

SINGULAR SOLUTIONS, GRADED MESHES, AND ADAPTIVITY FOR TOTAL-VARIATION REGULARIZED MINIMIZATION PROBLEMS

SÖREN BARTELS^{1,*}, ROBERT TOVEY² AND FRIEDRICH WASSMER¹

Abstract. Recent quasi-optimal error estimates for the finite element approximation of total-variation regularized minimization problems require the existence of a Lipschitz continuous dual solution. We discuss the validity of this condition and devise numerical methods using locally refined meshes that lead to improved convergence rates despite the occurrence of discontinuities. It turns out that linear convergence is possible on suitably constructed meshes.

Mathematics Subject Classification. 49M29, 65N15, 65N50.

Received February 7, 2022. Accepted June 23, 2022.

1. INTRODUCTION

In this article we consider the finite element discretization of the Rudin–Osher–Fatemi (ROF) model from [30] which serves as a model problem for general convex and nonsmooth minimization problems. This widely used model in image processing determines a function $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ via a minimization of

$$I(u) = |Du|(\Omega) + \frac{\alpha}{2} \|g - u\|^2,$$

where $|Du|(\Omega)$ denotes the total variation of $u \in \text{BV}(\Omega) \cap L^2(\Omega)$, $g \in L^2(\Omega)$ is the input data, for example a noisy image, and $\|g - u\|$ is the L^2 distance between the given image and its regularization. The fidelity parameter $\alpha > 0$ is also given and determines the balance between denoising and preserving the input image. For more information on analytical features, explicit solutions in particular examples, and numerical methods concerning this model we refer the reader to [2–4, 13, 16, 17, 19, 21–23, 26–28, 32]. Since this model allows for and preserves discontinuities of the input function g , *cf.* [18], continuous finite element methods are known to perform suboptimally, *cf.* [9, 11]. Recent results in [8, 9, 20] show that optimal convergence rates $O(h^{1/2})$ for discontinuous solutions on quasi-uniform triangulations can be obtained by using discontinuous, low order Crouzeix–Raviart finite elements from [24] or appropriate discontinuous Galerkin methods. These error estimates bound the error for approximating minimizers for I by minimizing the discrete functional

$$I_h(u_h) = \int_{\pi} |\nabla_h u_h| \, dx + \frac{\alpha}{2} \|\Pi_h(g - u_h)\|^2$$

Keywords and phrases. Nonsmooth minimization, graded meshes, adaptivity, total variation, error estimates.

¹ Abteilung für Angewandte Mathematik, Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Str. 10, 79104 Freiburg i. Br., Germany.

² INRIA de Paris, 2 Rue Simone IFF, 75012 Paris, France.

*Corresponding author: bartels@mathematik.uni-freiburg.de

over piecewise affine functions $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ that are continuous at midpoints of element sides. Here ∇_h denotes the elementwise gradient and Π_h is the projection onto elementwise constant functions on the triangulation \mathcal{T}_h . Note that the functional I_h defines a nonconforming approximation of I , as, *e.g.*, jump terms of u_h across interelement sides are not included. The quasi-optimal rate applies if the dual problem, given by a maximization of

$$D(z) = -\frac{1}{2\alpha} \|\operatorname{div} z + \alpha g\|^2 + \frac{\alpha}{2} \|g\|^2 - I_{K_1(0)}(z)$$

over vector fields $z \in W_N^2(\operatorname{div}; \Omega)$ admits a Lipschitz continuous solution. The indicator functional $I_{K_1(0)}$ of the closed unit ball centered at the origin enforces the pointwise constraint $|z| \leq 1$. Although Lipschitz continuity is known to be true in some settings, following an idea from [31] we show that the condition is not satisfied in general and lower convergence rates have to be expected. Our simple counterexample uses the difference of two characteristic functions of balls $B_r^\pm = B_r(\pm r, 0)$ with radius $r > 0$ that touch at the origin, *i.e.*,

$$g = \chi_{B_r^+} - \chi_{B_r^-}.$$

Precise characterizations, *cf.* [18], of dual solutions along the jump set of the primal solution, given by

$$u = c_{r,\alpha} g, \quad c_{r,\alpha} = \max\{1 - d/(\alpha r), 0\},$$

imply that Lipschitz continuity of dual solutions fails at the origin if $\alpha r > d$. Surprisingly, this singularity does not appear to affect the convergence rates of approximations on sequences of uniform triangulations.

To obtain the rate $O(h)$ of the quasi-interpolation of a weakly differentiable function we investigate two numerical methods that construct locally refined meshes. The first approach uses the fact that the jump set J_u of the primal solution u is contained in the jump set J_g of the given function g , *i.e.*,

$$J_u \subset J_g$$

cf. [18]. Reduced convergence rates are related to the suboptimal approximation of jumps and therefore our idea is to refine triangulations in a neighborhood of the jump set J_g . Since we aim at preserving shape regularity of triangulations, the grading strength is limited and it turns out that the minimal mesh-size h_{\min} used at the discontinuity set cannot be smaller than h^β with $\beta \leq 2$ and the average mesh-size h . Our numerical analysis shows that a quadratic grading is, under suitable conditions on a piecewise regular solution, the correct refinement strength to obtain a linear convergence rate, *i.e.*, we have

$$\|u - u_h\| = O(h),$$

where u_h is the Crouzeix–Raviart finite element solution. In one-dimensional situations it coincides with the $P1$ approximation.

The approach of using graded meshes can only be efficiently applied if the given function g is piecewise regular and the jump set J_g is sufficiently simple. A more general concept uses refinement indicators based on *a posteriori* error estimates that bound the approximation error of an approximation u_h , *e.g.*, a continuous $P1$ approximation, by computable quantities. These depend on an approximation u_h of the discrete primal problem and an admissible vector field q for the continuous dual problem, *i.e.*, we have

$$\frac{\alpha}{2} \|u - u_h\|^2 \leq I(u_h) - I(u) \leq I(u_h) - D(q) = \eta_h^2(u_h, q),$$

which follows from coercivity properties of the functional I and the duality principle $I(u) \geq D(q)$ for every admissible vector field q . A simple calculation shows that if u_h is weakly differentiable and q satisfies $|q| \leq 1$ in Ω , the error estimator is given as a sum of local, nonnegative quantities, *i.e.*,

$$\eta_h^2(u_h, q) = \int_{\Omega} |\nabla u_h| - \nabla u_h \cdot q \, dx + \frac{1}{2\alpha} \int_{\Omega} (\operatorname{div} q - \alpha(u_h - g))^2 \, dx.$$

By using the partitioning of the domain Ω given by the triangulation \mathcal{T}_h we obtain refinement indicators $\eta_T^2(u_h, q)$ that are used to refine elements $T \in \mathcal{T}_h$. The reliable error estimator can only be efficient if the vector field q is a nearly optimal approximation of a dual solution z . To avoid the expensive solution of a discretization of the dual problem we use the observation that an approximation can be obtained *via* post-processing the Crouzeix–Raviart approximation, *cf.* [9]. In particular, this provides a maximizing vector field for the discrete dual functional

$$D_h(z_h) = -\frac{1}{2\alpha} \|\operatorname{div} z_h + \alpha \Pi_h g\|^2 + \frac{\alpha}{2} \|\Pi_h g\|^2 - I_{K_1(0)}(\Pi_h z_h),$$

defined on Raviart–Thomas vector fields $z_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$. For this definition we have the discrete duality principle $I_h(u_h) \geq D_h(z_h)$. However, the unit length constraint $|z| \leq 1$ is only imposed at midpoints so that the optimal z_h is in general inadmissible in the continuous dual functional D . Since only the midpoint values of z_h and its elementwise constant divergence enter the error estimator it nevertheless appears to be a reasonable way to define error indicators $(\eta_T)_{T \in \mathcal{T}_h}$ although the error bound may fail to hold; obvious corrections of z_h do not seem to lead to efficient error estimators. Our numerical experiments based on a related adaptive mesh refinement algorithm lead to improved experimental convergence rates that are lower than the ones obtained for graded meshes. Our explanation for this is that the graded meshes are optimal for the L^2 error while the coercivity estimate leading to the *a posteriori* error estimate controls a stronger error quantity.

This article is organized as follows. In Section 2 we introduce the used notation and define the relevant finite element spaces. The example with a non-Lipschitz dual solution for the ROF model is investigated in Section 3. In Section 4 the graded grid approaches are devised and analyzed. In Section 5 the primal-dual error estimator is defined and the construction of an optimal discrete dual vector field *via* discrete duality relations is shown. Numerical experiments are presented in Section 6.

2. NOTATION AND FINITE ELEMENT SPACES

Given a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ we use standard notation for Sobolev spaces $W^{s,p}(\Omega)$ and abbreviate the norm in $L^2(\Omega)$ by

$$\|\cdot\| = \|\cdot\|_{L^2(\Omega)}.$$

The space of functions of bounded variation $\operatorname{BV}(\Omega)$ consists of all $u \in L^1(\Omega)$ such that its total variation

$$|Du|(\Omega) = \sup_{\phi \in C_c^\infty(\Omega; \mathbb{R}^d), |\phi| \leq 1} - \int_\Omega u \operatorname{div} \phi \, dx$$

is bounded. We refer the reader to [2, 3] for properties of the space and note here that distributional gradients Du of functions $u \in \operatorname{BV}(\Omega)$ can be decomposed into a regular, a jump, and a Cantor part *via*

$$Du = \nabla u \otimes dx - \llbracket un \rrbracket \otimes ds|_{J_u} + C_u.$$

Vector fields $w \in L^q(\Omega; \mathbb{R}^d)$ that have a weak divergence $\operatorname{div} w \in L^q(\Omega)$ and whose normal component $w \cdot n$ vanishes on the boundary part $\Gamma_N = \partial\Omega \setminus \Gamma_D$ are contained in the set $W_N^q(\operatorname{div}; \Omega)$, we omit the subindex if $\Gamma_N = \emptyset$. Note that in the standard ROF model we have $\Gamma_D = \emptyset$, occasionally we assume $\Gamma_D = \partial\Omega$.

