Error estimates for total-variation regularized minimization problems with singular dual solutions

Sören Bartels*[1] and Alex Kaltenbach†[2]

[1]Institute of Applied Mathematics, Albert–Ludwigs–University Freiburg, Hermann–Herder–Straße 10, 79104 Freiburg
[2]Institute of Applied Mathematics, Albert–Ludwigs–University Freiburg, Ernst–Zermelo–Straße 1, 79104 Freiburg

January 12, 2022

Abstract

Recent quasi-optimal error estimates for the finite element approximation of total-variation regularized minimization problems using the Crouzeix–Raviart finite element require the existence of a Lipschitz continuous dual solution, which is not generally given. We provide analytic proofs showing that the Lipschitz continuity of a dual solution is not necessary, in general. Using the Lipschitz truncation technique, we, in addition, derive error estimates that depend directly on the Sobolev regularity of a given dual solution.

Keywords: Error estimates, Lipschitz truncation, image processing, total variation.
AMS MSC (2010): 26A45, 65N15, 65N30, 68U10.

1. Introduction

In this article, we examine the finite element discretization of the Rudin–Osher–Fatemi (ROF) model from [32], which serves as a model problem for general convex and non-smooth minimization problems. This image processing model determines a function $u \in BV(\Omega)\cap L^2(\Omega)$ via minimizing $I : BV(\Omega)\cap L^2(\Omega) \to \mathbb{R}$, defined by

$$I(u) := |\mathrm{D}u|(\Omega) + \frac{\alpha}{2}\|u - g\|^2_{L^2(\Omega)} \tag{1.1}$$

for all $u \in BV(\Omega) \cap L^2(\Omega)$, where $|\mathrm{D}u|(\Omega)$ denotes the total variation, $g \in L^2(\Omega)$ is the input data, e.g., a noisy image, and $\|u-g\|^2_{L^2(\Omega)}$ is the so-called fidelity term. In addition, the fidelity parameter $\alpha > 0$ is a given constant, which determines the balance between de-noising and preserving the input image. For a more in-depth analysis of this model, concerning its analytical properties, explicit solutions, and numerical methods, we refer to [19, 4, 26, 16, 17, 18, 33, 6, 28, 5, 11, 25, 20, 9, 12]. Since this model allows for and preserves discontinuities of the input data g, cf. [16],

*Email: `bartels@mathematik.uni-freiburg.de`
†Email: `alex.kaltenbach@mathematik.uni-freiburg.de`

continuous finite element methods are known to perform sub-optimally, cf. [11, 9]. Recent contributions, cf. [20, 8, 9], reveal that the quasi-optimal convergence rate $\mathcal{O}(h^{\frac{1}{2}})$ for discontinuous solutions on quasi-uniform triangulations can be obtained using discontinuous, low-order Crouzeix–Raviart finite elements introduced in [21]. More precisely, these error estimates yield a bound for the error for the approximation of minimizers of $I : BV(\Omega)\cap L^2(\Omega) \to \mathbb{R}$ via minimizing the discrete functional $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, defined by

$$I_h(u_h) := \|\nabla_h u_h\|_{L^1(\Omega;\mathbb{R}^d)} + \frac{\alpha}{2}\|\Pi_h(u_h - g)\|^2_{L^2(\Omega)}$$

for all $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, where $\mathcal{S}^{1,cr}(\mathcal{T}_h)$ is the Crouzeix–Raviart finite element space, i.e., the space of piece-wise affine functions that are continuous at the midpoints of element sides, $\nabla_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathcal{L}^0(\mathcal{T}_h)^d$ denotes the element-wise gradient, and $\Pi_h : L^2(\Omega) \to \mathcal{L}^0(\mathcal{T}_h)$ is the L^2–projection operator onto element-wise constant functions. Note that the family of discrete functionals $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, $h > 0$, defines a non-conforming approximation of the functional $I : BV(\Omega)\cap L^2(\Omega) \to \mathbb{R}$, as, e.g., jump terms of u_h across inter-element sides are not included. For this family recently a Γ–convergence result with respect to strong convergence in $L^1(\Omega)$ or distributional convergence has been established under general assumptions, i.e., that $g \in L^2(\Omega)$, cf. [20, Propositon 3.1]. However, the quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ till now only holds if the dual problem given via maximizing $D : W^2_N(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R}\cup\{-\infty\}$, defined by

$$D(z) := -\frac{1}{2\alpha}\|\mathrm{div}(z) + \alpha g\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|g\|^2_{L^2(\Omega)} - I_{K_1(0)}(z) \tag{1.2}$$

for all $z \in W^2_N(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$, where $I_{K_1(0)} : L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$ is for $z \in L^\infty(\Omega;\mathbb{R}^d)$ defined by $I_{K_1(0)}(z) := 0$ if $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1$ and $I_{K_1(0)}(z) := +\infty$ else, admits a Lipschitz continuous solution. Unfortunately, the Lipschitz continuity of a maximum of $D : W^2_N(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R}\cup\{-\infty\}$ is not generally given, as [12, Section 3] clarified. Without imposing the existence of a Lipschitz continuous solution to (1.2), but that $g \in BV(\Omega)\cap L^\infty(\Omega)$, in [20, Section 5.2], the sub-optimal convergence rate $\mathcal{O}(h^{\frac{1}{4}})$ has been established. The approach of [20, Section 5.2] consists in a convolution of a maximum $z \in W^\infty_N(\mathrm{div};\Omega)$ of (1.2) in order to comply with the crucial Lipschitz continuity property at least in an approximate sense. We use an alternative regularization approach, which operates highly at a local level, the celebrated Lipschitz truncation technique. Its basic purpose is to approximate Sobolev functions $u \in W^{1,p}(\Omega)$ by λ–Lipschitz functions $u_\lambda \in W^{1,\infty}(\Omega)$, $\lambda > 0$. The original approach of this technique traces back to Acerbi and Fusco, cf. [1, 2, 3]. Since then, the Lipschitz truncation technique is used in various areas of analysis: In the calculus of variations, in the existence theory of partial differential equations, and in regularity theory. For a longer list of references, we refer the reader to [22]. To the best of the authors knowledge, this article provides the first deployment of the Lipschitz truncation technique in the field of image processing. More precisely, for the application in this article, the main advantage of the Lipschitz truncation technique in comparison to convolution is that not only u and u_λ coincide up to a set of small measure, but equally ∇u and ∇u_λ do. By deploying the Lipschitz truncation technique, we arrive at error estimates whose resulting rates directly depend on the respective Sobolev regularity of a given maximum $z \in W^{1,p}(\Omega;\mathbb{R}^d)$ of (1.2). If only $g \in L^\infty(\Omega)$ and one additionally has that, e.g., $z \in W^{1,p}(\Omega;\mathbb{R}^d)$ for $p \ge 3$, then the results of this article yield the sub-optimal rate $\mathcal{O}(h^{\frac{1}{4}})$. In this manner, we intend to fill the gap between the optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ for $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$ and $g \in L^\infty(\Omega)$ and the rate $\mathcal{O}(h^{\frac{1}{4}})$ for $z \in W_N^\infty(\mathrm{div};\Omega)$ and $g \in L^\infty(\Omega) \cap BV(\Omega)$.

As a maximum of (1.2) is not necessarily in a Sobolev space, but in $W_N^2(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$, we also study the case of a non-existence of Sobolev solutions to (1.2). It turns out that if a maximum $z \in W_N^2(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$ of (1.2) is element-wise Lipschitz continuous, i.e., the discontinuity set J_z is resolved by the triangulations, or at least in an approximate sense with the rate $\mathcal{O}(h)$, cf. Remark 4.8, then the optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ can be expected. Beyond that, we find that the optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ is attained if a dual solution fulfills $|z| < 1$ along its discontinuity set J_z while, simultaneously, its jump $[\![z]\!]$ over its discontinuity set J_z remains small. Some of these conditions apply, e.g., to the setting described in [12, Section 3] with a suitable triangulation $\mathcal{T}_h$, $h > 0$, of the domain Ω, for which the optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ could be reported without giving an analytical explanation. This article's purpose is to give – at least for special cases – a missing analytical explanation.

This article is organized as follows: In Section 2, we introduce the employed notation, define the relevant finite element spaces and give a brief review of the continuous and discretized ROF model. In Section 3, using the Lipschitz truncation technique, we establish error estimates that depend directly on the Sobolev regularity of a maximum of (1.2). In Section 4, we prove quasi-optimal error estimates without explicitly imposing that a Lipschitz continuous maximum of (1.2) exists. In Section 5, we confirm our theoretical findings via numerical experiments.

2. Preliminaries

Throughout the article, if not otherwise specified, we denote by $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, a bounded polyhedral Lipschitz domain, whose boundary is disjointly divided into a Dirichlet part Γ_D and a Neumann part Γ_N, i.e., $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\emptyset = \Gamma_D \cap \Gamma_N$.

