal{H}^{m-1}(J_u \setminus \bigcup_{i=1}^\infty \Gamma_i) = 0$. If the conclusion of the theorem does not hold, that is $\mathcal{H}^{m-1}(J_u \setminus J_f) > 0$, then for some $i \in \mathbb{Z}^+$ and $\Gamma := \Gamma_i$, also $\mathcal{H}^{m-1}((\Gamma \cap J_u) \setminus J_f) > 0$. We will show that this leads to a contradiction. Since R satisfies Assumption 3.1(ii), it is partially double-Lipschitz comparable in the direction z_Γ for u at almost every $x_0 \in (\Gamma \cap J_u) \setminus J_f$. In particular, since $\mathcal{H}^{m-1}(Z_u) = 0$, we may choose a point $x_0 \in (\Gamma \cap J_u) \setminus (J_f \cup Z_u)$, where R is also partially double-Lipschitz comparable in the direction z_Γ for u. We let v be the function given by Definition 3.4, and pick arbitrary $\epsilon > 0$, $\theta \in (0,1)$. Then for some $r_1 > 0$, every $U \subset B(x_0, r)$, $0 < r < r_1$ and $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega, U, z_\Gamma)$, the estimate holds

$$R(\overline{\gamma}_\# [\![u,v]\!]) + R(\underline{\gamma}_\# [\![u,v]\!]) - 2R(u) \le R^a T_{\overline{\gamma},\underline{\gamma}} |D(u-v)|(\operatorname{cl} U) + (T_{\overline{\gamma},\underline{\gamma}}^{1/2} + r)\epsilon r^m/(1-\theta). \tag{3.7}$$

The overall idea in adapting the proof of the corresponding Theorem in Part 1 now is to apply Lemma 3.3 on the function $q := u - v$ with data $g := f - v$ for v. Indeed

$$\widehat{u}_{\rho,r} := \theta u + (1-\theta)\gamma_\# [\![u,v]\!] = \big(\theta(u-v) + (1-\theta)\gamma_\#(u-v)\big) + v = \bar{q}_{\rho,r} + v,$$

where $\bar{q}_{\rho,r}$ is defined by (3.5). It is important here that $v \in W^{1,1}(\Omega)$ and $x_0 \notin S_v$, so $J_g = J_f$ modulo a $\mathcal{H}^{m-1}$-null set and $(u-v)^+(x_0) - (u-v)^-(x_0) = u^+(x_0) - u^-(x_0)$. Thus by Lemma 3.3 there exist $\theta \in (0,1)$, $r_2 > 0$, and a constant $C = C(\phi, u^\pm(x_0), \widetilde{f}(x_0), v) > 0$, such that whenever $0 < r < r_2$ and $0 < \rho < 1$, then

$$\begin{aligned}\int_\Omega \phi(\widehat{u}_{\rho,r}(x) - f(x))\,dx + &\int_\Omega \phi(\widehat{u}_{-\rho,r}(x) - f(x))\,dx \\ &- 2\int_\Omega \phi(u(x) - f(x))\,dx \le -C\rho r^m.\end{aligned}$$

By convexity, obviously

$$R(\widehat{u}_{\rho,r}) + R(\widehat{u}_{-\rho,r}) - 2R(u) \le (1-\theta)\left(R(\gamma_{\rho,r\#}[\![u,v]\!]) + R(\gamma_{-\rho,r\#}[\![u,v]\!]) - 2R(u)\right),$$

Since the transformations $\gamma_{\rho,r\#}[\![u,v]\!] \in \mathcal{F}(\Omega, U, z_\Gamma)$, and $U_r \subset B(x, \kappa r)$ for some $\kappa > 0$, it follows from (3.7), for $0 < r < r_1/\kappa$ that

$$R(\widehat{u}_{\rho,r}) + R(\widehat{u}_{-\rho,r}) - 2R(u) \le (1-\theta)R^a T_{\gamma_{\rho,r},\gamma_{-\rho,r}} |D(u-v)|(\operatorname{cl} U_r) + (T_{\gamma_{\rho,r},\gamma_{-\rho,r}}^{1/2} + \kappa r)\epsilon(\kappa r)^m. \tag{3.8}$$

Since $x_0 \in J_u \setminus Z_u$, we have

$$|D(u-v)|(\operatorname{cl} U_r) \le 2|u^+(x_0) - u^-(x_0)|\omega_{m-1}(\kappa r)^{m-1}$$

for $0 < r < r_3$ and some $r_3 > 0$. Lemma 3.2 gives

$$T_{\gamma_{-\rho,r},\gamma_{-\rho,r}} \le C'''\rho^2$$

for some constant $C''' > 0$. Setting

$$G(u) := \int_\Omega \phi(u(x) - f(x))\,dx + R(u)$$

and summing (3.6) with (3.8), we observe for some constants $C', C'' > 0$ and every $0 < r < \min\{r_1/\kappa, r_2, r_3\}$ and $0 < \rho < 1$ that

$$G(\widehat{u}_{\rho,r}(x)) + G(\widehat{u}_{-\rho,r}(x)) - 2G(u) \le C'\rho^2 r^{m-1} + C''\rho\epsilon r^m - C\rho r^m + \epsilon(\kappa r)^{m+1}.$$

To see how to make the right hand side negative, let us set $\rho = \bar{\rho} r^m$. Then we get

$$G(\widehat{u}_{\rho,r}(x)) + G(\widehat{u}_{-\rho,r}(x)) - 2G(u) \le (C'\bar{\rho}^2 + C''\bar{\rho}\epsilon - C\bar{\rho} + \epsilon\kappa^{m+1})r^{m+1}.$$

We first pick $\bar{\rho}$ small enough that $C'\bar{\rho} < C/4$. Then we pick $\epsilon > 0$ small enough that $C''\epsilon < C/4$ and $\epsilon\kappa^{m+1} < \bar{\rho}C/4$. This will force $r > 0$ small, but will give

$$G(\widehat{u}_{\rho,r}(x)) + G(\widehat{u}_{-\rho,r}(x)) - 2G(u) \le -C\bar{\rho}r^{m+1}/4,$$

which is negative. This says that $\min\{G(\widehat{u}_{\rho,r}(x)), G(\widehat{u}_{-\rho,r}(x))\} < G(u)$. Thus we produce a contradiction to u minimising G. ∎

4. Approximation results

In this section, we study study two aspects of approximation. The first is how well we can approximate functions of bounded deformation (or variation, for the matter) by differentials ∇v of functions $v \in W^{1,1}(\Omega)$. These approximations form the basis of proving partial double-Lipschitz comparability. The second aspect that we study is the approximation of a function $u \in \mathrm{BD}(\Omega)$ in terms of TGV-strict convergence, or generally convergence such that $u^i \to u$ weakly* in $\mathrm{BV}(\Omega)$ and $\|Du^i - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \to \|Du^i - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}$ for $w \in L^1(\Omega)$.

4.1. Local approximation in BD(Ω)

One of our most critical concepts is stated in the following definition.

Definition 4.1. We say that $w \in \mathrm{BD}(\Omega)$ is *BV-differentiable* at $x \in \Omega$ if there exists $\widehat{w}_x \in \mathrm{BV}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ such that

$$\lim_{r \searrow 0} \fint_{B(x,r)} \frac{\|w(y) - \widehat{w}_x(y)\|}{r}\, dy = 0.$$

Remark 4.1. Clearly w is BV-differentiable if actually $w \in \mathrm{BV}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$. On a related note, the $\mathrm{BV}_{\mathrm{loc}}$ assumption was also required in [2] for the study of traces of another function u with respect to $|D^s w|$.

Proposition 4.1. *Every $u \in \mathrm{BD}(\Omega)$ is BV-differentiable at $\mathcal{L}^m$-a.e. $x \in \Omega$.*

Proof. We know from [1, 23] that w is *approximately differentiable* at $\mathcal{L}^m$-a.e. $x \in \Omega$ in the sense of existence of $L = \nabla u(x) \in \mathbb{R}^{m \times m}$ such that

$$\lim_{r \searrow 0} \fint_{B(x,r)} \frac{\|w(y) - w(x) - L(y-x)\|}{r}\, dy = 0.$$

It therefore suffices to set $\widehat{w}_x(y) := w(x) + L(y-x)$. □

Remark 4.2. The domain of BV-differentiability is however potentially larger than approximate differentiability. A simple piece of evidence for this is the fact that any $w \in \mathrm{BV}_{\mathrm{loc}}(\Omega;\mathbb{R}^m)$ is BV-differentiable everywhere, but not approximately differentiable on the jump set J_w.

That we can show $\mathcal{L}^m$-a.e. approximate differentiability is not entirely satisfying. We would prefer to have the property $\mathcal{H}^{m-1} \llcorner J_w$-a.e. Whether this can be achieved at least for w a solution to (1.3), remains an interesting open question.

We will need the following simple result for our main application of BV-differentiability stated after it.

Lemma 4.1. *Suppose $w \in \mathrm{BD}(\Omega)$ is BV-differentiable at $x \in J_w$. Then $x \in J_{\hat{w}_x}$ with $w^+(x) = \widehat{w}_x^+(x)$ and $w^-(x) = \widehat{w}_x^-(x)$.*

Proof. By the definition of BV-differentiability

$$\lim_{r \searrow 0} \fint_{B(x,r)} \|w(y) - \widehat{w}_x(y)\|\, dy \le \lim_{r \searrow 0} \fint_{B(x,r)} \frac{\|w(y) - \widehat{w}_x(y)\|}{r}\, dy = 0.$$

This implies that w and $\widehat{w}_x$ have the same one-sided limits at x. □

The next lemma provides one of the most important ingredients of our approach to proving partial double-Lipschitz comparability for higher-order regularisation functionals.

Lemma 4.2. *Let $w \in \mathrm{BD}(\Omega)$, $x \in \Omega$, and $\Gamma \ni x$ be a C^1 $(m-1)$-graph. Suppose that w has traces $w^\pm(x)$ from both sides of Γ at x, and $P_{z_\Gamma}^\perp(w^+(x) - w^-(x)) = 0$. Then there exists $v \in W^{1,1}_{\mathrm{loc}}(\Omega)$ with $x \notin S_v$, satisfying*

$$\lim_{r \searrow 0} \fint_{B(x,r)} \|w - \nabla v\|\, dy = 0. \tag{4.1}$$

If w is moreover BV-differentiable at x, then given $\epsilon > 0$, there exists $r_\epsilon > 0$ such that every $U \subset B(x,r)$, $0 < r < r_\epsilon$, and $\gamma \in \mathcal{F}(\Omega, U, z_\Gamma)$ satisfy

$$\int_U \|\gamma_\#(w - \nabla v) - (w - \nabla v)\| \, dy \le (\overline{M}_\gamma + r)\epsilon r^m. \tag{4.2}$$

If $x \in \Omega \setminus S_w$, then we may take $\gamma \in \mathcal{F}(\Omega, U)$ (without any specification of Γ).