In the following $(\mathcal{T}_h)_{h>0}$ denotes a sequence of regular, *i.e.*, uniformly shape regular and conforming, triangulations of the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$. The set \mathcal{S}_h contains the sides of elements. The parameter h refers to an average mesh-size $h \sim (|\Omega|/N)^{1/d}$, where N is the number of vertices of \mathcal{T}_h . We furthermore let $h_T = \operatorname{diam}(T)$ for $T \in \mathcal{T}_h$ and

$$h_{\max} = \max_{T \in \mathcal{T}_h} h_T, \quad h_{\min} = \min_{T \in \mathcal{T}_h} h_T.$$

On quasi-uniform triangulations all mesh-sizes are comparable, *i.e.*, $h_{\min} \sim h_{\max} \sim h$. We let $\mathcal{P}_k(T)$ denote the set of polynomials of maximal total degree k on $T \in \mathcal{T}_h$ and define the set of discontinuous, elementwise polynomial functions or vector fields as

$$\mathcal{L}^k(\mathcal{T}_h)^\ell = \{w_h \in L^\infty(\Omega, \mathbb{R}^\ell) : w_h|_T \in P^k(T)^\ell \text{ for all } T \in \mathcal{T}_h\}.$$

Barycenters of elements and sides will be referred to as midpoints denoted by x_T for all $T \in \mathcal{T}_h$ and x_S for all $S \in \mathcal{S}_h$. The L^2 projection onto piecewise constant functions or vector fields is given by elementwise averages and denoted by

$$\Pi_h : L^1(\Omega, \mathbb{R}^\ell) \rightarrow \mathcal{L}^0(\mathcal{T}_h)^\ell.$$

Note that we have $v_h(x_T) = \Pi_h v_h|_T$ for all $T \in \mathcal{T}_h$ and $v_h \in \mathcal{L}^1(\mathcal{T}_h)$.

We next collect some elementary properties of standard and nonstandard finite elements and refer the reader to [7, 14, 15, 24, 29] for further details. The $P1$ -finite element space is defined *via*

$$\mathcal{S}^1(\mathcal{T}_h) = \{v_h \in \mathcal{L}^1(\mathcal{T}_h) : v_h \text{ continuous in } \bar{\Omega}\}.$$

A low order Crouzeix–Raviart finite element space is given by

$$\mathcal{S}^{1,cr}(\mathcal{T}_h) = \{v_h \in \mathcal{L}^1(\mathcal{T}_h) : v_h \text{ continuous in } x_S \text{ for all } S \in \mathcal{S}_h\}.$$

We let $\mathcal{S}_D^1(\mathcal{T}_h)$ and $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ denote the subspaces of functions satisfying boundary conditions in vertices or in midpoints of sides on Γ_D , respectively. The elementwise gradient $\nabla_h v_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ of a function $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ is defined *via*

$$(\nabla_h v_h)|_T = \nabla(v_h|_T)$$

for all $T \in \mathcal{T}_h$. A low order Raviart–Thomas finite element space is given by

$$\mathcal{RT}^0(\mathcal{T}_h) = \{y_h \in W^1(\text{div}; \Omega) : y_h|_T(x) = a_T + b_T(x - x_T), \text{ for all } T \in \mathcal{T}_h, a_T \in \mathbb{R}^d, b_T \in \mathbb{R}\}.$$

Vector fields $q_h \in \mathcal{RT}^0(\mathcal{T}_h)$ belong to $W^1(\text{div}; \Omega)$ and have continuous constant normal components on sides of elements, we set $\mathcal{RT}_N^0(\mathcal{T}_h) = \mathcal{RT}^0(\mathcal{T}_h) \cap W_N^1(\text{div}; \Omega)$. The spaces $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $\mathcal{RT}_N^0(\mathcal{T}_h)$ are connected *via* the integration-by-parts formula

$$\int_{\Omega} v_h \operatorname{div} q_h \, dx = - \int_{\Omega} \nabla_h v_h \cdot q_h \, dx$$

for all $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $q_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$.

Given a function $v \in \text{BV}(\Omega)$ there exists a sequence $(v_{\varepsilon})_{\varepsilon > 0} \subset W^{1,1}(\Omega) \cap C(\bar{\Omega})$ such that $v_{\varepsilon} \rightarrow v$ in $L^1(\Omega)$ and $\|\nabla v_{\varepsilon}\|_{L^1(\Omega)} \rightarrow |Dv|(\Omega)$, *cf.* [2, 3]. With this we define an extension of the nodal interpolation operator $\mathcal{I}_h : C(\bar{\Omega}) \rightarrow \mathcal{S}^1(\mathcal{T}_h)$ *via*

$$\mathcal{I}_h v = \lim_{\varepsilon \rightarrow 0} \mathcal{I}_h v_{\varepsilon},$$

possibly after selection of a subsequence. If $d = 1$ then we have the nodal interpolation estimates

$$\|v - \mathcal{I}_h v\|_{L^1(T)} \leq ch_T^r \|v^{(r)}\|_{L^1(T)} \quad (2.1)$$

for $T \in \mathcal{T}_h$, $v \in W^{r,p}(T)$, and $r \in \{1, 2\}$, *cf.* [15]. *Via* a limit passage for $\varepsilon \rightarrow 0$ it follows that the estimate holds for $r = 1$ and $v \in \text{BV}(T)$ and the right-hand side $ch_T|Dv|(T)$. A straightforward calculation reveals the total-variation diminishing property of the nodal interpolation operator and its extension to $\text{BV}(\Omega)$, *i.e.*,

$$\|(\mathcal{I}_h v)'\|_{L^1(\Omega)} \leq |Dv|(\Omega)$$

for $v \in \text{BV}(\Omega)$ if $\Omega \subset \mathbb{R}$. This estimate fails in higher dimensional situations, *cf.* [11]. As a consequence of Jensen's inequality, the Crouzeix–Raviart quasi-interpolation operator $\mathcal{J}_h^{cr} : W^{1,1}(\Omega) \rightarrow \mathcal{S}^{1,cr}(\mathcal{T}_h)$ satisfies the discrete variant

$$\|\nabla_h \mathcal{J}_h^{cr} v\|_{L^1(\Omega)} \leq \|\nabla v\|_{L^1(\Omega)}$$

for every $v \in W^{1,1}(\Omega)$. Note that the left-hand side of the inequality does not coincide with the total variation of $\mathcal{J}_h^{cr} v$ as jump terms are excluded. From a Poincaré inequality we deduce that

$$\|v - \mathcal{J}_h^{cr} v\|_{L^1(T)} \leq ch_T \|\nabla v\|_{L^1(T)} \quad (2.2)$$

for all $T \in \mathcal{T}_h$. The operator and estimates can be extended to $v \in \text{BV}(\Omega)$.

3. IRREGULAR SOLUTION

The construction of a function $g \in L^2(\Omega)$ that leads to dual solutions which are not Lipschitz continuous uses an idea from [31] and the function

$$g = \chi_{B_r^+} - \chi_{B_r^-}$$

in a sufficiently large domain Ω that is symmetric with respect to the x_1 coordinate. Without making use of the co-area formula, we show here that the example can be understood by using uniqueness of minimizers for I and resulting antisymmetry properties. Hence, the problem reduces to a simpler problem for which the solution can be explicitly derived. We impose Dirichlet conditions on u which eliminates explicit boundary conditions from the dual problem.

