Total Generalized Variation for Piecewise Constant Functions on Triangular Meshes with Applications in Imaging[*]

Lukas Baumgärtner[†]   Ronny Bergmann[‡]   Roland Herzog[§]   Stephan Schmidt[†]

José Vidal-Núñez[¶]

We propose a novel discrete concept for the total generalized variation (TGV), which has originally been derived to reduce the staircasing effect in classical total variation (TV) regularization, in image denoising problems. We describe discrete, second-order TGV for piecewise constant functions on triangular meshes, thus allowing the TGV functional to be applied to more general data structures than pixel images, and in particular in the context of finite element discretizations. Particular attention is given to the description of the kernel of the TGV functional, which, in the continuous setting, consists of linear polynomials. We discuss how to take advantage of this kernel structure using piecewise constant functions on triangular meshes. Numerical experiments include denoising and inpainting problems for images defined on non-standard grids, including data from a 3D scanner.

Keywords. total generalized variation, imaging, finite elements, surfaces

AMS subject classifications (MSC2010). 94A08, 68U10, 49M29, 65K05

[†]Institut für Mathematik, Humboldt University of Berlin, 10099 Berlin, Germany (lukas.baumgaertner@hu-berlin.de, https://www.mathematik.hu-berlin.de/en/people/mem-vz/1693318, ORCID 0000-0003-1007-4815, s.schmidt@hu-berlin.de, https://www.mathematik.hu-berlin.de/en/people/mem-vz/1693090, ORCID 0000-0002-4888-0794).

[‡]Norwegian University of Science and Technology, Department of Mathematical Sciences, NO-7041 Trondheim, Norway (ronny.bergmannn@ntnu.no, https://www.ntnu.edu/employees/ronny.bergmann, ORCID 0000-0001-8342-7218).

[§]Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany (roland.herzog@iwr.uni-heidelberg.de, https://scoop.iwr.uni-heidelberg.de, ORCID 0000-0003-2164-6575).

[¶]University of Alcalá, Department of Physics and Mathematics, 28801 Alcalá de Henares, Spain (j.vidal@uah.es, https://www.uah.es/es/estudios/profesor/Jose-Vidal-Nunez/, ORCID 0000-0002-1190-6700).

arXiv:2206.12331v2 [math.NA] 26 Sep 2022

1 Introduction

In recent years, research areas such as computer vision, remote sensing and medical imaging have been moving towards sophisticated applications requiring complex geometries. This in turn often necessitates the use of data structures that are more general than regular pixel meshes; see, e. g., John, Wilscy, 2016; Jensen, 2015; López Pérez, 2006 for some applications. Other areas where unstructured data appears naturally are 3D scanning and geophysics.

Often, triangular meshes are convenient data structures for these purposes, since they can model two-dimensional geometries and surfaces in a flexible way. We mention that triangular meshes are also underlying many finite element discretizations, which are convenient for the solution of inverse problems involving partial differential equations (PDEs); see, e. g., Chan, Tai, 2004; Bachmayr, Burger, 2009; Clason, Kruse, Kunisch, 2018. Even in classical mathematical imaging, some new camera devices leverage non-rectangular sub-pixel configurations for spatially varying exposure (SVE) sensors for high-dynamic-range (HDR) imaging, see Li et al., 2016, further motivating the use of data structures other than regular grids even in traditional imaging problems.

Common to the above problems is their inverse nature, and hence they usually require regularization due to the unavoidable noise in the data acquisition process. The total variation regularizer, introduced in Rudin, Osher, Fatemi, 1992 for imaging denoising, is a common choice due to its capacity of removing random noise while preserving discontinuities of data across edges. We recall that the total variation (TV)-seminorm of an L^1 function u on a bounded domain $\Omega \subset \mathbb{R}^2$ is defined as

$$\mathrm{TV}(u) = \sup \left\{ \int_\Omega u \operatorname{div} \boldsymbol{v} \, \mathrm{d}x \,\middle|\, \boldsymbol{v} \in C_c^1(\Omega, \mathbb{R}^2),\ \|\boldsymbol{v}\|_{L^\infty(\Omega)} \le 1 \right\}, \tag{1.1}$$

where $C_c^1(\Omega, \mathbb{R}^2)$ denotes the set of continuously differentiable functions with compact support in Ω. We refer the reader to Attouch, Buttazzo, Michaille, 2014; Ambrosio, Fusco, Pallara, 2000 for more on functions of bounded variation, i. e., functions with finite TV-seminorm.

Before recalling the concept of total *generalized* variation, we mention that the literature considering classical TV regularization with piecewise constant or piecewise linear functions on triangular meshes is already quite rich; see, e. g., Feng, Prohl, 2003; Elliott, Smitheman, 2009; Wu et al., 2012; Bartels, 2012; Stamm, Wihler, 2015; Bartels, Nochetto, Salgado, 2015; Alkämper, Langer, 2017; Berkels, Effland, Rumpf, 2017; Lee, Park, Park, 2019; Clason, Kruse, Kunisch, 2018; Herrmann, Herzog, Kröner, et al., 2018; Chambolle, Pock, 2021 for applications in denoising and inpainting of images as well as optimization problems in the coefficient of differential equations. Moreover Herrmann, Herzog, Schmidt, et al., 2019 established a fully discrete analogue of (1.1) for *higher-order* finite element spaces, which utilizes a Raviart-Thomas finite element space for the "test" functions $\boldsymbol{v}$ (1.1). Recently, Chambolle, Pock, 2020 proposed a different discrete approximation of the TV-seminorm over the Crouzeix-Raviart finite element space for the data functions u; see also Bartels, 2021. Furthermore, the discrete total variation of the piecewise constant normal vector on triangulated meshes in 3D was investigated as a regularizer for shape optimization problems in Zhang et al., 2015; Bergmann, Herrmann, et al., 2020.

To motivate the need for higher-order TV models, we recall that the TV-seminorm as a regularizer, e. g., in the ROF image denoising model Rudin, Osher, Fatemi, 1992, leads to the so-called "staircasing effect", i. e., the formation of islands of equal function values. To counteract this phenomenon, several

extensions have been proposed in the literature, often revolving around concepts of total variation of second order

$$\text{TV}^2(u) = \sup\left\{\int_\Omega u \operatorname{div}\operatorname{Div} V \, dx \,\middle|\, V \in C_c^2(\Omega, \mathbb{R}^{2\times2}), \; \|V\|_{L^\infty(\Omega,\mathbb{R}^{2\times2})} \le \alpha_0, \right\}, \tag{1.2}$$

and variations thereof; see Chambolle, Lions, 1997; Chan, Esedoglu, Park, 2010; Papafitsoros, Schönlieb, 2014 and the discussion in Bredies, Kunisch, Pock, 2010. In (1.2), div Div stands for the row-wise divergence (Div) applied to the matrix-valued function V, followed by the classical divergence (div) applied to the resulting vector field.

However, in the present work, our focus will be on the second-order total generalized variation (TGV) regularizer as introduced in Bredies, Kunisch, Pock, 2010, which is a state-of-the-art improvement over the TV-seminorm (1.1). Given parameters $\alpha_0, \alpha_1 > 0$, the second-order TGV-seminorm reads

$$\text{TGV}^2_{(\alpha_0,\alpha_1)}(u)$$
$$= \sup\left\{\int_\Omega u \operatorname{div}\operatorname{Div} V \, dx \,\middle|\, V \in C_c^2(\Omega, \mathbb{R}^{2\times2}_{\text{sym}}), \; \|V\|_{L^\infty(\Omega,\mathbb{R}^{2\times2})} \le \alpha_0, \; \|\operatorname{Div} V\|_{L^\infty(\Omega,\mathbb{R}^2)} \le \alpha_1 \right\}. \tag{1.3}$$

Here $\mathbb{R}^{2\times2}_{\text{sym}}$ denotes the space of symmetric 2×2-matrices. As shown in Bredies, Kunisch, Pock, 2010, this regularizer prefers piecewise linear reconstructions rather than piecewise constant ones, thereby preventing the staircasing effect while still preserving discontinuities. Second-order TGV results in a certain balancing of first- and second-order TV. This can be seen by rewriting (1.3) by means of Fenchel's duality theorem, see Bredies, Holler, 2014, to obtain

$$\text{TGV}^2_{(\alpha_0,\alpha_1)}(u) = \min_{w\in\mathcal{M}(\Omega,\mathbb{R}^2)} \alpha_1 \|\nabla u - w\|_{\mathcal{M}(\Omega,\mathbb{R}^2)} + \alpha_0 \|\mathcal{E}w\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times2}_{\text{sym}})}. \tag{1.4}$$

Here $\mathcal{E} = \frac{1}{2}(\nabla + \nabla^\mathsf{T})$ is the symmetric (distributional) Jacobian of vector-valued functions, and $\mathcal{M}(\Omega, X)$ denotes the Banach space of finite signed Radon measures taking values in the Banach space X. As will be discussed in Section 2.1, Equation (1.3) reduces to $\alpha_1 \text{TV}(u)$ for piecewise constant functions u. Therefore, the verbatim use of (1.3) has no advantage over first-order TV (1.1) for piecewise constant functions. In order to exploit the additional features of second-order TGV in a discrete setting, several authors have considered different discrete interpretations of (1.4) and its variations.

In their original publication, Bredies, Kunisch, Pock, 2010, proposed a discretization of (1.4) using one-sided finite differences to evaluate discrete gradients of piecewise constant functions on regular grids, allowing the use of TGV in imaging problems. These discrete gradients can be captured by the auxiliary variable w located in the grid points. The symmetric Jacobian of w is then evaluated using backward finite differences. This discrete formulation has proven very successful for image denoising problems on regular pixel grids; see, e. g., Bredies, Kunisch, Pock, 2010; Knoll et al., 2010.

As part of the present work, we propose a novel discrete interpretation of the second-order TGV-seminorm (1.3)–(1.4), termed FETGV^2, which allows it to be used for piecewise constant functions on general triangular meshes in 2D. Our approach significantly extends the applicability of the TGV in imaging and other inverse problems, possibly involving PDEs, dispensing with the need to work on regular Cartesian grids. In addition, it can be extended to surface meshes in 3D in a straightforward way, as we demonstrate also in this paper.

To the best of our knowledge, there is only one approach with a similar purpose in the literature so far, see Gong et al., 2018. In that paper, the authors utilize the dual graph of the triangular mesh and apply the graph-based definition of TGV from Ono, Yamada, Kumazawa, 2015. We defer further discussions of the relation to our work to Section 2.

Our paper is structured as follows. Section 2 reviews two formulations of TGV, (1.3) and its non-symmetric variant, from Bredies, Kunisch, Pock, 2010 and discusses its relation to the infimal convolution approach. Furthermore, background material on finite element spaces is collected and related work on discrete versions of TGV is discussed in more detail, particularly Gong et al., 2018; Ono, Yamada, Kumazawa, 2015. Our proposal for a discrete version of the second-order TGV-seminorm for piecewise constant functions on general triangular meshes, termed FETGV^2, is detailed in Section 3. It is based on lowest-order discontinuous Lagrange and Raviart-Thomas finite elements. We discuss its properties and show that its kernel consists of functions arising from the interpolation of linear functions in the triangle circumcenters. Moreover, we demonstrate that its 1D analogue coincides with the 1D discretization of second-order TGV proposed in Bredies, Kunisch, Pock, 2010. Section 4 of this paper is devoted to the numerical realization of optimization algorithms involving the non-smooth FETGV^2 functional. We employ an alternating direction method of multipliers (ADMM), specifically the split Bregman algorithm for this purpose. Numerical results for image denoising and inpainting problems are presented in Section 5 and compared to those obtained with the discrete TGV proposals from Bredies, Kunisch, Pock, 2010; Gong et al., 2018. The paper finishes with a conclusion in Section 6.

2 Literature Review and Background Material

Throughout the rest of this paper, we assume that $\Omega \subset \mathbb{R}^2$ is a bounded domain unless otherwise noted.

2.1 Continuous Formulations of Second-Order TGV

In this subsection we discuss the second-order total generalized variation

$$\begin{aligned}
&\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) \\
&= \sup \left\{ \int_\Omega u \operatorname{div} \operatorname{Div} V \, \mathrm{d}x \;\middle|\; V \in C^2_c(\Omega, \mathbb{R}^{2\times 2}_{\mathrm{sym}}), \; \|V\|_{L^\infty(\Omega, \mathbb{R}^{2\times 2})} \le \alpha_0, \; \|\operatorname{Div} V\|_{L^\infty(\Omega, \mathbb{R}^2)} \le \alpha_1 \right\}, \quad (2.1)
\end{aligned}$$

and its non-symmetric variant

$$\begin{aligned}
&\neg\operatorname{sym}\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) \\
&= \sup \left\{ \int_\Omega u \operatorname{div} \operatorname{Div} V \, \mathrm{d}x \;\middle|\; V \in C^2_c(\Omega, \mathbb{R}^{2\times 2}), \; \|V\|_{L^\infty(\Omega, \mathbb{R}^{2\times 2})} \le \alpha_0, \; \|\operatorname{Div} V\|_{L^\infty(\Omega, \mathbb{R}^2)} \le \alpha_1 \right\}, \quad (2.2)
\end{aligned}$$

both proposed in Bredies, Kunisch, Pock, 2010. As was mentioned in the introduction, (2.1) and (2.2) can be rewritten by Fenchel's duality theorem, yielding

$$\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) = \min_{w \in \mathcal{M}(\Omega,\mathbb{R}^2)} \alpha_1 \|\nabla u - w\|_{\mathcal{M}(\Omega,\mathbb{R}^2)} + \alpha_0 \|\mathcal{E}w\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times 2}_{\mathrm{sym}})} \tag{2.3}$$

as well as

$$\neg \operatorname{symTGV}^2_{(\alpha_0, \alpha_1)}(u) = \min_{\mathbf{w} \in \mathcal{M}(\Omega, \mathbb{R}^2)} \alpha_1 \, \|\nabla u - \mathbf{w}\|_{\mathcal{M}(\Omega, \mathbb{R}^2)} + \alpha_0 \, \|\nabla \mathbf{w}\|_{\mathcal{M}(\Omega, \mathbb{R}^{2 \times 2})}$$

$$= \min_{\mathbf{w} \in \mathcal{M}(\Omega, \mathbb{R}^2)} \alpha_1 \, \|\nabla u - \mathbf{w}\|_{\mathcal{M}(\Omega, \mathbb{R}^2)} + \alpha_0 \operatorname{TV}(\mathbf{w}) \tag{2.4}$$

respectively. Again, $\mathcal{E} = \frac{1}{2}(\nabla + \nabla^\mathsf{T})$ denotes the symmetric (distributional) Jacobian of vector-valued functions, and $\mathcal{M}$ is the Banach space of finite signed Radon measures η taking values in the Banach space X, equipped with the norm

$$\|\eta\|_{\mathcal{M}(\Omega, X)} := \sup \left\{ \int_\Omega v \, \mathrm{d}\eta \, \middle| \, v \in C_c(\Omega, X), \|v\|_{L^\infty(\Omega, X)} \leq 1 \right\}. \tag{2.5}$$

Specifically, the cases $X = \mathbb{R}^2$ and $X = \mathbb{R}^{2 \times 2}$ will be relevant for us. We refer the reader to Bredies, Holler, 2014 for details on the space $\mathcal{M}(\Omega, X)$.