Proof. We first prove the results for $\gamma \in \mathcal{F}(\Omega, U, z_\Gamma)$ with Γ specified. We denote for short $z := z_\Gamma$, and let

$$V^\pm := \{x \in \mathbb{R}^m \mid \pm\langle z, x\rangle > 0\}.$$

We define the transformation $\psi : \mathbb{R}^m \to \mathbb{R}^m$ by

$$\psi(y) := y + f_\Gamma(P_z^\perp y)z,$$

Then $\psi(V_\Gamma) = g_\Gamma(V_\Gamma) = \Gamma$. We observe also that

$$\psi^{-1}(y) = y - f_\Gamma(P_z^\perp y)z. \tag{4.3}$$

Therefore

$$\nabla\psi^{-1}(y) = I - (P_z^\perp)^* \nabla f_\Gamma(P_z^\perp y) \otimes z,$$

Since $\langle z, (P_z^\perp)^* \nabla f_\Gamma(P_z^\perp y)\rangle = 0$ we find that $\nabla\psi^{-1}(x)$ is invertible. Because ∇f_Γ is by assumption continuous, this implies that

$$\Psi(y) := \nabla\psi^{-1}(y)\nabla\psi(\psi^{-1}(x)) = \nabla\psi^{-1}(y)[\nabla\psi^{-1}(x)]^{-1}$$

is continuous with $\Psi(x) = I$. More precisely for any $\epsilon > 0$, for suitable $r_\epsilon > 0$,

$$\|\Psi(y) - I\| \le \epsilon, \quad (\|P_z^\perp(y - x)\| \le r_\epsilon). \tag{4.4}$$

We then let

$$\begin{aligned}\bar{v}(y) := &\langle [\nabla\psi(\psi^{-1}(x))]^* w^+(x), y - \psi^{-1}(x)\rangle \chi_{V^+}(y) \\ &+ \langle [\nabla\psi(\psi^{-1}(x))]^* w^-(x), y - \psi^{-1}(x)\rangle \chi_{V^-}(y),\end{aligned}$$

and

$$v := \psi_\# \bar{v}.$$

Recalling (4.3), we observe that v is continuous and differentiable, $v \in W^{1,1}_{\mathrm{loc}}(\Omega) \cap C(\Omega)$ and $x \notin S_v$. This is the only place where we need the assumption $P_{z_\Gamma}^\perp(w^+(x) - w^-(x)) = 0$. Defining

$$w_0(y) := w^+(x)\chi_{\Gamma^+}(y) + w^-(x)\chi_{\Gamma^-}(y),$$

we get

$$\nabla v(y) = \Psi(y) w_0(y).$$

Moreover, given $\epsilon > 0$, by the definition of the one-sided limits $w^\pm(x)$, we have for some $r_\epsilon > 0$ that

$$⨍_{B(x,r)} \|w(y) - w_0(y)\| \, dy \le \epsilon, \quad (0 < r < r_\epsilon). \tag{4.5}$$

Thus with $C = 2\max\{w^+(x), w^-(x)\}$, recalling (4.4), we obtain

$$⨍_{B(x,r)} \|w_0(y) - \nabla v(y)\| \, dy \le ⨍_{B(x,r)} \|\Psi(y) - I\| \|w_0(y)\| \, dy \le C\epsilon, \quad (0 < r < r_\epsilon). \tag{4.6}$$

Combined (4.5) and (4.6) give

$$⨍_{B(x,r)} \|w(y) - \nabla v(y)\| \, dy \le (1 + C)\epsilon, \quad (0 < r < r_\epsilon).$$

Since $\epsilon > 0$ was arbitrary, we conclude that (4.1) holds.

We now have to prove (4.2), assuming that w is BV-differentiable at x. We begin by observing that $q := w - w_0$ is then BV-differentiable with $\widehat{q}_x = \widehat{w}_x - w_0$. Moreover

$$\begin{aligned}\int_U \|\gamma_\#(\widehat{w}_x - w_0)(y) - (\widehat{w}_x(y) - w_0(y))\|\, dy &\le \overline{M}_\gamma \operatorname{diam}(U)|D(\widehat{w}_x - w_0)|(U). \\ &\le C_m \overline{M}_\gamma r |D(\widehat{w}_x - w_0)|(U),\end{aligned} \tag{4.7}$$

for some dimensional constant C_m needed to apply (3.2) to vector fields. By assumption $\gamma \in \mathcal{F}(\Omega, U, z)$. Thus $P_z^\perp \gamma^{-1}(y) = P_z^\perp y$, which implies

$$\Psi \circ \gamma^{-1} = \Psi. \tag{4.8}$$

Consequently

$$\begin{aligned}\gamma_\#(\nabla v - w_0) - (\nabla v - w_0) &= \gamma_\#([\Psi - I]w_0) - [\Psi - I]w_0 \\ &= [\Psi - I](\gamma_\# w_0 - w_0).\end{aligned}$$

Using (4.4) again,

$$\begin{aligned}&\int_U \|\gamma_\#(\nabla v - w_0)(y) - (\nabla v - w_0)(y)\|\, dy \\ &\qquad \le \epsilon \int_U \|\gamma_\# w_0(y) - w_0(y)\|\, dy \\ &\qquad \le \epsilon C_m \overline{M}_\gamma r |Dw_0|(U) \le \overline{M}_\gamma C' \epsilon r^m, \quad (0 < r < r_\epsilon),\end{aligned} \tag{4.9}$$

for suitable $r_\epsilon > 0$ and some constant $C' = C'(\Gamma, w^\pm(x_0))$. Choosing $r_\epsilon > 0$ small enough, we may now for $0 < r < r_\epsilon$ finally approximate

$$\begin{aligned}&\int_U \|(\gamma_\# w(y) - \gamma_\# \nabla v(y)) - (w(y) - \nabla v(y))\|\, dy \\ &\qquad \le \int_U \|(\gamma_\# w_0(y) - \gamma_\# \nabla v(y)) - (w_0(y) - \nabla v(y))\|\, dy \\ &\qquad\quad + \int_U \|(\gamma_\# \widehat{w}_x(y) - \gamma_\# w_0(y)) - (\widehat{w}_x(y) - w_0(y))\|\, dy \\ &\qquad\quad + \int_U \|\gamma_\# \widehat{w}_x(y) - \gamma_\# w(y)\|\, dy + \int_U \|\widehat{w}_x(y) - w(y)\|\, dy \\ &\qquad \le \overline{M}_\gamma r |D(\widehat{w}_x - w_0)|(U) + \overline{M}_\gamma C' \epsilon r^m + \epsilon r^{m+1}.\end{aligned} \tag{4.10}$$

For the final inequality, we have used (4.7) and (4.9) for the two first terms on the left hand side, and the definition of BV-differentiability with the area formula for the last two terms. Referring to Lemma 4.1 and (2.2), we now observe for suitable $r_\epsilon > 0$ that

$$|D(\widehat{w}_x - w_0)|(U) \le \epsilon r^{m-1}, \quad (0 < r < r_\epsilon).$$

The arbitrariness of $\epsilon > 0$ allows us to get rid of the constant factors in (4.10), and thus conclude the proof of (4.2) in the case that $\gamma \in \mathcal{F}(\Omega, U, z_\Gamma)$.

If $x \in \Omega \setminus S_w$, and $\gamma \in \mathcal{F}(\Omega, U)$ (without any specification of Γ), we set

$$\psi(y) := y.$$

Then

$$\bar{v}(y) = v(y) = w_0(y) \equiv \widetilde{w}(x).$$

Also

$$\Psi(y) = I,$$

so we get

$$\gamma_\#(\nabla v - w_0) - (\nabla v - w_0) = 0,$$

and do not need the property (4.8), which is the only place where we used the fact that $\gamma \in \mathcal{F}(\Omega, U, z_\Gamma)$. Indeed, instead of (4.4), we have the stronger property

$$\int_U \|\gamma_\#(\nabla v - w_0)(y) - (\nabla v - w_0)(y)\|\, dy = 0.$$

The rest follows as before. ☐

Remark 4.3. Our main reason for introducing the notion of BV-differentiability is to be able to perform the pushforward approximation in (4.2). For this it would also suffice to require the existence of $\widehat{w}^{\perp}_{x,z} \in \mathrm{BV}_{\mathrm{loc}}(\Omega; z^{\perp})$ satisfying

$$\lim_{r\searrow 0} \fint_{B(x,r)} \frac{\|P_z^{\perp} w(y) - \widehat{w}^{\perp}_{x,z}(y)\|}{r}\, dy = 0. \tag{4.11}$$

Here $z := z_\Gamma$. This holds because the slice $u_y^z(t) := \langle z, w(y+tz)\rangle \in \mathrm{BV}(\Omega_y^z)$ for $\Omega_y^z := \{t \in \mathbb{R} \mid y + tz \in \Omega\}$, and $\mathcal{H}^{m-1}$-a.e. $y \in P_z^{\perp}\Omega$ [1]. Unfortunately, we do not know much about the slices $t \mapsto \langle e, w(y+tz)\rangle$ for $e \perp z$, and therefore need to assume BV-differentiability. Combined with (4.11), the assumption $P_{z_\Gamma}^{\perp}(w^+(x) - w^-(x)) = 0$ that we required reduces into $x \notin S_{\hat{w}^{\perp}_{x,z}}$. In other words, we an approximate $P_z^{\perp} w$ by a BV function for which x is a Lebesgue point. What all this means is that we need to assume extra regularity from w in directions parallel to the plane $z_\Gamma^{\perp}$, but do not need to assume anything beyond $w \in \mathrm{BD}(\Omega)$ along z_Γ.

4.2. TGV-strict smooth approximation

We now study alternative forms of strict convergence in $\mathrm{BV}(\Omega)$.

Theorem 4.1. *Suppose $\Omega \subset \mathbb{R}^m$ is open and let $(u, w) \in \mathrm{BV}(\Omega) \times \mathrm{BD}(\Omega)$. Then there exists a sequence $\{(u^i, w^i)\}_{i=1}^{\infty} \in C^{\infty}(\Omega) \times C^{\infty}(\Omega; \mathbb{R}^m)$ with*

$$\begin{aligned} &u^i \to u \text{ in } L^1(\Omega), && \|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \to \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}, \\ &w^i \to w \text{ in } L^1(\Omega;\mathbb{R}^m), && \|Ew^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} \to \|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}. \end{aligned}$$

If only $w \in L^1(\Omega; \mathbb{R}^m)$, then we get the three first converges, but not the fourth one.

Proof. The proof follows the outlines of the approximation of $u \in \mathrm{BV}(\Omega)$ in terms of strict convergence in $\mathrm{BV}(\Omega)$, see [3, Theorem 3.9]. We just have to add a few extra steps to deal with w, which is approximated similarly.

To start with the proof, given a positive integer m, we set $\Omega_0 = \emptyset$ and

$$\Omega_k := B(0, k+m) \cap \{x \in \Omega \mid \inf_{y \in \partial\Omega} \|x - y\| \ge 1/(m+k))\}.$$

We pick m large enough that

$$|Du - w|(\Omega \setminus \Omega_1) < 1/i \quad \text{and} \quad |Ew|(\Omega \setminus \Omega_1) < 1/i. \tag{4.12}$$

With

$$V_k := \Omega_{k+1} \setminus \mathrm{cl}\, \Omega_{k-1},$$

each $x \in \Omega$ belongs to at most four sets V_k. We may then find a partition of unity $\{\zeta_k\}_{k=1}^{\infty}$ with $\zeta_k \in C_c^{\infty}(V_k)$, $0 \le \zeta_k \le 1$ and $\sum_{k=1}^{\infty} \zeta_k \equiv 1$ on Ω.

With $\{\rho_\epsilon\}_{\epsilon>0}$ a family of mollifiers, and $\epsilon_k > 0$, we let

$$u_k := \rho_{\epsilon_k} * (u\zeta_k), \quad \text{and} \quad w_k := \rho_{\epsilon_k} * (w\zeta_k).$$

We select $\epsilon_k > 0$ small enough that $\operatorname{supp} u_k, \operatorname{supp} w_k \subset V_k$ (doable because $\zeta_k \in C_c^{\infty}(V_k)$), and such that the estimates

$$\|u_k - u\zeta_k\| \le 1/(2^k i), \quad \text{and} \quad \|\rho_{\epsilon_k} * (u\nabla\zeta_k) - u\nabla\zeta_k\| \le 1/(2^k i), \tag{4.13}$$

hold, as do

$$\|w_k - w\zeta_k\| \le 1/(2^k i), \quad \text{and} \quad \|\rho_{\epsilon_k} * (w \otimes \nabla\zeta_k) - w \otimes \nabla\zeta_k\| \le 1/(2^k i). \tag{4.14}$$

We then let

$$u^i(x) := \sum_{k=1}^{\infty} u_k(x), \quad \text{and} \quad w^i(x) := \sum_{k=1}^{\infty} w_k(x).$$

By the construction of the partition of unity, for every $x \in \Omega$ there exists a neighbourhood of x such the these sums have only finitely many non-zero terms. Hence $u^i \in C^{\infty}(\Omega)$. Moreover, as $u = \sum_{k=1}^{\infty} \zeta_k u$, (4.13) gives

$$\|u - u^i\| \le \sum_{k=1}^{\infty} \|u_k - u\zeta_k\| < 1/i.$$

Thus $u^i \to u$ in $L^1(\Omega)$ as $i \to \infty$. Completely analogously $w^i \to w$ in $L^1(\Omega;\mathbb{R}^m)$ as $i \to \infty$.