Proposition 3.1. *Assume that $\Omega \subset \mathbb{R}^d$ is symmetric with respect to the x_1 coordinate, i.e., $\Omega = \Omega^+ \cup L_0 \cup \Omega^-$, where $x = (x_1, x_2, \dots, x_d) \in \Omega^+$ if and only if $\hat{x} = (-x_1, x_2, \dots, x_d) \in \Omega^-$ and $L_0 \subset \{0\} \times \mathbb{R}^{d-1}$. Assume that $r > 0$ is such that $\overline{B_r^\pm} \subset \Omega$. Then for $g = \chi_{B_r^+} - \chi_{B_r^-}$ the minimizer $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ for I subject to $u|_{\partial\Omega} = 0$ is given by*

$$u = c_{r,\alpha}g, \quad c_{r,\alpha} = \max\{1 - d/(\alpha r), 0\}.$$

Proof. Given a minimizer $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ we define its antisymmetric reflection by $\hat{u}(x_1, x_2, \dots, x_d) = -u(-x_1, x_2, \dots, x_d)$ and note that by the antisymmetry of g we have $I(u) = I(\hat{u})$. By convexity of I we have for $\tilde{u} = (u + \hat{u})/2$ that $I(\tilde{u}) \leq I(u)$ and conclude by uniqueness that $u = \tilde{u} = \hat{u}$, i.e., u is antisymmetric in x_1 . Since the jump set J_u of u is contained in the jump set J_g of g , cf. [18], we find that u is continuous with value 0 in $L_0 \setminus \{0\}$. It thus suffices to consider the reduced minimization of I^+ on Ω^+ with $g^+ = \chi_{B_r^+}$ subject to $u^+|_{\partial\Omega^+} = 0$. The solution is given by $u^+ = c_{r,\alpha}g^+$, cf. e.g., [6, 22], which implies the assertion. \square

The explicit representation of the solution u implies properties of dual solutions. The idea for proving the failure of Lipschitz continuity is illustrated in Figure 1. We make use of the fact that every dual solution equals a normal along the jump set of a piecewise constant primal solution. This follows from the characterization of $|Du|(\Omega)$ via a supremum, the strong duality property, and an integration by parts in the resulting optimality relation $|Du|(\Omega) = -(u, \text{div } z)$. We refer the reader to [1, 18] for related results. More specifically, we have that $z \cdot Du = |Du|$ pointwise almost everywhere and Du is parallel to the normal of the jump set, cf. Equation (26) and Theorem 4 of [22].

Corollary 3.2. *Assume that $\alpha r > d$. Any dual solution $z \in W^2(\text{div}; \Omega)$ for the setting considered in Proposition 3.1 satisfies $\text{div } z = \alpha(u - g)$, i.e., $\text{div } z = -(d/r)g$, and $z \in \partial|Du|$, i.e., on $\partial B_r^+ \cup \partial B_r^-$ we have*

$$z = \mp \nu^\pm$$

with the outer unit normals ν^\pm on ∂B_r^\pm . In particular, z is not θ -Hölder continuous at $x = 0$ for every $\theta > 1/2$.

Proof. A piecewise integration by parts in the representation of $|Du|(\Omega)$ via z yields the relation $z = \mp \nu^\pm$ on the jump set $\partial B_r^+ \cup \partial B_r^-$. Given an angle $\varphi \in (0, \pi/2)$ we define points $x^\pm \in \partial B_r^\pm$ by

$$x^+ = r(-\cos \varphi + 1, \sin \varphi, 0), \quad x^- = r(\cos \varphi - 1, \sin \varphi, 0).$$

Any dual solution z satisfies

$$z(x^+) = (\cos \varphi, -\sin \varphi, 0), \quad z(x^-) = (\cos \varphi, \sin \varphi, 0).$$

We show that the modulus of the θ -difference quotient L_φ for z defined with x^+ and x^- is unbounded as $\varphi \rightarrow 0$, i.e., we have

$$L_\varphi = \frac{|z(x^+) - z(x^-)|}{|x^+ - x^-|^\theta} = 2^{1-\theta} \frac{\sin \varphi}{r^\theta (1 - \cos \varphi)^\theta}.$$

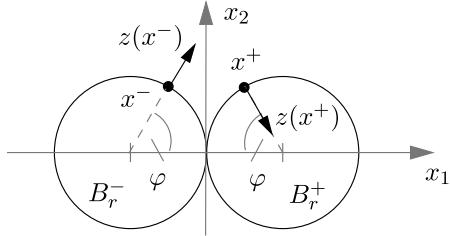


FIGURE 1. Failure of Lipschitz continuity of a dual solution z at $x = 0$ resulting from $|x^+ - x^-| \approx r\varphi^2$ and $|z(x^+) - z(x^-)| \approx 2\varphi$ for $\varphi \rightarrow 0$.

Using the Taylor expansions $\sin \varphi = \varphi + O(\varphi^3)$ and $1 - \cos \varphi = \varphi^2/2 + O(\varphi^4)$ or l'Hospital's rule we find that $L_\varphi \rightarrow \infty$ as $\varphi \rightarrow 0$ whenever $\theta > 1/2$. \square

Remark 3.3. If we repeat the proof of Corollary 3.2 in the more general case where two levelsets meet with a rate $|x^+ - x^-| = O(\varphi^p)$ for $p \geq 1$, we get Hölder exponent $\theta \leq 1 - \frac{1}{p}$. In particular, if the levelsets meet with vanishing curvature, i.e., $p > 2$, then a larger Hölder exponent is possible; Lipschitz continuity can only be achieved if $p = \infty$. We thus conjecture that the example of Proposition 3.1, i.e., $p = 2$, is a worst-case for TV minimisation problems. If g contains a singularity $p < 2$, then the minimiser u must smooth the levelsets at this point. We refer the reader to [1] for results concerning the invariance of jump sets in the minimization of I .

4. GRADED MESHES

To improve convergence results for discontinuous solutions, we use a graded grid approach. This capitalizes on the precise observations about the approximation of a discontinuity by linear finite elements. We motivate the use of graded meshes with the following canonical example.

Proposition 4.1. *For $\ell > 0$ let $\Omega = (-\ell, \ell) \subset \mathbb{R}$ and $u(x) = \text{sign}(x) + v(x)$ with a function $v \in W^{1,2}(\Omega)$. Let $u_h^* = P_h u \in \mathcal{S}^1(\mathcal{T}_h)$ be the L^2 projection of u onto $\mathcal{S}^1(\mathcal{T}_h)$. Assume that the partitioning \mathcal{T}_h is symmetric with respect to $x = 0$. Then we have*

$$c_1 h_0^{1/2} \leq \|u - u_h^*\| \leq c_2 (h_0^{1/2} + h_{\max} \|v'\|),$$

where h_0 is the length of the elements containing the origin.

Proof. We decompose $v = a + s$ into antisymmetric and symmetric, continuous parts $a, s \in W^{1,2}(\Omega)$ with $a(0) = 0$. We then have by L^2 orthogonality that

$$\|u - u_h^*\|^2 = \min_{a_h, s_h \in \mathcal{S}^1(\mathcal{T}_h)} \|\text{sign} + a - a_h\|^2 + \|s - s_h\|^2,$$

where a_h, s_h are antisymmetric and symmetric, respectively. In particular, it suffices to consider the interval $(0, \ell)$ and the restriction $a_h(0) = 0$. The upper bound for the approximation error is obtained by defining a_h via the nodal values of $1 + a$ except for the origin and choosing s_h as the nodal interpolant of s . We then find that

$$\begin{aligned} \|u - u_h^*\| &\leq 2^{1/2} h_0^{1/2} + \|a - \mathcal{I}_h a\| + \|s - \mathcal{I}_h s\| \\ &\leq 2^{1/2} h_0^{1/2} + c_{\mathcal{I}} h_{\max} (\|a'\| + \|s'\|). \end{aligned}$$

To verify the lower bound we consider the contribution from the element $T_0 = [0, h_0]$ to the antisymmetric part. Since $a_h(0) = 0$, we have $a_h(x) = cx$ on T_0 and find that

$$\|\operatorname{sign} +a - a_h\|^2 \geq \|1 + a - cx\|_{L^2(0, h_0)}^2.$$

For the optimal slope c we have Pythagoras' identity

$$\|1 + a - cx\|_{L^2(0, h_0)}^2 = \int_0^{h_0} (1 + a)^2 \, dx - \|cx\|_{L^2(0, h_0)}^2.$$

Since $1 + a$ and x are linearly independent there exists $0 \leq \gamma_a < 1$ with

$$c\|x\|_{L^2(0, h_0)}^2 = \int_0^{h_0} (1 + a)x \, dx \leq \gamma_a \|1 + a\|_{L^2(0, h_0)} \|x\|_{L^2(0, h_0)}.$$

The continuity of a and $a(0) = 0$ show that $\|1 + a\|_{L^2(0, h_0)} \geq c_a h_0^{1/2}$ which implies that we have $\|1 + a - cx\|_{L^2(0, h_0)}^2 \geq (1 - \gamma_a^2) c_a^2 h_0$. \square

Meshes with a grading towards a given point are obtained from affine mappings of a graded grid of the reference interval to a macro element.

Definition 4.2. A β -graded grid of the reference element $\hat{T} = [0, 1]$ is for $J \in \mathbb{N}$ and $\beta \geq 1$ given by the points

$$0 = \xi_0 < \xi_1 < \dots < \xi_J = 1$$

with $\xi_j = (j/J)^\beta$. The length of the interval $[\xi_{j-1}, \xi_j]$ is bounded by $J^{-\beta} \beta j^{\beta-1}$, in particular, we have $h_{\min} = J^{-\beta}$, $h_{\max} = \beta J^{-1}$, and $h = J^{-1}$.

Proposition 4.1 implies that, for a grading strength $\beta = 2$ towards the origin, so that $h_0 = h^2$ and $h_{\max} \leq ch$, we obtain the linear rate

$$\|u - u_h^*\| \leq ch.$$

For the approximation of the ROF problem a similar result can be obtained.