2.1 *Function spaces*

For $p \in [1,\infty]$ and $l \in \mathbb{N}$, we employ the standard notations

$$W_D^{1,p}(\Omega;\mathbb{R}^l) := \big\{u \in L^p(\Omega;\mathbb{R}^l) \mid \nabla u \in L^p(\Omega;\mathbb{R}^{l\times d}),\ \mathrm{tr}(u) = 0 \text{ in } L^p(\Gamma_D;\mathbb{R}^l)\big\},$$

$$W_N^p(\mathrm{div};\Omega) := \big\{z \in L^p(\Omega;\mathbb{R}^d) \mid \mathrm{div}(z) \in L^p(\Omega),\ \mathrm{tr}(z)\cdot n = 0 \text{ in } W^{-\frac{1}{p},p}(\Gamma_N)\big\},$$

$W^{1,p}(\Omega;\mathbb{R}^l) := W_D^{1,p}(\Omega;\mathbb{R}^l)$ if $\Gamma_D = \emptyset$, and $W^p(\mathrm{div};\Omega) := W_N^p(\mathrm{div};\Omega)$ if $\Gamma_N = \emptyset$, where $\mathrm{tr} : W^{1,p}(\Omega;\mathbb{R}^l) \to L^p(\partial\Omega)$ and $\mathrm{tr}(\cdot)\cdot n : W^p(\mathrm{div};\Omega) \to (W^{1,p'}(\Omega))^*$ denote the trace and normal trace operator. In particular, we predominantly omit $\mathrm{tr}(\cdot)$ in this context. Apart from that, we fall back on the abbreviations $L^p(\Omega) := L^p(\Omega;\mathbb{R}^1)$ and $W^{1,p}(\Omega) := W^{1,p}(\Omega;\mathbb{R}^1)$. Let $|\mathrm{D}(\cdot)|(\Omega) : L^1_{\mathrm{loc}}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, defined by[1]

$$|\mathrm{D}u|(\Omega) := \sup_{\phi \in C_c^\infty(\Omega;\mathbb{R}^d),\, \|\phi\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1} -\int_\Omega u \,\mathrm{div}(\phi)\,\mathrm{d}x$$

for all $u \in L^1_{\mathrm{loc}}(\Omega)$, denote the total variation. Then, the space of functions of bounded variation is defined by $BV(\Omega) := \big\{u \in L^1(\Omega) \mid |\mathrm{D}u|(\Omega) < \infty\big\}$.

$$BV(\Omega) := \big\{u \in L^1(\Omega) \mid |\mathrm{D}u|(\Omega) < \infty\big\}.$$

[1] Here, $C_c^\infty(\Omega;\mathbb{R}^d)$ denotes the space of smooth and in Ω compactly supported vector fields.

2.2 Triangulations

In what follows, we let $(\mathcal{T}_h)_{h>0}$ be a sequence of regular, i.e., uniformly shape regular and conforming, triangulations of $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, cf. [14]. The sets $\mathcal{S}_h$ and $\mathcal{N}_h$ contain the sides and vertices, resp., of the elements. The parameter $h > 0$ refers to the maximal mesh-size of $\mathcal{T}_h$. More precisely, if we define $h_T := \mathrm{diam}(T)$ for all $T \in \mathcal{T}_h$, then we have that $h = \max_{T\in\mathcal{T}_h} h_T$. For any $k \in \mathbb{N}$ and $T \in \mathcal{T}_h$, we let $\mathcal{P}_k(T)$ denote the set of polynomials of maximal total degree k on T. Then, the set of element-wise polynomial functions or vector fields, resp., is defined by

$$\mathcal{L}^k(\mathcal{T}_h)^l := \big\{ v_h \in L^\infty(\Omega;\mathbb{R}^l) \mid v_h|_T \in \mathcal{P}_k(T) \text{ for all } T \in \mathcal{T}_h \big\}.$$

For any $T \in \mathcal{T}_h$ and $S \in \mathcal{S}_h$, we let $x_T := \frac{1}{d+1}\sum_{z\in\mathcal{N}_h\cap T} z$ and $x_S := \frac{1}{d}\sum_{z\in\mathcal{N}_h\cap S} z$ denote the midpoints (barycenters) of T and S, resp. The L^2–projection operator onto piece-wise constant functions or vector fields, resp., is denoted by

$$\Pi_h : L^1(\Omega;\mathbb{R}^l) \to \mathcal{L}^0(\mathcal{T}_h)^l.$$

For $v_h \in \mathcal{L}^1(\mathcal{T}_h)^l$, it holds $\Pi_h v_h|_T = v_h(x_T)$ for all $T \in \mathcal{T}_h$. Moreover, for $p \in [1,\infty]$, there exists a constant $c_\Pi > 0$ such that for all $v \in L^p(\Omega;\mathbb{R}^l)$, cf. [23], we have that

(L0.1) $\|\Pi_h v\|_{L^p(\Omega;\mathbb{R}^l)} \le \|v\|_{L^p(\Omega;\mathbb{R}^l)}$,
(L0.2) $\|v - \Pi_h v\|_{L^p(\Omega;\mathbb{R}^l)} \le c_\Pi h \|\nabla v\|_{L^p(\Omega;\mathbb{R}^{l\times d})}$ if $v \in W^{1,p}(\Omega;\mathbb{R}^l)$.

2.3 Crouzeix–Raviart finite elements

A particular instance of a larger class of non-conforming finite element spaces, introduced in [21], is the Crouzeix–Raviart finite element space, which consists of piece-wise affine functions that are continuous at the midpoints of element sides, i.e.,

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

The element-wise application of the gradient to $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defines an element-wise constant vector field $\nabla_h v_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ via $\nabla_h v_h|_T := \nabla(v_h|_T)$ for all $T \in \mathcal{T}_h$. Crouzeix–Raviart finite element functions that vanish at midpoints of boundary element sides that correspond to the Dirichlet boundary Γ_D are contained in the space

$$\mathcal{S}^{1,cr}_D(\mathcal{T}_h) := \big\{ v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h) \mid v_h(x_S) = 0 \text{ for all } S \in \mathcal{S}_h \text{ with } S \subseteq \Gamma_D \big\}.$$

In particular, we have that $\mathcal{S}^{1,cr}_D(\mathcal{T}_h) = \mathcal{S}^{1,cr}(\mathcal{T}_h)$ if $\Gamma_D = \emptyset$. A basis of $\mathcal{S}^{1,cr}(\mathcal{T}_h)$ is given by the functions $\varphi_S \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying the Kronecker property $\varphi_S(x_{S'}) = \delta_{S,S'}$ for all $S, S' \in \mathcal{S}_h$. A basis of $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ is given by $(\varphi_S)_{S\in\mathcal{S}_h;S\not\subseteq\Gamma_D}$. For any $p \in [1,\infty]$, the quasi-interpolation operator $I_{cr} : W^{1,p}_D(\Omega) \to \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, for all $v \in W^{1,p}_D(\Omega)$ defined by

$$I_{cr} v := \sum_{S\in\mathcal{S}_h} v_S \varphi_S, \quad v_S := \fint_S v \,\mathrm{d}s \tag{2.1}$$

preserves averages of gradients, i.e., $\nabla_h(I_{cr} v) = \Pi_h(\nabla v)$ in $\mathcal{L}^0(\mathcal{T}_h)^d$ for $v \in W^{1,p}_D(\Omega)$. Moreover, for $p \in [1,\infty]$, there exits a constant $c_{cr} > 0$ such that for all $v \in W^{1,p}_D(\Omega)$, cf. [13], we have that

(CR.1) $\|\nabla_h(I_{cr} v)\|_{L^p(\Omega;\mathbb{R}^d)} \le \|\nabla v\|_{L^p(\Omega;\mathbb{R}^d)}$,
(CR.2) $\|v - I_{cr} v\|_{L^p(\Omega)} \le c_{cr} h \|\nabla v\|_{L^p(\Omega;\mathbb{R}^d)}$,
(CR.3) $\|I_{cr} v\|_{L^\infty(\Omega)} \le c_d \|v\|_{L^\infty(\Omega)}$, where $c_d := (d+1)(d-1)$, if $v \in L^\infty(\Omega)$.

For $p=1$, due to the density of $C^\infty(\Omega)\cap BV(\Omega)$ in $BV(\Omega)$, cf. [5], the operator and *(CR.1)*–*(CR.3)* can be extended to $v \in BV(\Omega)$, losing the representation (2.1).

2.4 Raviart–Thomas finite elements

The lowest order Raviart–Thomas finite element space, introduced in [31], consists of piece-wise affine vector fields that possess weak divergences, i.e.,

$$\mathcal{RT}^0(\mathcal{T}_h) := \big\{ z_h \in \mathcal{L}^1(\mathcal{T}_h)^d \mid z_h|_T \cdot n_T = -z_h|_{T'} \cdot n_{T'} \text{ on } T \cap T' \text{ if } T \cap T \in \mathcal{S}_h \big\},$$

where $n_T : \partial T \to \mathbb{S}^{d-1}$ for all $T \in \mathcal{T}_h$ denotes the unit normal vector field to T pointing outward. Raviart–Thomas finite element functions that have vanishing normal components on the Neumann boundary Γ_N are contained in the space

$$\mathcal{RT}^0_N(\mathcal{T}_h) := \big\{ z_h \in \mathcal{RT}^0(\mathcal{T}_h) \mid z_h \cdot n = 0 \text{ on } \Gamma_N \big\}.$$

In particular, we have that $\mathcal{RT}^0_N(\mathcal{T}_h) = \mathcal{RT}^0(\mathcal{T}_h)$ if $\Gamma_N = \emptyset$. A basis of $\mathcal{RT}^0(\mathcal{T}_h)$ is given by the vector fields $\psi_S \in \mathcal{RT}^0(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying the Kronecker property $\psi_S|_{S'} \cdot n_{S'} = \delta_{S,S'}$ on S' for all $S \in \mathcal{S}_h$, where n_S for all $S \in \mathcal{S}_h$ denotes the unit normal vector on S that points from T_- to T_+ if $S = \partial T_- \cap \partial T_+ \in \mathcal{S}_h$. A basis of $\mathcal{RT}^0_N(\mathcal{T}_h)$ is given by $\psi_S \in \mathcal{RT}^0_N(\mathcal{T}_h)$, $S \in \mathcal{S}_h \setminus \Gamma_N$. The quasi-interpolation operator $I_{\mathcal{RT}} : V^{\mathrm{div}}(\Omega) := \{ y \in L^p(\Omega;\mathbb{R}^d) \mid \mathrm{div}(y) \in L^q(\Omega)\} \to \mathcal{RT}^0_N(\mathcal{T}_h)$, where $p > 2$ and $q > \frac{2d}{d+2}$, for all $z \in V^{\mathrm{div}}(\Omega)$ defined by

$$I_{\mathcal{RT}} z := \sum_{S \in \mathcal{S}_h} z_S \psi_S, \quad z_S := \fint_S z \cdot n_S \,\mathrm{d}s \tag{2.2}$$

preserves averages of divergences, i.e., $\mathrm{div}(I_{\mathcal{RT}} z) = \Pi_h(\mathrm{div}(z))$ in $\mathcal{L}^0(\mathcal{T}_h)$ for all $z \in V^{\mathrm{div}}(\Omega)$. Moreover, for $p \in [1,\infty]$, there exists a constant $c_{\mathcal{RT}} > 0$ such that for all $z \in V^{\mathrm{div}}(\Omega)$, cf. [24], we have that

(RT.1) $\|z - I_{\mathcal{RT}} z\|_{L^p(\Omega;\mathbb{R}^d)} \le c_{\mathcal{RT}} h \|\nabla z\|_{L^p(\Omega;\mathbb{R}^{d\times d})}$ if $z \in W^{1,p}(\Omega;\mathbb{R}^d)$,

(RT.2) $\|I_{\mathcal{RT}} z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le c_{\mathcal{RT}} \|z\|_{L^\infty(\Omega;\mathbb{R}^d)}$ if $z \in L^\infty(\Omega;\mathbb{R}^d)$.