By lower semicontinuity of the total variation, we have

$$\|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \leq \liminf_{i\to\infty} \|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}, \quad \text{and}$$
$$\|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} \leq \liminf_{i\to\infty} \|Ew^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}.$$

It therefore only remains to prove the opposite inequalities. Let $\varphi \in C_c^1(\Omega;\mathbb{R}^m)$ with $\sup_{x\in\Omega} |\varphi(x)| \leq 1$. We have

$$\begin{aligned}
\int_\Omega \operatorname{div}\varphi(x) u_k(x) &= \int_\Omega \operatorname{div}\varphi(x)(\rho_{\epsilon_k} * \zeta_k u)(x)\,dx \\
&= \int_\Omega \operatorname{div}(\rho_{\epsilon_k} * \varphi)(x)\zeta_k(x)u(x)\,dx \\
&= \int_\Omega \operatorname{div}[\zeta_k(\rho_{\epsilon_k} * \varphi)](x)u(x)\,dx - \int_\Omega \langle \nabla\zeta_k(x), (\rho_{\epsilon_k} * \varphi)(x)\rangle u(x)\,dx \\
&= \int_\Omega \operatorname{div}[\zeta_k(\rho_{\epsilon_k} * \varphi)](x)u(x)\,dx \\
&\quad - \int_\Omega \langle \varphi(x), (\rho_{\epsilon_k} * (u\nabla\zeta_k))(x) - (u\nabla\zeta_k)(x)\rangle\,dx \\
&\quad - \int_\Omega \langle \varphi(x), (u\nabla\zeta_k)(x)\rangle\,dx.
\end{aligned}$$

Since $\sum_{k=1}^\infty \nabla\zeta_k = 0$, we have

$$\sum_{k=1}^\infty \int_\Omega \langle \varphi(x), (u\nabla\zeta_k)(x)\rangle\,dx = 0.$$

Thus using (4.13), we get

$$\begin{aligned}
\int_\Omega \operatorname{div}\varphi(x) u^i(x) &= \sum_{k=1}^\infty \int_\Omega \operatorname{div}\varphi(x) u_k(x) \\
&= \sum_{k=1}^\infty \int_\Omega \operatorname{div}[\zeta_k(\rho_{\epsilon_k} * \varphi)](x)u(x)\,dx \\
&\quad - \sum_{k=1}^\infty \left(\int_\Omega \langle \varphi(x), (\rho_{\epsilon_k} * (u\nabla\zeta_k))(x) - (u\nabla\zeta_k)(x)\rangle\,dx\right) \\
&\leq \sum_{k=1}^\infty \int_\Omega \operatorname{div}\varphi_k(x)u(x)\,dx + 1/i.
\end{aligned}$$

In the final step, we have set

$$\varphi_k := \zeta_k(\rho_{\epsilon_k} * \varphi).$$

By the definition of w^i, we also have

$$\begin{aligned}
\int_\Omega \langle \varphi(x), w^i(x)\rangle\,dx &= \sum_{k=1}^\infty \int_\Omega \langle \varphi(x), [\rho_{\epsilon_k} * (w\zeta_k)](x)\rangle\,dx \\
&= \sum_{k=1}^\infty \int_\Omega \langle \varphi_k(x), w(x)\rangle\,dx.
\end{aligned}$$

Observing that $-1 \leq \varphi_k \leq 1$, and using the fact that $\sum_{k=1}^\infty \chi_{V_k} \leq 4$, we further get

$$\begin{aligned}
\int_\Omega \varphi(x)\,d[Du^i - w^i](x) &= -\int_\Omega \operatorname{div}\varphi(x)u^i(x) + \langle \varphi(x), w^i(x)\rangle\,dx \\
&\leq -\int_\Omega \operatorname{div}\varphi_1(x)u(x) + \langle \varphi_1(x), w(x)\rangle\,dx \\
&\quad - \sum_{k=2}^\infty \int_\Omega \operatorname{div}\varphi_k(x)u(x) + \langle \varphi_k(x), w(x)\rangle\,dx + 1/i \\
&\leq |Du - w|(\Omega) + \sum_{k=2}^\infty |Du - w|(V_k) + 1/i \\
&\leq |Du - w|(\Omega) + 4|Du - w|(\Omega \setminus \Omega_1) + 1/i \\
&\leq |Du - w|(\Omega) + 5/i.
\end{aligned}$$

In the final step we have used (4.12). This shows that $\|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \to \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}$.

Next we recall that

$$|Ew|(\Omega) = \sup_{\varphi \in C_c^\infty(\Omega;\mathrm{Sym}^{n\times n}) : \|\varphi(x)\|_\infty \le 1} \int_\Omega \langle \operatorname{div} \varphi(x), w(x)\rangle \, dx$$

with the divergence taken columnwise. Therefore, arguments analogous to the ones above show that $\|Ew^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} \to \|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}$ if $w \in \mathrm{BD}(\Omega)$. If only $w \in L^1(\Omega;\mathbb{R}^m)$, then we do not get this converges, but the proof of the other converges did not depend on $w \in \mathrm{BD}(\Omega)$ at all. This concludes the proof. □

For our present needs, the most important corollary of the above theorem is the following.

Corollary 4.1. *Suppose $\Omega \subset \mathbb{R}^m$ is open and let $(u, w) \in \mathrm{BV}(\Omega) \times L^1(\Omega;\mathbb{R}^m)$. Then there exists a sequence $\{u^i\}_{i=1}^\infty \in C^\infty(\Omega)$ with*

$$u^i \to u \text{ in } L^1(\Omega) \quad \text{and} \quad \|Du^i - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \to \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}, \tag{4.15}$$

as well as $Du^i \stackrel{*}{\rightharpoonup} Du$ weakly in $\mathcal{M}(\Omega;\mathbb{R}^m)$.*

Proof. Let $\{(u^i, w^i)\}_{i=1}^\infty \in C^\infty(\Omega) \times C^\infty(\Omega;\mathbb{R}^m)$ be given by Theorem 4.1. Then

$$\begin{aligned} \lim_{i\to\infty} \|Du^i - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} &\le \lim_{i\to\infty} \left(\|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \|w^i - w\|_{2,L^1(\Omega;\mathbb{R}^m)}\right) \\ &= \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}. \end{aligned}$$

Analogously we deduce

$$\lim_{i\to\infty} \|Du^i - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \ge \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}.$$

This gives (4.15). Clearly, by moving to a subsequence of the original bounded sequence, we may further force $Du^i \stackrel{*}{\rightharpoonup} Du$ weakly* in $\mathcal{M}(\Omega;\mathbb{R}^m)$. □

The following corollary shows the approximability of $u \in \mathrm{BV}(\Omega)$ in terms of TGV^2-strict convergence. It is of course easy to extend to TGV^k for $k > 2$.

Corollary 4.2. *Suppose $\Omega \subset \mathbb{R}^m$ is open and let $u \in \mathrm{BV}(\Omega)$. Then there exists a sequence $\{u^i\}_{i=1}^\infty \in C^\infty(\Omega)$ with $u^i \to u$ in $L^1(\Omega)$, $Du^i \stackrel{*}{\rightharpoonup} Du$ weakly* in $\mathcal{M}(\Omega;\mathbb{R}^m)$, and $\mathrm{TGV}^2_{(\beta,\alpha)}(u^i) \to \mathrm{TGV}^2_{(\beta,\alpha)}(u)$ for any $\alpha, \beta > 0$.*

Proof. Let w achieve the minimum in the differentiation cascade of definition (1.3) of TGV^2, the minimiser existing by [10]. Let then the sequence $\{(u^i, w^i)\}_{i=1}^\infty \in C^\infty(\Omega) \times C^\infty(\Omega;\mathbb{R}^m)$ be given by Theorem 4.1. As in the proof of Corollary 4.1, we may assume that $Du^i \stackrel{*}{\rightharpoonup} Du$ weakly* in $\mathcal{M}(\Omega;\mathbb{R}^m)$.

To see the convergence of $\mathrm{TGV}^2_{(\beta,\alpha)}(u^i)$ to $\mathrm{TGV}^2_{(\beta,\alpha)}(u)$, we observe that by definition

$$\mathrm{TGV}^2_{(\beta,\alpha)}(u^i) \le \alpha\|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|Ew^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}$$

Moreover

$$\lim_{i\to\infty} \alpha\|Du^i - w^i\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|Ew^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} = \mathrm{TGV}^2_{(\beta,\alpha)}(u).$$

Since the TGV^2 functional is lower semicontinuous with respect to weak* convergence in $\mathrm{BV}(\Omega)$ ([9], see also Lemma 5.1 below), the claim follows. □

5. Higher-order regularisers

We now study partial double-Lipschitz comparability of second- and higher-order regularisers. We start in Section 5.1 with TGV, after which in Section 5.2 we consider variants of TGV^2 for which we have stronger results than TGV^2 itself. We finish in Section 5.3 with infimal convolution TV.

5.1. Second-order total generalised variation

Total generalised variation was introduced in [9] as a higher-order extension of TV that avoids the stair-casing effect. Following the *differentiation cascade* formulation of [11, 10], it may be defined for $u \in \mathrm{BV}(\Omega)$ and $\vec{\alpha} = (\beta, \alpha)$ as

$$\mathrm{TGV}^2_{\vec{\alpha}}(u) := \min_{w \in L^1(\Omega;\mathbb{R}^m)} \alpha \|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta \|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}, \tag{5.1}$$

with a minimising $w \in \mathrm{BD}(\Omega)$ existing. Clearly

$$\mathrm{TGV}^2_{\vec{\alpha}}(u) \leq \alpha \mathrm{TV}(u).$$

Moreover, $\mathrm{TGV}^2_{\vec{\alpha}}$ is a seminorm. In fact, it turns out that the norms $\|u\|_{L^1} + \mathrm{TV}(u)$ and $\|u\|_{L^1} + \mathrm{TGV}^2_{\vec{\alpha}}(u)$ are equivalent, as shown in [11, 10]. In other words $\mathrm{TGV}^2_{\vec{\alpha}}$ induces the same topology in $\mathrm{BD}(\Omega)$ as TV does, but different geometry, as can be witnessed from often much improved behaviour in practical image processing tasks.

Lemma 5.1. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Then there exist constants $c, C > 0$, dependent on Ω, such that for all $u \in L^1(\Omega)$ it holds*

$$c\big(\|u\|_{L^1(\Omega)} + \|Du\|_{\mathcal{M}(\Omega;\mathbb{R}^m)}\big) \leq \|u\|_{F,1} + \mathrm{TGV}^2_{(\beta,\alpha)}(u) \leq C\big(\|u\|_{L^1(\Omega)} + \|Du\|_{\mathcal{M}(\Omega;\mathbb{R}^m)}\big). \tag{5.2}$$

Moreover, the functional $\mathrm{TGV}^2_{\vec{\alpha}}$ is lower semicontinuous with respect to weak convergence in $\mathrm{BV}(\Omega)$.*

Proof. Lower semicontinuity is proved in [9] for the original dual ball formulation. Equivalence to the differentiation cascade formulation presented here is proved in [11, 10], where the norm equivalence is also proved. □

The following proposition states what we can say about partial double-Lipschitz comparability of standard TGV^2. Unfortunately, we cannot prove Assumption 3.1(ii) quite exactly, only for $\mathcal{H}^{m-1}$-a.e. $x \in \Gamma \cap D_w \cap O^\Gamma_w$, where we denote by $D_w \subset \Omega$ the set of points where w is BV-differentiable, and by O^Γ_w the set of points $x \in \Gamma$ where $P^\perp_{z_\Gamma}(w^+(x) - w^-(x)) = 0$.