The dual formulations (2.3) and (2.4) are in fact closely related to the approach of Chambolle, Lions, 1997, which balances first- and second-order total variation by their weighted infimal convolution,

$$\operatorname{ICTV}(u) = \min_v \alpha_1 \operatorname{TV}(u - v) + \alpha_0 \operatorname{TV}^2(v)$$

$$= \min_v \alpha_1 \, \|\nabla(u - v)\|_{\mathcal{M}(\Omega, \mathbb{R}^2)} + \alpha_0 \, \|\nabla^2 v\|_{\mathcal{M}(\Omega, \mathbb{R}^{2 \times 2})}. \tag{2.6}$$

As pointed out, e. g., in Bergmann, Fitschen, et al., 2017, the difference between (2.4) and (2.6) is that the infimal convolution formulation (2.6) uses a decomposition of u into v and a rest, rather than a decomposition of ∇u into $\mathbf{w}$ and a rest as in (2.4). These two formulations are generally not the same since $\mathbf{w}$ in (2.4) is not necessarily a gradient field.

The minimization involved in both, the TGV (2.4) and the infimal convolution (2.6) formulations can be seen as an optimal additive decomposition of u or ∇u, respectively. Informally, this leads to a balancing between two priors, weighted by α_1 and α_0, respectively. The first summand in (2.4), i. e., the first-order total variation prior, is zero if and only if the minimizer $\mathbf{w}$ satisfies $\mathbf{w} = \nabla u$. In case of (2.6), the respective first summand is zero if and only if $v = u + c$ holds for some constant function c. Hence in these cases (2.4) and (2.6) capture only a second-order total variation part. On the other hand, the second term in (2.4) vanishes if and only if $\mathbf{w}$ is constant, while the second term in (2.6) vanishes if and only if v is linear. Note that consequently both (2.4) and (2.6) vanish altogether when u is a linear function. In that case we obtain $\nabla u = \mathbf{w}$ as well as $v = u + c$ for arbitrary constant functions c. One thus concludes that precisely the linear functions span the kernel of (2.4) and (2.6).

It was already observed in Bredies, Holler, 2014, Theorem 3.5 that the auxiliary variable $\mathbf{w}$ must have more regularity and can, in fact, not be measure-valued at the minimizer in (2.3) and (2.4). This is due to the fact that measures generally have unbounded (distributed) gradients themselves. Consequently, at a minimizer, $\mathbf{w}$ must have a bounded symmetric (or non-symmetric) distributed Jacobian since otherwise the term involving α_0 would equal $+\infty$. Therefore, instead of $\mathcal{M}(\Omega, \mathbb{R}^2)$, the space of bounded deformations

$$BD(\Omega, \mathbb{R}^2) = \left\{ u \in L^1(\Omega, \mathbb{R}^2) \, \middle| \, \mathcal{E}u \in \mathcal{M}(\Omega, \mathbb{R}^{2 \times 2}_{\mathrm{sym}}) \right\} \tag{2.7}$$

respectively of bounded variation

$$BV(\Omega, \mathbb{R}^2) = \left\{ u \in L^1(\Omega, \mathbb{R}^2) \, \middle| \, \nabla u \in \mathcal{M}(\Omega, \mathbb{R}^{2 \times 2}) \right\} \tag{2.8}$$

can be used for w in (2.3) and (2.4). Notice that w is thus a Radon measure with a Lebesgue density, i. e.,

$$\int_\Omega v \, \mathrm{d}w \coloneqq \int_\Omega v \cdot w \, \mathrm{d}x \tag{2.9}$$

holds.

For piecewise polynomial functions u on a conforming mesh with polyhedral cells T and interior edges E, (2.4) evaluates to

$$\begin{aligned}
\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) & \tag{2.10} \\
&= \min_{w \in BV(\Omega,\mathbb{R}^2)} \alpha_1 \left\| \nabla u - w \right\|_{\mathcal{M}(\Omega,\mathbb{R}^2)} + \alpha_0 \left\| \nabla w \right\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times2})} \\
&= \min_{w \in BV(\Omega,\mathbb{R}^2)} \alpha_1 \sup \left\{ \int_\Omega u \operatorname{div} v + w \cdot v \, \mathrm{d}x \; \middle| \; v \in C_c^1(\Omega,\mathbb{R}^2), \; \|v\|_{L^\infty(\Omega,\mathbb{R}^2)} \le 1 \right\} \\
&\qquad + \alpha_0 \left\| \nabla w \right\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times2})} \\
&= \min_{w \in BV(\Omega,\mathbb{R}^2)} \alpha_1 \sum_E \int_E \left| [\![u]\!] \right| \mathrm{d}x + \alpha_1 \sum_T \int_T |\nabla u - w|_2 \, \mathrm{d}x + \alpha_0 \left\| \nabla w \right\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times2})}.
\end{aligned}$$

Here $[\![u]\!]$ denotes the scalar jump of u across an edge; see (2.18). Moreover, $|\cdot|_2$ denotes the Euclidean norm of a vector. Since for piecewise constant functions u, ∇u is zero inside each triangle, so will be the minimizer for w. Then one is left with $\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) = \alpha_1 \mathrm{TV}(u)$ since all terms except for the first vanish. This shows that in order to exploit the additional features of the TGV-seminorm in a discrete setting, an appropriate discrete interpretation of (2.4) is necessary. We review existing approaches in the following subsection.

Remark 2.1 (Piecewise linear, continuous functions). *When u is a piecewise linear, continuous function ($u \in C\mathcal{G}_1(\Omega)$), then (2.10) reduces to*

$$\mathrm{TGV}^2_{(\alpha_0,\alpha_1)}(u) = \min_{w \in BV(\Omega,\mathbb{R}^2)} \alpha_1 \sum_T \int_T |\nabla u - w|_2 \, \mathrm{d}x + \alpha_0 \left\| \nabla w \right\|_{\mathcal{M}(\Omega,\mathbb{R}^{2\times2})}. \tag{2.11}$$

In contrast with the case that u is piecewise constant, this formulation does not reduce to first-order total variation. It is therefore not necessary to introduce a discrete interpretation of (2.11) beyond a choice for the discretization of w. In view of the fact that the gradient maps $C\mathcal{G}_1(\Omega)$ surjectively onto curl-free functions in the lowest-order Nédélec space $\mathcal{N}_0(\Omega)$, at least on simply connected domains, it appears natural to choose $w \in \mathcal{N}_0(\Omega)$; see for instance Arnold, Falk, Winther, 2006; Christiansen, Hu, Hu, 2018. However, we leave those investigations to future work.

2.2 Discrete Formulations of Second-Order TGV

2.2.1 Discretization on Cartesian Grids

The original discrete interpretation of (2.3) from Bredies, Kunisch, Pock, 2010 was derived for a two-dimensional Cartesian grid with grid size h,

$$\Omega_h = \{ (ih, jh) \mid (0,0) \le (i,j) < (M,N) \}.$$

For a grid function v_h defined through its nodal values $(v_h)_{i,j}$, define the horizontal and vertical forward differences as

$$(\delta^h_{x_+} v_h)_{i,j} = \begin{cases} (v^h_{i+1,j} - v^h_{i,j})/h & \text{if } 0 \le i < M - 1, \\ 0 & \text{if } i = M - 1, \end{cases}$$

$$(\delta^h_{y_+} v_h)_{i,j} = \begin{cases} (v^h_{i,j+1} - v^h_{i,j})/h & \text{if } 0 \le j < N - 1, \\ 0 & \text{if } j = N - 1, \end{cases}$$

and the backward differences according to

$$(\delta^h_{x_-} v_h)_{i,j} = \begin{cases} -v^h_{i-1,j}/h & \text{if } i = M - 1, \\ (v^h_{i,j} - v^h_{i-1,j})/h & \text{if } 0 < i < M - 1, \\ v^h_{i,j}/h & \text{if } i = 0, \end{cases}$$

$$(\delta^h_{y_-} v_h)_{i,j} = \begin{cases} -v^h_{i,j-1}/h & \text{if } j = N - 1, \\ (v^h_{i,j} - v^h_{i,j-1})/h & \text{if } 0 < j < N - 1, \\ v^h_{i,j}/h & \text{if } j = 0. \end{cases} \tag{2.12}$$

Further, denote by

$$\nabla^h v_h = \begin{pmatrix} \delta^h_{x_+} v^h \\ \delta^h_{y_+} v^h \end{pmatrix}$$

the discrete forward gradient of v_h and by

$$\mathcal{E}^h \begin{pmatrix} \boldsymbol{w}^h_1 \\ \boldsymbol{w}^h_2 \end{pmatrix} = \begin{pmatrix} \delta^h_{x_-} \boldsymbol{w}^h_1 & \frac{\delta^h_{y_-} \boldsymbol{w}^h_1 + \delta^h_{x_-} \boldsymbol{w}^h_2}{2} \\ \frac{\delta^h_{y_-} \boldsymbol{w}^h_1 + \delta^h_{x_-} \boldsymbol{w}^h_2}{2} & \delta^h_{y_-} \boldsymbol{w}^h_2 \end{pmatrix}$$

the discrete symmetric Jacobian of a discrete nodal vector field $\boldsymbol{w}^h$ on Ω_h. The discretization of (2.3) originally proposed in Bredies, Kunisch, Pock, 2010 can now be written as

$$\text{TGV}^h_{(\alpha_0,\alpha_1)}(u^h) = \min_{\boldsymbol{w}^h} \alpha_1 \sum_{i,j} \big|(\nabla^h u^h)_{i,j} - (\boldsymbol{w}^h)_{i,j}\big|_2 + \alpha_0 \sum_{i,j} \big|(\mathcal{E}\boldsymbol{w}^h)_{i,j}\big|_F. \tag{2.13}$$

Here $|\cdot|_F$ denotes the Frobenius norm of a matrix.

2.2.2 TGV on Graphs

One of the first approaches to use the concept of TGV outside the realm of the regular pixel grid structures used in imaging was devised by Ono, Yamada, Kumazawa, 2015, where total generalized variation for graph signals was introduced. Suppose that $G = (\mathcal{V}, \mathcal{E})$ is an undirected graph with vertex sets $\mathcal{V}$ and edge sets $\mathcal{E}$ of finite cardinalities $\#\mathcal{V}$ and $\#\mathcal{E}$. Let $\tilde{u} \in \mathbb{R}^{\#\mathcal{V}}$ be a scalar-valued function on $\mathcal{V}$. According to Chan, Osher, Shen, 2001, the total variation of such a graph signal is measured using the help of a jump operator $\mathbb{R}^{\#\mathcal{V}} \to \mathbb{R}^{\#\mathcal{E}}$, represented by a matrix $J \in \mathbb{R}^{\#\mathcal{E} \times \#\mathcal{V}}$, which

maps the nodal value vector $\tilde{u}$ to the vector of differences of adjacent values (with arbitrary but fixed orientation) across the incident edges. We then have the definition

$$\text{G-TV}(\tilde{u}) := |J\tilde{u}|_1 = \sum_{e_{ij} \in \mathcal{E}} w_{ij} |\tilde{u}_i - \tilde{u}_j|, \tag{2.14}$$

where $w_{ij} > 0$ is a weight associated with the edge e_{ij}. To obtain second-order derivatives on graphs one usually employs the differential operator $J^\mathsf{T} J$. Notice that the jump operator corresponds to the gradient in the continuous case. The adjoint J^T corresponds to a divergence, making $J^\mathsf{T} J$ correspond to $\operatorname{div} \nabla = \Delta$, the Laplace operator. In fact, $J^\mathsf{T} J$ is the graph Laplacian matrix. Leveraging again the idea of balancing first- and second-order derivatives, the total generalized variation for data on graphs was introduced in Ono, Yamada, Kumazawa, 2015 as

$$\text{G-TGV}^2_{(\alpha_0,\alpha_1)}(\tilde{u}) := \min_{q \in \mathbb{R}^{\#\mathcal{E}}} \alpha_1 |J\tilde{u} - q|_1 + \alpha_0 |J^\mathsf{T} q|_1. \tag{2.15}$$

The corresponding continuous formulation

$$\text{lapTGV}^2_{(\alpha_0,\alpha_1)}(u) := \min_{\boldsymbol{w} \in \mathcal{M}(\Omega,\mathbb{R}^2)} \alpha_1 \|\nabla u - \boldsymbol{w}\|_{\mathcal{M}(\Omega,\mathbb{R}^2)} + \alpha_0 \|\operatorname{div} \boldsymbol{w}\|_{\mathcal{M}(\Omega)} \tag{2.16}$$

differs from (2.3) in the α_0-term, in which the symmetric Jacobian has been replaced by a divergence. The effects of this change in the differential operator have been studied in Brinkmann, Burger, Grah, 2018. In the continuous setting, (2.16) is zero for all functions satisfying $\Delta u = \operatorname{div}(\nabla u) = 0$, i.e., its kernel is infinite-dimensional. This regularizer, which we refer to as $\text{lapTGV}^2_{(\alpha_0,\alpha_1)}$, thus promotes piecewise harmonic reconstructions. This is in contrast to (2.3), whose kernel contains precisely the linear functions and thus is three-dimensional on connected domains in $\mathbb{R}^2$. In the finite difference implementation of Brinkmann, Burger, Grah, 2018 for image denoising, artifacts occurred for (2.16), particularly at locations where jumps are to be reconstructed, see, e.g., Figure 5.1e. This can be explained by the oscillatory nature of solutions to the Laplace equation without boundary condition, e.g., $\sin(kx) \sinh(ky)$ for arbitrary k.