Proposition 4.3. Let $\Omega = (a, b) \subset \mathbb{R}$, $g \in L^\infty(\Omega)$, and assume that the minimizer $u \in \operatorname{BV}(\Omega) \cap L^2(\Omega)$ of I is a piecewise $W^{2,1}$ function. Then if \mathcal{T}_h is graded with strength β towards the jumps J_u of u we have for the P1 finite element minimizer $u_h \in \mathcal{S}_D^1(\mathcal{T}_h)$ for I that

$$\|u - u_h\|^2 \leq c(h^\beta |Du|(\Omega) + \beta^2 h^2 \|u''\|_{L^1(\Omega \setminus J_u)}) \|g\|_{L^\infty(\Omega)}.$$

Proof. By minimality of u and coercivity properties of I , we find that for every $\tilde{u}_h \in \mathcal{S}_D^1(\mathcal{T}_h)$ we have

$$\frac{\alpha}{2} \|u - u_h\|^2 \leq I(u_h) - I(u) \leq I(\tilde{u}_h) - I(u).$$

Via regularization we set $\tilde{u}_h = \mathcal{I}_h u$ and use the total-variation diminishing property of \mathcal{I}_h to deduce with a binomial formula that

$$\begin{aligned} \frac{\alpha}{2} \|u - u_h\|^2 &\leq \int_{\Omega} (\tilde{u}_h - g)^2 - (u - g)^2 \, dx \\ &\leq \|\tilde{u}_h - u\|_{L^1(\Omega)} \|\tilde{u}_h + u - 2g\|_{L^\infty(\Omega)}. \end{aligned}$$

By properties of nodal interpolation we have $\|\tilde{u}_h\|_{L^\infty(\Omega)} \leq \|u\|_{L^\infty(\Omega)}$. Moreover, truncating u at the essential extrema \underline{g} and \bar{g} of g decreases its energy and by uniqueness implies $\|u\|_{L^\infty(\Omega)} \leq \|g\|_{L^\infty(\Omega)}$. We decompose the first factor into elementwise contributions. If $T \cap J_u \neq \emptyset$ we have

$$\|\tilde{u}_h - u\|_{L^1(T)} \leq ch_{\min} |Du|(T).$$

Otherwise, standard interpolation estimates show that

$$\|\tilde{u}_h - u\|_{L^1(T)} \leq ch_{\max}^2 \|u''\|_{L^1(T)}.$$

A summation over the elements leads to the asserted estimate. \square

Remark 4.4. If the solution u is piecewise linear then the estimate can be improved to the convergence rate $O(h^{\beta/2})$ for every $\beta \geq 1$.

Because the total-variation diminishing property is not satisfied, the generalization of Proposition 4.3 to a higher-dimensional setting requires additional assumptions on a dual solution. A key difficulty is that the quasi-interpolation of a dual solution in the space $\mathcal{RT}_N^0(\mathcal{T}_h)$ leads to a local violation of the constraint $|z_h(x_T)| \leq 1$ of order $O(h_T)$ independently of additional regularity properties. For piecewise constant primal solutions we assume the strict inequality $|z(x)| < 1$ with linear decay away from discontinuities so that no violation occurs, while in a neighborhood of the jump set we have a quadratic violation $O(h^2)$ on suitably graded meshes.

Proposition 4.5. *For $\Omega \subset \mathbb{R}^2$ let $g \in L^\infty(\Omega)$ and assume that the primal solution $u \in \text{BV}(\Omega) \cap L^\infty(\Omega)$ is piecewise constant with piecewise regular jump set J_u , and there exists a dual solution $z \in W_N^1(\text{div}; \Omega) \cap W^{1,\infty}(\Omega; \mathbb{R}^2)$ with Lipschitz constant $L \geq 0$ and such that there exists $\ell_z > 0$ with*

$$|z(x)| \leq 1 - \ell_z d_{J_u}(x)$$

where $d_{J_u}(x) = \inf_{y \in J_u} |x - y|$. Let $(\mathcal{T}_h)_{h>0}$ be a sequence of quadratically graded triangulations towards J_u , i.e., with $d_{J_u}(T) = \inf_{x \in T} d_{J_u}(x)$ there exists $c_d > 0$ such that for all $T \in \mathcal{T}_h$ we have

$$h_T \leq c_d \max \left\{ h d_{J_u}(T)^{1/2}, h^2 \right\}$$

Then, the Crouzeix–Raviart finite element solution $u_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ of the ROF model satisfies

$$\|\Pi_h(u - u_h)\| \leq h M(\alpha, u, z, g).$$

The factor $M(\alpha, u, z, g)$ depends on generic quantities related to \mathcal{T}_h and Ω as well as the indicated quantities that are well defined by the stated assumptions.

Proof. We follow the strategies of [9, 20] that make use of projection operators related to the spaces $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $\mathcal{RT}_N^0(\mathcal{T}_h)$ and Jensen's inequality as well as binomial formulas. Letting $g_h = \Pi_h g$ this shows that there exists a Crouzeix–Raviart quasi-interpolant $\tilde{u}_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ of u such that

$$I_h(\tilde{u}_h) \leq I(u) + \frac{\alpha}{2} \|\Pi_h \tilde{u}_h - u\|_{L^1(\Omega)} 4 \|g\|_{L^\infty(\Omega)} - \frac{\alpha}{2} \|g - g_h\|^2.$$

If $\gamma_h \geq 1$ is such that for the corrected Raviart–Thomas quasi-interpolant $\tilde{z}_h = \gamma_h^{-1} \mathcal{J}_{RT} z \in \mathcal{RT}_N^0(\mathcal{T}_h)$ of z we have $|\tilde{z}_h(x_T)| \leq 1$ for all $T \in \mathcal{T}_h$ then we have

$$D_h(\tilde{z}_h) \geq D(z) - (1 - \gamma_h^{-1}) \|g\| \|\text{div } z\| - \frac{\alpha}{2} \|g - g_h\|^2.$$

Using the coercivity of I_h , the minimality of u_h , the discrete duality relation $I_h(u_h) \geq D_h(\tilde{z}_h)$, and the strong duality relation $I(u) = D(z)$, we find that

$$\begin{aligned} \frac{\alpha}{2} \|\Pi_h(\tilde{u}_h - u_h)\|^2 &\leq I_h(\tilde{u}_h) - I_h(u_h) \leq I_h(\tilde{u}_h) - D_h(\tilde{z}_h) \\ &\leq \frac{\alpha}{2} \|\Pi_h \tilde{u}_h - u\|_{L^1(\Omega)} 4 \|g\|_{L^\infty(\Omega)} + (1 - \gamma_h^{-1}) \|g\| \|\text{div } z\|. \end{aligned}$$

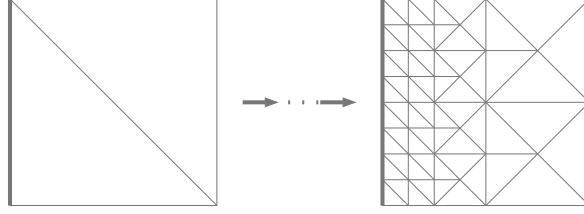


FIGURE 2. Triangulation \mathcal{T}_3 in Example 4.7 obtained by refinements of \mathcal{T}_0 with a grading towards the side $\{0\} \times [0, \ell]$.

To bound the correction factor γ_h we estimate $\gamma_T = |\mathcal{J}_{RT}z(x_T)|$ for every $T \in \mathcal{T}_h$ using that $\gamma_T \leq |z(x_T)| + c_{RT}h_T L$.

We seek to prove $\gamma_T \leq 1 + c_\gamma h^2$ for some $c_\gamma > 0$. From the decay condition we have

$$\gamma_T \leq 1 - \ell_z d_{J_u}(T) + c_{RT}Lh_T.$$

If $h_T \leq c_d h^2$, then $c_\gamma \geq c_d$ suffices. Otherwise, we focus on the case $h_T \leq c_d h d_{J_u}(T)^{1/2}$ and let $C = c_{RT}c_d L/(2\ell_z)$. Then, using $(Ch - d_{J_u}(T)^{1/2})^2 \geq 0$ we have

$$\gamma_T \leq 1 + \ell_z \left(2Ch d_{J_u}(T)^{1/2} - d_{J_u}(T) \right) \leq 1 + \ell_z C^2 h^2.$$

We conclude that $\gamma_h \leq 1 + c_\gamma h^2$, so $(1 - \gamma_h^{-1}) \leq c_\gamma h^2$. To bound the term $\|\Pi_h \tilde{u}_h - u\|_{L^1(\Omega)}$, we note that if $T \cap J_u \neq \emptyset$ we obtain with (2.2) that

$$\|u - \Pi_h \tilde{u}_h\|_{L^1(T)} \leq \|u - \tilde{u}_h\|_{L^1(T)} + \|\tilde{u}_h - \Pi_h \tilde{u}_h\|_{L^1(T)} \leq ch_T |Du|(T)$$

with $h_T \leq c_d h^2$. Otherwise, if $T \cap J_u = \emptyset$ we have that u is constant and $u = \Pi_h \tilde{u}_h$ on T . The estimate of the proposition follows from a combination of the estimates and the triangle inequality, noting that by Jensen's and Hölder's inequalities we have $\|\Pi_h(u - \tilde{u}_h)\|^2 \leq \|u - \tilde{u}_h\|_{L^1(\Omega)} \|u - \tilde{u}_h\|_{L^\infty(\Omega)}$. \square

The assumptions of the proposition apply to certain settings with piecewise constant solutions.