Proposition 5.1. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Then $\mathrm{TGV}^2_{\vec{\alpha}}$ is an admissible regularisation functional on $L^1(\Omega)$ and satisfies Assumption 3.1(i). Moreover, for any $u \in \mathrm{BV}(\Omega)$, a minimiser $w \in \mathrm{BD}(\Omega)$ of (5.1), and any Lipschitz $(m-1)$-graph $\Gamma \subset \Omega$, the following holds.*

(ii') $\mathrm{TGV}^2_{\vec{\alpha}}$ is partially double-Lipschitz comparable for u in the direction z_Γ at $\mathcal{H}^{m-1}$-a.e. $x \in \Gamma \cap D_w \cap O^\Gamma_w$.

The basic idea of the proof is similar to the proof double-Lipschitz comparability of TV in Part 1, but we need to deal with w as well. This adds significant extra complications. One of them is the use of the symmetrised gradient Ew, which does not allow us to use estimates of the type in Lemma 3.1. We need the BV-differentiability of Section 4.1 here. Also, the variable w alone is problematic in the expression $D\gamma_\# u - w$ for the use of the area formula. In order to deal with it, we have to take something, ∇v, out w, and shift this into u. Finally, we need to be careful with the jump set of w, also removing it from some estimates.

Proof. We know from Lemma 5.1 that $\mathrm{TGV}^2_{\vec{\alpha}}$ is lower semi-continuous with respect to weak* convergence in $\mathrm{BV}(\Omega)$, and that (3.1) holds. It therefore only remains to prove Assumption 3.1(i) and (ii'), that is partial double-Lipschitz comparability $\mathcal{L}^m$-a.e. , and, for any given Lipschitz graph Γ, in the direction z_Γ at $\mathcal{H}^{m-1}$-a.e. $x \in \Gamma \cap D_w \cap O^\Gamma_w$.

We pick arbitrary $w \in \mathrm{BD}(\Omega)$ achieving the minimum in (5.1). Regarding Assumption 3.1(i), we first of all observe that $\mathcal{L}^m(\Omega \setminus Q) = 0$ for $Q := D_w \setminus S_w$. We claim that $\mathrm{TGV}^2_{\vec{\alpha}}$ is partially double-Lipschitz comparable for u at every $x \in Q$. Regarding (ii'), in order to apply Lemma 4.2, we need a C^1 $(m-1)$-graph. Indeed, as a consequence of the Whitney extension theorem [22, 3.1.14] and Lusin's theorem applied to f_Γ, we may cover Γ by C^1 $(m-1)$-graphs $\{\Lambda_i\}_{i=1}^\infty$ satisfying $z_{\Lambda_i} = z_\Gamma$, and $\mathcal{H}^{m-1}(\Gamma \setminus \bigcup_{i=1}^\infty \Lambda_i) = 0$. If we show that $\mathrm{TGV}^2_{\vec{\alpha}}$ is partially double-Lipschitz comparable for u in the direction z_Γ at $\mathcal{H}^{m-1}$-a.e. $x \in \Lambda_i \cap D_w \cap O^\Gamma_w$, for every $i \in \mathbb{Z}^+$, the claim will follow.

To show Assumption 3.1(i), we apply Lemma 4.2 at a point $x \in Q$. To show (ii'), we apply the lemma at a point $x \in \Lambda_i \cap D_w \cap O^\Gamma_w$, $(i \in \mathbb{Z}^+)$, where the traces $w^\pm(x)$ from both sides of Λ_i exist and $P^\perp_{z_\Gamma}(w^+(x) - w^-(x)) = 0$. This set, which we denote Q_i, satisfies $\mathcal{H}^{m-1}((\Lambda_i \cap D_w \cap O^\Gamma_w) \setminus Q_i) = 0$ for each $i \in \mathbb{Z}^+$. This is exactly what we need.

We fix x and let $U \subset B(x,r)$ for suitable $r>0$. Lemma 4.2 then gives us $v \in W^{1,1}(\Omega)$ with $x \notin S_v$, and for each $\epsilon > 0$ for $0 < r < r_\epsilon$ the estimates

$$\int_U \|\nabla v - w\|\,dy \le \epsilon r^m, \tag{5.3}$$

and

$$\int_U \|\gamma_\#(w-\nabla v) - (w - \nabla v)\|\,dy \le (\overline{M}_\gamma + r)\epsilon r^m/2. \tag{5.4}$$

We define

$$u_\gamma := \gamma_\#[\![u,v]\!] = \gamma_\#(u-v) + v, \quad (\gamma = \overline{\gamma}, \underline{\gamma}).$$

If we also set

$$G(u,w) := \alpha\|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|Ew\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}, \tag{5.5}$$

then $\mathrm{TGV}^2_{\vec{\alpha}}(u') \le G(u',w')$ for all $(u',w') \in \mathrm{BV}(\Omega)\times \mathrm{BD}(\Omega)$. To prove partial double-Lipschitz comparability for u at x, it therefore suffices to prove for any $\epsilon > 0$ the existence of $r_0 > 0$ such that for any $0 < r < r_0$, $U \subset B(x,r)$, and $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega, U)$, resp. $\overline{\gamma}, \underline{\gamma} \in \mathcal{F}(\Omega, U, z_\Gamma)$, that

$$G(u_{\overline{\gamma}}, w) + G(u_{\underline{\gamma}}, w) - 2G(u,w) \le T_{\overline{\gamma},\underline{\gamma}}|(Du - v)|(\mathrm{cl}\, U) + (T_{\overline{\gamma},\underline{\gamma}}^{1/2} + r)\epsilon r^m, \tag{5.6}$$

for w achieving the minimum in (5.1) for u.

We suppose first that $u \in W^{1,1}(\Omega)$. With $\gamma = \overline{\gamma}, \underline{\gamma}$, by a lemma in Part 1, we have $\gamma_\#\nabla u = \nabla\gamma^{-1}\gamma_\#\nabla u$. Thus we may expand

$$\begin{aligned}
\nabla u_\gamma - w &= \nabla\gamma_\#(u-v) + \nabla v - w \\
&= \nabla\gamma^{-1}\gamma_\#(\nabla u - \nabla v) + \gamma_\#(\nabla v - w) \\
&\quad - [\gamma_\#(\nabla v - w) - (\nabla v - w)] \\
&= \nabla\gamma^{-1}\gamma_\#(\nabla u - w) + (I - \nabla\gamma^{-1})\gamma_\#(\nabla v - w) \\
&\quad - [\gamma_\#(\nabla v - w) - (\nabla v - w)].
\end{aligned}$$

It follows

$$\begin{aligned}
\int_U \|\nabla u_\gamma - w\|\,dy \le &\int_U \|\nabla\gamma^{-1}(\gamma)(\nabla u - w)\|\mathcal{J}_m\gamma\,dy \\
&+ \int_U \|(I - \nabla\gamma^{-1}(\gamma))(\nabla v - w)\|\mathcal{J}_m\gamma\,dy \\
&+ \int_U \|\gamma_\#(\nabla v - w) - (\nabla v - w)\|\,dy.
\end{aligned} \tag{5.7}$$

With $\gamma = \overline{\gamma}, \underline{\gamma}$, using (5.3) we get

$$\begin{aligned}
\int_U \|(I - \nabla\gamma^{-1}(\gamma))(\nabla v - w)\|\mathcal{J}_m\gamma\,dy &\le D_\gamma \int_U \|\nabla v - w\|\mathcal{J}_m\gamma\,dy \\
&\le D_\gamma(J_\gamma + 1)\int_U \|\nabla v - w\|\,dy \\
&\le D_\gamma(J_\gamma + 1)\epsilon r^m.
\end{aligned} \tag{5.8}$$

Using (5.8) and (5.4) in (5.7), we see that

$$\begin{aligned}
\int_U \|\nabla u_\gamma - w\|\,dy \le &\int_U \|A_\gamma(\nabla u - w)\|\,dy + D_\gamma(J_\gamma + 1)\epsilon r^m \\
&+ (\overline{M}_\gamma + r)\epsilon r^m/2.
\end{aligned} \tag{5.9}$$

Also by (5.3)

$$\int_U \|\nabla u - w\|\,dy \le \int_U \|\nabla u - \nabla v\|\,dy + \epsilon r^m.$$

Summing (5.9) for $\gamma = \overline{\gamma}, \underline{\gamma}$, and subtracting $2\int_U \|\nabla u - w\|\,dy$, we thus obtain

$$\begin{aligned}
\int_U \|\nabla u_{\overline{\gamma}} - w\|\,dy + \int_U \|\nabla u_{\underline{\gamma}} - w\|\,dy - 2\int_U \|\nabla u - w\|\,dy& \\
\le G_{\overline{\gamma},\underline{\gamma}} \int_U \|\nabla u - \nabla v\|\,dy + (C_{\overline{\gamma},\underline{\gamma}} + r)\epsilon r^m&,
\end{aligned} \tag{5.10}$$

where

$$C_{\overline{\gamma},\underline{\gamma}} := G_{\overline{\gamma},\underline{\gamma}} + D_{\overline{\gamma}}(J_{\overline{\gamma}}+1) + D_{\underline{\gamma}}(J_{\underline{\gamma}}+1) + \overline{M}_{\overline{\gamma}} + \overline{M}_{\underline{\gamma}}.$$

Under the assumption $T_{\overline{\gamma},\underline{\gamma}} < 1$ contained in Definition 3.4, this can be made less than a constant times $T_{\overline{\gamma},\underline{\gamma}}^{1/2}$. Since $\epsilon > 0$ was arbitrary, we can get rid of any extra constant factors, proving (5.6) if $u \in W^{1,1}(\Omega)$.