2.3 Discrete Formulations of Second-Order TGV using Finite Element Spaces

In this section, we review the (scarce) literature on discrete formulations of the total generalized variation with finite elements. Before that, we use the opportunity to recap the required finite element spaces of discontinuous Lagrange as well as Raviart-Thomas type.

2.3.1 Discontinuous Lagrange and Raviart-Thomas Finite Element Spaces

Suppose that Ω is a two-dimensional polygonal domain covered by a mesh of non-degenerate triangular cells T and interior edges E. We denote the discontinuous Lagrange finite element spaces of order $r \in \mathbb{N}_0$ (the non-negative integers) on such a mesh by

$$\mathcal{DG}_r(\Omega) := \left\{ u \in L^2(\Omega) \,\middle|\, u|_T \in P_r(T) \right\}.$$

Here $P_r(T)$ denotes the space of bivariate polynomials of degree at most r. We will use the space $\mathcal{DG}_r$ for $r = 0$ (piecewise constant functions) and $r = 1$ (piecewise linear). We represent elements of $P_0(T)$

by their values in the cell center and elements of $P_1(T)$ by their values in the three vertices $\{X_{T,k}\}$ of T. The corresponding linear Lagrange polynomials are denoted by $\{\Phi_{T,k}\}$ with $k = 1, 2, 3$, and we have $\Phi_{X,k}(X_{T,\ell}) = \delta_{k\ell}$.

Analogously to the above, we define the discontinuous Lagrange finite element spaces of order $r \in \mathbb{N}_0$ on the skeleton (the union of the interior edges, $\cup E$) of the mesh as

$$\mathcal{DG}_r(\cup E) := \left\{ u \in L^2(\cup E) \,\middle|\, u|_E \in P_r(E) \right\}.$$

Here $P_r(E)$ denotes the space of univariate polynomials on an interior edge E of the mesh. Similarly as before, elements of $P_1(E)$ are represented by their values in the two vertices $X_{E,k}$ of E, $k = 1, 2$. The corresponding linear Lagrange polynomials are denoted by $\Phi_{E,k}$. We also define the interpolation operator

$$\mathcal{I}_E\{v\} := \sum_{k=1}^{2} v(X_{E,k}) \, \Phi_{E,k} \tag{2.17}$$

for continuous functions v.

A function $u \in \mathcal{DG}_r(\Omega)$ will in general exhibit discontinuities across interior edges. In order to evaluate these discontinuities, we choose an arbitrary but fixed orientation for each interior edge E, such that one of the adjacent triangles is denoted by T_+ and the other one by T_-. Define u_+ and u_- as the value of u restricted to T_+ and T_-, respectively, and evaluated on the common edge E. Then the jump operator is given by

$$[\![u]\!] := u_+ - u_-. \tag{2.18}$$

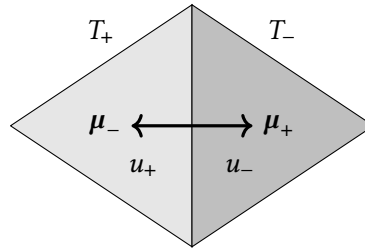

Figure 2.1: Visualization of some notation for functions $u \in \mathcal{DG}_r(\Omega)$.

Recall the Sobolev space

$$H(\mathrm{div}; \Omega) := \left\{ \boldsymbol{v} \in L^2(\Omega; \mathbb{R}^2) \,\middle|\, \mathrm{div}\, \boldsymbol{v} \in L^2(\Omega) \right\}$$

as the space of square-integrable, vector-valued functions having a weak divergence in $L^2(\Omega)$. A piecewise polynomial function $\boldsymbol{w} \in \mathcal{DG}_r(\Omega; \mathbb{R}^2)$ belongs to $H(\mathrm{div}; \Omega)$ if and only if the normal component of $\boldsymbol{w}$ is continuous across each edge. We denote by $\boldsymbol{\mu}_+$ and $\boldsymbol{\mu}_-$ the outward unit normal vectors on E as seen from the cells T_+ and T_-, respectively. Similarly as in the scalar case, the jump of $\boldsymbol{w}$ across E is defined as $[\![\boldsymbol{w}]\!] := \boldsymbol{w}_+ - \boldsymbol{w}_-$, and the continuity of the normal component can be expressed as $[\![\boldsymbol{w}]\!] \cdot \boldsymbol{\mu}_+ = 0$ or, equivalently, as $[\![\boldsymbol{w}]\!] \cdot \boldsymbol{\mu}_- = 0$.

We denote the lowest-order Raviart-Thomas finite element space by

$$\mathcal{RT}_0(\Omega) := \left\{ \boldsymbol{v} \in H(\mathrm{div}; \Omega) \,\middle|\, \boldsymbol{v}|_T \in P_0(T)^2 + \begin{pmatrix} x \\ y \end{pmatrix} P_0(T) \right\},$$

We refer to Herrmann, Herzog, Schmidt, et al., 2019 as well as Logg, Mardal, Wells, 2012 for more details on this $H(\text{div}; \Omega)$-conforming finite element space. Notice that for $\boldsymbol{p} \in \mathcal{RT}_0$, the following degrees of freedom are going to be used in this paper:

$$\int_E \boldsymbol{p}_+ \cdot \boldsymbol{\mu}_+ \, \mathrm{d}S. \tag{2.19}$$

Finally, let $\mathcal{RT}_0^0(\Omega)$ be the subspace of $\mathcal{RT}_0(\Omega)$ of functions with vanishing normal component on the outer boundary of Ω.

2.3.2 Discrete Total Variation with Finite Elements

It is well known that for piecewise constant and piecewise linear functions, the total variation (1.1) simplifies to

$$\mathrm{TV}(u) = \sum_E \int_E \big| [\![u]\!] \big| \, \mathrm{d}S \qquad\qquad \text{for } u \in \mathcal{DG}_0, \tag{2.20}$$

$$\mathrm{TV}(u) = \sum_T \int_T |\nabla u|_2 \, \mathrm{d}x + \sum_E \int_E \big| [\![u]\!] \big| \, \mathrm{d}S \quad \text{for } u \in \mathcal{DG}_1. \tag{2.21}$$

The *discrete* total variation proposed in Herrmann, Herzog, Schmidt, et al., 2019 agrees with (2.20) in case of $u \in \mathcal{DG}_0$ but differs from (2.21) in case of $u \in \mathcal{DG}_1$. In that case, instead of integrating the nonlinear term $\big| [\![u]\!] \big|$ on an edge E, it was proposed to integrate its linear interpolant, i. e., (2.21) is replaced by

$$\mathrm{DTV}(u) = \sum_T \int_T |\nabla u|_2 \, \mathrm{d}x + \sum_E \int_E \mathcal{I}_E \big\{ \big| [\![u]\!] \big| \big\} \, \mathrm{d}S \quad \text{for } u \in \mathcal{DG}_1. \tag{2.22}$$

2.3.3 Total Generalized Variation for Finite Element Functions via the Dual Graph

As was already mentioned, the first proposal to define the second-order total generalized variation for piecewise constant finite element functions was made by Gong et al., 2018. There, the total generalized variation concept for graph signals from Ono, Yamada, Kumazawa, 2015, see Section 2.2.2, was used on the dual graph of the finite element mesh. We remind the reader that the vertices of the dual graph are the triangle centers and each pair of neighboring triangles generates an edge; see Figure 2.2 for an illustration.

To retrace the approach by Gong et al., 2018, let us denote by $G = (\mathcal{V}, \mathcal{E})$ be the dual graph of a triangulated mesh with vertices $\mathcal{V}$ and edges $\mathcal{E}$, with cardinalities $\#\mathcal{V}$ and $\#\mathcal{E}$, respectively. It is then possible to represent piecewise constant data $u \in \mathcal{DG}_0(\Omega)$ equivalently by vertex data $\tilde{u} \in \mathbb{R}^{\#\mathcal{V}}$ on the graph. Notice that the total variation for $u \in \mathcal{DG}_0(\Omega)$ agrees with the total variation of the data $\tilde{u}$ on the dual graph, provided that the edge lengths $|E|$ in the mesh are used as weights in J, see (2.14):

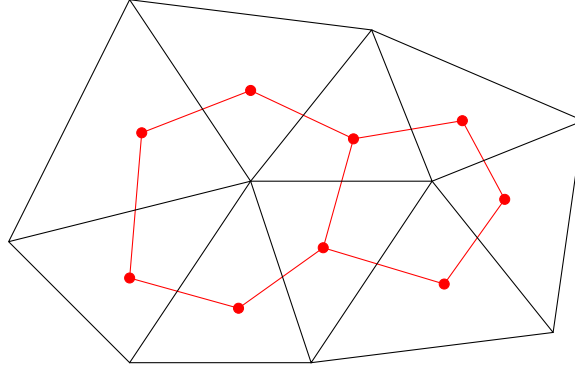$$\text{G-TV}(\tilde{u}) \coloneqq |J\tilde{u}|_1 = \sum_E |E| \big| [\![u]\!] \big|. \tag{2.23}$$

Figure 2.2: Finite element mesh (black) and corresponding dual graph (red).

The total generalized variation G-TGV$^2_{(\alpha_0,\alpha_1)}(\tilde{u})$ of Gong et al., 2018 was recalled in (2.15) as

$$\text{G-TGV}^2_{(\alpha_0,\alpha_1)}(\tilde{u}) = \min_{q \in \mathbb{R}^{\#\mathcal{E}}} \alpha_1 \, |J\tilde{u} - q|_1 + \alpha_0 \, |J^\mathsf{T}q|_1.$$

We can now interpret this expression in terms of the finite element mesh corresponding to the dual graph. Using the previously defined degrees of freedom (2.19) for the Raviart-Thomas space $\mathcal{RT}_0(\Omega)$, the coefficient vector representation of the following problem,

$$\text{lapFETGV}^2_{(\alpha_0,\alpha_1)}(u) := \min_{\boldsymbol{w} \in \mathcal{RT}_0^0(\Omega)} \alpha_1 \sum_E \left\| [\![u]\!] + \frac{1}{|E|}\boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+ \right\|_{L^1(E)} + \alpha_0 \sum_T \|\text{div}\,\boldsymbol{w}\|_{L^1(T)} \qquad (2.24)$$

matches (2.15) after the minor substitution of the auxiliary variable $\tilde{\boldsymbol{w}} = |E|q$, where $\tilde{\boldsymbol{w}}$ denotes the coefficient vector of $\boldsymbol{w}$. This emphasizes that the total generalized variation on a graph corresponds to a discrete version of lapTGV$^2_{(\alpha_0,\alpha_1)}$ from (2.16).

3 Novel Discrete Formulation of Second-Order TGV for Piecewise Constant Elements on Unstructured Meshes

In this section we propose a novel discretization of the non-symmetric TGV functional (2.4). Our formulation has a number of advantages. It is tailored to triangulated grids, i.e., data structures more flexible than regular pixel grids. An earlier approach for triangulated grids, here denoted by lapFETGV$^2_{(\alpha_0,\alpha_1)}$, see (2.16) and equivalently (2.24), was developed in Gong et al., 2018 but it shows artifacts, as confirmed by the independent implementation due to Brinkmann, Burger, Grah, 2018. We attribute those artifacts to the observation that (2.24) favors piecewise discrete harmonic functions rather than piecewise linear. We refer the reader to Section 5 for numerical experiments reflecting this observation.

Our approach starts from the continuous formulation of non-symmetric TGV, (2.4). There, the α_0-term measures the norm of the distributional gradient $\nabla \boldsymbol{w}$, i.e., the total variation of the auxiliary variable $\boldsymbol{w}$, rather than the norm of the divergence $\text{div}\,\boldsymbol{w}$. In the discrete case, where $\boldsymbol{w}$ is an $\mathcal{RT}_0$-finite element function, a vector-valued variant of the discrete total variation (2.22) for $\mathcal{DG}_1$ functions from Herrmann,

Herzog, Schmidt, et al., 2019 can be used for this purpose, noting that $\mathcal{RT}_0(\Omega) \subset \mathcal{DG}_1(\Omega; \mathbb{R}^2)$. We therefore combine ideas from Gong et al., 2018 and Herrmann, Herzog, Schmidt, et al., 2019 and propose the following discrete approximation to (2.4) for piecewise constant $u \in \mathcal{DG}_0(\Omega)$:

$$\mathrm{FETGV}^2_{(\alpha_0, \alpha_1)}(u) := \min_{w \in \mathcal{RT}_0(\Omega)} \alpha_1 \sum_E \big\| [\![u]\!] + h_E \, w \cdot \mu_+ \big\|_{L^1(E)}$$

$$+ \alpha_0 \sum_T \int_T |\nabla w|_F \, dx + \alpha_0 \sum_E \int_E \mathcal{I}_E \big\{ \big| [\![w]\!] \big|_2 \big\} \, dS. \quad (3.1)$$

Recall from (2.17) that $\mathcal{I}_E\big\{\big|[\![w]\!]\big|_2\big\}$ denotes the linear interpolation of the pointwise 2-norm of the linear function $[\![w]\!] = w_+ - w_-$ onto the space of linear functions along E. The interpolation points are the end points of E. Just like the continuous formulation (2.16), the above (3.1) reduces to $\alpha_1 \, \mathrm{TV}(u)$ in case the minimizing w is zero. In comparison to the term $\nabla u - w$ in (2.16) however, one can notice the extra scaling factor h_E in front of $w \cdot \mu_+$.