Example 4.6. If $g = \chi_{B_r(0)}$ for $r > 0$ with $\overline{B_r(0)} \subset \Omega$ and if Dirichlet conditions on u are imposed on $\Gamma_D = \partial\Omega$, then we have $u = c_{r,\alpha}g$ and

$$z(x) = -c'_{r,\alpha} \begin{cases} r^{-1}x, & |x| \leq r, \\ rx/|x|^2, & |x| \geq r, \end{cases}$$

for every $x \in \Omega$, where $c'_{r,\alpha} = \min\{1, r\alpha/d\}$, cf. [1, 6, 22].

A quadratic grading is the optimal grading strength to locally refine a two-dimensional triangulation towards a one-dimensional subset.

Example 4.7. Let $\Omega = (0, \ell)^2 \subset \mathbb{R}^2$ with $\ell > 0$ with initial triangulation $\mathcal{T}_0 = \{T_1, T_2\}$. We inductively define \mathcal{T}_{k+1} by first applying a red refinement to all elements in \mathcal{T}_k that intersect the x_2 -axis and then refining further elements to avoid hanging nodes by a red-green-blue refinement strategy as, e.g., in Algorithm 4.2 of [7], cf. Figure 2.

We define the (asymptotic) *grading strength* of a sequence of regular triangulations $(\mathcal{T}_k)_{k \geq 0}$ as the logarithmic relation of the minimal and average mesh-size, i.e.,

$$\beta = \lim_{k \rightarrow \infty} \frac{\log(h_{k,\min})}{\log(h_k)}.$$

We note that the speed of convergence depends on the diameter of Ω , *e.g.*, for the triangulations defined in Example 4.7 we have that $h_{k,\min}$ and h_k are proportional to ℓ , which is irrelevant in the limit passage. For the generic setting of Example 4.7 we identify a quadratic grading strength.

Proposition 4.8. *Let $(\mathcal{T}_k)_{k=0,1,\dots}$ be a sequence of triangulations of $(0,\ell)^2$ such that triangles along the side $\{0\} \times [0,\ell]$ are k -times refined with $h_T \sim h_{k,\min} \sim q^k h_{k,\max}$ and for triangles along the side $\{\ell\} \times [0,\ell]$ we have $h_T = h_{k,\max}$. If $k^{-1} \log(h_{k,\max}) \rightarrow 0$ then the graduation strength is $\beta = 2$.*

Proof. To determine the average mesh-size we note that the refinement process defines after k -steps a partitioning into stripes S_j , $j = 0, 1, \dots, k$, with $n_j \sim q^{-j} h_{\max}$ elements. A summation shows that \mathcal{T}_k contains $N_k \sim q^{-k} h_{\max}$ elements so that $h_k \sim q^{k/2}$ and

$$\beta = \lim_{k \rightarrow \infty} \frac{\log(cq^k h_{k,\max})}{\log(c'q^{k/2} h_{k,\max})} = \lim_{k \rightarrow \infty} \frac{\log(ch_{k,\max}) + k \log(q)}{\log(c'h_{k,\max}) + (k/2) \log(q)} = 2,$$

which proves the assertion. \square

5. PRIMAL-DUAL GAP ESTIMATOR VIA DISCRETE DUALITY

We next devise a strategy that leads to an automatic and adaptive local mesh refinement algorithm. To illustrate the main ideas we consider a general convex minimization problem

$$\tilde{I}(u) = \int_{\Omega} \phi(\nabla u) + \psi(x, u) \, dx$$

defined on a Sobolev space $X = W_{\text{D}}^{1,p}(\Omega)$, $1 < p < \infty$, or on $X = \text{BV}(\Omega)$, whose dual is given by the maximization of

$$\tilde{D}(z) = - \int_{\Omega} \phi^*(z) + \psi^*(x, \text{div } z) \, dx$$

on a space of vector fields $W = W_{\text{N}}^{p'}(\text{div}; \Omega)$. Here, $\phi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ and $\psi : \Omega \times \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ are convex functionals and ϕ^* and ψ^* are their convex conjugates. The duality relation $\tilde{I}(u) \geq \tilde{D}(z)$ in combination with coercivity properties of \tilde{I} described by a functional $\sigma_{\tilde{I}}$ imply, for the minimizer $u \in X$ and arbitrary $v \in X$ and $q \in W$ that

$$\sigma_{\tilde{I}}^2(u, v) \leq \tilde{I}(v) - \tilde{I}(u) \leq \tilde{I}(v) - \tilde{D}(q) =: \eta_h^2(v, q). \quad (5.1)$$

If $v = u_h$ for an approximation $u_h \in X$ of u , then $\eta_h(u_h, q)$ provides a computable bound on the approximation error $\sigma(u, u_h)$ whenever an admissible q is explicitly given. We use the following extended result from [5].

Proposition 5.1. *Let $u \in \text{BV}(\Omega) \cap L^2(\Omega)$ be the minimizer for the ROF functional I and $u_h \in \mathcal{S}_{\text{D}}^1(\mathcal{T}_h)$ an approximation. We then have, for every $q \in W_{\text{N}}^2(\text{div}; \Omega)$ with $|q| \leq 1$ in Ω , that*

$$\|u - u_h\| \leq (2/\alpha)^{1/2} \eta_h(u_h, q) + \|g - \Pi_h g\|,$$

where

$$\eta_h^2(u_h, q) = \int_{\Omega} |\nabla u_h| - \nabla u_h \cdot \Pi_h q \, dx + \frac{1}{2\alpha} \int_{\Omega} (\text{div } q - \alpha(u_h - g_h))^2 \, dx.$$

Proof. We define $g_h = \Pi_h g$ and let \tilde{I} be the ROF functional with g replaced by g_h whose minimizer we denote by $\tilde{u} \in \text{BV}(\Omega) \cap L^2(\Omega)$. By the strong convexity of the L^2 term in I we find that $\|u - \tilde{u}\| \leq \|g - g_h\|$, *cf. e.g.*, [6]. We apply the error estimate (5.1) to \tilde{I} and obtain that

$$\frac{\alpha}{2} \|\tilde{u} - u_h\|^2 \leq \tilde{I}(u_h) - \tilde{D}(q) - \int_{\Omega} \nabla u_h \cdot q \, dx - \int_{\Omega} u_h \text{div } q \, dx.$$

A straightforward calculation, the fact that ∇u_h is elementwise constant, and the triangle inequality lead to the formula for $\eta_h(u_h, q)$. \square

Remark 5.2. If the estimate is derived for a Crouzeix–Raviart approximation $u_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, then jumps across sides occur on the right-hand side of the identity for $\eta_h^2(u_h, q)$.

The optimal estimator $\eta_h(u_h, q)$ requires an exact solution of the dual problem or a numerical approximation of sufficient accuracy, *cf.* [10]. Since the numerical solution of the dual problem is computationally expensive, we aim at the construction of a nearly optimal approximation at a computational cost that is comparable to the numerical solution of the discretized primal problem. For this we use a reconstruction of a discrete dual solution from the Crouzeix–Raviart approximation of the primal problem from [9].

Proposition 5.3 ([9], Prop. 3.1). *Let \tilde{I}_h and \tilde{D}_h be defined on $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $\mathcal{RT}_N^0(\mathcal{T}_h)$ with $\psi_h(\cdot, a) = \Pi_h \psi(\cdot, a)$ for all $a \in \mathbb{R}$ via*

$$\begin{aligned}\tilde{I}_h(u_h) &= \int_{\Omega} \phi(\nabla_h u_h) + \psi_h(x, \Pi_h u_h) \, dx, \\ \tilde{D}_h(z_h) &= - \int_{\Omega} \phi^*(\Pi_h z_h) + \psi_h^*(x, \operatorname{div} z_h) \, dx.\end{aligned}$$

We then have the duality relation $\tilde{I}_h(u_h) \geq \tilde{D}_h(z_h)$. If $s \mapsto \phi(s)$ and $a \mapsto \psi_h(x, a)$ are continuously differentiable and if u_h is minimal for \tilde{I}_h then a maximizing element z_h for \tilde{D}_h is given by

$$z_h = \phi'(\nabla_h u_h) + d^{-1} \psi_h'(\cdot, \Pi_h u_h)(\cdot - x_{\mathcal{T}}),$$

where $x_{\mathcal{T}} = \Pi_h \operatorname{id}$, and strong duality $\tilde{I}_h(u_h) = \tilde{D}_h(z_h)$ applies.

To apply the result to the discretized ROF functional, we consider for $\varepsilon > 0$ the regularization $\phi(a) = |a|_{\varepsilon} = (|a|^2 + \varepsilon^2)^{1/2}$ of the non-differentiable modulus. We then obtain the reconstruction $z_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ given by

$$z_h = \frac{\nabla u_h}{|\nabla_h u_h|_{\varepsilon}} + \frac{\alpha}{d} \Pi_h(u_h - g)(\cdot - x_{\mathcal{T}}).$$

The vector field z_h satisfies $|z_h(x_T)| \leq 1$ for all $T \in \mathcal{T}_h$, but in general not $|z_h(x)| \leq 1$ for almost every $x \in \Omega$. We have for every $T \in \mathcal{T}_h$ that

$$|z_h|_T \leq 1 + \frac{\alpha}{d} |(u_h - g_h)(x_T)| \frac{d}{d+1} h_T = \gamma_T.$$

The globally re-scaled vector field $\hat{z}_h = (\max_{T \in \mathcal{T}_h} \gamma_T)^{-1} z_h$ satisfies $|\hat{z}_h| \leq 1$ but does not lead to an efficient error estimator. Our experiments reported below indicate that also the scaling $\tilde{z}_h = \tilde{\gamma}_h^{-1} z_h$ with a continuous function $\tilde{\gamma}_h$ satisfying $\tilde{\gamma}_h|_T \geq \gamma_T$ for all $T \in \mathcal{T}_h$ does not lead to an efficient estimator.