For general $u \in \mathrm{BV}(\Omega)$ we use an analogous smoothing argument as in Part 1 for TV. Namely, we use Corollary 4.1 to approximate u by a sequence $\{u^i\}_{i=1}^\infty \in C^\infty(\Omega)$ with

$$u^i \to u \text{ in } L^1(\Omega) \quad \text{and} \quad \|D(u^i - v)\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} \to \|D(u-v)\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)}, \tag{5.11}$$

as well as $Du^i \overset{*}{\rightharpoonup} Du$. Observe that $\epsilon > 0$ in (5.10) does not depend on u itself, and neither does $r_0 > 0$ nor the sets Q and Q_i, $(i \in \mathbb{Z}^+)$. Therefore (5.10) holds in a uniform sense for the sequence $\{u^i\}_{i=1}^\infty$. In particular

$$G(u^i_{\overline{\gamma}}, w) + G(u^i_{\underline{\gamma}}, w) - 2G(u^i, w) \le G_{\overline{\gamma},\underline{\gamma}} |D(u^i - v)|(U) + c \quad (i \in \mathbb{Z}^+) \tag{5.12}$$

for the small nuisance variable $c := (T_{\overline{\gamma},\underline{\gamma}}^{1/2} + r)\epsilon r^m$, independent of i. Since (5.11) bounds he right hand side, we deduce

$$\mathrm{TGV}^2_{\vec{\alpha}}(u^i_{\overline{\gamma}}) + \mathrm{TGV}^2_{\vec{\alpha}}(u^i_{\underline{\gamma}}) \le G(u^i_{\overline{\gamma}}, w) + G(u^i_{\underline{\gamma}}, w) < \infty.$$

By the BV-coercivity in (5.2), we may therefore extract a subsequence, unrelabelled, such that both $\{u^i_{\overline{\gamma}}\}_{i=1}^\infty$ and $\{u^i_{\underline{\gamma}}\}_{i=1}^\infty$ are convergent weakly* to some $\overline{u} \in \mathrm{BV}(\Omega)$ and $\underline{u} \in \mathrm{BV}(\Omega)$, respectively. Moreover, by (5.11), (5.12), and the lower semicontinuity of the Radon norm with respect to weak* convergence, we find that

$$G(\overline{u}, w) + G(\underline{u}, w) - 2G(u, w) \le \liminf_{i\to\infty} G_{\overline{\gamma},\underline{\gamma}} |D(u^i - v)|(U) + c.$$

Let us pick an open set $U' \supset U$ such that $|Du|(\partial U') = 0$. Then $|D(u^i - v)|(U') \to |D(u-v)|(U')$ because $D(u^i - v) \to D(u-v)$ strictly in $\mathcal{M}(\Omega;\mathbb{R}^m)$; see [3, Proposition 1.62]. It follows

$$G(\overline{u}, w) + G(\underline{u}, w) - 2G(u, w) \le G_{\overline{\gamma},\underline{\gamma}} |D(u - v)|(U') + c.$$

By taking the intersection over all admissible $U' \supset U$, we deduce

$$G(\overline{u}, w) + G(\underline{u}, w) - 2G(u, w) \le G_{\overline{\gamma},\underline{\gamma}} |D(u - v)|(\operatorname{cl} U) + c. \tag{5.13}$$

This is almost (5.6) just have to show that $\overline{u} = \overline{\gamma}_\# [\![u, v]\!]$ and $\underline{u} = \underline{\gamma}_\# [\![u, v]\!]$. Indeed

$$\begin{aligned} \int_\Omega |\overline{u}(x) - \overline{\gamma}_\# [\![u, v]\!]|\, dx &\le \int_\Omega |\overline{u}(x) - \overline{\gamma}_\# [\![u^i, v]\!]|\, dx + \int_\Omega |\overline{\gamma}_\# [\![u, v]\!] - \overline{\gamma}_\# [\![u^i, v]\!]|\, dx \\ &\le \int_\Omega |\overline{u}(x) - \overline{\gamma}_\# [\![u^i, v]\!]|\, dx + C \int_\Omega |u(x) - u^i(x)|\, dx \end{aligned} \tag{5.14}$$

for

$$C := \left(\sup_x \mathcal{J}_m \overline{\gamma}(x) \right) \le (\operatorname{lip} \overline{\gamma})^m < \infty.$$

The integrals on the right hand side of (5.14) moreover tend to zero by the strict convergence of u^i to u and the weak* convergence of $\overline{\gamma}_\# u^i$ to $\overline{u}$. It follows that $\overline{u} = \overline{\gamma}_\# u$. Analogously we show that $\underline{u} = \underline{\gamma}_\# u$. The bound (5.6) is now immediate from (5.13). □

Remark 5.1. Since w is kept fixed throughout, the proof of Proposition 5.1 trivially extends to the differentiation cascade formulation of TGV^k, $(k \ge 3)$, defined for the parameter vector $\vec{\alpha} = (\alpha_1, \ldots, \alpha_k) > 0$ as

$$\mathrm{TGV}^k_{\vec{\alpha}}(u) = \inf_{\substack{u_\ell \in L^1(\Omega;\mathrm{Sym}^\ell(\mathbb{R}^m));\\ \ell=1,\ldots,k-1;\, u_0=u,\, u_k=0}} \sum_{\ell=1}^k \alpha_{k-\ell} \|Eu_{\ell-1} - u_\ell\|.$$

The extension of the proof of this formulation in [11] for $k = 2$ to $k > 2$ may be found in [8].

5.2. Variants of TGV^2

As we have seen, we are unable to prove jump set containment for TGV^2 unless we assume that the minimising $w \in \mathrm{BD}(\Omega)$ in (5.1) actually satisfies $w \in \mathrm{BV}_{\mathrm{loc}}(\Omega)$ and $P_{z_\Gamma}^\perp(w^+(x) - w^-(x)) = 0$ for any Lipschitz graph Γ. Of course, we also have to assume that $u \in L^\infty_{\mathrm{loc}}(\Omega)$. Whether we can prove any of these properties, we leave as a fascinating question for future studies. Here we consider a couple of variants of TGV^2 for which at least $w \in \mathrm{BV}_{\mathrm{loc}}(\Omega)$, and even $P_{z_\Gamma}^\perp(w^+(x) - w^-(x)) = 0$, which we recall having denoted by $x \in O_w^\Gamma$.

The first modification, already considered in [9], is the non-symmetric variant, which may be defined as

$$\mathrm{nsTGV}^2_{\vec\alpha}(u) := \min_{w \in \mathrm{BV}(\Omega;\mathbb{R}^m)} \alpha\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|Dw\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}. \tag{5.15}$$

It is not difficult to extend Lemma 5.1 to this this functional, and then repeat the proof of Proposition 5.1 to obtain the following.

Proposition 5.2. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Then $\mathrm{nsTGV}^2_{\vec\alpha}$ is an admissible regularisation functional on $\mathrm{BV}(\Omega)$ satisfying Assumption 3.1(i). Moreover, for any $u \in \mathrm{BV}(\Omega)$, a minimiser $w \in \mathrm{BD}(\Omega)$ of (5.15), and any Lipschitz $(m-1)$-graph $\Gamma \subset \Omega$, the following holds.*

(ii") $\mathrm{nsTGV}^2_{\vec\alpha}$ is partially double-Lipschitz comparable for u in the direction z_Γ at $\mathcal{H}^{m-1}$-a.e. $x \in O_w^\Gamma$.

In fact, since the proof keeps w fixed, we can do a little bit more.

Proposition 5.3. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Suppose $\Psi : \mathrm{BV}(\Omega) \to \mathbb{R}$ is convex and lower semicontinuous with respect to weak* convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, and satisfies for some constant $C > 0$ the inequality*

$$\|Dw\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \le C(1 + \Psi(w)). \tag{5.16}$$

For any $\vec\alpha = (\beta, \alpha) > 0$, define

$$F_\Psi(u) := \inf_{w \in L^1(\Omega;\mathbb{R}^m)} \alpha\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\Psi(w), \quad (u \in \mathrm{BV}(\Omega)). \tag{5.17}$$

Then F_Ψ is an admissible regularisation functional on $L^1(\Omega)$ and satisfies Assumption 3.1(i) and (ii").

Proof. Minding (5.16), it is not difficult to see that a minimising sequence $\{w^i\}_{i=1}^\infty$ for the expression of $F_\Psi(u)$ in (5.17) is bounded in $\mathrm{BV}(\Omega)$. The existence of a minimising $w \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ for $F_\Psi(u)$ therefore follows from the lower semicontinuity of Ψ. If now $u^i \to u$ weakly* in $\mathrm{BV}(\Omega)$, with corresponding minimisers w^i to the expression of $F_\Psi(u^i)$ in (5.17), then we may again deduce that $\{w^i\}_{i=1}^\infty$ is bounded in $\mathrm{BV}(\Omega;\mathbb{R}^m)$. Therefore, we may extract a subsequence, unrelabelled, such that also $\{w^i\}_{i=1}^\infty$ converge weakly* to some $u \in \mathrm{BV}(\Omega)$. But the functional

$$G(u, w) := \alpha\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\Psi(w), \tag{5.18}$$

is clearly lower semicontinuous with respect to weak* convergence of both variables. Since $F_\Psi(u) \le G(u, w)$, and $F_\Psi(u^i) = G(u^i, w^i)$, we deduce that F_Ψ is weak* lower semicontinuous.

To see the coercivity property (3.1), we use the fact that

$$\|Du\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} \le C\big(\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \|Dw\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} + \|u\|_{L^1(\Omega)}\big).$$

This follows from the Poincaré inequality and an argument by contradiction; for details see [11, 10]. Plugging in (5.16) immediately proves (3.1).

Finally, F_Ψ is clearly convex, so the above considerations show that it is admissible. To prove Assumption 3.1, we adapt the proof of Proposition 5.1, replacing G defined by (5.5) by that in (5.18). Now w is BV-differentiable everywhere, $D_w = \Omega$, so this part of the complications with TGV^2 does not arise. □

As we recall from Korn's inequality, functions with bounded symmetrised gradient in L^q for $q > 1$ are much better behaved than for $q = 1$. We now want to exploit this to define variants of TGV^2 with stronger double-Lipschitz comparability properties.

Corollary 5.1. *Suppose $\Omega \subset \mathbb{R}^m$ is a bounded open set with Lipschitz boundary. For $1 < q < \infty$, let*

$$\Psi(w) := \begin{cases} \|\mathcal{E}w\|_{F,L^q(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}, & w \in W_0^{1,q}(\Omega;\mathbb{R}^m), \\ \infty, & \textit{otherwise.} \end{cases}$$

Then $\mathrm{TGV}^{2,q}_{\vec{\alpha},0} := F_\Psi$ is an admissible regularisation functional on $L^1(\Omega)$, satisfying Assumption 3.1.

Proof. The condition (5.16) is an immediate consequence of Korn's inequality (2.3). For weak* lower semicontinuity, we have to establish that any BV-weak* limit point w of a sequence $\{w^i\}_{i=1}^\infty \subset W_0^{1,q}(\Omega;\mathbb{R}^m)$ with

$$\sup_i \|w\|_{2,L^1(\Omega;\mathbb{R}^m)} + \|\mathcal{E}w\|_{F,L^q(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} \leq C < \infty, \tag{5.19}$$

also satisfies $w \in W_0^{1,q}(\Omega;\mathbb{R}^m)$. The lower semicontinuity of $\|\mathcal{E}\,\cdot\,\|_{2,L^q(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}$ itself is standard. By the Gagliardo-Nirenberg-Sobolev inequality, Korn's inequality (2.3), and approximation in $C_c^\infty(\Omega;\mathbb{R}^m)$, we also discover

$$\|w^i\|_{2,L^q(\Omega;\mathbb{R}^m)} + \|\nabla w^i\|_{F,L^q(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \leq C'\|\mathcal{E}w^i\|_{F,L^q(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))} \leq C'C.$$

We may therefore assume $\{w^i\}_{i=1}^\infty$ convergent weakly in $W^{1,q}(\Omega;\mathbb{R}^m)$, necessarily to w. It follows that $w \in W^{1,q}(\Omega;\mathbb{R}^m)$. But $w^i \in W_0^{1,q}(\Omega;\mathbb{R}^m)$ and $W_0^{1,q}(\Omega;\mathbb{R}^m)$ is strongly closed within $W^{1,q}(\Omega;\mathbb{R}^m)$, hence weakly closed as a convex set. Therefore $w \in W_0^{1,q}(\Omega;\mathbb{R}^m)$. This establishes BV-weak* lower semicontinuity of Ψ. Finally, we employ Proposition 5.3, noting that $\mathcal{H}^{m-1}(\Gamma \setminus O_w^\Gamma) = 0$ because $J_w = \emptyset$ and the (equal) one-sided traces exist $\mathcal{H}^{m-1}$-a.e. on Γ (by the BV trace theorem or $\mathcal{H}^{m-1}(S_w \setminus J_w) = 0$). ☐

In Figure 1, we have a simple comparison of the effect of the exponent q with fidelity $\phi(t) = t^2/$. For $q = 1$, we have chosen the base parameters $\alpha = 25$ and $\beta = 250$ on the image domain $\Omega := [1, 256]^2$. For other values of q, namely $q = 1.5$ and $q = 2$, we have scaled β by the factor $256^{2(q-1)/q}$. This is what the the Cauchy-Schwarz inequality gives as the factor for the q-norm to dominate the 1-norm on an image with 256^2 pixels. We also include the TV result for comparison. The PSNR for variants of TGV^2 with different q values is always the same, 29.2, while TV has PSNR 28.0. There is also visually no discernible difference between the different q-values, whereas TV clearly exhibits the staircasing effect in the background sky. It therefore seems reasonable to also employ in practise this kind of variants of TGV^2, for which we have stronger theoretical results now, only lacking a proof of the local boundedness of u to complete the proof of the property $\mathcal{H}^{m-1}(J_u \setminus J_f)$.