Let $m_+, m_- \in \mathbb{R}^2$ be the circumcenters of two adjacent triangles sharing the edge E. We then choose the factor $h_E := |m_+ - m_-|_2$ so that we have

$$\mu_- \, h_E = -\mu_+ \, h_E = [\![m]\!] = m_+ - m_-. \quad (3.2)$$

Notice that (3.2) is sensible since $[\![m]\!]$ always intersects the midpoint of the edge E orthogonally and is therefore parallel to μ_+ and μ_-, see Figure 3.1.
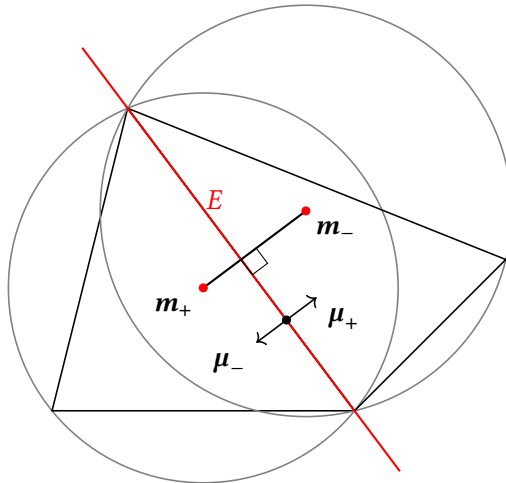

Figure 3.1: Illustration of (3.2): the factor h_E scales μ_+ and μ_- such that both their lengths equal the distance of the circumcenters m_+ and m_-.

By interpreting the jump $[\![u]\!]$ across the edge E as a directional derivative of u in the direction of $[\![m]\!]$, i.e.,

$$[\![u]\!] \approx \nabla u \cdot [\![m]\!], \quad (3.3)$$

we are offered with an alternative motivation for the formulation (3.1). Because only edge-concentrated (and directional) information on the derivative of u is available, the α_1-term in (3.1) is used to couple the directional derivative of u (the jump in the sense of (3.3)) and the directional component of w, both in the direction of $[\![m]\!] = h_E \, (-\mu_+)$. The Raviart-Thomas space is then in fact a very natural choice

for the auxiliary variable w, because it has exactly the normal component as a degree of freedom. It furthermore propagates the edge-concentrated information into the triangle so that the total variation of w can be measured through the α_0-terms in (3.1).

Notice that there is a structural difference between the proposed formulation (3.1), the finite difference discretization (2.13) from Bredies, Kunisch, Pock, 2010, and the lapFETGV (2.24) from Gong et al., 2018 with regards to the treatment of the auxiliary variable w on the boundary. For instance, (2.13) uses one-sided finite differences and thus requires some values of w^h on or outside the boundary. These values are taken to be zero, see (2.12) for $i = 0$ and $j = 0$ as well as for $i = M - 1$ and $j = N - 1$. Similarly in (2.24), an essential boundary condition $w \cdot n = 0$ is assumed. In both cases, these boundary conditions also impact the solution w in the interior of the domain. In comparison with the continuous formulation (2.3), either boundary condition on w appears artificial and we wish to avoid them.

3.1 1D ANALOGUE

In order to get a better understanding, we study the 1D analogue of the proposed formulation (3.1) in this subsection. In the 1D case, edges are replaced by vertices as the loci of intersection of adjacent mesh entities, which are now intervals instead of triangles. Given an interval Ω consisting of subintervals I in between vertices V, the analogue of (3.1) reads

$$\text{FETGV}^2_{(\alpha_0,\alpha_1)}(u) := \min_{w \in C\mathcal{G}_1(I)} \alpha_1 \sum_V \left\| [\![u]\!] - h_V w \right\|_{L^1(V)} + \alpha_0 \sum_I \|\nabla w\|_{L^1(I)}, \tag{3.4}$$

where $h_V = m_+ - m_- = \frac{|I_+| + |I_-|}{2}$ is the distance between the midpoints of the subintervals adjacent to the vertex V. Notice that we take here, w.l.o.g., the "+" subinterval to be the one to the right of any vertex and thus μ_+ becomes -1. Notice also that the Raviart-Thomas space $\mathcal{RT}_0(\Omega)$ coincides with the continuous Lagrange space $C\mathcal{G}_1(\Omega)$. By dividing $[\![u]\!] - h_V w$ in (3.4) by h_V it can be observed that the finite difference $[\![u]\!]/h_V$ across a vertex now involves only the single degree of freedom of w located at that vertex.
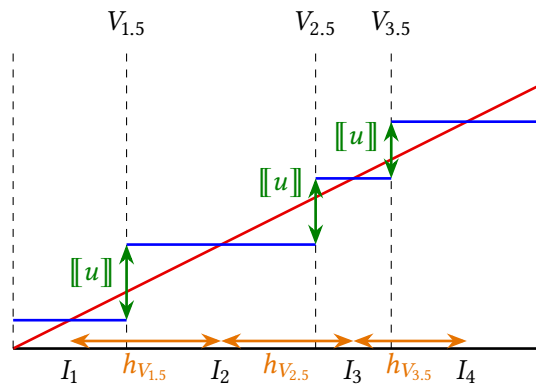

Figure 3.2: Visualization of a linear function (red) and its $\mathcal{DG}_0$ discretization (blue). The quotient of $[\![u]\!]/h_V$ is exactly the slope of a linear function.

We are now in the position to analyze the kernel of the semi-norm (3.4). The first condition for a piecewise constant function $u \in \mathcal{DG}_0(\Omega)$ to belong to the kernel is that $[\![u]\!] - h_V w = 0$ holds. The

second condition is that $w \in CG_1(\Omega)$ has zero gradient, i.e., w is constant. Both conditions together imply that $[\![u]\!]/h_V \equiv c$ is constant. We conclude that $\mathrm{FETGV}^2_{(\alpha_0,\alpha_1)}(u) = 0$ holds if any only if u interpolates a linear function on Ω with slope c at the midpoints of subintervals on the domain Ω; see Figure 3.2. The correct choice of the factor h_V is critical in achieving this result.

In the particular case that $n+1$ vertices are equally distributed and form n subintervals of length 1, the 1D formulation (3.4) coincides with the 1D finite difference discretization used in Bredies, Kunisch, Pock, 2010, except that in Bredies, Kunisch, Pock, 2010, $w = 0$ is assumed on or outside of Ω, depending which side of the domain the boundary is on. In terms of the degrees of freedom of the finite element functions, our formulation (3.4) in the setting at hand reads

$$\min_{w \in \mathbb{R}^{n+1}} \alpha_1 \sum_{i=1}^{n-1} \left| (u_{i+1} - u_i) - w_{i+\frac{1}{2}} \right| + \alpha_0 \sum_{i=1}^{n} \left| w_{i+\frac{1}{2}} - w_{i-\frac{1}{2}} \right|, \tag{3.5}$$

whereas the formulation (2.13) from Bredies, Kunisch, Pock, 2010 would read

$$\min_{w \in \mathbb{R}^{n}} \alpha_1 \sum_{i=1}^{n-1} \left| (u_{i+1} - u_i) - w_i \right| + \alpha_1 |w_n| + \alpha_0 |w_0| + \alpha_0 \sum_{i=2}^{n-1} |w_i - w_{i-1}| + \alpha_0 |-w_{n-1}|, \tag{3.6}$$

where the additional terms arise from the particular choice of finite difference scheme. Using a simple index shift, (3.5) and (3.6) coincide in the interior of the domain but not on the boundary. Our approach (3.5) avoids the introduction of artificial boundary conditions for w and thus the additional terms that arise in (3.6).

3.2 Properties of the Kernel in 2D

As discussed in the introduction, the kernel of TGV in the continuous settings contains precisely the linear functions on Ω. In this section we discuss the kernel of the proposed formulation (3.1) in 2D for piecewise constant functions. As a first observation, we note that a necessary condition for $\mathrm{FETGV}^2_{(\alpha_0,\alpha_1)}(u)$ to be zero, the minimizer for w must be globally constant since otherwise the α_0-terms will be non-zero. Furthermore, the minimizer $u \in \mathcal{DG}_0(\Omega)$ must satisfy

$$[\![u]\!] = -h_E w_+ \cdot \mu_+ = w_+ \cdot [\![m]\!] \tag{3.7}$$

on all interior edges E. Again m_+, m_- denote the circumcenters of the respective triangles. The kernel of this set of equations in $\mathcal{DG}_0(\Omega)$, one for each interior edge, consists of exactly the constant functions, hence u can be recovered from (3.7) only up to a constant. Since $w \in \mathcal{RT}_0(\Omega)$ is vector-valued and constant, it is determined by two degrees of freedom. Altogether, we infer that the kernel of (3.1) is a three-dimensional subspace of $\mathcal{DG}_0(\Omega)$. Notice that by (3.7), the jumps $[\![u]\!]$ will be zero on all edges parallel to w, i.e., where we have $w_+ \cdot \mu_+ = 0$. Furthermore, for edges perpendicular to w, $[\![u]\!]/h_E$ is equal to $\pm w_+ \cdot \mu_+$. This is exactly the slope of w in the direction of $\pm \mu_+$, thus the direction of the jump.

We now argue that the kernel of the novel formulation (3.1) includes precisely those functions in $\mathcal{DG}_0(\Omega)$ that interpolate linear functions at the circumcenters of each triangle; see Figure 3.3 for a visualization. This can be seen directly from (3.7) since for a linear function $f: \Omega \to \mathbb{R}$ one has

$$[\![f(m)]\!] = f(m_+) - f(m_-) = \nabla f \cdot [\![m]\!] \tag{3.8}$$

and setting $u_\pm = f(\boldsymbol{m}_\pm)$ and $\boldsymbol{w} = \nabla f$ yields (3.7).

Notice that for triangles with a right angle, the circumcenter lies on the hypotenuse. Therefore a kernel element always has the same function values $u_\pm$ on any two triangles $T_\pm$ which form a rectangle and share their hypotenuses; see Figure 3.3.
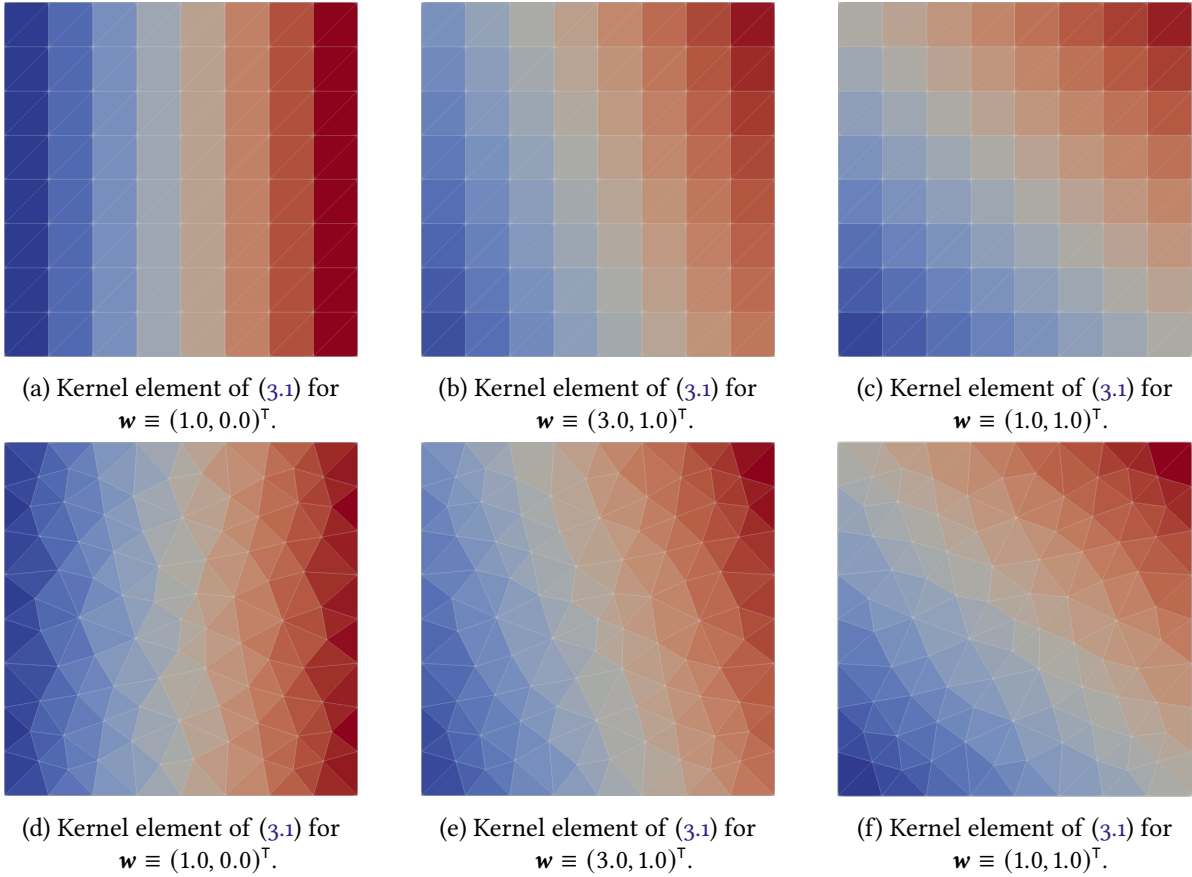
(a) Kernel element of (3.1) for $\boldsymbol{w} \equiv (1.0, 0.0)^\mathsf{T}$.

(b) Kernel element of (3.1) for $\boldsymbol{w} \equiv (3.0, 1.0)^\mathsf{T}$.

(c) Kernel element of (3.1) for $\boldsymbol{w} \equiv (1.0, 1.0)^\mathsf{T}$.

(d) Kernel element of (3.1) for $\boldsymbol{w} \equiv (1.0, 0.0)^\mathsf{T}$.