For $p \in [1,\infty)$, due to the density of $C^\infty(\Omega;\mathbb{R}^d) \cap W^p_N(\mathrm{div};\Omega)$ in $W^p_N(\mathrm{div};\Omega)$, the operator and *(RT.2)* can be extended to $z \in W^p_N(\mathrm{div};\Omega)$, losing the representation (2.2).

2.5 The continuous Rudin–Osher–Fatemi (ROF) model

Given $g \in L^2(\Omega)$ and $\alpha > 0$, the Rudin–Osher–Fatemi (ROF) model, cf. [32], determines a function $u \in BV(\Omega) \cap L^2(\Omega)$ that is minimal for $I : BV(\Omega) \cap L^2(\Omega) \to \mathbb{R}$, defined by

$$I(v) := |\mathrm{D}v|(\Omega) + \frac{\alpha}{2} \|v - g\|^2_{L^2(\Omega)} \tag{2.3}$$

for all $v \in BV(\Omega) \cap L^2(\Omega)$. In [7, Theorem 10.5 & Theorem 10.6], it is established that for every $g \in L^2(\Omega)$, there exists a unique minimizer $u \in BV(\Omega) \cap L^2(\Omega)$ of $I : BV(\Omega) \cap L^2(\Omega) \to \mathbb{R}$. If $g \in L^\infty(\Omega)$, then $u \in L^\infty(\Omega)$ with $\|u\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}$ (cf. [7, Proposition 10.2]). In [26, Theorem 2.2], it is shown that the corresponding dual problem to (2.3) determines a vector field $z \in W^2_N(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$, where $\Gamma_N = \partial\Omega$, that is maximal for $D : W^2_N(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R} \cup \{-\infty\}$, defined by

$$D(y) := -\frac{1}{2\alpha} \|\mathrm{div}(y) + g\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|g\|^2_{L^2(\Omega)} - I_{K_1(0)}(y) \tag{2.4}$$

for all $y \in W^2_N(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$, where $I_{K_1(0)} : L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R} \cup \{\infty\}$ is defined by $I_{K_1(0)}(y) := 0$ if $y \in L^\infty(\Omega;\mathbb{R}^d)$ with $\|y\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1$ and $I_{K_1(0)}(y) := \infty$ else. Apart from that, in [26, Theorem 2.2], it is shown that (2.4) possesses a maximizer $z \in W^2_N(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$, which satisfies the strong duality principle

$$I(u) = D(z). \tag{2.5}$$

The strong duality principle (2.5), appealing to [7, Proposition 10.4] and referring to standard convex optimization arguments, is equivalent to the optimality relations

$$\operatorname{div}(z) = \alpha(u - g) \quad \text{in } L^2(\Omega), \qquad |\mathrm{D}u|(\Omega) = -(u, \operatorname{div}(z)). \tag{2.6}$$

2.6 The discretized Rudin–Osher–Fatemi (ROF) model

Given some $g \in L^2(\Omega)$ and $\alpha > 0$, with $g_h := \Pi_h g \in \mathcal{L}^0(\mathcal{T}_h)$, the discretized ROF model proposed by [20] determines a Crouzeix–Raviart function $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ that is minimal for $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, defined by

$$I_h(v_h) := \|\nabla_h v_h\|_{L^1(\Omega;\mathbb{R}^d)} + \frac{\alpha}{2}\|\Pi_h v_h - g_h\|^2_{L^2(\Omega)} \tag{2.7}$$

for all $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$. In [20] and [9], it has been shown that the corresponding dual problem to (2.7) determines a Raviart–Thomas vector field $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$, where $\Gamma_N = \partial\Omega$, that is maximal for $D_h : \mathcal{RT}^0_N(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, defined by

$$D_h(y_h) := -\frac{1}{2\alpha}\|\operatorname{div}(y_h) + g_h\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|g_h\|^2_{L^2(\Omega)} - I_{K_1(0)}(\Pi_h y_h) \tag{2.8}$$

for all $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Apart from that, in [20] and [9], it has been established that a discrete weak duality principle holds, i.e., it holds

$$\inf_{v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)} I_h(v_h) \geq \sup_{y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)} D_h(y_h), \tag{2.9}$$

which is a cornerstone of the error analysis for (2.7). In particular, note that for the validity of (2.9) the L^2–projection operator Π_h in (2.7) and (2.8) plays a key role.

2.7 Piece-wise Lipschitz, but not globally Lipschitz, continuous solution to (2.4)

In [12, Section 3], the construction of an input data $g \in BV(\Omega) \cap L^\infty(\Omega)$ that leads to a solution $z \in W^\infty_N(\operatorname{div};\Omega)$ to (2.4) such that $z \notin W^{1,\infty}(\Omega;\mathbb{R}^2)$, in essence, is based on the asymmetry of the function

$$g := \chi_{B^2_r(re_1)} - \chi_{B^2_r(-re_1)} \in BV(\Omega) \cap L^\infty(\Omega) \tag{2.10}$$

defined on a domain that is symmetric with respect to the $\mathbb{R}e_2$–axis.[2] More precisely, using this asymmetry property, it is possible to reduce the minimization problem (2.3) on Ω into two independent minimization problems on $\Omega^+ := \Omega \cap (\mathbb{R}_{>0} \times \mathbb{R})$ and $\Omega^- := \Omega \cap (\mathbb{R}_{<0} \times \mathbb{R})$ for which explicit solutions $u^\pm \in BV(\Omega^\pm) \cap L^\infty(\Omega^\pm)$ exist. In this way, the following result could be derived.

Proposition 2.1. *Let $\Omega \subseteq \mathbb{R}^2$ be symmetric with respect to the $\mathbb{R}e_2$–axis and let $r > 0$ be such that $B^2_r(\pm re_1) \subset\subset \Omega$. Then, for (2.10) and $\alpha > 0$, the minimizer of $I : \mathcal{A} \to \mathbb{R}$, where $\mathcal{A} := \{u \in BV(\Omega) \cap L^\infty(\Omega) \mid \operatorname{tr}(u) = 0 \text{ in } L^1(\partial\Omega)\}$, is given via*

$$u = \max\left\{0, 1 - \frac{2}{\alpha r}\right\} g \in \mathcal{A}. \tag{2.11}$$

Proof. See [12, Proposition 3.1]. □

Combining the representation formula (2.11) and the optimality conditions (2.6), it turns out that there exists no Lipschitz continuous dual solution to the setting described in Proposition 2.1.

[2] For every $i = 1, \ldots, d$, we denote by $e_i \in \mathbb{S}^{d-1}$, the i–th. unit vector.

Corollary 2.2. *Let the assumptions of Proposition 2.1 be satisfied with $\alpha r > 2$. Then, any dual solution $z \in W^{\infty}(\mathrm{div};\Omega)$ to (2.11) is not θ–Hölder continuous if $\theta > \frac{1}{2}$.*

Proof. See [12, Corollary 3.2]. □

An example of a not Lipschitz continuous dual solution to (2.11) is the following, which is separately Lipschitz continuous on Ω^+ and Ω^-, resp., and jumps over the $\mathbb{R}e_2$–axis. We will resort to this dual solution to derive optimal error estimates for the setting described in Proposition 2.1, which has already been reported in [12, Example 6.1].

Proposition 2.3. *Let $\Omega \subseteq \mathbb{R}^2$ and $r > 0$ be such as in Proposition 2.1. Moreover, let $\alpha > 0$ be such that $\alpha r > 2$. Then, the vector field $z : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, defined by*

$$z(x) := \begin{cases} \mp\frac{1}{r}(x \mp re_1) & \text{if } |x \mp re_1| < r \\ \mp\frac{r}{|x\mp re_1|^2}(x \mp re_1) & \text{if } |x \mp re_1| \geq r \end{cases}$$

for all $x \in \Omega$, satisfies $z \in W^{\infty}(\mathrm{div};\Omega)$, $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \leq 1$, $|\mathrm{D}u|(\Omega) = -(u, \mathrm{div}(z))_{L^2(\Omega)}$ and $\mathrm{div}(z) = \alpha(u-g)$ in $L^\infty(\Omega)$, where $u \in \mathcal{A}$ is defined by (2.11), i.e., $z \in W^{\infty}(\mathrm{div};\Omega)$ is a dual solution to (2.11).