Remark 5.4. The error estimator $\eta_h(u_h, q)$ controls the approximation error $\sigma_{\tilde{I}}(u, u_h)$ defined by the maximal coercivity of the functional \tilde{I} . For the ROF functional I the scaled squared L^2 norm is a lower bound for this quantity and the error estimator controls a stronger error quantity. For the regularized ROF functional I^{ε} and a minimizers u , *i.e.*, $\delta I^{\varepsilon}(u) = 0$, a Taylor expansion formally yields with a suitable function ξ that

$$I^{\varepsilon}(u_h) = I^{\varepsilon}(u) + \int_{\Omega} \phi''_{\varepsilon}(\nabla \xi)[\nabla(u - u_h), \nabla(u - u_h)] \, dx + \frac{\alpha}{2} \|u - u_h\|^2,$$

where the convex function $\phi_{\varepsilon} = |\cdot|_{\varepsilon}$ has a positive definite Hessian, *cf.* [25].

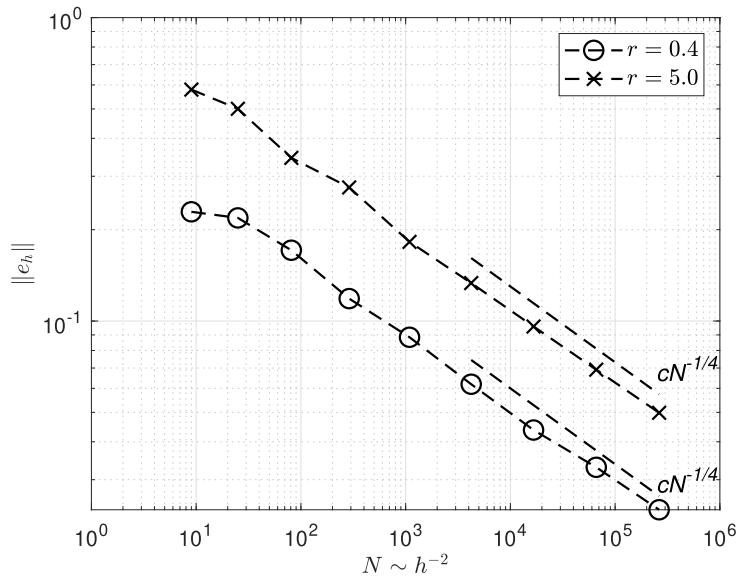


FIGURE 3. Experimental convergence rates $h^{1/2} \sim N^{-1/4}$ for Crouzeix–Raviart finite element approximations of the ROF model on sequences of uniform triangulations for a solution with non-Lipschitz continuous dual solutions defined in Example 6.1 and different magnifications of the irregular region at the origin.

6. NUMERICAL EXPERIMENTS

We verify in this section the theoretical results and investigate the performance of numerical methods beyond their validity. Our computations are based on the use of the regularized ROF functional

$$I^\varepsilon(u) = \int_{\Omega} |\nabla u|_\varepsilon \, dx + \frac{\alpha}{2} \|u - g\|^2$$

with the regularized modulus $|a|_\varepsilon = (|a|^2 + \varepsilon^2)^{1/2}$ for $a \in \mathbb{R}^d$ and $\varepsilon > 0$. Owing to the bounds $0 \leq |a|_\varepsilon - |a| \leq \varepsilon$, the error estimates and identified convergence rates remain valid provided that $\varepsilon = O(h^\sigma)$ with $\sigma = 1$ or $\sigma = 2$ to obtain an L^2 error $O(h^{\sigma/2})$ on uniform and locally refined meshes. The iterative minimization of I^ε was realized with the unconditionally stable semi-implicit L^2 gradient flow from [12]. We always use the step-size $\tau = 1$ but different stopping criteria $\|u^k - u^{k-1}\| \leq \varepsilon_{\text{stop}}$.

6.1. Irregular solution

We investigate the numerical approximation of the example from Section 3 to verify whether the failure of Lipschitz continuity of dual solutions affects the convergence rate $O(h^{1/2})$ for the Crouzeix–Raviart method on uniform triangulations. We use a coordinate transformation to avoid superconvergence phenomena related to mesh symmetries.

Example 6.1 (Non-Lipschitz dual). Let $\Omega = (-1, 1)^2 \subset \mathbb{R}^2$, $\alpha = 10$, and $\tilde{g} = \chi_{B_r^+} - \chi_{B_r^-}$ for $r \in \{0.4, 5.0\}$ and $g = \tilde{g} \circ \Phi$, where $\Phi(x) = Qx + b$ realizes a rotation by $\phi = 70^\circ$ and shift by $b = (0.1, 0)^\top$. Dirichlet conditions $u_D = u|_{\partial\Omega}$ from the solution $u = c_{r,\alpha}g$, $c_{r,\alpha} = 1 - 2/(r\alpha)$, are imposed.

The experimental convergence rates shown in Figure 3 are obtained on k -times red-refined triangulations \mathcal{T}_k of an initial triangulation \mathcal{T}_0 with four elements. We have $h_k = 2^{-k}$ and use $\varepsilon_{\text{stop}} = h_k/20$. The optimal

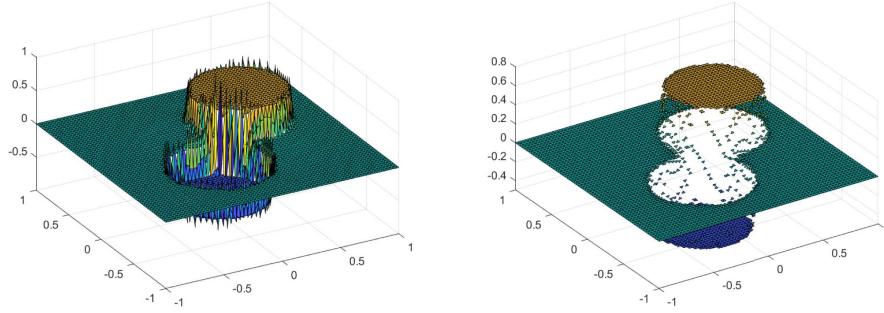


FIGURE 4. Numerical solution $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_6)$ (left) and its projection $\Pi_h u_h$ (right) in Example 6.1 for $r = 0.4$. Large discrete gradients occur near the origin where dual solutions are not Lipschitz continuous.

convergence rate $O(h^{1/2})$ is observed for both choices of $r = 0.4$ and $r = 5.0$, despite the lack of a Lipschitz continuous dual solution. Also the scaling of the problem towards the singular point obtained by increasing the radius r does not affect the experimental convergence rates. In Figure 4 the numerical solution u_h on the triangulation \mathcal{T}_6 and its projection onto elementwise constant functions are displayed for the parameter $r = 0.4$. Large gradients occur near the origin, the midpoint values do not, however, show artifacts.

6.2. Mesh grading in one dimension

We next confirm our theoretical findings for the use graded meshes for the approximation of the ROF model in one-dimensional settings. The problem specification leads to a multiple of the sign function as exact solution.

Example 6.2 (1D sign function). Let $\Omega = (-1, 1)$, $\alpha = 10$, and define $g(x) = \text{sign}(x)$. The minimizer for the ROF functional subject to Dirichlet boundary conditions is given by $u = c_{r,\alpha}g$, $c_{r,\alpha} = (1 - 1/(r\alpha))$, for $r = 2$.

In our experiments we choose the regularization $\varepsilon = h^\beta$ so that the corresponding error contribution is of the same order as the discretization error. We note that the stopping criterion has to be carefully chosen and we used $\varepsilon_{\text{stop}} = h/20$ for $\beta = 1$ and the finer tolerance $\varepsilon_{\text{stop}} = h^{\beta+1}/20$ for non-uniform meshes with grading strength $\beta > 1$. The experimental convergence rates obtained with these settings for a $P1$ method are given in Figure 5, typical numerical solutions are displayed in Figure 6.

6.3. Mesh grading in two dimensions

We experimentally investigate the performance of finite element approximations for a standard example using mesh grading based on the discontinuity set of the given function g .

Example 6.3 (Single disc phantom). Let $\Omega = (-1, 1)^2$, $\alpha = 10$, and $g = \chi_{B_r(0)}$ for $r = 1/2$. For homogeneous Dirichlet boundary conditions the minimizer of the ROF model is given by $u = c_{r,\alpha}g$ with $c_{r,\alpha} = 1 - 2/(r\alpha)$.

Our initial triangulation \mathcal{T}_0 consists of two right triangles that partition Ω and we iteratively define a sequence of regular triangulations $(\mathcal{T}_k)_{k=0,1,\dots}$ by performing a red refinement for all triangles in \mathcal{T}_k that have a non-empty intersection with the discontinuity set $J_g = \partial B_r(0)$ of g and then carrying out a red-green-blue refinement procedure to avoid hanging nodes. We verified that this leads to a quadratic grading strength. To allow for a nearly linear experimental convergence rate, we choose $\varepsilon_{\text{stop}} = h^2/20$ and $\varepsilon = h^2$. For the approximations obtained with the Crouzeix–Raviart method we observe a nearly linear experimental convergence rate. This is not the case for approximations obtained with less flexible $P1$ finite elements, as can be seen in Figure 7. We also illustrated the convergence behavior of the error estimator from Section 5 and observe that it serves as a reliable but non-efficient error bound. An explanation for this observation is that the graded meshes are optimal for the L^2 error, but not necessarily for the error quantity controlled by the estimator, *cf.* Remark 5.4.