5.3. Infimal convolution TV

Let $v \in W^{1,1}(\Omega)$ and $\nabla v \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. Define the second-order total variation by

$$\mathrm{TV}^2(w) = \|D\nabla w\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}.$$

Then second-order infimal convolution TV of $u \in \mathrm{BV}(\Omega)$, first introduced in [14], is written

$$\mathrm{ICTV}_{\vec{\alpha}}(u) := (\alpha\mathrm{TV} \,\square\, \beta\mathrm{TV}^2)(u) := \inf_{u=v^1+v^2} \bigl(\alpha\mathrm{TV}(v^1) + \beta\mathrm{TV}^2(v^2)\bigr), \tag{5.20}$$

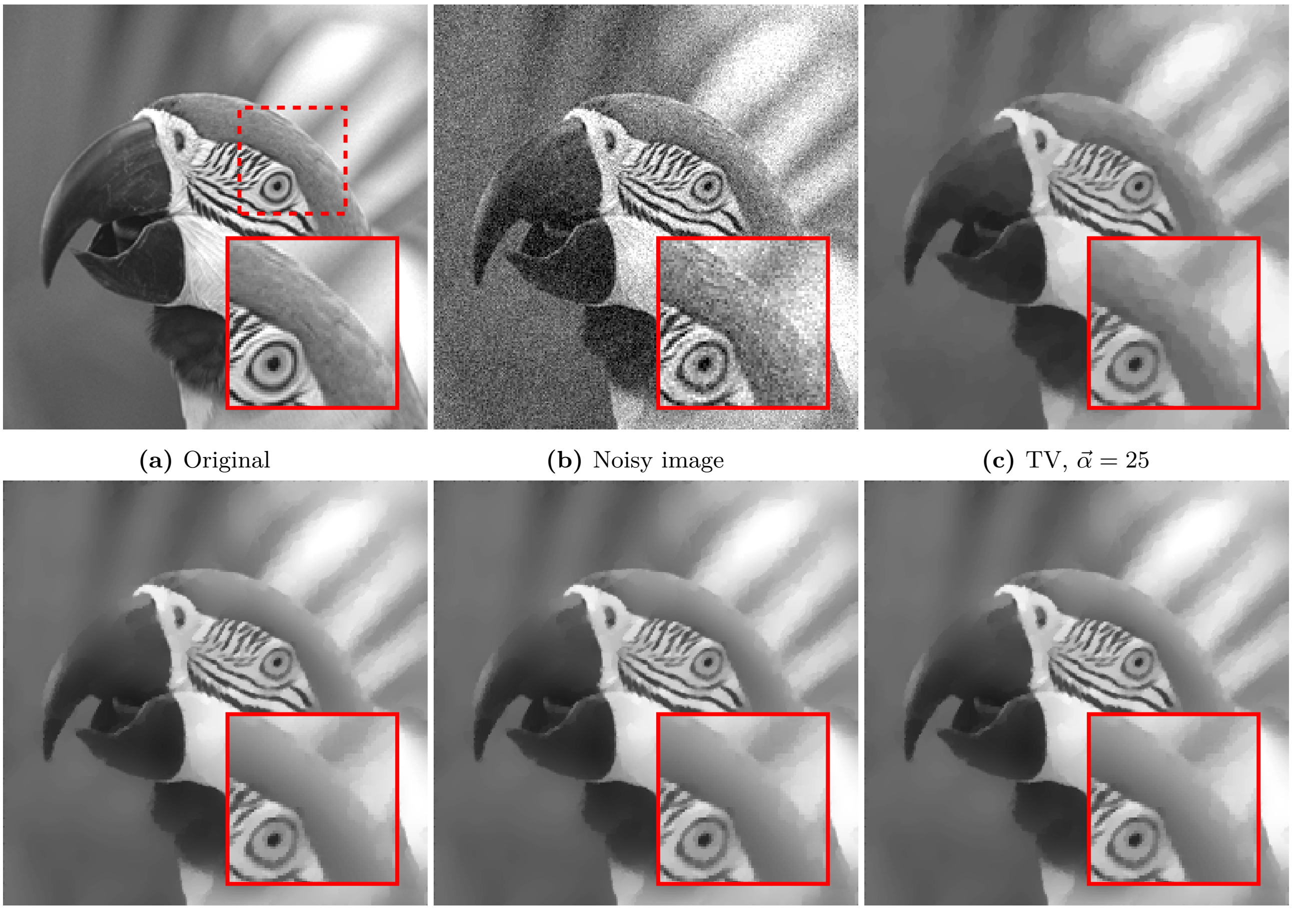
where necessarily $w \in W^{1,1}(\Omega)$, $\nabla w \in \mathrm{BV}(\Omega;\mathbb{R}^m)$, and $v \in \mathrm{BV}(\Omega)$. Clearly we have

$$\mathrm{TGV}^2_{\vec{\alpha}}(u) \leq \mathrm{nsTGV}^2_{\vec{\alpha}}(u) \leq \mathrm{ICTV}_{\vec{\alpha}}(u) \leq \alpha\mathrm{TV}(u). \tag{5.21}$$

It has been observed that while ICTV is better at avoiding the stair-casing effect than TV, it fares worse than TGV^2 [4].

We did not find a proof of the weak* lower semi-continuity of ICTV in the literature, so we provide one below. Then we show that $\mathrm{ICTV}_{\vec{\alpha}}$ is admissible and partially double-Lipschitz comparable as required by Assumption 3.1. As already observed in the Introduction, we remark, however, that the the jump set containment $\mathcal{H}^{m-1}(J_u \setminus J_f) = 0$ can be proved for ICTV using the result for TV.

Lemma 5.2. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Then $\mathrm{ICTV}_{\vec{\alpha}}$ is lower semi-continuous with respect to weak* convergence in $\mathrm{BV}(\Omega)$.*



(a) Original **(b)** Noisy image **(c)** TV, $\vec{\alpha} = 25$

(d) $q = 1$, $\vec{\alpha} = (250, 25)$ **(e)** $q = 1.5$, $\vec{\alpha} = (10079, 25)$ **(f)** $q = 2.0$, $\vec{\alpha} = (64000, 25)$

Figure 1: Effect of the of exponent q in the norm $\|\mathcal{E}w\|_{F,L^q(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}$ in variants of TGV^2 together with fidelity $\phi(t) = t^2/2$. The β factor has been scaled from the case $q = 1$ with the help of the Cauchy-Schwarz inequality. There is no discernible difference between the results for different q, all having PSNR 29.2, while the TV comparison has PSNR 28.0 and exhibits the stair-casing effect in the sky that TGV^2 variants do not.

Proof. Let $u^i \overset{*}{\rightharpoonup} u$ weakly* in $\mathrm{BV}(\Omega)$, $(i = 0, 1, 2, \ldots)$. We may then without loss of generality assume that $\{\|u^i\|_{L^1(\Omega)} + \|Du^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}\}_{i=0}^{\infty}$ is bounded. Let $v_1^i \in \mathrm{BV}(\Omega)$ and $v_2^i \in W^{1,1}(\Omega)$ with $\nabla w^i \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ be such that

$$\alpha\|Dv_1^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|D\nabla v_2^i\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \leq \mathrm{ICTV}_{\vec{\alpha}}(u^i) + 1/i, \quad (i = 0, 1, 2, \ldots).$$

Observe that we may take each v_1^i such that

$$\bar{v}_1^i := \int_\Omega v_1^i(x)\,dx = 0, \tag{5.22}$$

since the infimum in (5.20) is independent of the mean of v^1 and v^2.

If $\limsup_i \mathrm{ICTV}_{\vec{\alpha}}(u^i) = \infty$, there is nothing to prove, so we may assume that $\sup_i \mathrm{ICTV}_{\vec{\alpha}}(u^i) < \infty$. It follows that both the sequence $\{\|Dv_1^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}\}_{i=0}^{\infty}$ and the sequence $\{\|D\nabla v_2^i\|_{F,\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))}\}_{i=0}^{\infty}$ are bounded. Minding (5.22) and the assumption that Ω has Lipschitz boundary, the Poincaré inequality now shows the existence of a constant $C > 0$ such that

$$\|v_1^i\|_{L^1(\Omega)} = \|v_1^i - \bar{v}_1^i\|_{L^1(\Omega)} \leq C\|Dv_1^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}, \quad (i = 0, 1, 2, \ldots),$$

Consequently $\{v_1^i\}_{i=0}^{\infty}$ admits a subsequence, unrelabelled, weakly* convergent in $\mathrm{BV}(\Omega)$ to some $v \in \mathrm{BV}(\Omega)$. By the boundedness of the sequence $\{\|v_1^i\|_{L^1(\Omega)} + \|Dv_1^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}\}\}_{i=0}^{\infty}$ and of $\{\|u^i\|_{L^1(\Omega)} + \|Du^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}\}\}_{i=0}^{\infty}$ it follows from $u^i = v_1^i + v_2^i$, moreover, that $\{\|v_2^i\|_{L^1(\Omega)} + \|Dv_2^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}\}\}_{i=0}^{\infty}$ is bounded. Consequently, moving to a further subsequence, we may assume that $v_2^i \to v_2$ strongly in $W^{1,1}(\Omega)$ and $\nabla v_2^i \overset{*}{\rightharpoonup} \nabla v_2$ weakly* in $\mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))$, for some $v_2 \in W^{1,1}(\Omega)$ with $\nabla v_2 \in \mathcal{M}(\Omega;\mathrm{Sym}^2(\mathbb{R}^m))$. We clearly have $u = \lim u^i = \lim(v_1^i + v_2^i) = v_1 + v_2$. Hence, by the lower semicontinuity of the Radon norm with respect to weak* convergence of measures, we obtain

$$\begin{aligned}\mathrm{ICTV}_{\vec{\alpha}}(u) &\le \alpha\|Dv_1\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|D\nabla v_2\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}\\ &\le \liminf_{i\to\infty} \alpha\|Dv_1^i\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|D\nabla v_2^i\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}\\ &\le \liminf_{i\to\infty}\big(\mathrm{ICTV}_{\vec{\alpha}}(u^i) + 1/i\big)\\ &= \liminf_{i\to\infty} \mathrm{ICTV}_{\vec{\alpha}}(u^i).\end{aligned}$$

Thus $\mathrm{ICTV}_{\vec{\alpha}}$ is weak* lower semi-continuous, as claimed. □

Proposition 5.4. *Let $\Omega \subset \mathbb{R}^m$ be a bounded domain with Lipschitz boundary. Then $\mathrm{ICTV}_{\vec{\alpha}}$ is an admissible regularisation functional on $L^1(\Omega)$, and satisfies Assumption 3.1.*

Proof. We have already proved that $\mathrm{TGV}^2_{\vec{\alpha}}$ satisfies the coercivity criterion (3.1). It immediately follows from (5.21) that $\mathrm{ICTV}_{\vec{\alpha}}$ also satisfies this. By Lemma 5.2 ICTV is weak* lower semicontinuous in $\mathrm{BV}(\Omega)$. The rest of the conditions of Definition 3.1 are obvious. Thus $\mathrm{ICTV}_{\vec{\alpha}}$ is admissible. Assumption 3.1 can be proved following the proof of Proposition 5.1 as follows. Instead of (5.5), we define

$$G(u,w) := \alpha\|Du - w\|_{2,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|Dw\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}.$$

With $u = v_1 + v_2$ a minimising decomposition in (5.20), we set $w = \nabla v_2$ and $v = v_2$, observing that the conclusions of Lemma (4.2) hold trivially for $v = v_2$ at every Lebesgue point x of v_2. Since $v_2 \in W^{1,1}(\Omega)$, this is in particular the case for $\mathcal{H}^{m-1}$-a.e. $x \in \Gamma$ for any given Lipschitz $(m-1)$-graph Γ. Thus we have no complications as in the case of TGV^2. It follows that Assumption 3.1 holds. □

6. Limiting behaviour of L^p-TGV^2

Having studied qualitatively the behaviour of the jump set J_u, and obtained good results for variants of TGV^2 although not TGV^2 itself, we now want to study it quantitatively. We let the second regularisation parameter β of TGV^2 go to zero, and see what happens to $D^s u$ for u solution to the L^p-TGV^2 regularisation problem, namely

$$\min_{u\in \mathrm{BV}(\Omega)} \|f-u\|^p_{L^p(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}(u). \tag{6.1}$$

The next proposition states our findings.