(e) Kernel element of (3.1) for $\boldsymbol{w} \equiv (3.0, 1.0)^\mathsf{T}$.

(f) Kernel element of (3.1) for $\boldsymbol{w} \equiv (1.0, 1.0)^\mathsf{T}$.

Figure 3.3: Kernel elements of the proposed formulation (3.1) on different meshes.

4 Using the Discrete Total Generalized Variation

In this section we present algorithmic details for optimization problems on triangulated meshes using the novel discrete formulation $\mathrm{FETGV}^2_{(\alpha_0, \alpha_1)}$ from (3.1) of second-order TGV. Numerical results will follow in Section 5, where we compare $\mathrm{FETGV}^2_{(\alpha_0, \alpha_1)}$ to the graph-based approach $\mathrm{lapFETGV}^2_{(\alpha_0, \alpha_1)}$ of Gong et al., 2018, see (2.24), as well as the first-order total variation semi-norm (2.20).

We solve all problems using a split Bregman algorithm; see Goldstein, Osher, 2009. We detail its realization for an image denoising problem using the novel discrete formulation (3.1). The case of (2.24) is similar and (2.20) is simpler and well known, so we skip the details. We assume that f is a

given $\mathcal{DG}_0(\Omega)$ data function. We wish to solve the discrete TGV-L^2 denoising problem

$$\text{Minimize} \quad \frac{1}{2}\sum_T |T|(u-f)^2 + \alpha_1 \sum_E |E|\big|[\![u]\!] + h_E\,\boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+\big|$$

$$+ \alpha_0 \sum_T |T||\nabla\boldsymbol{w}|_F + \alpha_0 \sum_E \sum_{i=1}^2 \frac{|E|}{2}\big|[\![\boldsymbol{w}]\!](X_{E,i})\big|_2, \tag{4.1}$$

$$\text{where} \quad (u,\boldsymbol{w}) \in \mathcal{DG}_0(\Omega) \times \mathcal{RT}_0(\Omega).$$

In (4.1), we spelled out the norms and integrals in (3.1). As before, $|E|$ denotes the length of an interior edge E, and $|T|$ denotes the area of a triangle T. For simplicity, we denote the constant value of u on T simply by u, and similarly for f and $\nabla\boldsymbol{w}$. A similar convention applies to the quantities $[\![u]\!]$, $\boldsymbol{w}_+$, $\boldsymbol{\mu}_+$ and $[\![\boldsymbol{w}]\!]$ on edges E. The quantity $[\![\boldsymbol{w}]\!](X_{E,i})$ is the evaluation of the jump of $\boldsymbol{w}$ at the end points $X_{E,i}$ ($i = 1, 2$) of the edge E. It is apparent that (4.1) is not differentiable due to the presence of absolute values and norms.

The split Bregman method introduced in Goldstein, Osher, 2009 is an alternating direction method of multipliers (ADMM) tailored to problems involving absolute values and norms. To derive this method for problem (4.1), three new variables are introduced: the scalar-valued $d_0 \in \mathcal{DG}_0(\cup E)$ on the skeleton, the matrix-valued $D_1 \in \mathcal{DG}_0(\Omega;\mathbb{R}^{2\times2})$ in the domain Ω and the vector-valued $\boldsymbol{d}_2 \in \mathcal{DG}_1(\cup E;\mathbb{R}^2)$ again on the skeleton. Recall that $\cup E$ denotes the skeleton of the mesh, i. e., the union of interior edges. The variables are coupled to the original unknowns $u, \boldsymbol{w}$ through the constraints

$$d_0|_E = [\![u]\!] + h_E\,\boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+ \in P_0(E),$$
$$D_1|_T = \nabla\boldsymbol{w} \in P_0(T)^{2\times2},$$
$$\boldsymbol{d}_2|_E = [\![\boldsymbol{w}]\!] \in P_1(E)^2$$

for all interior edges E and triangles T, respectively. Since $\boldsymbol{d}_2$ and $[\![\boldsymbol{w}]\!]$ are linear on every edge, we only need to enforce the respective constraint in the two Lagrange nodes used to describe $\mathcal{DG}_1(E)$, which we take to be the end points. The augmented Lagrangian function for problem (4.1) with penalty parameters $\lambda_0, \lambda_1, \lambda_2 > 0$, respectively, reads

$$\text{Minimize} \quad \frac{1}{2}\sum_T |T|(u-f)^2 + \alpha_1 \sum_E |E||d_0| + \alpha_0 \sum_T |T||D_1|_F + \alpha_0 \sum_E \sum_{i=1}^2 \frac{|E|}{2}\big|\boldsymbol{d}_2(X_{E,i})\big|_2$$

$$+ \frac{\lambda_0}{2}\sum_E |E|\big(d_0 - ([\![u]\!] + h_E\,\boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+) - b_0\big)^2 + \frac{\lambda_1}{2}\sum_T |T|\big|D_1 - \nabla\boldsymbol{w} - B_1\big|_F^2$$

$$+ \frac{\lambda_2}{2}\sum_E \sum_{i=1}^2 \frac{|E|}{2}\big|\boldsymbol{d}_2(X_{E,i}) - [\![\boldsymbol{w}]\!](X_{E,i}) - \boldsymbol{b}_2(X_{E,i})\big|_2^2, \tag{4.2}$$

$$\text{where} \quad (u,\boldsymbol{w}) \in \mathcal{DG}_0(\Omega) \times \mathcal{RT}_0(\Omega)$$
$$\text{and} \quad (d_0, D_1, \boldsymbol{d}_2) \in \mathcal{DG}_0(\cup E) \times \mathcal{DG}_0(\Omega;\mathbb{R}^{2\times2}) \times \mathcal{DG}_1(\cup E;\mathbb{R}^2).$$

The scaled Lagrange multiplier estimates b_0, B_1 and $\boldsymbol{b}_2$ are taken from the same spaces as d_0, D_1 and $\boldsymbol{d}_2$, respectively.

The split Bregman method proceeds by minimizing over $(u,\boldsymbol{w})$, then over $(d_0, D_1, \boldsymbol{d}_2)$, and by updating

$(b_0, B_1, \boldsymbol{b}_2)$, in cyclic repetitions. The minimization w.r.t. $(u, \boldsymbol{w})$, i.e.,

$$
\begin{aligned}
\text{Minimize} \quad & \frac{1}{2} \sum_T |T|(u-f)^2 + \frac{\lambda_0}{2} \sum_E |E| \big(d_0 - (\llbracket u \rrbracket + h_E \, \boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+) - b_0\big)^2 \\
& + \frac{\lambda_1}{2} \sum_T |T| \big|D_1 - \nabla \boldsymbol{w} - B_1\big|_F^2 \\
& + \frac{\lambda_2}{2} \sum_E \sum_{i=1}^2 \frac{|E|}{2} \big|\boldsymbol{d}_2(X_{E,i}) - \llbracket \boldsymbol{w} \rrbracket(X_{E,i}) - \boldsymbol{b}_2(X_{E,i})\big|_2^2, \\
\text{where} \quad & (u, \boldsymbol{w}) \in \mathcal{DG}_0(\Omega) \times \mathcal{RT}_0(\Omega),
\end{aligned}
\tag{4.3}
$$

is a uniformly convex quadratic problem and thus its solution is given by a linear system, governed by a sparse matrix. Details are given below.

The minimization w.r.t. $(d_0, D_1, \boldsymbol{d}_2)$ fully decouples into the following subproblems formulated in terms of their coefficients (denoted by d_0, D_1, $\boldsymbol{d}_{2,1}$ and $\boldsymbol{d}_{2,2}$) on each interior edge or triangle.

$$
\begin{aligned}
\text{Minimize} \quad & \alpha_1 |d_0| + \frac{\lambda_0}{2} \big(d_0 - (\llbracket u \rrbracket + h_E \, \boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+) - b_0\big)^2 \quad \text{w.r.t. } d_0 \in \mathbb{R}, \\
\text{Minimize} \quad & \alpha_0 |D_1|_F + \frac{\lambda_1}{2} \big|D_1 - \nabla \boldsymbol{w} - B_1\big|_F^2 \quad \text{w.r.t. } D_1 \in \mathbb{R}^{2\times 2}, \\
\text{Minimize} \quad & \alpha_0 \big|\boldsymbol{d}_{2,i}\big|_2 + \frac{\lambda_2}{2} \big|\boldsymbol{d}_{2,i} - \llbracket \boldsymbol{w} \rrbracket(X_{E,i}) - \boldsymbol{b}_{2,i}\big|_2^2 \quad \text{w.r.t. } \boldsymbol{d}_{2,i} \in \mathbb{R}^2, \ i = 1, 2.
\end{aligned}
$$

Notice that $\boldsymbol{d}_{2,i}$, $i = 1, 2$ denotes the two $\mathbb{R}^2$-valued degrees of freedom of $\boldsymbol{d}_2$, at the endpoints of the interior edge E. It is well known that each of the above problems has a unique, closed-form solution, given by a soft shrinkage operation. Specifically, the optimal $d_0, D_1, \boldsymbol{d}_2$ on each edge, triangle and vertex are given by

$$
\text{shrink}(x, \ \delta_j) := \begin{cases} \frac{x}{|x|_*} \max\{|x|_* - \delta_j, \ 0\} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}
$$

where $\delta_j \in \mathbb{R}$ is some data and $|x|_*$ stands either for the absolute value in case of $j = 0$, the Frobenius norm in $\mathbb{R}^{2\times 2}$ in case of $j = 1$, or the Euclidean norm in $\mathbb{R}^2$ in case of $j = 2$; see Algorithm 1 for full details.

The specification of the complete split Bregman algorithm is given as Algorithm 1. To avoid a cluttered notation, we do not endow the variables with an iteration index. The convergence of a class of methods for convex functions with linear constraints comprising Algorithm 1 is studied in Boyd et al., 2010. To measure the current state of convergence one often considers the primal residual, i.e., the error in the constraint, and the dual residual, which is the gradient of the non-augmented Lagrangian with respect to the $(u, \boldsymbol{w})$ variable. We stop when both the primal and dual residual norms fall below a certain threshold.

With suitable substitutions, Algorithm 1 can also solve variants of (4.1) with the regularization term replaced by $\text{lapFETGV}^2_{(\alpha_0, \alpha_1)}$ (2.24) or the first-order total variation function $\alpha_1 \text{TV}$ (2.20). Moreover, it can be adapted to the pixel-based discretization (2.13) of the continuous second-order TGV semi-norm (2.4) from Bredies, Kunisch, Pock, 2010; compare He et al., 2014.

Algorithm 1 Split Bregman method for the solution of (4.1).

Input: noisy image data $f \in \mathcal{DG}_0(\Omega)$
Output: approximate solution of (4.1)
1: Set $k := 0$
2: Set $(d_0, D_1, \boldsymbol{d}_2) := (0, 0, 0) \in \mathcal{DG}_0(\cup E) \times \mathcal{DG}_0(\Omega; \mathbb{R}^{2\times 2}) \times \mathcal{DG}_1(\cup E; \mathbb{R}^2)$
3: Set $(b_0, B_1, \boldsymbol{b}_2) := (d_0, D_1, \boldsymbol{d}_2)$
4: **while** not converged **do**
5:     Set $(u, \boldsymbol{w})$ to the solution of (4.3)
6:     Set $d_0 := \text{shrink}\left(\llbracket u \rrbracket + h_E\, \boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+ + b_0,\ \frac{\alpha_1}{\lambda_0}\right)$ on each interior edge E
7:     Set $D_1 := \text{shrink}\left(\nabla \boldsymbol{w} + B_1,\ \frac{\alpha_0}{\lambda_1}\right)$ on each triangle T
8:     Set $\boldsymbol{d}_{2,i} := \text{shrink}\left(\llbracket \boldsymbol{w} \rrbracket(X_{E,i}) + \boldsymbol{b}_{2,i},\ \frac{\alpha_0}{\lambda_2}\right)$ on each interior edge E for $i = 1, 2$
9:     Set $b_0 := b_0 + \llbracket u \rrbracket + h_E\, \boldsymbol{w}_+ \cdot \boldsymbol{\mu}_+ - d_0$ on each interior edge E
10:    Set $B_1 := B_1 + \nabla \boldsymbol{w} - D_1$ on each triangle T
11:    Set $\boldsymbol{b}_{2,i} := \boldsymbol{b}_{2,i} + \llbracket \boldsymbol{w} \rrbracket(X_{E,i}) - \boldsymbol{d}_{2,i}$ on each interior edge E for $i = 1, 2$
12:    $k := k + 1$
13: **end while**

An implementation of Algorithm 1 was carried out in the finite element framework FEniCS (version 2019.2), see Logg, Mardal, Wells, 2012. This framework offers a considerable variety of finite elements including $\mathcal{DG}$, $\mathcal{CG}$ and $\mathcal{RT}$ on simplicial meshes, and in particular on triangular meshes in 2D. Notice, however, that the convention for the lowest-order Raviart-Thomas space in FEniCS is $\mathcal{RT}_1$, not $\mathcal{RT}_0$. It also features the automatic generation of code from variational forms based on a programming paradigm involving the unified form language UFL Alnæs, 2012.

Subproblem (4.3) is solved in the following way. We formulate its objective using the UFL and differentiate it to obtain a sparse linear system. We then employ a sparse direct solver to store the LU factorization of the system matrix, which is constant throughout the algorithm. Concerning the implementation of the soft shrinkage operations in steps 6 to 8 in Algorithm 1, which are formulated in terms of the coefficient vectors of finite element functions, we cast the latter into NumPy arrays. The implementation of the soft shrinkage functions is then straightforward. Likewise, the Lagrange multiplier updates in steps 9 to 11 are implemented in terms of coefficient vectors.

5 Numerical Results for Image Reconstruction Problems

In this section we present numerical results for image denoising and inpainting problems on triangulated meshes using the novel discrete formulation $\text{FETGV}^2_{(\alpha_0,\alpha_1)}$ from (3.1) of second-order TGV. We also compare it to the graph-based approach $\text{lapFETGV}^2_{(\alpha_0,\alpha_1)}$ of Gong et al., 2018, see (2.24), as well as the first-order total variation semi-norm (2.20). As detailed in Section 4, we solve all three problems using a split Bregman algorithm.