Proof. Apparently, we have that $z \in L^\infty(\Omega;\mathbb{R}^d)$ with $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \leq 1$. In addition, it is not difficult to see that $z|_{\Omega^\pm} \in W^{1,\infty}(\Omega^\pm;\mathbb{R}^d)$. Since $z|_{\Omega^+} \cdot n_{\Omega^+} = -z|_{\Omega^-} \cdot n_{\Omega^-}$ on $\mathbb{R}e_2 \cap \Omega$, we find that $z \in W^2(\mathrm{div};\Omega)$. It is well-known, cf. [7, Example 10.4], that

$$|\mathrm{D}u|(\Omega^\pm) = -(u, \mathrm{div}(z))_{L^2(\Omega^\pm)}, \qquad \mathrm{div}(z) = \alpha(u-g) \quad \text{ in } L^\infty(\Omega^\pm).$$

Thus, we have that $\mathrm{div}(z) = \alpha(u-g)$ in $L^\infty(\Omega)$, which implies that $z \in W^\infty(\mathrm{div};\Omega)$. Apart from that, using that $u = 0$ continuously in $\mathbb{R}e_2 \cap \Omega$, we finally conclude that $|\mathrm{D}u|(\Omega) = -(u, \mathrm{div}(z))_{L^2(\Omega)}$, i.e., $z \in W^\infty(\mathrm{div};\Omega)$ is a dual solution to (2.11). □

3. Error estimates depending on Sobolev regularity

The validity of quasi-optimal error estimates for the finite element approximation of total-variation regularized minimization problems by means of the Crouzeix–Raviart element in the case of an existing Lipschitz continuous solution to (2.4) in [20, 9], in essence, is based on four results: The discrete weak duality principle (2.9), the discrete strong coercivity of $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, i.e.,

$$\frac{\alpha}{2}\|\Pi_h(v_h - u_h)\|^2_{L^2(\Omega)} \leq I_h(v_h) - I_h(u_h) \tag{3.1}$$

for all $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, where $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ is the minimum of $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, the strong duality principle (2.5), and the existence of appropriate primal and dual quasi-interpolants, guaranteed through the following two lemmas:

For the benefit of readability and without loss of generality, we assume for the remainder of this article, if not otherwise specified, that $\alpha = 1$.

Lemma 3.1 (Primal quasi-interpolant). *For every $u \in BV(\Omega)\cap L^\infty(\Omega)$, there exists a Crouzeix–Raviart function $\tilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ with the following properties:*

(P.1) $\|\nabla_h \tilde{u}_h\|_{L^1(\Omega;\mathbb{R}^d)} \leq |\mathrm{D}u|(\Omega)$,
(P.2) $\|u-\tilde{u}_h\|_{L^1(\Omega)} \leq c_{cr} h |\mathrm{D}u|(\Omega)$,
(P.3) $\|\tilde{u}_h\|_{L^\infty(\Omega)} \leq c_d \|u\|_{L^\infty(\Omega)}$,
(P.4) $I_h(\tilde{u}_h) \leq I(u) + 2c_d c_{cr}\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega)h - \frac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}$.

Proof. See [9, Lemma 4.4] or [20, Section 5]. □

Lemma 3.2 (Dual quasi–interpolant). *For every $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)\cap W^2_N(\mathrm{div};\Omega)$ such that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \leq 1$, there exists a Raviart–Thomas vector field $\tilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ with the following properties:*

(D.1) $\|\Pi_h \tilde{z}_h\|_{L^\infty(\Omega;\mathbb{R}^d)} \leq 1$,
(D.2) $D_h(\tilde{z}_h) \geq D(z) - c_{\mathcal{RT}}\|\nabla z\|_{L^\infty(\Omega;\mathbb{R}^{d\times d})}\|g\|_{L^2(\Omega)}\|\mathrm{div}(z)\|_{L^2(\Omega)}h - \frac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}$.

Proof. See [9, Lemma 4.5] or [20, Section 5]. □

While Lemma 3.1 does not impose restrictive assumptions on the minimum $u \in BV(\Omega)\cap L^2(\Omega)$, since already $u \in L^\infty(\Omega)$ if $g \in L^\infty(\Omega)$ (cf. [7, Proposition 10.2]), the required Lipschitz continuity of a solution $z \in W^\infty_N(\mathrm{div};\Omega)$ to (2.4) in Lemma 3.2 is often not fulfilled, cf. [12] or Section 2.7. We resort to the Lipschitz truncation technique to fulfill the Lipschitz continuity requirement on a solution to (2.4) in Lemma 3.2 at least in an approximate sense and, in this way, derive error estimates that depend directly on the Sobolev regularity of a solution $z \in W^{1,p}(\Omega;\mathbb{R}^d)$ to (2.4). The main advantage of this approach is that the Lipschitz truncation technique is based on local arguments, while regularization by convolution as in [20, Section 5.2], for example, operates highly non-local and, therefore, wipes out point-wise and/or local properties of a solution to (2.4) that potentially could have been incorporated. To be more precise, in [20, Section 5.2], the requirement $g \in BV(\Omega)\cap L^\infty(\Omega)$ was needed to estimate $\|\mathrm{div}(z)-\mathrm{div}(z_\varepsilon)\|_{L^1(\Omega)}$, where $z_\varepsilon := z\circ\omega_\varepsilon \in C^\infty(\mathbb{R}^d;\mathbb{R}^d)$, $\varepsilon>0$, denotes the convolution with a suitably scaled kernel $\omega_\varepsilon \in C^\infty_0(\mathbb{R}^d)$, by $\varepsilon|\mathrm{D}g|(\Omega)$. In contrast to that, if $z_\lambda \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$, $\lambda>0$, denotes the Lipschitz truncation of a suitable extension $\overline{z} \in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)$ of $z \in W^{1,p}(\Omega;\mathbb{R}^d)$, then we can exploit the particular properties $\nabla z_\lambda = \nabla\overline{z}$ in $\{z_\lambda = \overline{z}\}$ and $|\{z_\lambda \neq \overline{z}\}| \leq |\{\mathcal{M}(\nabla\overline{z}) > \lambda\}|$[3], to conclude that $\|\mathrm{div}(z)-\mathrm{div}(z_\lambda)\|_{L^1(\Omega)} \leq c\lambda^{1-p}\|\nabla z\|_{L^p(\Omega;\mathbb{R}^{d\times d})}$. In this way, we obtain the same rate $\mathcal{O}(h^{\frac{1}{4}})$ in [20, Section 5.2] without the assumption $g \in BV(\Omega)$ but need to require $z \in W^{1,3}(\Omega;\mathbb{R}^d)\cap L^\infty(\Omega;\mathbb{R}^d)$ instead of only $z \in W^\infty_N(\mathrm{div};\Omega)$.

Theorem 3.3 (Lipschitz truncation technique). *Let $z \in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)$, $p \in [1,\infty)$, and $\theta,\lambda>0$. Then, there is a Lipschitz continuous vector field $z_{\theta,\lambda} \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$ and a constant $c_{\mathrm{LT}}>0$, which does not depend on $p \in [1,\infty)$ and $\theta,\lambda>0$, such that the following statements apply:*

(LT.1) $\|z_{\theta,\lambda}\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)} \leq \theta$,
(LT.2) $\|\nabla z_{\theta,\lambda}\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^{d\times d})} \leq c_{\mathrm{LT}}\lambda$,
(LT.3) $|\{z_{\theta,\lambda} \neq z\}| \leq |\{\mathcal{M}(z) > \theta\}| + |\{\mathcal{M}(\nabla z) > \lambda\}|$,
(LT.4) $\nabla z_{\theta,\lambda} = \nabla z$ *in* $\{z_{\theta,\lambda} = z\}$.

Proof. See the first part of the proof of [22, Theorem 2.3] or [30, Section 1.3.3]. □

[3]Here, $\mathcal{M}\colon L^p(\mathbb{R}^d;\mathbb{R}^l)\to L^p(\mathbb{R}^d;\mathbb{R}^l)$, $d,l\in\mathbb{N}$, defined by $\mathcal{M}(f)(x) := \sup_{r>0} \fint_{B_r^d(x)} |f(y)|\,\mathrm{d}y$ for a.e. $x \in \mathbb{R}^d$ and all $f \in L^p(\mathbb{R}^d;\mathbb{R}^l)$, denotes the Hardy–Littlewood–Maximal operator.

A crucial property of the Lipschitz truncation technique for this article is that, similar to regularization by convolution, it does not increase the maximal length of a vector field.

Remark 3.4 (Maximal length preservation of the Lipschitz truncation technique)**.** *If $z\in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d;\mathbb{R}^d)$, $p\in[1,\infty)$, for some $\theta>0$ has the property $\|z\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)}\le\theta$, then*

$$\mathcal{M}(z)(x)=\sup_{r>0}\fint_{B_r^d(x)}|z(y)|\,dy\le\|z\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)}\le\theta\qquad\text{for a.e. } x\in\mathbb{R}^d,$$

i.e., $|\{\mathcal{M}(z)>\theta\}|=0$, which (cf. Theorem 3.3, (LT.3)) for arbitrary $\lambda>0$ yields

$$|\{z_{\theta,\lambda}\neq z\}|\le|\{\mathcal{M}(\nabla z)>\lambda\}|. \tag{3.2}$$

Through the combination of the p–type Tschebyscheff–Markoff–inequality, i.e., $|\{\mathcal{M}(\nabla z)>\lambda\}|\le\lambda^{-p}\|\mathcal{M}(\nabla z)\|^p_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}$, and the strong type (p,p)–estimate of the Hardy–Littlewood–Maximal operator (cf. [30, Theorem 1.22]), i.e., for $c_{\mathcal{M}}>0$[4]*, $\|\mathcal{M}(\nabla z)\|^p_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}\le c_{\mathcal{M}}\|\nabla z\|^p_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}$, we deduce from (3.2) that*

$$|\{z_{\theta,\lambda}\neq z\}|\le c_{\mathcal{M}}\lambda^{-p}\|\nabla z\|^p_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}. \tag{3.3}$$

By means of (3.3), also using Theorem 3.3, (LT.2) & (LT.4), we, then, deduce that

$$\begin{aligned}\|\nabla z_{\theta,\lambda}\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}&=\|\nabla z\chi_{\{z_{\theta,\lambda}=z\}}\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}+\|\nabla z_{\theta,\lambda}\chi_{\{z_{\theta,\lambda}\neq z\}}\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}\\&\le\|\nabla z\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}+c_{\mathrm{LT}}\lambda|\{z_{\theta,\lambda}\neq z\}|^{\frac{1}{p}}\\&\le\big(1+c_{\mathcal{M}}^{\frac{1}{p}}c_{\mathrm{LT}}\big)\|\nabla z\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}.\end{aligned}\tag{3.4}$$

Through the combination of Lemma 3.2, Theorem 3.3 and Remark 3.4, we arrive at the following result providing an admissible dual quasi-interpolant whose particular properties depend directly on the Sobolev regularity of a solution to (2.4).