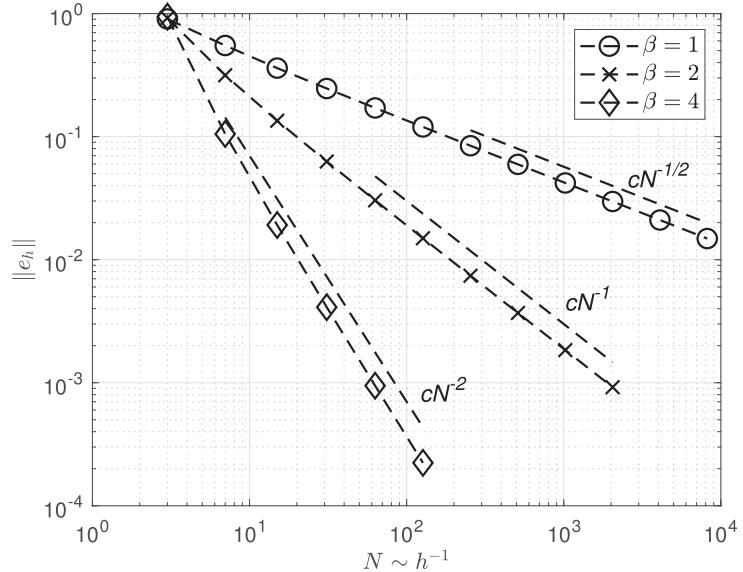


FIGURE 5. Experimental convergence rates in approximating the one-dimensional ROF model defined in Example 6.2 on meshes with a grading towards the discontinuity with different grading strengths β leading to convergence rates $h^{\beta/2} \sim N^{-\beta/2}$.

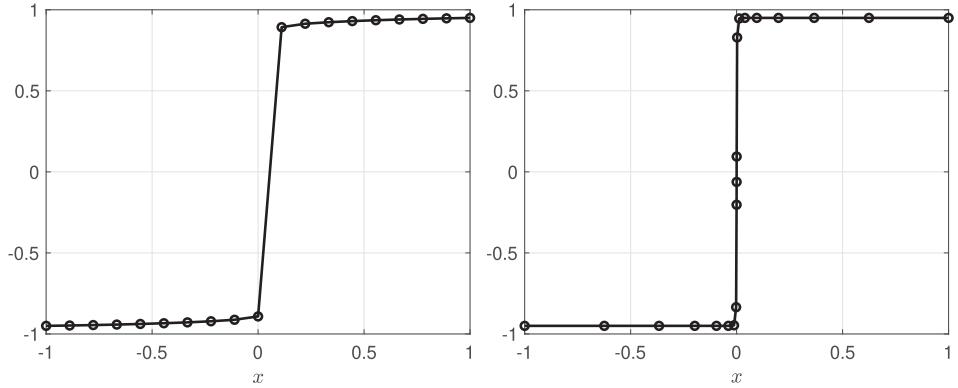


FIGURE 6. Numerical solutions on a uniform and a graded mesh in the one-dimensional setting with piecewise constant solution specified in Example 6.2. The strong grading with $\beta = 4$ (right) leads to a high accuracy in comparison with the uniform grid corresponding to $\beta = 1$ (left).

6.4. Adaptive mesh refinement

We finally investigate the automatic generation of locally refined triangulations based on the *a posteriori* error estimate provided by Proposition 5.1. We use the reconstructed, unscaled approximation z_h of the dual problem provided by the Crouzeix–Raviart approximation u_h^{cr} for the primal problem. This defines the error estimator

$$E_{\text{est}} = \left(\frac{2}{\alpha} \sum_{T \in \mathcal{T}_h} \eta_{h,T}^2(u_h, z_h) \right)^{1/2} + \left(\sum_{T \in \mathcal{T}_h} \|g - \Pi_h g\|_{L^2(T)}^2 \right)^{1/2},$$

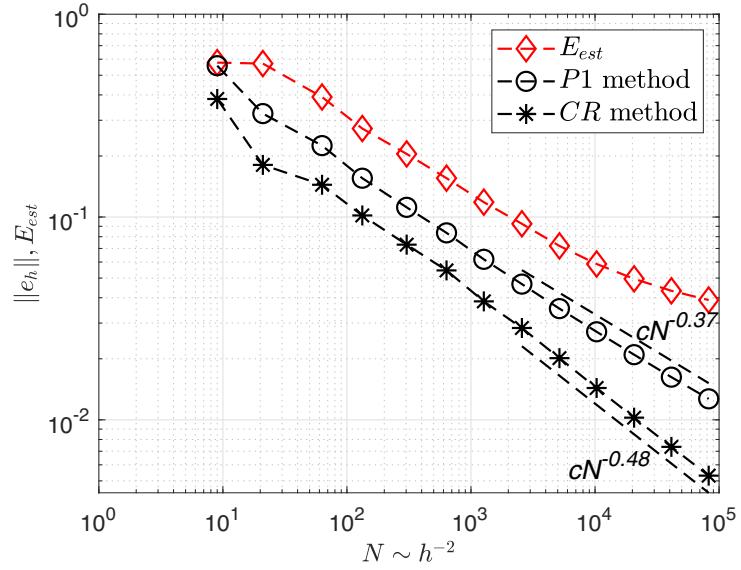


FIGURE 7. Experimental convergence rates on quadratically graded triangulations in the approximation of the two-dimensional ROF-problem with piecewise constant solution specified in Example 6.3. Crouzeix–Raviart approximations lead to nearly linear convergence of the L^2 error.

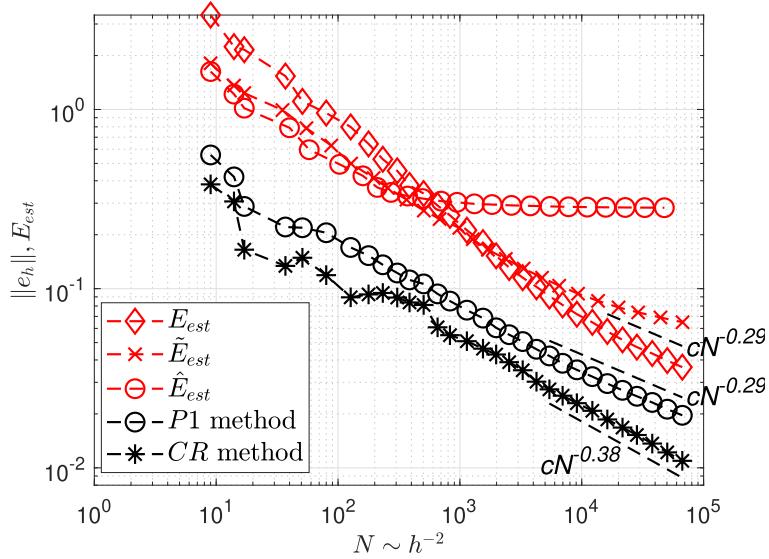


FIGURE 8. Experimental convergence rates in the adaptive approximation of the ROF-problem defined in Example 6.4 using the primal-dual-gap error estimator E_{est} . Different experimental convergence rates are observed for Crouzeix–Raviart and $P1$ finite element approximations. The estimators \hat{E}_{est} and \tilde{E}_{est} obtained from globally and locally scaled dual variables are inefficient.

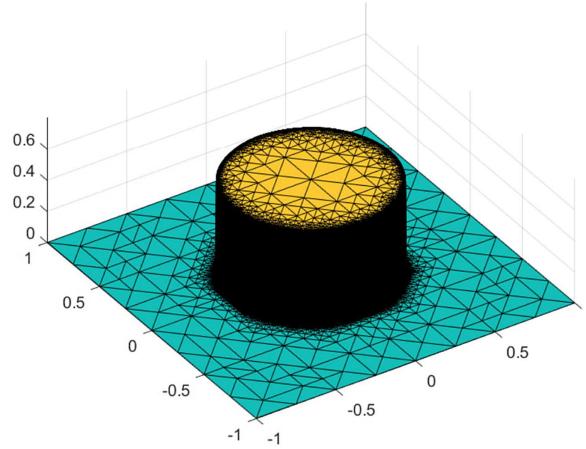


FIGURE 9. Adaptively generated $P1$ approximation in Example 6.4. The automatic mesh refinement procedure leads to a local refinement in a neighborhood of the discontinuity set.