To make a reasonable comparison between the above formulations, we need to choose the regularization parameters α_0 and α_1 for each approach independently and in an optimal way. We use the mean

---

structural similarity measure (MSSIM) introduced in Wang et al., 2004 for this purpose. In each case, we seek parameters α_0 and α_1 which maximize the MSSIM between the reconstruction u and the original image. These optimal parameters are determined using an interval bisection method.

5.1 MSSIM for Unstructured Data

The MSSIM combines similarity measurements in luminance (l), contrast (c) and structure (s) over various windows, i. e., subdomains, of a pair of images. Given two images u, v and a common window centered at a point x, one usually uses

$$l(x) := \frac{2\,\mu_u(x)\,\mu_v(x) + C_1}{\mu_u(x)^2 + \mu_v(x)^2 + C_1},$$

$$c(x) := \frac{2\,\sigma_u(x)\,\sigma_v(x) + C_2}{\sigma_u(x)^2 + \sigma_v(x)^2 + C_2},$$

$$s(x) := \frac{\sigma_{uv}(x) + C_3}{\sigma_u(x)\,\sigma_v(x) + C_3},$$

where C_1, C_2, C_3 are image independent constants, $\mu_u(x)$ and $\mu_v(x)$ denote the respective averages, $\sigma_u(x)^2$ and $\sigma_v(x)^2$ denote the respective variances and $\sigma_{uv}(x)$ denotes the covariance of (u, v) over the window around the point x. For details on their computation on pixels we refer the reader to Wang et al., 2004. The structural similarity measure at x is then given by

$$\text{SSIM}(x) := l(x)^\alpha\, c(x)^\beta\, s(x)^\gamma. \tag{5.1}$$

Usually, $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$ are chosen, resulting in

$$\text{SSIM}(x) = \frac{(2\,\mu_u(x)\,\mu_v(x) + C_1)(2\,\sigma_{uv}(x) + C_2)}{(\mu_u(x)^2 + \mu_v(x)^2 + C_1)(\sigma_u(x)^2 + \sigma_v(x)^2 + C_2)}. \tag{5.2}$$

The mean SSIM, or MSSIM, is then computed by averaging (5.2) over the whole domain. When the image has values in $[0, 1]$, then often $C_1 = 0.01$ and $C_2 = 0.03$ are chosen, see Nilsson, Akenine-Möller, 2020. Also, it appears to be common to choose windows of size 11×11.

Obviously, this procedure requires some modifications when we work with unstructured images defined over meshes instead of rectangular grids. For unstructured meshes, we propose to use windows for the evaluation of averages and variances which include all cells within a certain distance of the vertex representing the triangle T of interest in the dual graph of the mesh. For instance, a radius of 2 would mean to include all neighbors and neighbors of neighbors of T. We compute the average, variance and covariance by means of integrals over the ensuing subdomain Ω' with total area $|\Omega'|$ consisting of triangles $T \in \Omega'$ as follows:

$$\mu_u(T) = \frac{1}{|\Omega'|} \int_{\Omega'} u(x)\,\mathrm{d}x = \frac{\sum_{T \in \Omega'} |T|\, u}{\sum_{T \in \Omega'} |T|},$$

$$\sigma_u(T) = \frac{1}{|\Omega'|} \int_{\Omega'} (u(x) - \mu_u(T))^2\,\mathrm{d}x = \frac{\sum_{T \in \Omega'} |T|(u - \mu_u(T))^2}{\sum_{T \in \Omega'} |T|},$$

$$\sigma_{uv}(T) = \frac{1}{|\Omega'|} \int_{\Omega'} (u(x) - \mu_u(T))(v(x) - \mu_v(T))\,\mathrm{d}x = \frac{\sum_{T \in \Omega'} |T|(u - \mu_u(T))(v - \mu_v(T))}{\sum_{T \in \Omega'} |T|}.$$

As before in (4.1), we denote the values of u and v on a triangle T in any of the sums simply by u and v, respectively. The average and variance of the second image v are computed in the same way as $\mu_u(T)$ and $\sigma_u(T)$. The SSIM at T is then evaluated as in (5.2) and MSSIM follows by averaging SSIM(T) over all triangles T, weighted by the area $|T|$.

Notice that, as in the case of pixel images, the MSSIM measure depends on the window sizes and shape. Therefore, the MSSIM scores of different regularizers are only comparable when the same windows are used.

5.2 Results for Image Denoising

We show three test cases for the TGV-L^2 denoising problem in this section. We begin with a grayscale image on a classical pixel grid, then consider a grayscale image on a triangulated mesh, and finally a real-world test case on an unstructured surface mesh obtained from 3D scanning a real object. In all test cases, normally distributed noise with standard deviation of 5% of the data range is added to the original image data.

5.2.1 Case 1: Grayscale Gradient Image on a Pixel Grid

We employ the often-used grayscale gradient image with an inverted area in the middle; see e. g. Setzer, Steidl, Teuber, 2011 and Figure 5.1. This image has size 128×128 pixels and it features both sharp edges as well as linear transitions. Due to the regular pixel structure, we can also include the finite difference discretization $\text{TGV}^h_{(\alpha_0,\alpha_1)}$ of Bredies, Kunisch, Pock, 2010, see (2.13), into the comparison. For the purpose of our formulation (3.1) and $\text{lapFETGV}^2_{(\alpha_0,\alpha_1)}$ (2.24), we subdivide each pixel into two triangles. For comparison of first-order TV and second-order TGV, we also show results using the first-order total variation functional α_1 TV on triangular meshes, see (2.20).

The results are shown in Figure 5.1. They were obtained using Algorithm 1 with penalty parameters $\lambda_0 = \lambda_1 = \lambda_2 = 10$. The MSSIM quantities for each image were evaluated using a window size of 11×11 square pixels.

As expected, all discrete second-order TGV models outperform the first-order TV regularization in Figure 5.1c. Moreover, the results obtained with the novel formulation (3.1) closely resemble those with (2.13) from Bredies, Kunisch, Pock, 2010, which can be considered the reference implementation for square pixel images. Notice that Figure 5.1f is obtained on a triangular discretization with twice as many degrees of freedom compared to Figure 5.1d. In addition, we observe artifacts in the $\text{lapFETGV}^2_{(\alpha_0,\alpha_1)}$ method (2.24). In fact, similar artifacts can already be found in Gong et al., 2018, Fig. 3 and its implementation for square pixels from Brinkmann, Burger, Grah, 2018, Fig. 2.
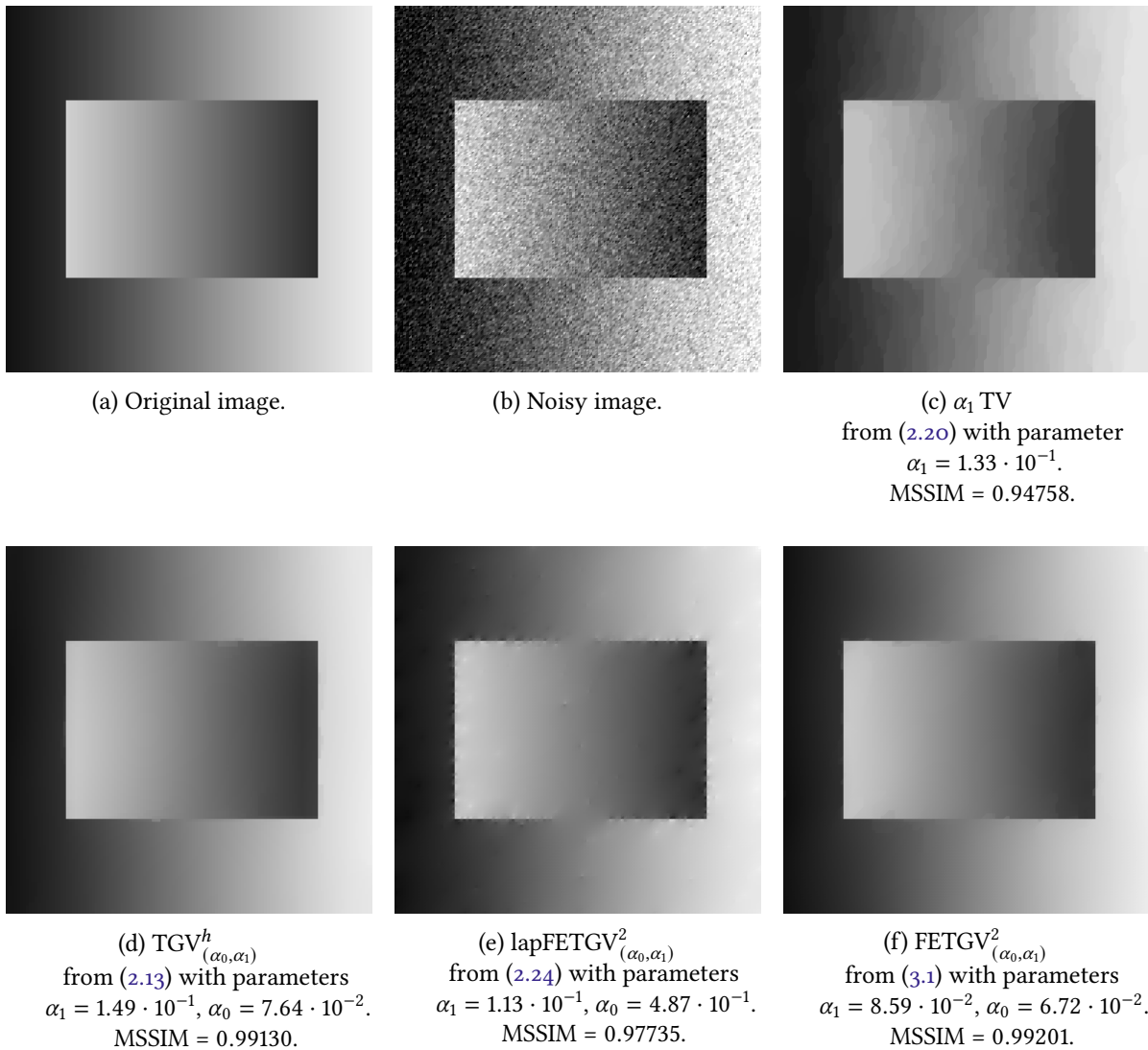
(a) Original image.

(b) Noisy image.

(c) α_1 TV
from (2.20) with parameter
$\alpha_1 = 1.33 \cdot 10^{-1}$.
MSSIM = 0.94758.

(d) $\mathrm{TGV}^h_{(\alpha_0, \alpha_1)}$
from (2.13) with parameters
$\alpha_1 = 1.49 \cdot 10^{-1}$, $\alpha_0 = 7.64 \cdot 10^{-2}$.
MSSIM = 0.99130.

(e) $\mathrm{lapFETGV}^2_{(\alpha_0, \alpha_1)}$
from (2.24) with parameters
$\alpha_1 = 1.13 \cdot 10^{-1}$, $\alpha_0 = 4.87 \cdot 10^{-1}$.
MSSIM = 0.97735.

(f) $\mathrm{FETGV}^2_{(\alpha_0, \alpha_1)}$
from (3.1) with parameters
$\alpha_1 = 8.59 \cdot 10^{-2}$, $\alpha_0 = 6.72 \cdot 10^{-2}$.
MSSIM = 0.99201.

Figure 5.1: Denoising results for the grayscale gradient image (case 1).

5.2.2 Case 2: Grayscale Gradient Image on an Unstructured Grid

Test case 2 is set on an unstructured mesh and it is taken from Herrmann, Herzog, Schmidt, et al., 2019. The results are shown in Figure 5.2. They were obtained using Algorithm 1 with penalty parameters $\lambda_0 = \lambda_1 = \lambda_2 = 0.1$. The MSSIM for each image was evaluated using a window including each triangle's neighbors up to a distance of 10. As for case 1, we observe artifacts in the lapFETGV$^2_{(\alpha_0,\alpha_1)}$ method (2.24) but not in our approach (3.1).

5.2.3 Case 3: Image Denoising on an Unstructured Surface Grid

As we mentioned in the introduction, our approach (3.1) can be generalized to images on surface meshes embedded in 3D in a straightforward way. For this purpose, the following two minor modifications are necessary. First, the scalar factor h_E, which is defined in the planar case according to (3.2) as the distance between the circumcenters $\boldsymbol{m}_\pm$ of the triangles adjacent to an edge E, has to be computed as the surface-intrinsic distance. This is achieved using the midpoint $\boldsymbol{m}_E$ of the edge E and the formula

$$h_E := \boldsymbol{\mu}_+ \cdot (\boldsymbol{m}_E - \boldsymbol{m}_+) + \boldsymbol{\mu}_- \cdot (\boldsymbol{m}_E - \boldsymbol{m}_-).$$

Furthermore, the jump of $\boldsymbol{w}$ also has to be computed intrinsically. This is done by replacing $[\![\boldsymbol{w}]\!]$ by $\boldsymbol{t}\,(\boldsymbol{t} \cdot [\![\boldsymbol{w}]\!])$ in the last term of (3.1), where $\boldsymbol{t}$ is a unit vector tangent to the edge with arbitrary orientation. The graph-based approach lapFETGV$^2_{(\alpha_0,\alpha_1)}$ (2.24), as well as the first-order total variation semi-norm (2.20) can be used for surface images as well with similar modifications.

We are choosing the flower test case from Artec Europe[1]. To denoise the color image, which is given by $\mathcal{DG}_0$ data on a mesh of 389 632 triangles, we are applying the denoising on each RGB channel using the same regularization parameters. For this purpose we are using the above described modifications in Algorithm 1 with penalty parameters $\lambda_0 = \lambda_1 = \lambda_2 = 1$. The MSSIM is computed with a window size of 20 neighbors. Results are shown in Figure 5.3.