Lemma 3.5 (Dual quasi–interpolant depending on Sobolev regularity for $\Gamma_N=\emptyset$)**.** *Let $g\in L^\infty(\Omega)$ and let $z\in W^{1,p}(\Omega;\mathbb{R}^d)\cap W^\infty(\mathrm{div};\Omega)$, $p\in[2,\infty)$, be such that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)}\le 1$. Then, there exists a Raviart–Thomas vector field $\tilde{z}_h\in\mathcal{RT}^0(\mathcal{T}_h)$ with the following properties:*

(D_p.1) $\|\Pi_h\tilde{z}_h\|_{L^\infty(\Omega;\mathbb{R}^d)}\le 1$.
(D_p.2) $D_h(\tilde{z}_h)\ge D(z)-c_p h^{\frac{p-2}{p-1}}-\frac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}$, *where*

$$\begin{aligned}c_p(z):=&\,2c_{\mathcal{RT}}\|g\|_{L^2(\Omega)}d^{\frac{1}{2}}\big(1+c_{\mathcal{M}}^{\frac{1}{2}}c_{\mathrm{LT}}\big)c_{\mathrm{E}}\|\nabla z\|_{L^2(\Omega;\mathbb{R}^{d\times d})}c_{\mathrm{LT}}\\&+8dc_{\mathrm{LT}}^2c_{\mathcal{M}}c_{\mathrm{E}}^p\|\nabla z\|^p_{L^p(\Omega;\mathbb{R}^{d\times d})}.\end{aligned}$$

Here, $c_{\mathrm{E}}>0$ is the Lipschitz constant of the lower-order extension operator $P:W^{1,q}(\Omega;\mathbb{R}^l)\to W^{1,q}(\mathbb{R}^d;\mathbb{R}^l)$, $q\in[1,\infty]$, constructed in [15, Section 9.2], which does not depend on $q\in[1,\infty]$.

Remark 3.6. *(i) The arguments remain valid for $\Gamma_D\neq\partial\Omega$ if for $z\in W^{1,p}(\Omega;\mathbb{R}^d)\cap W^\infty_N(\mathrm{div};\Omega)$, $p\in[2,\infty)$, such that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)}\le 1$, there exists an extension $\overline{z}\in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)$ with $\overline{z}|_\Omega=z$ and $\|\overline{z}\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)}\le 1$ and if for this extension, the Lipschitz truncation $\overline{z}_{1,\lambda}\in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$ from Theorem 3.3 satisfies $\overline{z}_{1,\lambda}\cdot n=0$ in Γ_N.*
(ii) In general, the constant $c_p>0$ deteriorates as $p\to\infty$, i.e., $c_p\to\infty$ $(p\to\infty)$.

[4]More precisely, one has $c_{\mathcal{M}}=2\big(\frac{p}{p-1}\big)^{\frac{1}{p}}5^{\frac{d}{p}}$, implying the limit behavior $c_{\mathcal{M}}\to 2$ for $(p\to\infty)$.

Proof. (of Lemma 3.5) Resorting to a lower-order extension operator, as, e.g., in [15, Theorem 9.7], we get some $\overline{z}\in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)$ with $\overline{z}|_\Omega = z$ and $\|\overline{z}\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)}\leq 1$. Denote by $z_\lambda := \overline{z}_{1,\lambda}\in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$, i.e., for $\theta=1$, the Lipschitz truncation of $\overline{z}\in W^{1,p}(\mathbb{R}^d;\mathbb{R}^d)$ in the sense of Theorem 3.3. Then, also using Remark 3.4, we get:

(α) $\|z_\lambda\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^d)}\leq 1$,
(β) $\|\nabla z_\lambda\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^{d\times d})}\leq c_{\mathrm{LT}}\lambda$,
(γ) $|\{z_\lambda\neq \overline{z}\}|\leq |\{\mathcal{M}(\nabla\overline{z})>\lambda\}|$,
(δ) $\nabla z_\lambda = \nabla\overline{z}$ in $\{z_\lambda=\overline{z}\}$.

For $z_\lambda|_\Omega\in W^{1,\infty}(\Omega;\mathbb{R}^d)$ we obtain, in analogy with Lemma 3.2, i.e., introducing $z_h^\lambda := (\gamma_h^\lambda)^{-1} I_{\mathcal{R}T} z_\lambda\in \mathcal{R}T^0(\mathcal{T}_h)$, where we define $\gamma_h^\lambda := 1 + c_{\mathcal{R}T}\|\nabla z_\lambda\|_{L^\infty(\Omega;\mathbb{R}^{d\times d})} h$, a dual quasi-interpolant $z_h^\lambda\in\mathcal{R}T^0(\mathcal{T}_h)$ such that both $\|\Pi_h z_h^\lambda\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq 1$ and

$$\begin{aligned} D_h(z_h^\lambda) \geq D(z_\lambda) - c_{\mathcal{R}T}\|g\|_{L^2(\Omega)}\|\mathrm{div}(z_\lambda)\|_{L^2(\Omega)}\|\nabla z_\lambda\|_{L^\infty(\Omega;\mathbb{R}^{d\times d})} h \\ - \tfrac{1}{2}\|g-g_h\|_{L^2(\Omega)}^2 . \end{aligned} \tag{3.5}$$

Then, on the basis of *(β)* and (3.4), we find that

$$\begin{aligned} \|\mathrm{div}(z_\lambda)\|_{L^2(\Omega)}\|\nabla z_\lambda\|_{L^\infty(\Omega;\mathbb{R}^{d\times d})} &\leq d^{\frac{1}{2}}\|\nabla z_\lambda\|_{L^2(\mathbb{R}^d;\mathbb{R}^{d\times d})}\|\nabla z_\lambda\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^{d\times d})} \\ &\leq d^{\frac{1}{2}}\big(1+c_{\mathcal{M}}^{\frac{1}{2}}c_{\mathrm{LT}}\big)\|\nabla\overline{z}\|_{L^2(\mathbb{R}^d;\mathbb{R}^{d\times d})} c_{\mathrm{LT}}\lambda \\ &\leq d^{\frac{1}{2}}\big(1+c_{\mathcal{M}}^{\frac{1}{2}}c_{\mathrm{LT}}\big)c_{\mathrm{E}}\|\nabla z\|_{L^2(\Omega;\mathbb{R}^{d\times d})} c_{\mathrm{LT}}\lambda . \end{aligned} \tag{3.6}$$

Using *(β)*, *(δ)* and (3.3), also assuming that $d^{\frac{1}{2}}c_{\mathrm{LT}}\lambda > \|\mathrm{div}(z)\|_{L^\infty(\Omega)} + 2\|g\|_{L^\infty(\Omega)}$, we further deduce that

$$\begin{aligned} |D(z)-D(z_\lambda)| &\leq \|(\mathrm{div}(z_\lambda)-\mathrm{div}(z))\chi_{\{z_\lambda\neq\overline{z}\}\cap\Omega}\|_{L^1(\Omega)}\|\mathrm{div}(z_\lambda)+\mathrm{div}(z)-2g\|_{L^\infty(\Omega)} \\ &\leq (d^{\frac{1}{2}}c_{\mathrm{LT}}\lambda + \|\mathrm{div}(z)\|_{L^\infty(\Omega)} + 2\|g\|_{L^\infty(\Omega)})^2 |\{z_\lambda\neq\overline{z}\}| \\ &\leq 4dc_{\mathrm{LT}}^2\lambda^2 c_{\mathcal{M}}\lambda^{-p}\|\nabla\overline{z}\|_{L^p(\mathbb{R}^d;\mathbb{R}^{d\times d})}^p \\ &\leq 4dc_{\mathrm{LT}}^2 c_{\mathcal{M}} c_{\mathrm{E}}^p\|\nabla z\|_{L^p(\Omega;\mathbb{R}^{d\times d})}^p\lambda^{2-p} . \end{aligned} \tag{3.7}$$

Therefore, on combining (3.5), (3.6) and (3.7), we observe that

$$\begin{aligned} D_h(z_h^\lambda) \geq D(z) - 4dc_{\mathrm{LT}}^2 c_{\mathcal{M}} c_{\mathrm{E}}^p\|\nabla z\|_{L^p(\Omega;\mathbb{R}^{d\times d})}^p\lambda^{2-p} - \tfrac{1}{2}\|g-g_h\|_{L^2(\Omega)}^2 \\ - c_{\mathcal{R}T}\|g\|_{L^2(\Omega)} d^{\frac{1}{2}}\big(1+c_{\mathcal{M}}^{\frac{1}{2}}c_{\mathrm{LT}}\big)c_{\mathrm{E}}\|\nabla z\|_{L^2(\Omega;\mathbb{R}^{d\times d})} c_{\mathrm{LT}}\lambda h . \end{aligned} \tag{3.8}$$

For $c_p>0$ defined as above and $\lambda = h^{-s}$, where $s>0$ is arbitrary, (3.8) yields

$$D_h(z_h^\lambda) \geq D(z) - \tfrac{c_p}{2}(h^{s(p-2)} + h^{1-s}) - \tfrac{1}{2}\|g-g_h\|_{L^2(\Omega)}^2 . \tag{3.9}$$