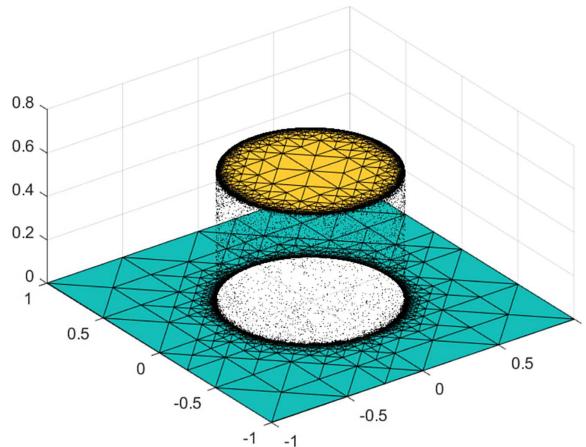


FIGURE 10. Projection $\Pi_h u_h$ of the Crouzeix–Raviart approximation on the quadratically graded triangulation \mathcal{T}_{13} in Example 6.3. The localized refinement of the jump set leads to a high accuracy.

where the second sum contains data oscillation terms and the first one the local refinement indicators $\eta_{h,T}$ which are given by the element residuals

$$\eta_{h,T}^2(u_h, z_h) = \int_T |\nabla u_h| - \nabla u_h \cdot \Pi_h z_h \, dx + \frac{1}{2\alpha} \int_T (\operatorname{div} z_h - \alpha(u_h - g_h))^2 \, dx.$$

We follow established strategies in adaptive mesh refinement methods and select a minimal subset $M_h \subset \mathcal{T}_h$ for refinement, *cf.* [7], that constitutes 50% of the total error estimator. We used the regularization parameter $\varepsilon = h^2$ to allow for an overall linear convergence rate, as stopping criterion we used $\varepsilon_{\text{stop}} = h^2/20$. We again use a setting with a piecewise constant solution.

Example 6.4 (Piecewise constant solution). Let $\Omega = (-1, 1)^2$, $\alpha = 10$, and $g = \chi_{B_r(0)}$ for $r = 1/2$. For homogeneous Dirichlet boundary conditions the minimizer of the ROF model is given by $u = c_{r,\alpha}g$ with $c_{r,\alpha} = 1 - 2/(r\alpha)$.

The experimental convergence rates for $P1$ and Crouzeix–Raviart finite element approximations on adaptively generated triangulations in Example 6.4 are shown in Figure 8. For both methods we observe an improvement over the optimal rate $O(h^{1/2})$ on sequences of uniform triangulations. The Crouzeix–Raviart method leads to the experimental convergence rate $O(h^{0.76})$ while for the $P1$ method we obtain the lower rate $O(h^{0.58})$. Our explanation for this is the good compatibility of the Crouzeix–Raviart method specified by the projection property of the quasi-interpolation operator and the resulting discrete total-variation diminishing property. As addressed in Remark 5.4 the error estimator E_{est} cannot be expected to lead to meshes that are optimal for the L^2 approximation error. The error estimator E_{est} converges with nearly the same rate as the $P1$ approximation error indicating good reliability and efficiency properties. A $P1$ finite element approximation obtained with the adaptive mesh refinement strategy is shown in Figure 9. We observe an automatic local mesh refinement towards the discontinuity set of the solution but a weaker grading of approximately $\beta \approx 1.7$ in comparison with Figure 10. The reliable estimators \hat{E}_{est} and \tilde{E}_{est} , obtained from using the globally and locally scaled vector fields \hat{z}_h and \tilde{z}_h lead to meshes on which these estimators converge suboptimally, cf. Figure 8. The L^2 error converged with similar rates reported above for meshes constructed with \hat{E}_{est} but not with \tilde{E}_{est} .

Acknowledgements. The authors thank Ricardo H. Nochetto for stimulating discussions on various aspects of the results of this article. The first author acknowledges support by the DFG via the priority programme SPP 1962 *Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization*. The second author acknowledges support from the ANR CIPRESSI project grant ANR-19-CE48-0017-01 of the French Agence Nationale de la Recherche. The authors also thank the anonymous reviewers for their valuable comments.

REFERENCES

- [1] F. Alter, V. Caselles and A. Chambolle, A characterization of convex calibrable sets in \mathbb{R}^N . *Math. Ann.* **332** (2005) 329–366.
- [2] L. Ambrosio, N. Fusco and D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York (2000).
- [3] H. Attouch, G. Buttazzo and G. Michaille, Variational Analysis in Sobolev and BV Spaces – Applications to PDEs and Optimization. Vol. 4 of *MPS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics, Philadelphia (2006).
- [4] S. Bartels, Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.* **50** (2012) 1162–1180.
- [5] S. Bartels, Error control and adaptivity for a variational model problem defined on functions of bounded variation. *Math. Comput.* **84** (2015) 1217–1240.
- [6] S. Bartels, Numerical Methods for Nonlinear Partial Differential Equations. Vol. 47 of *Springer Series in Computational Mathematics*. Springer, Cham (2015).
- [7] S. Bartels, Numerical Approximation of Partial Differential Equations. Vol. 64 of *Texts in Applied Mathematics*. Springer, Berlin, Heidelberg (2016).
- [8] S. Bartels, Error estimates for a class of discontinuous Galerkin methods for nonsmooth problems via convex duality relations. *Math. Comput.* **90** (2021) 2579–2602.
- [9] S. Bartels, Nonconforming discretizations of convex minimization problems and precise relations to mixed methods. *Comput. Math. App.* **93** (2021) 214–229.
- [10] S. Bartels and M. Milicevic, Primal-dual gap estimators for *a posteriori* error analysis of nonsmooth minimization problems. *ESAIM: Math. Model. Numer. Anal.* **54** (2020) 1635–1660.
- [11] S. Bartels, R.H. Nochetto and A.J. Salgado, A total variation diminishing interpolation operator and applications. *Math. Comput.* **84** (2015) 2569–2587.
- [12] S. Bartels, L. Diening and R.H. Nochetto, Unconditional stability of semi-implicit discretizations of singular flows. *SIAM J. Numer. Anal.* **56** (2018) 1896–1914.
- [13] B. Berkels, A. Effland and M. Rumpf, A posteriori error control for the binary Mumford–Shah model. *Math. Comput.* **86** (2017) 1769–1791.
- [14] S.C. Brenner, Forty years of the Crouzeix–Raviart element. *Numer. Methods Part. Differ. Equ.* **31** (2015) 367–396.
- [15] S.C. Brenner and L. Ridgway Scott, The Mathematical Theory of Finite Element Methods. Vol. 15 of *Texts in Applied Mathematics*, 3rd edition. Springer, New York (2008).

- [16] M. Burger, Bregman distances in inverse problems and partial differential equations. In: *Advances in Mathematical Modeling, Optimization and Optimal Control*. Vol. 109 of *Springer Optim. Appl.* Springer, Cham (2016) 3–33.
- [17] C. Caillaud and A. Chambolle, Error estimates for finite differences approximations of the total variation. HAL preprint nr. 02539136 (2020).
- [18] V. Caselles, A. Chambolle and M. Novaga, The Discontinuity Set of Solutions of the TV Denoising Problem and Some Extensions. *Multiscale Model. Simul.* **6** (2007) 879–894.
- [19] A. Chambolle and P.L. Lions, Image recovery via total variation minimization and related problems. *Numer. Math.* **76** (1997) 167–188.
- [20] A. Chambolle and T. Pock, Crouzeix–Raviart approximation of the total variation on simplicial meshes. *J. Math. Imaging Vision* **62** (2020) 872–899.
- [21] A. Chambolle and T. Pock, Approximating the total variation with finite differences or finite elements. In: *Handbook of Numerical Analysis: Geometric Partial Differential Equations II*. Elsevier (2021).
- [22] A. Chambolle, V. Caselles, D. Cremers, M. Novaga and T. Pock, An introduction to total variation for image analysis. In: *Theoretical Foundations and Numerical Methods for Sparse Recovery*. Vol. 9 of *Radon Ser. Comput. Appl. Math.* Walter de Gruyter, Berlin (2010) 263–340.
- [23] A. Chambolle, S.E. Levine and B.J. Lucier, An upwind finite-difference method for total variation-based image smoothing. *SIAM J. Imaging Sci.* **4** (2011) 277–299.
- [24] M. Crouzeix and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *R.A.I.R.O.* **7** (1973) 33–75.
- [25] F. Fierro and A. Veeser, A posteriori error estimators for regularized total variation of characteristic functions. *SIAM J. Numer. Anal.* **41** (2003) 2032–2055.
- [26] M. Herrmann, R. Herzog, S. Schmidt, J. Vidal-Núñez and G. Wachsmuth, Discrete total variation with finite elements and applications to imaging. *J. Math. Imaging Vision* **61** (2019) 411–431.
- [27] M. Hintermüller and K. Kunisch, Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM J. Appl. Math.* **64** (2004) 1311–1333.
- [28] M.-J. Lai and L. Matamba Messi, Piecewise linear approximation of the continuous Rudin–Osher–Fatemi model for image denoising. *SIAM J. Numer. Anal.* **50** (2012) 2446–2466.
- [29] P.A. Raviart and J.M. Thomas, A mixed finite element method for 2-nd order elliptic problems. In: *Mathematical Aspects of Finite Element Methods*, edited by I. Galligani and E. Magenes. Springer, Berlin, Heidelberg (1977) 292–315.
- [30] L.I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. D: Nonlinear Phenom.* **60** (1992) 259–268.
- [31] R. Tovey, *Mathematical challenges in electron microscopy*. Ph.D. thesis, University of Cambridge (2020).
- [32] J. Wang and B.J. Lucier, Error bounds for finite-difference methods for Rudin–Osher–Fatemi image smoothing. *SIAM J. Numer. Anal.* **49** (2011) 845–868.

Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a **Subscribe-to-Open** model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/math-s2o-programme>