Again both versions of TGV prevent most of the staircasing effect. In addition to the appearance of artifacts, the lapFETGV$^2_{(\alpha_0,\alpha_1)}$ method (2.24) also suffers from smoothing of jumps for the optimal choice of regularization parameters α_1 and α_0. To some extent, this effect can also be seen in the other test cases, however only in regions where the jumps are very small, see Figure 5.1e (case 1) and Figure 5.2d (case 2). The smoothing effect through the particular choice of parameters made in the present test case 3 does not only smooth out jumps but also leads to a reduction of artifact in areas without any jumps. Since these areas are very large compared to the regions that are affected from the smoothing of jumps, this choice of parameters yields a better MSSIM score.

5.3 Case 4: Results for Joint Image Inpainting and Denoising

Finally we present an example of joint image denoising and inpainting (case 4). This is achieved through a modification of the first term in (4.1) so that the sum extends only over a subset of triangles

---

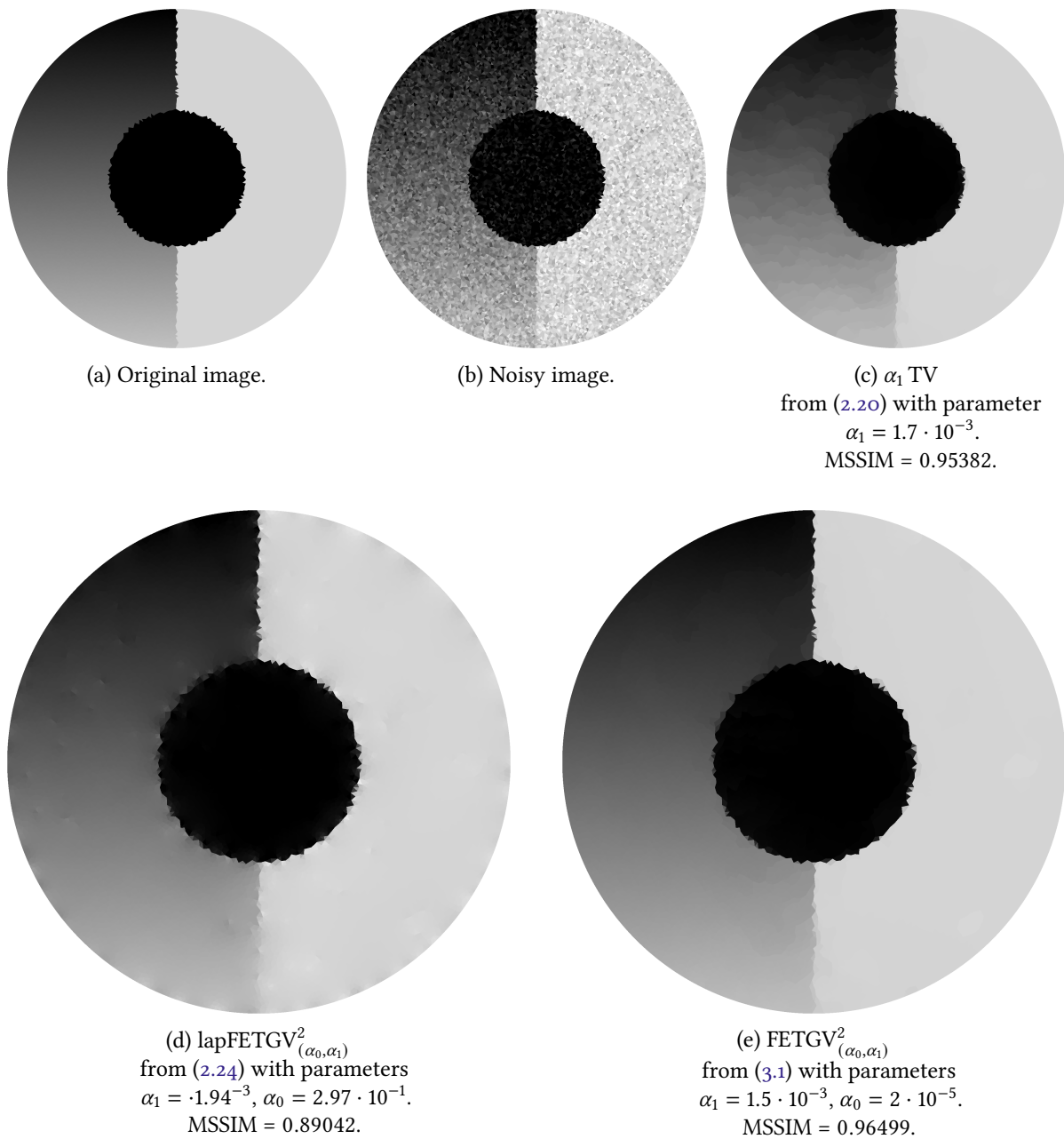[1] available under a Creative Commons 3.0 license from https://www.artec3d.com/de/3d-models/flower

(a) Original image.

(b) Noisy image.

(c) α_1 TV
from (2.20) with parameter
$\alpha_1 = 1.7 \cdot 10^{-3}$.
MSSIM = 0.95382.

(d) lapFETGV$^2_{(\alpha_0,\alpha_1)}$
from (2.24) with parameters
$\alpha_1 = \cdot 1.94^{-3}$, $\alpha_0 = 2.97 \cdot 10^{-1}$.
MSSIM = 0.89042.

(e) FETGV$^2_{(\alpha_0,\alpha_1)}$
from (3.1) with parameters
$\alpha_1 = 1.5 \cdot 10^{-3}$, $\alpha_0 = 2 \cdot 10^{-5}$.
MSSIM = 0.96499.

Figure 5.2: Denoising results for a grayscale gradient image on an unstructured mesh (case 2).
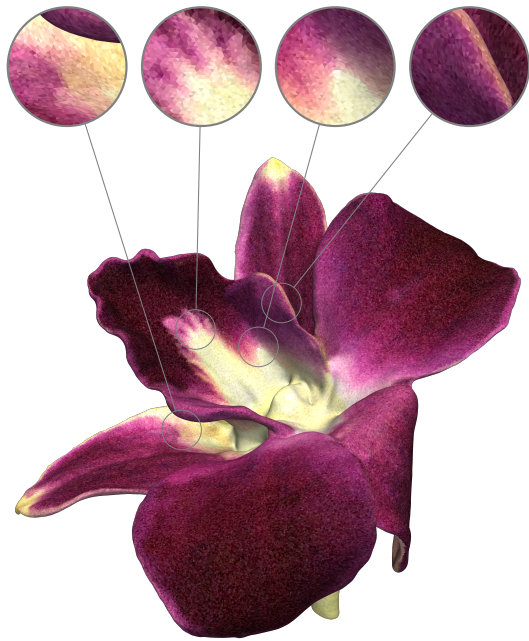
(a) Noisy surface image.

(b) α_1 TV
from (2.20) with parameter
$\alpha_1 = 5.3 \cdot 10^{-3}$.
MSSIM = 0.94321.
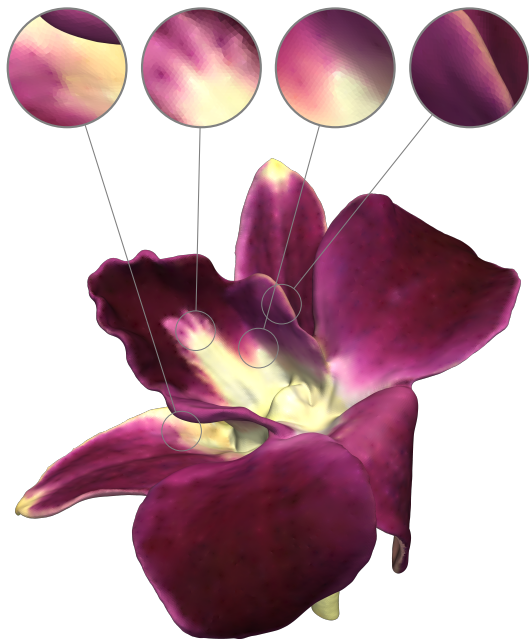
(c) lapFETGV$^2_{(\alpha_0,\alpha_1)}$
from (2.24) with parameters
$\alpha_1 = 8.1 \cdot 10^{-3}$, $\alpha_0 = 5.1 \cdot 10^{-2}$.
MSSIM = 0.94123.
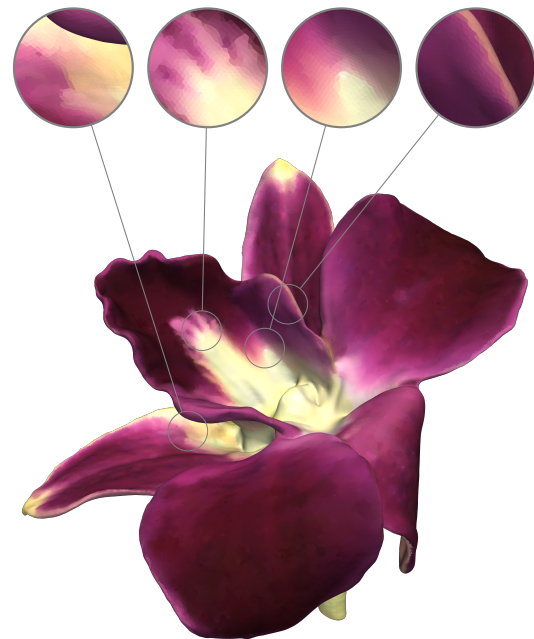
(d) FETGV$^2_{(\alpha_0,\alpha_1)}$
from (3.1) with parameters
$\alpha_1 = 5 \cdot 10^{-3}$, $\alpha_0 = 1 \cdot 10^{-3}$.
MSSIM = 0.94546.

Figure 5.3: Denoising results for a color image on an unstructured surface mesh embedded in 3D (case 3).

where data is available. Our test case is the well-known peppers image taken from the USC-SIPI Image Database[2]. As before, the pixel image is discretized on a triangulated mesh by letting two triangles form a pixel. A text overlay serves as the region representing missing data. The results are shown in Figure 5.4. They were obtained using Algorithm 1 with penalty parameters $\lambda_0 = \lambda_1 = \lambda_2 = 10$. The MSSIM for each image was evaluated using a window including each triangle's neighbors up to a distance of 10.

Due to the kernel of the proposed discretization (3.1) consisting precisely of interpolations of linear functions, we see a linear fill-in in regions with missing data. As before, the reconstruction using lapFETGV$^2_{(\alpha_0,\alpha_1)}$ from (2.24) shows a lot of artifacts due to the (piecewise constant) harmonic functions in the kernel of this regularizer, see Section 2.2.2. The artifacts occur in particularly in the fill-in regions since no other form of tracking is available there. These additional oscillations impact the MSSIM score more than the staircasing effect (clearly visible for the first-order TV with results shown in Figure 5.4b) and thus make lapFETGV$^2_{(\alpha_0,\alpha_1)}$ perform worse than the first-order total-variation regularizer for this test case.

6 Conclusion

In this paper we have introduced a novel discrete formulation (3.1) for the second-order total generalized variation (1.3) for piecewise constant data on triangulated meshes. Our formulation extends (2.13) introduced in Bredies, Kunisch, Pock, 2010 from pixel grids to arbitrary triangular meshes, which can be planar or embedded in 3D. A particular feature is the use of the standard lowest-order Raviart-Thomas finite element space for the discretization of the auxiliary variable $\boldsymbol{w}$.

We have shown that the kernel of (3.1) consists precisely of interpolants of linear functions. Compared to the alternative formulation (2.24) from Gong et al., 2018, our proposal is thus closer to the continuous formulation (1.3) and does not exhibit the artifacts observed using (2.24). We also gave an implementation of the split Bregman method suitable for solving image denoising and inpainting problems using (3.1) as regularizer.

References

Alkämper, M.; A. Langer (2017). "Using DUNE-ACFem for non-smooth minimization of bounded variation functions". *Archive of Numerical Software* 5.1, pp. 3–19. DOI: 10.11588/ans.2017.1.27475.

Alnæs, M. S. (2012). "UFL: a finite element form language". *Automated Solution of Differential Equations by the Finite Element Method*. Ed. by A. Logg; K.-A. Mardal; G. N. Wells. Vol. 84. Lecture Notes in Computational Science and Engineering. Springer. Chap. 17. DOI: 10.1007/978-3-642-23099-8_17.

Ambrosio, L.; N. Fusco; D. Pallara (2000). *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York.

Arnold, D. N.; R. S. Falk; R. Winther (2006). "Finite element exterior calculus, homological techniques, and applications". *Acta Numerica* 15, pp. 1–155. DOI: 10.1017/s0962492906210018.