We have that $s(p-2) = 1-s$ if and only if $s=\frac{1}{p-1}=\frac{p'}{p}$. Thus, for $\lambda=h^{-s}$, $s=\frac{1}{p-1}$ and $\tilde{z}_h := z_h^\lambda\in\mathcal{R}T^0(\mathcal{T}_h)$, from (3.9), it follows that both *(D_p.1)* and *(D_p.2)* hold. □

Theorem 3.7 (Error estimate depending on the Sobolev regularity for $\Gamma_N=\emptyset$)**.** *Let $g\in L^\infty(\Omega)$, let $z\in W^{1,p}(\Omega;\mathbb{R}^d)\cap W^\infty(\mathrm{div};\Omega)$, $p\in[2,\infty)$, with $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq 1$ be maximal for $D: W^\infty(\mathrm{div};\Omega)\to\mathbb{R}\cup\{-\infty\}$, let $u\in\mathcal{A} := \{v\in BV(\Omega)\cap L^\infty(\Omega)\mid v=0$ in $L^1(\partial\Omega)\}$ be minimal for $I:\mathcal{A}\to\mathbb{R}$, and let $u_h\in\mathcal{S}^{1,cr}(\mathcal{T}_h)$ be minimal for $I_h:\mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}$. Then, there holds*

$$\|u-\Pi_h u_h\|_{L^2(\Omega)}^2 \leq c h^{\frac{p-2}{p-1}},$$

where $c>0$ depends only on the quantities c_{cr}, c_d, c_p, $\|u\|_{L^\infty(\Omega)}$ and $|\mathrm{D}u|(\Omega)$.

Proof. Combining the discrete strong coercivity of $I_h\colon\mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}$, i.e., (3.1), and the discrete weak duality principle $I_h(u_h)\geq D_h(\tilde{z}_h)$ for all $\tilde{z}_h\in\mathcal{R}T^0(\mathcal{T}_h)$ (cf. (2.9)), we obtain for all $\tilde{u}_h\in\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $\tilde{z}_h\in\mathcal{R}T^0(\mathcal{T}_h)$

$$\frac{1}{2}\|\Pi_h(\tilde{u}_h-u_h)\|^2_{L^2(\Omega)}\leq I_h(\tilde{u}_h)-I_h(u_h)\leq I_h(\tilde{u}_h)-D_h(\tilde{z}_h). \tag{3.10}$$

Resorting to Lemma 3.1, we obtain a function $\tilde{u}_h\in\mathcal{S}^{1,cr}_D(\mathcal{T})$ satisfying *(P.1)*–*(P.4)*. In addition, Lemma 3.5 yields a vector field $\tilde{z}_h\in\mathcal{R}T^0(\mathcal{T}_h)$ with *(D_p.1)* and *(D_p.2)*. Combining *(P.4)*, *(D_p.2)* and the strong duality principle $I(u)=D(z)$ (cf. (2.5)), we deduce from (3.10) that

$$\frac{1}{2}\|\Pi_h(\tilde{u}_h-u_h)\|^2_{L^2(\Omega)}\leq 2c_dc_{cr}\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega)h+c_ph^{\frac{p-2}{p-1}}, \tag{3.11}$$

where $c_p>0$ is as in Lemma 3.5. Since $\tilde{u}_h-\Pi_h\tilde{u}_h=\nabla_h\tilde{u}_h\cdot(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})$ in $\mathcal{L}^1(\mathcal{T}_h)$, using *(P.1)*, *(P.3)* and $\|\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d}\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq h$, we find that

$$\begin{aligned}\|\tilde{u}_h-\Pi_h\tilde{u}_h\|^2_{L^2(\Omega)}&\leq 2\|\tilde{u}_h\|_{L^\infty(\Omega)}\|\nabla_h\tilde{u}_h\|_{L^1(\Omega;\mathbb{R}^d)}\|\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d}\|_{L^\infty(\Omega;\mathbb{R}^d)}\\&\leq 2c_d\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega)h.\end{aligned} \tag{3.12}$$

Using *(P.2)*, *(P.3)*, *(L0.1)* and proceeding as for (3.12), we further obtain that

$$\begin{aligned}\|u-\Pi_h\tilde{u}_h\|^2_{L^2(\Omega)}&\leq\|u-\Pi_h\tilde{u}_h\|_{L^\infty(\Omega)}\big(\|u-\tilde{u}_h\|_{L^1(\Omega)}+\|\tilde{u}_h-\Pi_h\tilde{u}_h\|_{L^1(\Omega)}\big)\\&\leq(1+c_d)\|u\|_{L^\infty(\Omega)}\big(c_{cr}+1\big)|\mathrm{D}u|(\Omega)h.\end{aligned} \tag{3.13}$$

Finally, combining (3.11), (3.12) and (3.13), we conclude the claimed error bound. □

4. Error estimates for discontinuous dual solutions

In this section, we prove error estimates for the ROF model without explicitly imposing that the dual solution possesses Sobolev regularity. Recall that, in general, a solution of the dual ROF model only needs to satisfy $z\in W^2_N(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$ with $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq 1$. The following lemma gives general assumptions on the dual solution for which it is still possible to construct a suitable dual quasi-interpolant.

Lemma 4.1 (Dual quasi-interpolant for non–Sobolev vector fields)**.** *Let $g\in L^2(\Omega)$ and $z\in W^2_N(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$. If $\|\Pi_hI_{\mathcal{R}T}z\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq 1+\kappa(h)$ for $\kappa(h)\geq 0$, then the re-scaled vector field $\tilde{z}_h:=\frac{1}{\gamma_h}I_{\mathcal{R}T}z\in\mathcal{R}T^0_N(\mathcal{T}_h)$, where $\gamma_h:=1+\kappa(h)>0$, has the following properties:*

(D.1)* $\|\Pi_h\tilde{z}_h\|_{L^\infty(\Omega;\mathbb{R}^d)}\leq 1$.
(D.2)* $D_h(\tilde{z}_h)\geq D(z)-\kappa(h)\|g\|_{L^2(\Omega)}\|\mathrm{div}(z)\|_{L^2(\Omega)}-\frac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}$.

Proof. Claim *(D.1*)* is evident. Resorting to $\mathrm{div}(I_{\mathcal{R}T}z)=\Pi_h(\mathrm{div}(z))$ in $\mathcal{L}^0(\mathcal{T}_h)$, we deduce that $\mathrm{div}(\tilde{z}_h)+g_h=\Pi_h(\frac{1}{\gamma_h}\mathrm{div}(z)+g)$ in $\mathcal{L}^0(\mathcal{T}_h)$ and, hence, also using $\|g-g_h\|^2_{L^2(\Omega)}=\|g\|^2_{L^2(\Omega)}-\|g_h\|^2_{L^2(\Omega)}$, $I_{K_1(0)}(\Pi_h\tilde{z}_h)=0$ and Jensen's inequality, that

$$\begin{aligned}D_h(\tilde{z}_h)&=-\tfrac{1}{2}\big\|\Pi_h(\tfrac{1}{\gamma_h}\mathrm{div}(z)+g)\big\|^2_{L^2(\Omega)}-\tfrac{1}{2}\|g_h\|^2_{L^2(\Omega)}\\&\geq-\tfrac{1}{2}\big\|\tfrac{1}{\gamma_h}\mathrm{div}(z)+g\big\|^2_{L^2(\Omega)}+\tfrac{1}{2}\|g\|^2_{L^2(\Omega)}-\tfrac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}\\&\geq-\tfrac{1}{2}\tfrac{1}{\gamma_h^2}\|\mathrm{div}(z)\|^2_{L^2(\Omega)}+\tfrac{1}{\gamma_h}(g,\mathrm{div}(z))_{L^2(\Omega)}-\tfrac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}\\&\geq D(z)-(1-\tfrac{1}{\gamma_h})(g,\mathrm{div}(z))_{L^2(\Omega)}-\tfrac{1}{2}\|g-g_h\|^2_{L^2(\Omega)}.\end{aligned}$$

Finally, using that $\frac{1}{\gamma_h^2}\leq 1$ and $1-\frac{1}{\gamma_h}\leq\kappa(h)$, we conclude that *(D.2*)* holds. □

Theorem 4.2 (Error estimate for discontinuous dual solution)**.** *Let $g \in L^\infty(\Omega)$, let $z \in W_N^\infty(\mathrm{div};\Omega)$ be maximal for $D : W_N^\infty(\mathrm{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$ with the same properties as in Lemma 4.1, let $u \in BV(\Omega)\cap L^\infty(\Omega)$ minimal for $I : BV(\Omega)\cap L^2(\Omega) \to \mathbb{R}$, and let $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ minimal for $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$. Then, we have that*

$$\|u - \Pi_h u_h\|_{L^2(\Omega)}^2 \le c \max\{\kappa(h), h\}.$$

where $c > 0$ depends only on the quantities c_{cr}, c_d, $\|u\|_{L^\infty(\Omega)}$, and $|\mathrm{D}u|(\Omega)$.

Proof. Using the discrete strong coercivity of $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, i.e., (3.1), and the discrete weak duality principle $I_h(u_h) \ge D_h(\tilde{z}_h)$ for all $\tilde{z}_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ (cf. (2.9)), we obtain for all $\tilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $\tilde{z}_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$

$$\frac{1}{2}\|\Pi_h(\tilde{u}_h - u_h)\|_{L^2(\Omega)}^2 \le I_h(\tilde{u}_h) - I_h(u_h) \le I_h(\tilde{u}_h) - D_h(\tilde{z}_h). \tag{4.1}$$

Resorting to Lemma 3.1, we obtain a function $\tilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T})$ satisfying *(P.1)*–*(P.4)*. In addition, Lemma 4.1 yields a vector field $\tilde{z}_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ with *(D.1*)* and *(D.2*)*. Then, using *(P.4)*, *(D.2*)* and the strong duality principle $I(u) = D(z)$ (cf. (2.5)), we deduce from (4.1) that

$$\frac{1}{2}\|\Pi_h(\tilde{u}_h - u_h)\|_{L^2(\Omega)}^2 \le 2c_d c_{cr}\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega)h + \kappa(h)\|g\|_{L^2(\Omega)}\|\mathrm{div}(z)\|_{L^2(\Omega)}.$$

Hence, incorporating (3.12) and (3.13), we conclude the claimed error bound. □

A sufficient condition for a solution to (2.4) to guarantee the quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ is element-wise Lipschitz continuity. In addition, if a solution to (2.4) is only element-wise α–Hölder continuous, it is, however, possible to derive the rate $\mathcal{O}(h^{\frac{\alpha}{2}})$.