Proposition 6.1. *Let $\alpha > 0$, $1 \le p < \infty$, and $f \in L^p(\Omega)$. Then for every $\epsilon > 0$ there exists $\beta_0 > 0$ such that any solution u to (6.1) satisfies*

$$\|D^s u\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} < \epsilon \quad \text{for } \beta \in (0,\beta_0). \tag{6.2}$$

Proof. Let $\{\rho_\tau\}_{\tau>0}$ be the standard family of mollifiers on $\mathbb{R}^m$, and use the notation

$$u_\tau := \rho_\tau * u, \quad \text{and} \quad w_\tau := \rho_\tau * w$$

for mollified functions, where w minimises (5.1) for u. Then

$$\begin{aligned}\|f-u_\tau\|_{L^p(\Omega)} &\le \|f_\tau - u_\tau\|_{L^p(\Omega)} + \|f - f_\tau\|_{L^p(\Omega)}\\ &\le \|f-u\|_{L^p(\Omega)} + \|f-f_\tau\|_{L^p(\Omega)}.\end{aligned}$$

Let $\delta > 0$ be arbitrary. Since $\|f - f_\tau\|_{L^p(\Omega)} \to 0$, we deduce the existence of $\tau_\delta > 0$ such that

$$\|f-u_\tau\|^p_{L^p(\Omega)} \le \|f-u\|^p_{L^p(\Omega)} + \delta, \quad \text{for } \tau \in (0,\tau_\delta].$$

It can easily be shown by application of Green's identities that the symmetric differential operator E satisfies

$$Ew_\tau = \rho_\tau * Ew = \mathcal{E}\rho_\tau * w,$$

similarly to corresponding well known results on the operator D. With

$$\widetilde{w} := \nabla u_\tau$$

we thus obtain for some constant $C > 0$ the estimate

$$\begin{aligned}\|E\widetilde{w}\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} &\leq \|E\widetilde{w} - Ew_\tau\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} + \|Ew_\tau\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \\ &\leq \|\mathcal{E}\rho_\tau * (Du - w)\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} + \|Ew\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \\ &\leq C\tau^{-1}\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \|Ew\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))}.\end{aligned}$$

As

$$\|Du_\tau - \widetilde{w}\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} = 0,$$

it follows that

$$\begin{aligned}\|f - u_\tau\|^p_{L^p(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}(u_\tau) &\leq \|f - u_\tau\|^p_{L^p(\Omega)} + \alpha\|Du_\tau - \widetilde{w}\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \beta\|E\widetilde{w}\| \\ &\leq \|f - u\|^p_{L^p(\Omega)} + \delta + C\beta\tau^{-1}\|Du - w\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \\ &\quad + \beta\|Ew\|_{F,\mathcal{M}(\Omega;\mathcal{T}^2(\mathbb{R}^m))} \\ &\leq \|f - u\|^p_{L^p(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}(u) \\ &\quad + (C\beta\tau^{-1} - \alpha)\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} + \delta.\end{aligned}$$

Consequently u_τ provides a contradiction to u being a solution to (6.1) if

$$\delta < (\alpha - C\beta\tau^{-1})\|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)}.$$

Since

$$\|D^s u\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} \leq \|Du - w\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)},$$

it follows that for an optimal solution u, it must hold

$$\|D^s u\|_{F,\mathcal{M}(\Omega;\mathbb{R}^m)} \leq \delta/(\alpha - C\beta\tau_\delta^{-1})$$

Thus (6.2) holds if

$$\delta + C\beta\tau_\delta^{-1}\epsilon < \epsilon\alpha.$$

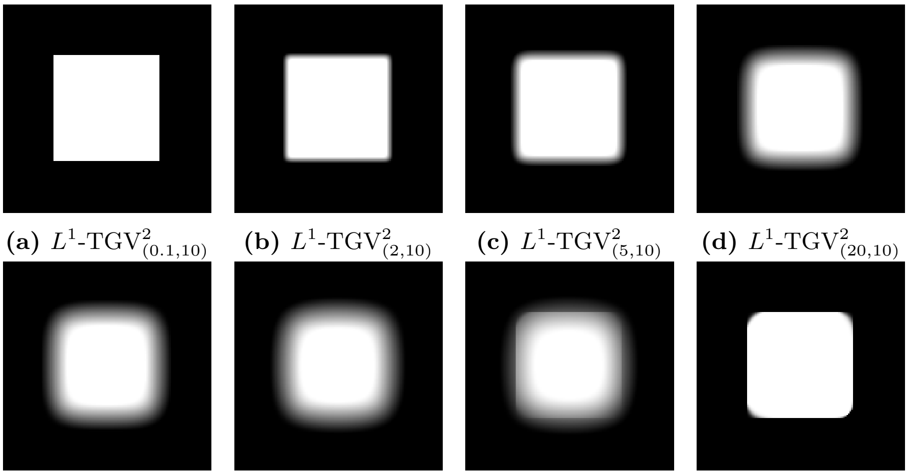
Choosing $\delta < \epsilon\alpha$, we find $\beta_0 > 0$ such that this is satisfied for $\beta \in (0, \beta_0)$. □

We illustrate numerically in Figure 2 to Figure 3 the implications of Proposition 6.1 and Theorem 3.2 on a very simple test image with a square in the middle. We did the experiments for fixed $\alpha = 10$ or $\alpha = 5$ and varying β, with fidelity $\phi(t) = t^p$ for $p = 1$ and $p = 2$. In all cases, as β goes down from a large value with good reconstruction, the image first starts to smooth out. This happens until the smallest β, for which we appear to have recovered f! This may seem a little counterintuitive, as Proposition 6.1 forbids big jumps for small β. But we should indeed have very steep gradients near the boundary. These are lost in the discretisation.

Besides this, the numerical experiments verify $\mathcal{H}^{m-1}(J_u \setminus J_f)$ for $p = 2$, and demonstrate the fact that it does not hold for $p = 1$. However, the set $J_u \setminus J_f$ has specific curvature dependent on the parameter α. For $p = 2$, we of course observe the well-known phenomenon of contrast loss. In the corners where $p = 1$ starts to produce new jumps, $p = 2$ starts to smooth out the solution, also not reconstructing the jumps of the corners.

7. Conclusion

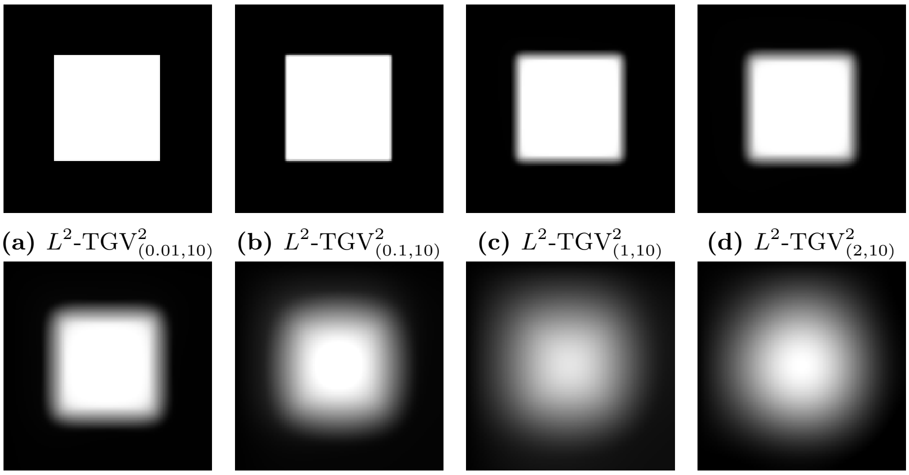
In this pair of papers, we have provided a new technique for studying the jump sets of a general class of regularisation functionals, not dependent on the co-area formula as existing results for TV are. In the case that the fidelity ϕ is p-increasing for $p > 1$, besides TV, we proved in Part 1 the property $\mathcal{H}^{m-1}(J_u \setminus J_f) = 0$ for u a solution to (P) for Huber-regularised total variation. We also demonstrated in Part 1 that the technique would apply to non-convex TV models and the Perona-Malik anisotropic diffusion, if these models were well-posed, and had solutions in $\mathrm{BV}(\Omega)$. For variants of TGV^2 using L^q, $(q > 1)$, energies for the second-order component, we proved that the jump set containment property holds if the solution u is locally bounded. For TGV^2 itself, we obtained much weaker results, depending additionally on the differentiability assumptions of Lemma 4.2 on the minimising second-order variable w. The two most important further questions that these studies pose are whether the assumptions of Lemma 4.2 on w can be proved for TGV^2, and whether the local boundedness of the solution u can be proved for higher-order regularisers in general. In the first-order cases this was no work at all.



(a) L^1-TGV$^2_{(0.1,10)}$ (b) L^1-TGV$^2_{(2,10)}$ (c) L^1-TGV$^2_{(5,10)}$ (d) L^1-TGV$^2_{(20,10)}$

(e) L^1-TGV$^2_{(30,10)}$ (f) L^1-TGV$^2_{(40,10)}$ (g) L^1-TGV$^2_{(50,10)}$ (h) L^1-TGV$^2_{(60,10)}$

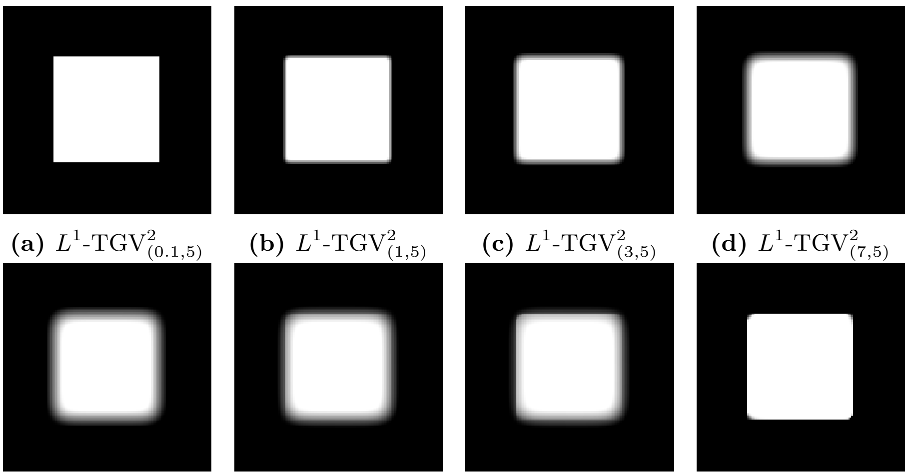
Figure 2: Illustration of varying β parameter for TGV^2 regularisation with L^1 fidelity on $f = \chi_{(-32,32)^2}$. For $\beta = 0.1$ in (a) it appears that we have recovered f. The apparent full recovery in (a) for $\beta = 0.1$ is an effect of the discretisation. For $\beta = 50$ in (h) due to numerical difficulties we have not fully recovered the corners (of curvature $1/\alpha = 0.2$) that should start to become sharp.