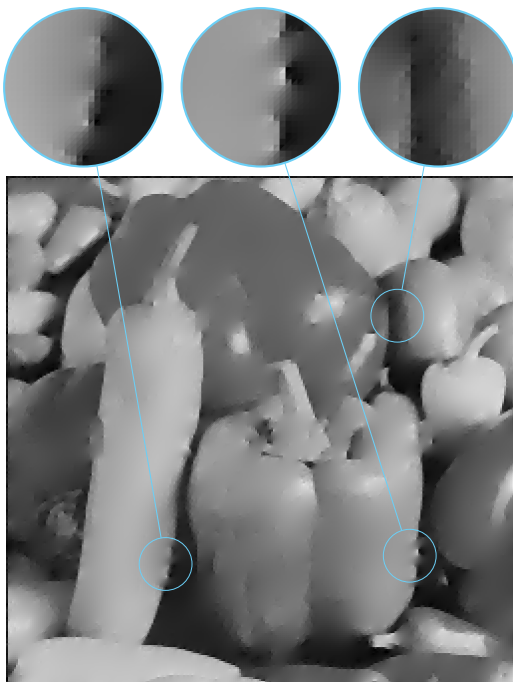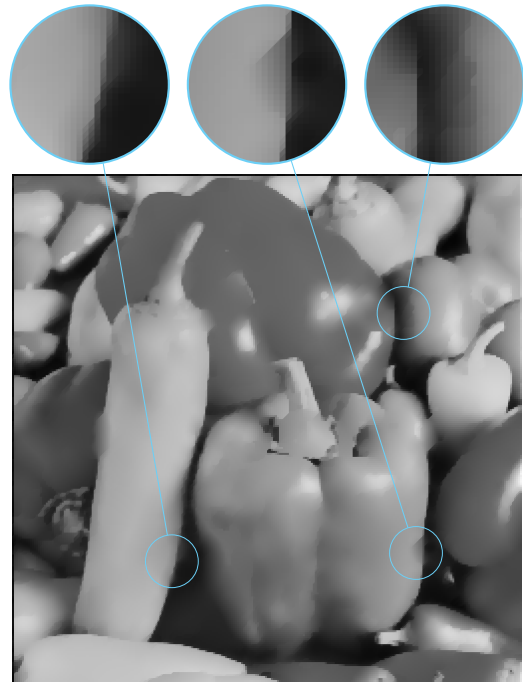[2]https://sipi.usc.edu/database/database.php?volume=misc&image=13#top
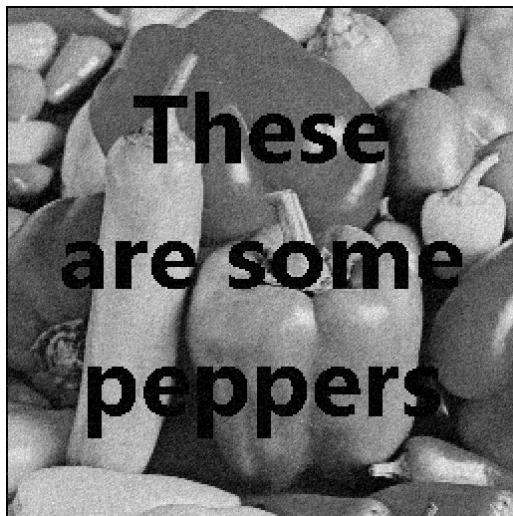
(a) Noisy image with data loss.
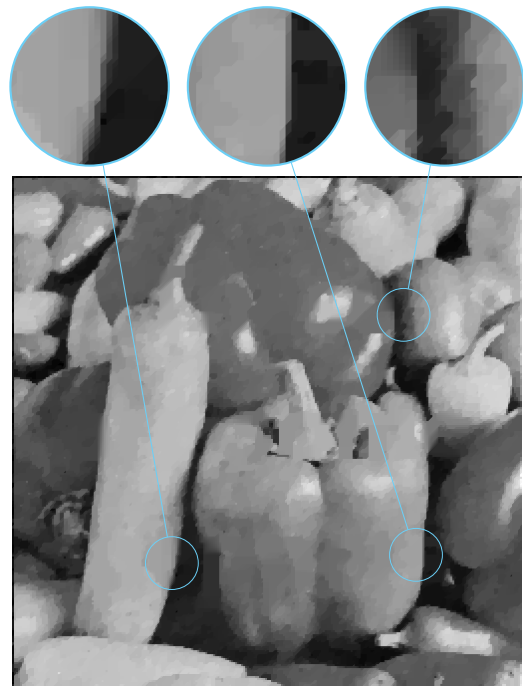
(b) α_1 TV
from (2.20) with parameter
$\alpha_1 = 2.2 \cdot 10^{-2}$.
MSSIM = 0.90192.

(c) lapFETGV$^2_{(\alpha_0, \alpha_1)}$
from (2.24) with parameters
$\alpha_1 = 9.1 \cdot 10^{-2}$, $\alpha_0 = 1.1 \cdot 10^{-0}$.
MSSIM = 0.89070.

(d) FETGV$^2_{(\alpha_0, \alpha_1)}$
from (3.1) with parameters
$\alpha_1 = 5.1 \cdot 10^{-2}$, $\alpha_0 = 4.5 \cdot 10^{-2}$.
MSSIM = 0.92697.

Figure 5.4: Combined denoising and inpainting results for a test image on a structured mesh (case 4).

Attouch, H.; G. Buttazzo; G. Michaille (2014). *Variational Analysis in Sobolev and BV Spaces*. 2nd ed. Vol. 17. MOS-SIAM Series on Optimization. Applications to PDEs and optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA. DOI: 10.1137/1.9781611973488.

Bachmayr, M.; M. Burger (2009). "Iterative total variation schemes for nonlinear inverse problems". *Inverse Problems* 25.10, pp. 105004, 26. DOI: 10.1088/0266-5611/25/10/105004.

Bartels, S. (2012). "Total variation minimization with finite elements: convergence and iterative solution". *SIAM Journal on Numerical Analysis* 50.3, pp. 1162–1180. DOI: 10.1137/11083277X.

Bartels, S. (2021). "Nonconforming discretizations of convex minimization problems and precise relations to mixed methods". *Computers & Mathematics with Applications* 93, pp. 214–229. DOI: 10.1016/j.camwa.2021.04.014. arXiv: 2002.02359.

Bartels, S.; R. H. Nochetto; A. J. Salgado (2015). "A total variation diminishing interpolation operator and applications". *Mathematics of Computation* 84.296, pp. 2569–2587. DOI: 10.1090/mcom/2942.

Bergmann, R.; J. H. Fitschen; J. Persch; G. Steidl (2017). *Priors with coupled first and second order differences for manifold-valued image processing*. arXiv: 1709.01343.

Bergmann, R.; M. Herrmann; R. Herzog; S. Schmidt; J. Vidal-Núñez (2020). "Discrete total variation of the normal vector field as shape prior with applications in geometric inverse problems". *Inverse Problems* 36.5, p. 054003. DOI: 10.1088/1361-6420/ab6d5c. arXiv: 1908.07916.

Berkels, B.; A. Effland; M. Rumpf (2017). "A posteriori error control for the binary Mumford-Shah model". *Mathematics of Computation* 86.306, pp. 1769–1791. DOI: 10.1090/mcom/3138.

Boyd, S.; N. Parikh; E. Chu; B. Peleato; J. Eckstein (2010). "Distributed optimization and statistical learning via the alternating direction method of multipliers". *Foundations and Trends in Machine Learning* 3.1, pp. 1–122. DOI: 10.1561/2200000016.

Bredies, K.; M. Holler (2014). "Regularization of linear inverse problems with total generalized variation". *Journal of Inverse and Ill-posed Problems* 22.6, pp. 871–913. DOI: 10.1515/jip-2013-0068.

Bredies, K.; K. Kunisch; T. Pock (2010). "Total generalized variation". *SIAM Journal on Imaging Sciences* 3.3, pp. 492–526. DOI: 10.1137/090769521.

Brinkmann, E.-M.; M. Burger; J. S. Grah (2018). "Unified models for second-order TV-type regularisation in imaging: a new perspective based on vector operators". *Journal of Mathematical Imaging and Vision* 61.5, pp. 571–601. DOI: 10.1007/s10851-018-0861-6.

Chambolle, A.; P.-L. Lions (1997). "Image recovery via total variation minimization and related problems". *Numerische Mathematik* 76.2, pp. 167–188. DOI: 10.1007/s002110050258.

Chambolle, A.; T. Pock (2020). "Crouzeix–Raviart approximation of the total variation on simplicial meshes". *Journal of Mathematical Imaging and Vision* 62.6-7, pp. 872–899. DOI: 10.1007/s10851-019-00939-3. HAL: hal-01787012.

Chambolle, A.; T. Pock (2021). "Approximating the total variation with finite differences or finite elements". *Geometric Partial Differential Equations - Part II*. Elsevier, pp. 383–417. DOI: 10.1016/bs.hna.2020.10.005. HAL: hal-02959358.

Chan, T.; S. Osher; J. Shen (2001). "The digital TV filter and nonlinear denoising". *IEEE Transactions on Image Processing* 10.2, pp. 231–241. DOI: 10.1109/83.902288.

Chan, T. F.; S. Esedoglu; F. Park (2010). "A fourth order dual method for staircase reduction in texture extraction and image restoration problems". *2010 IEEE International Conference on Image Processing*. IEEE. DOI: 10.1109/icip.2010.5653199.

Chan, T. F.; X.-C. Tai (2004). "Level set and total variation regularization for elliptic inverse problems with discontinuous coefficients". *Journal of Computational Physics* 193.1, pp. 40–66. DOI: 10.1016/j.jcp.2003.08.003.

Christiansen, S. H.; J. Hu; K. Hu (2018). "Nodal finite element de Rham complexes". *Numerische Mathematik* 139.2, pp. 411–446. DOI: `10.1007/s00211-017-0939-x`.

Clason, C.; F. Kruse; K. Kunisch (2018). "Total variation regularization of multi-material topology optimization". *ESAIM. Control, Optimisation and Calculus of Variations* 52.1, pp. 275–303. DOI: `10.1051/m2an/2017061`.

Elliott, C. M.; S. A. Smitheman (2009). "Numerical analysis of the TV regularization and H^{-1} fidelity model for decomposing an image into cartoon plus texture". *IMA Journal of Numerical Analysis* 29.3, pp. 651–689. DOI: `10.1093/imanum/drn025`.

Feng, X.; A. Prohl (2003). "Analysis of total variation flow and its finite element approximations". *M2AN. Mathematical Modelling and Numerical Analysis* 37.3, pp. 533–556. DOI: `10.1051/m2an:2003041`.

Goldstein, T.; S. Osher (2009). "The split Bregman method for $L1$-regularized problems". *SIAM Journal on Imaging Sciences* 2.2, pp. 323–343. DOI: `10.1137/080725891`.

Gong, B.; B. Schullcke; S. Krueger-Ziolek; F. Zhang; U. Mueller-Lisse; K. Moeller (2018). "Higher order total variation regularization for EIT reconstruction". *Medical & Biological Engineering & Computing* 56.8, pp. 1367–1378. DOI: `10.1007/s11517-017-1782-z`.

He, C.; C. Hu; X. Yang; H. He; Q. Zhang (2014). "An adaptive total generalized variation model with augmented Lagrangian method for image denoising". *Mathematical Problems in Engineering* 2014, pp. 1–11. DOI: `10.1155/2014/157893`.

Herrmann, M.; R. Herzog; S. Schmidt; J. Vidal-Núñez; G. Wachsmuth (2019). "Discrete total variation with finite elements and applications to imaging". *Journal of Mathematical Imaging and Vision* 61.4, pp. 411–431. DOI: `10.1007/s10851-018-0852-7`. arXiv: `1804.07477`.

Herrmann, M.; R. Herzog; H. Kröner; S. Schmidt; J. Vidal-Núñez (2018). "Analysis and an interior point approach for TV image reconstruction problems on smooth surfaces". *SIAM Journal on Imaging Sciences* 11.2, pp. 889–922. DOI: `10.1137/17M1128022`.

Jensen, J. R. (2015). *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th ed. Englewood Cliffs, N.J.: Prentice-Hall Inc.

John, J.; M. Wilscy (2016). "Image processing techniques for surface characterization of nanostructures". *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*. IEEE. DOI: `10.1109/iccpct.2016.7530317`.

Knoll, F.; K. Bredies; T. Pock; R. Stollberger (2010). "Second order total generalized variation (TGV) for MRI". *Magnetic Resonance in Medicine* 65.2, pp. 480–491. DOI: `10.1002/mrm.22595`.

Lee, C.-O.; E.-H. Park; J. Park (2019). "A finite element approach for the dual Rudin–Osher–Fatemi model and its nonoverlapping domain decomposition methods". *SIAM Journal on Scientific Computing* 41.2, B205–B228. DOI: `10.1137/18m1165499`. arXiv: `1805.02562`.

Li, J.; C. Bai; Z. Lin; J. Yu (2016). "Penrose high-dynamic-range imaging". *Journal of Electronic Imaging* 25.3, p. 033024. DOI: `10.1117/1.jei.25.3.033024`.

Logg, A.; K.-A. Mardal; G. N. Wells, eds. (2012). *Automated Solution of Differential Equations by the Finite Element Method*. Springer. DOI: `10.1007/978-3-642-23099-8`.

López Pérez, L. D. (2006). "Régularisation d'Images sur des Surfaces Non Planes". PhD thesis. Université de Nice-Sophia Antipolis. HAL: `tel-00141417`.

Nilsson, J.; T. Akenine-Möller (2020). *Understanding SSIM*. arXiv: `2006.13846`.

Ono, S.; I. Yamada; I. Kumazawa (2015). "Total generalized variation for graph signals". *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 5456–5460. DOI: `10.1109/icassp.2015.7179014`.

Papafitsoros, K.; C. B. Schönlieb (2014). "A combined first and second order variational approach for image reconstruction". *Journal of Mathematical Imaging and Vision* 48.2, pp. 308–338. DOI: `10.1007/s10851-013-0445-4`.

Rudin, L. I.; S. Osher; E. Fatemi (1992). "Nonlinear total variation based noise removal algorithms". *Physica D* 60.1–4, pp. 259–268. DOI: 10.1016/0167-2789(92)90242-F.

Setzer, S.; G. Steidl; T. Teuber (2011). "Infimal convolution regularizations with discrete ℓ_1-type functionals". *Communications in Mathematical Sciences* 9.3, pp. 797–827. DOI: 10.4310/CMS.2011.v9.n3.a7.

Stamm, B.; T. P. Wihler (2015). "A total variation discontinuous Galerkin approach for image restoration". *International Journal of Numerical Analysis and Modeling* 12.1, pp. 81–93.

Wang, Z.; A. Bovik; H. Sheikh; E. Simoncelli (2004). "Image quality assessment: from error visibility to structural similarity". *IEEE Transactions on Image Processing* 13.4, pp. 600–612. DOI: 10.1109/tip.2003.819861.

Wu, C.; J. Zhang; Y. Duan; X.-C. Tai (2012). "Augmented Lagrangian method for total variation based image restoration and segmentation over triangulated surfaces". *Journal of Scientific Computing* 50.1, pp. 145–166. DOI: 10.1007/s10915-011-9477-3.

Zhang, H.; C. Wu; J. Zhang; J. Deng (2015). "Variational mesh denoising using total variation and piecewise constant function space". *IEEE Transactions on Visualization and Computer Graphics* 21.7, pp. 873–886. DOI: 10.1109/TVCG.2015.2398432.