Lemma 4.3 (Dual quasi-interpolant for element-wise α–Hölder vector fields)**.** *Let $g \in L^2(\Omega)$ and let $z \in W_N^2(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$ be such that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1$. Furthermore, assume that there exist constants $\alpha \in [0,1]$ and $c_\alpha > 0$ such that for all $T \in \mathcal{T}_h$, it holds $z|_T \in C^{0,\alpha}(T;\mathbb{R}^d)$ with*

$$|z(x) - z(y)| \le c_\alpha |x - y|^\alpha \tag{4.2}$$

for all $x, y \in T$. Then, the assumptions in Lemma 4.1 are satisfied with $\kappa(h) = \mathcal{O}(h^\alpha)$.

Remark 4.4. *Lemma 4.3 is of particular interest if the discontinuity set J_z of a piece-wise regular (piece-wise Lipschitz or piece-wise α–Hölder continuous) vector field $z \in W_N^2(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$ is resolved by the triangulation, i.e., $J_z \subseteq \bigcup_{S\in\mathcal{S}_h} S$.*

Proof. (of Lemma 4.3) We need to check that $\|\Pi_h I_{\mathcal{RT}} z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1 + \kappa(h)$ for some $\kappa(h) > -1$ with $\kappa(h) = \mathcal{O}(h^\alpha)$. Note that $I_{\mathcal{RT}}(z(x_T)) = z(x_T)$ in T for all $T \in \mathcal{T}_h$, which results from $\mathrm{div}(I_{\mathcal{RT}}(z(x_T))) = \Pi_h(\mathrm{div}(z(x_T))) = 0$ in T for all $T \in \mathcal{T}_h$. Using this, *(RT.2)*, (4.2) and that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1$, we deduce that for all $T \in \mathcal{T}_h$

$$\begin{aligned} |(I_{\mathcal{RT}} z)(x_T)| &\le |I_{\mathcal{RT}}(z - z(x_T))(x_T)| + |z(x_T)| \\ &\le \|I_{\mathcal{RT}}(z - z(x_T))\|_{L^\infty(T;\mathbb{R}^d)} + 1 \\ &\le c_{\mathcal{RT}}\|z - z(x_T)\|_{L^\infty(T;\mathbb{R}^d)} + 1 \\ &\le c_{\mathcal{RT}} \sup_{x\in T}|x - x_T|^\alpha + 1 \\ &\le c_{\mathcal{RT}} c_\alpha h_T^\alpha + 1, \end{aligned} \tag{4.3}$$

i.e., setting $\kappa(h) := c_{\mathcal{RT}} c_\alpha h^\alpha$, we conclude that $\|\Pi_h I_{\mathcal{RT}} z\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1 + \kappa(h)$. □

Theorem 4.5 (Error estimate for element-wise α–Hölder dual solution)**.** *Let $z \in W_N^2(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d)$ be maximal for $D : W_N^2(\mathrm{div};\Omega)\cap L^\infty(\Omega;\mathbb{R}^d) \to \mathbb{R}\cup\{-\infty\}$ with the same properties as in Lemma 4.3, let $u \in BV(\Omega)\cap L^\infty(\Omega)$ minimal for $I : BV(\Omega)\cap L^2(\Omega) \to \mathbb{R}$ and let $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ minimal for $I_h : \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$. Then, we have that*

$$\|u - \Pi_h u_h\|_{L^2(\Omega)}^2 \leq c h^\alpha.$$

where $c>0$ depends only on the quantities c_{cr}, c_d, c_α, $\|u\|_{L^\infty(\Omega)}$, and $|\mathrm{D}u|(\Omega)$.

Proof. Follows from Theorem 4.2 by resorting to Lemma 4.3. □

Remark 4.6 (Comparison of Theorem 3.7 and Theorem 4.5)**.** *(i) If $\alpha = 1$, then Theorem 4.5 extends the results [9, Proposition 4.2] and [20, Sectioin 5.1.1] to the case of an existing element-wise Lipschitz continuous solution to* (2.4)*.*

(ii) If $p > d$ in Theorem 3.7, then $z \in W^{1,p}(\Omega;\mathbb{R}^d)$ satisfies $z \in C^{0,\alpha}(\overline{\Omega};\mathbb{R}^d)$ for $\alpha = 1 - \frac{d}{p}$ by Sobolev's embedding theorem [15, Corollary 9.14]. As a result, Theorem 4.5 in the particular case $\Gamma_N = \emptyset$ is applicable and yields the rate $\mathcal{O}(h^{\frac{\alpha}{2}})$. On the other hand, Theorem 3.7 yields the slightly improved rate $\mathcal{O}(h^{\frac{p-2}{2(p-1)}})$, which gives the impression that Theorem 3.7 is utterly superior to Theorem 4.5. Nevertheless, the major strength of Theorem 4.5 – and equally of Lemma 4.3 – is that it is also applicable when it is unclear whether a solution to (2.4) *with Sobolev regularity is available. This allows us to justify analytically the quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ for the setting in Section 2.7 at least for the particular case that the discontinuity set J_z is resolved by the triangulation, i.e., $J_z \subseteq \bigcup_{S\in\mathcal{S}_h} S$, cf. Example 5.2 and Example 5.3.*

If the discontinuity set of a solution to (2.4) is not resolved by the triangulation, then, apparently, Theorem 4.5 does not apply. In this case, however, the following argument applies, which exploits that for $g \in L^\infty(\Omega)$, we have that $\mathrm{div}(z) \in L^\infty(\Omega)$, which to some extent can serve as a substitute for $\nabla z \in L^\infty(\Omega;\mathbb{R}^{d\times d})$.

Remark 4.7 (Optimal dual quasi-interpolant for non–Lipschitz vector fields)**.** *Let $z \in W_N^\infty(\mathrm{div};\Omega)$ be such that $\|z\|_{L^\infty(\Omega;\mathbb{R}^d)} \leq 1$. Furthermore, assume that there exists a constant $\tilde{c}_z > 0$ such that for all $T \in \mathcal{T}_h$, there exists some $\widetilde{x}_T \in T$ such that*

$$|(I_{\mathcal{RT}} z)(\widetilde{x}_T)| \leq 1 + \tilde{c}_z h, \tag{4.4}$$

For each $T \in \mathcal{T}_h$, since $I_{\mathcal{RT}} z \in \mathcal{RT}_N^0(\mathcal{T}_h) \subseteq \mathcal{L}^1(\mathcal{T}_h)^d$, we have that

$$(I_{\mathcal{RT}} z)(x) = (I_{\mathcal{RT}} z)(x_T) + d^{-1}\mathrm{div}(I_{\mathcal{RT}} z)(x - x_T) \tag{4.5}$$

for all $x \in T$. Thus, resorting to $\mathrm{div}(I_{\mathcal{RT}} z) = \Pi_h(\mathrm{div}(z))$ in $\mathcal{L}^0(\mathcal{T}_h)$, also using (4.4) *and (L0.1) in* (4.5) *at $x = \widetilde{x}_T \in T$, we conclude that*

$$\begin{aligned} \|\Pi_h I_{\mathcal{RT}} z\|_{L^\infty(T;\mathbb{R}^d)} &\leq |(I_{\mathcal{RT}} z)(\tilde{x}_T)| + d^{-1}\|\Pi_h(\mathrm{div}(z))\|_{L^\infty(T)}|\tilde{x}_T - x_T| \\ &\leq 1 + \tilde{c}_z h + d^{-1}\|\mathrm{div}(z)\|_{L^\infty(T)} h_T, \end{aligned}$$

i.e., we have that $\|\Pi_h I_{\mathcal{RT}} z\|_{L^\infty(T;\mathbb{R}^d)} \leq 1 + \big(\tilde{c}_z + d^{-1}\|\mathrm{div}(z)\|_{L^\infty(\Omega)}\big)h$.

The following remark discusses particular sufficient conditions for (4.4) on a vector field $z \in W_N^\infty(\mathrm{div};\Omega)$ that is piece-wise Lipschitz continuous, such as, e.g., that its discontinuity set J_z is approximated by $\mathcal{T}_h$, $h > 0$, with rate $\mathcal{O}(h)$ or that $|z| < 1$ along J_z while, simultaneously, its jump $[\![z]\!]$ over J_z remains small.

On the other hand, this remark finds that (4.4) cannot be expected, in general, for piece-wise Lipschitz continuous vector fields, even in generic situations.

Remark 4.8 (Sufficient conditions for (4.4))**.** *Let $d=2$ and $z\in W^{\infty}_N(\mathrm{div};\Omega)$ with $\|z\|_{L^\infty(\Omega;\mathbb{R}^2)}\leq 1$ be piece-wise Lipschitz continuous, i.e., there exist open $\Omega_i\subseteq\Omega$, $i=1,\dots,m$, $m\in\mathbb{N}$, with $z|_{\Omega_i}\in W^{1,\infty}(\Omega_i;\mathbb{R}^2)$ for all $i=1,\dots,m$ and $\overline{\Omega}=\bigcup_{i=1}^m\overline{\Omega_i}$. Next, we fix an arbitrary $T\in\mathcal{T}_h$. Then, we need to distinguish two cases:*

(i) Assume that $T\subseteq\overline{\Omega_i}$ for some $i=1,\dots,m$. Then, we deduce along the lines of the proof of (4.3) that

$$|(I_{\mathcal{R}T}z)(x_T)|\leq 1+c_{\mathcal{R}T}\|\nabla z\|_{L^\infty(\Omega_i;\mathbb{R}^{2\times 2})}h_T,$$

i.e., $|(I_{\mathcal{R}T}z)(x_T)|\leq 1+c_zc_{\mathcal{R}T}h_T$, where $c_z:=\max_{i=1,\dots,m}\|\nabla(z|_{\Omega_i})\|_{L^\infty(\Omega;\mathbb{R}^{2\times 2})}$.