(a) L^2-TGV$^2_{(0.01,10)}$ (b) L^2-TGV$^2_{(0.1,10)}$ (c) L^2-TGV$^2_{(1,10)}$ (d) L^2-TGV$^2_{(2,10)}$

(e) L^2-TGV$^2_{(5,10)}$ (f) L^2-TGV$^2_{(20,10)}$ (g) L^2-TGV$^2_{(50,10)}$ (h) L^2-TGV$^2_{(10000,10)}$

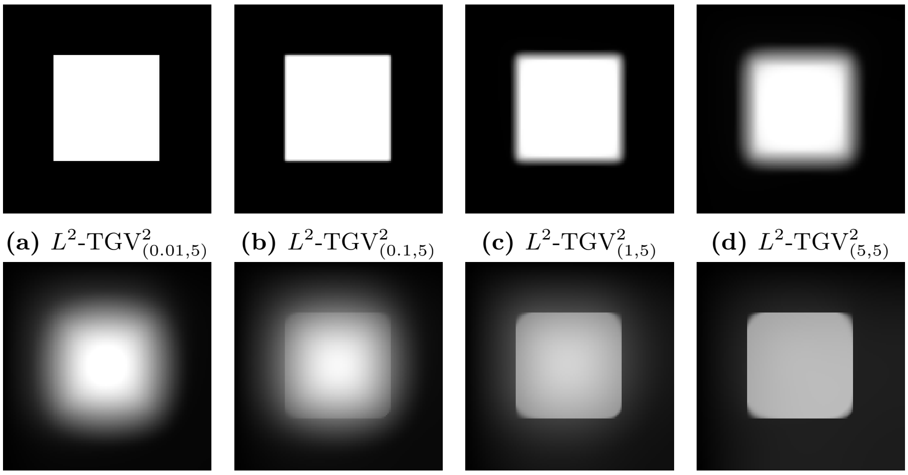
Figure 3: Illustration for $\alpha = 10$ of varying β parameter for TGV^2 regularisation with squared L^2 fidelity on $f = \chi_{(-32,32)^2}$, to compare with Figure 2 for L^1 fidelity.



(a) L^1-TGV$^2_{(0.1,5)}$ (b) L^1-TGV$^2_{(1,5)}$ (c) L^1-TGV$^2_{(3,5)}$ (d) L^1-TGV$^2_{(7,5)}$

(e) L^1-TGV$^2_{(12,5)}$ (f) L^1-TGV$^2_{(12.45,5)}$ (g) L^1-TGV$^2_{(12.55,5)}$ (h) L^1-TGV$^2_{(13,5)}$

Figure 4: Illustration for $\alpha = 5$ of varying β parameter for TGV^2 regularisation with L^1 fidelity on $f = \chi_{(-32,32)^2}$.



(a) L^2-TGV$^2_{(0.01,5)}$ (b) L^2-TGV$^2_{(0.1,5)}$ (c) L^2-TGV$^2_{(1,5)}$ (d) L^2-TGV$^2_{(5,5)}$

(e) L^2-TGV$^2_{(25,5)}$ (f) L^2-TGV$^2_{(35,5)}$ (g) L^2-TGV$^2_{(50,5)}$ (h) L^2-TGV$^2_{(100,5)}$

Figure 5: Illustration for $\alpha = 5$ of varying β parameter for TGV^2 regularisation with squared L^2 fidelity on $f = \chi_{(-32,32)^2}$.

Acknowledgements

This manuscript has been prepared over the course of several years, exploiting funding from various more short-term projects. While the author was at the Institute for Mathematics and Scientific Computing at the University of Graz, this work was financially supported by the SFB research program F32 "Mathematical Optimization and Applications in Biomedical Sciences" of the Austrian Science Fund (FWF). While at the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge, this work was financially supported by the King Abdullah University of Science and Technology (KAUST) Award No. KUK-I1-007-43 as well as the EPSRC / Isaac Newton Trust Small Grant "Non-smooth geometric reconstruction for high resolution MRI imaging of fluid transport in bed reactors", and the EPSRC first grant Nr. EP/J009539/1 "Sparse & Higher-order Image Restoration". At the Research Center on Mathematical Modeling (Modemat) at the Escuela Politécnica Nacional de Quito, this work has been supported by the Prometeo initiative of the Senescyt.

The author would like to express sincere gratitude to Simon Morgan, Antonin Chambolle, Kristian Bredies, and Carola-Bibiane Schönlieb for fruitful discussions.

References

1. L. Ambrosio, A. Coscia and G. Dal Maso, *Fine properties of functions with bounded deformation*, Archive for Rational Mechanics and Analysis **139** (1997), 201–238.
2. L. Ambrosio, G. Crippa and S. Maniglia, *Traces and fine properties of a BD class of vector fields and applications*, Annales de la faculté des sciences de Toulouse, Sér. 6 **14** (2005), 527–561.
3. L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, 2000.
4. M. Benning, C. Brune, M. Burger and J. Müller, *Higher-order TV methods—enhancement via Bregman iteration*, Journal of Scientific Computing **54** (2013), 269–310, doi:10.1007/s10915-012-9650-3.
5. A. L. Bertozzi and J. B. Greer, *Low-curvature image simplifiers: Global regularity of smooth solutions and laplacian limiting schemes*, Communications on Pure and Applied Mathematics **57** (2004), 764–790, doi:10.1002/cpa.20019.
6. R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds*, Dover Publications, 1980, Dover edition.
7. K. Bredies, *Symmetric tensor fields of bounded deformation*, Annali di Matematica Pura ed Applicata **192** (2013), 815–851, doi:10.1007/s10231-011-0248-4.
8. K. Bredies and M. Holler, *Regularization of linear inverse problems with total generalized variation*, SFB-Report 2013-009, University of Graz (2013).
9. K. Bredies, K. Kunisch and T. Pock, *Total generalized variation*, SIAM Journal on Imaging Sciences **3** (2011), 492–526, doi:10.1137/090769521.
10. K. Bredies, K. Kunisch and T. Valkonen, *Properties of L^1-TGV^2: The one-dimensional case*, Journal of Mathematical Analysis and Applications **398** (2013), 438–454, doi:10.1016/j.jmaa.2012.08.053.
URL `http://math.uni-graz.at/mobis/publications/SFB-Report-2011-006.pdf`
11. K. Bredies and T. Valkonen, *Inverse problems with second-order total generalized variation constraints*, in: *Proceedings of the 9th International Conference on Sampling Theory and Applications (SampTA) 2011, Singapore*, 2011.
URL `http://iki.fi/tuomov/mathematics/SampTA2011.pdf`

12. M. Burger, M. Franek and C.-B. Schönlieb, *Regularized regression and density estimation based on optimal transport*, Applied Mathematics Research eXpress **2012** (2012), 209–253, doi:10.1093/amrx/abs007.
13. V. Caselles, A. Chambolle and M. Novaga, *The discontinuity set of solutions of the TV denoising problem and some extensions*, Multiscale Modeling and Simulation **6** (2008), 879–894.
14. A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numerische Mathematik **76** (1997), 167–188.
15. T. Chan, A. Marquina and P. Mulet, *High-order total variation-based image restoration*, SIAM Journal on Scientific Computation **22** (2000), 503–516, doi:10.1137/S1064827598344169.
16. T. F. Chan, S. H. Kang and J. Shen, *Euler's elastica and curvature-based inpainting*, SIAM Journal on Applied Mathematics (2002), 564–592.
17. P. G. Ciarlet, *On korn's inequality*, Chinese Annals of Mathematics, Series B **31** (2010), 607–618, doi:10.1007/s11401-010-0606-3.
18. S. Conti, D. Faraco and F. Maggi, *A new approach to counterexamples to l 1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions*, Archive for Rational Mechanics and Analysis **175** (2005), 287–300, doi:10.1007/s00205-004-0350-5.
19. G. Dal Maso, I. Fonseca, G. Leoni and M. Morini, *A higher order model for image restoration: the one-dimensional case*, SIAM Journal on Mathematical Analysis **40** (2009), 2351–2391, doi:10.1137/070697823.
20. S. Didas, J. Weickert and B. Burgeth, *Properties of higher order nonlinear diffusion filtering*, Journal of Mathematical Imaging and Vision **35** (2009), 208–226, doi:10.1007/s10851-009-0166-x.
21. V. Duval, J. F. Aujol and Y. Gousseau, *The TVL1 model: A geometric point of view*, Multiscale Modeling and Simulation **8** (2009), 154–189.
22. H. Federer, *Geometric Measure Theory*, Springer, 1969.
23. P. Hajłasz, *On approximate differentiability of functions with bounded deformation*, Manuscripta Mathematica **91** (1996), 61–72, doi:10.1007/BF02567939.
24. M. Hintermüller, T. Valkonen and T. Wu, *Limiting aspects of non-convex TV^q models* (2014). In preparation.
25. M. Hintermüller and T. Wu, *Nonconvex TV^q-models in image restoration: Analysis and a trust-region regularization–based superlinearly convergent solver*, SIAM Journal on Imaging Sciences **6** (2013), 1385–1415.
26. M. Hintermüller and T. Wu, *A superlinearly convergent R-regularized Newton scheme for variational models with concave sparsity-promoting priors*, Computational Optimization and Applications **57** (2014), 1–25.
27. J. Huang and D. Mumford, *Statistics of natural images and models*, in: *IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, volume 1, IEEE, 1999, volume 1.
28. J. Lellmann, D. Lorenz, C.-B. Schönlieb and T. Valkonen, *Imaging with Kantorovich-Rubinstein discrepancy* (2014). Submitted, arXiv:1407.0221.
URL `http://iki.fi/tuomov/mathematics/krtv.pdf`
29. M. Lysaker, A. Lundervold and X.-C. Tai, *Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time*, IEEE Transactions on Image Processing **12** (2003), 1579 – 1590, doi:10.1109/TIP.2003.819229.
30. Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations*, American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.
31. P. Ochs, A. Dosovitskiy, T. Brox and T. Pock, *An iterated l1 algorithm for non-smooth non-convex optimization in computer vision*, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
32. K. Papafitsoros and K. Bredies, *A study of the one dimensional total generalised variation regularisation problem* (2013). Preprint.
33. K. Papafitsoros and C. Schönlieb, *A combined first and second order variational approach for image reconstruction*, Journal of Mathematical Imaging and Vision **48** (2014), 308–338, doi:10.1007/s10851-013-0445-4.
34. J. Shen, S. Kang and T. Chan, *Euler's elastica and curvature-based inpainting*, SIAM Journal on Applied Mathematics **63** (2003), 564–592, doi:10.1137/S0036139901390088.
35. R. Temam, *Mathematical problems in plasticity*, Gauthier-Villars, 1985.
36. T. Valkonen, *The jump set under geometric regularisation. Part 1: Basic technique and first-order denoising* (2014). Submitted, arXiv:1407.1531.
URL `http://iki.fi/tuomov/mathematics/jumpset.pdf`
37. T. Valkonen, K. Bredies and F. Knoll, *Total generalised variation in diffusion tensor imaging*, SIAM Journal on Imaging Sciences **6** (2013), 487–525, doi:10.1137/120867172.
URL `http://iki.fi/tuomov/mathematics/dtireg.pdf`
38. L. A. Vese and S. J. Osher, *Modeling textures with total variation minimization and oscillating patterns in image processing*, Journal of Scientific Computing **19** (2003), 553–572.