(ii) Assume that there exists an interface $\gamma=\partial\Omega_a\cap\partial\Omega_b$ for some $a,b=1,\dots,m$ such that $\mathrm{int}(T)\cap\gamma\neq\emptyset$[5]. As $z|_{\Omega_a}\in W^{1,\infty}(\Omega_a;\mathbb{R}^2)$ and $z|_{\Omega_b}\in W^{1,\infty}(\Omega_b;\mathbb{R}^2)$, without loss of generality, we may assume that $\gamma\subseteq b_\gamma+\mathbb{R}t_\gamma$ for some $b_\gamma\in\mathbb{R}^2$ and $t_\gamma\in\mathbb{S}^1$. Next, fix $x_\gamma\in\gamma$ and set $z_a:=(z|_{\overline{\Omega_a}})(x_\gamma)$, $z_b:=(z|_{\overline{\Omega_b}})(x_\gamma)\in\mathbb{R}^2$. Then, for $i\in\{a,b\}$, it holds

$$|z_i-(z|_{\overline{\Omega_i}})(x)|\leq\|\nabla(z|_{\Omega_i})\|_{L^\infty(\Omega_i;\mathbb{R}^{2\times 2})}|x_\gamma-x|\leq c_zh_T\quad\text{for all } x\in T\cap\overline{\Omega_i}.$$

Furthermore, if $n_\gamma\in\mathbb{S}^1$ denotes a unit normal to $t_\gamma\in\mathbb{S}^1$, i.e., $n_\gamma\cdot t_\gamma=0$, then, taking into account that $z\in W^\infty_N(\mathrm{div};\Omega)$, we find that $z_a\cdot n_\gamma=z_b\cdot n_\gamma$. Thus, if we define $z_T(x):=z_a$ for $x\in T\cap\overline{\Omega_a}$ and $z_T(x):=z_b$ for $x\in T\cap\overline{\Omega_b}$, cf. Figure 1, ($\alpha$), then $z_T\in W^\infty(\mathrm{div};T)$ and, owing to (RT.2),

$$\|I_{\mathcal{R}T}z_T-I_{\mathcal{R}T}z\|_{L^\infty(T;\mathbb{R}^2)}\leq c_{\mathcal{R}T}\|z_T-z\|_{L^\infty(T;\mathbb{R}^2)}\leq c_{\mathcal{R}T}c_zh_T,$$

i.e., we have that

$$\|I_{\mathcal{R}T}z\|_{L^\infty(T;\mathbb{R}^2)}\leq\|I_{\mathcal{R}T}z_T\|_{L^\infty(T;\mathbb{R}^2)}+c_{\mathcal{R}T}c_zh_T.\tag{4.6}$$

As a result of (4.6), it is sufficient to prove $\|I_{\mathcal{R}T}z_T\|_{L^\infty(T;\mathbb{R}^2)}\leq 1+\mathcal{O}(h)$ to conclude that $\|I_{\mathcal{R}T}z\|_{L^\infty(T;\mathbb{R}^2)}\leq 1+\mathcal{O}(h)$. Because, owing to Lemma 4.1 (ii), it holds

$$\mathrm{div}(I_{\mathcal{R}T}z_T)=\Pi_h(\mathrm{div}(z_T))=0\quad\text{ in } T,$$

where $I_{\mathcal{R}T}z_T:=\sum_{S\in\mathcal{S}_h;S\subseteq\partial T}z_T\cdot n_S\psi_S$, it even holds $I_{\mathcal{R}T}z_T\equiv\mathrm{const}$ in T. Next, we denote by $S_1\in\mathcal{S}_h$ a side of $T\in\mathcal{T}_h$ such that $S_1\cap\gamma\neq\emptyset$ and by $S_2\in\mathcal{S}_h$ the side of $T\in\mathcal{T}_h$ such that $S_2\cap\gamma=\emptyset$. Let $n_1,n_2\in\mathbb{S}^1$ denote the corresponding unit normal vectors to $S_1,S_2\in\mathcal{T}_h$, resp., cf. Figure 1, (α). Then, it holds

$$\left.\begin{aligned}I_{\mathcal{R}T}z_T\cdot n_1&=\int_{S_1}z_T\cdot n_1\,\mathrm{d}s=\frac{|S_1\cap\Omega_b|}{|S_1|}z_b\cdot n_1+\frac{|S_1\cap\Omega_a|}{|S_1|}z_a\cdot n_1,\\ I_{\mathcal{R}T}z_T\cdot n_2&=\int_{S_2}z_T\cdot n_2\,\mathrm{d}s=z_b\cdot n_2.\end{aligned}\right\}\tag{4.7}$$

Introducing $\rho:=|S_1\cap\Omega_b|/|S_1|\in[0,1]$ as well as $M_T:=(n_1,n_2)\in\mathbb{R}^{2\times 2}$, also exploiting that $z_a=z_b+((z_a-z_b)\cdot t_\gamma)t_\gamma$, where we used that $z_a\cdot n_\gamma=z_b\cdot n_\gamma$, the system (4.7) can be rewritten as

$$M_T^\top I_{\mathcal{R}T}z_T=M_T^\top z_b+(1-\rho)((z_a-z_b)\cdot t_\gamma)(t_\gamma\cdot n_1)e_1,$$

[5]Apparently, we should also take into account the case in which $T\in\mathcal{T}_h$ is intersected by two or more interfaces. However, for the benefit of readability, we limit ourselves to this simplified case.

i.e., since $M_T^\top \in \mathbb{R}^{2\times 2}$ is a regular matrix, we find that

$$I_{\mathcal{R}T} z_T = z_b + (1-\rho)((z_a - z_b)\cdot t_\gamma)(t_\gamma \cdot n_1) M_T^{-\top} e_1. \tag{4.8}$$

Resorting to the formula (4.8), we can derive special cases that imply (4.4):

(ii.a) If $t_\gamma \cdot n_1 = \mathcal{O}(h)$, i.e., $(b_\gamma + \mathbb{R}t_\gamma)\cap T$ approximates S_1 with rate $\mathcal{O}(h)$, cf. Figure 1, (β), then $|I_{\mathcal{R}T} z_T| \le |z_b| + \mathcal{O}(h) \le 1 + \mathcal{O}(h)$.

(ii.b) If $1-\rho = \mathcal{O}(h)$, i.e., $S_1 \cap \Omega_b$ approximates S_1 with rate $\mathcal{O}(h)$, cf. Figure 1, (γ), then $|I_{\mathcal{R}T} z_T| \le |z_b| + \mathcal{O}(h) \le 1 + \mathcal{O}(h)$.

Apparently, (ii.a) and (ii.b) describe the particular case in which the discontinuity set J_z is not resolved by the triangulation but approximated with rate $\mathcal{O}(h)$.

(ii.c) If we have that both $|z_b| < 1$ and $(z_a - z_b)\cdot t_\gamma$ is sufficiently small, i.e., such that $|(1-\rho)((z_a - z_b)\cdot t_\gamma) M_T^{-\top}(t_\gamma \cdot n_1)e_1| \le 1 - |z_b| + \mathcal{O}(h)$, then $|I_{\mathcal{R}T} z_T| \le |z_b| + 1 - |z_b| + \mathcal{O}(h) = 1 + \mathcal{O}(h)$.

(ii.d) If T is nearly right-angled, so that M_T is approximately an orthogonal matrix, i.e., $M_T^{-\top} = M_T + \mathcal{O}(h)$, and $t_\gamma = \pm n_1 + \mathcal{O}(h)$, then, using that $z_a \cdot n_2 = z_b \cdot n_2 + \mathcal{O}(h)$ because $n_\gamma = \pm n_2 + \mathcal{O}(h)$, we deduce that $z_b = (z_b\cdot n_1)n_1 + (1-\rho)(z_a\cdot n_2)n_2 + \rho(z_b \cdot n_2)n_2 + \mathcal{O}(h)$ and, thus, that

$$\begin{aligned} I_{\mathcal{R}T} z_T &= z_b + (1-\rho)((z_a - z_b)\cdot n_1)n_1 + \mathcal{O}(h) \\ &= \rho((z_b\cdot n_1)n_1 + \rho(z_b\cdot n_2)n_2) \\ &\quad + (1-\rho)((z_a\cdot n_1)n_1 + (z_a \cdot n_2)n_2) + \mathcal{O}(h) \\ &= (1-\rho)z_a + \rho z_b + \mathcal{O}(h), \end{aligned}$$

which implies that $|I_{\mathcal{R}T} z_T| \le (1-\rho)|z_a| + \rho|z_b| + \mathcal{O}(h) \le 1 + \mathcal{O}(h)$.

More generally, the sub-cases (ii.a)–(ii.d) can occur in combination so that the conclusion holds under significantly weaker conditions on the individual factors. On the other hand, the formula (4.8), simultaneously, demonstrates that $\|I_{\mathcal{R}T} z_T\|_{L^\infty(T;\mathbb{R}^2)} \le 1 + \mathcal{O}(h)$ and, therefore, also $\|I_{\mathcal{R}T} z\|_{L^\infty(T;\mathbb{R}^2)} \le 1 + \mathcal{O}(h)$ cannot be expected in general, even in generic situations.

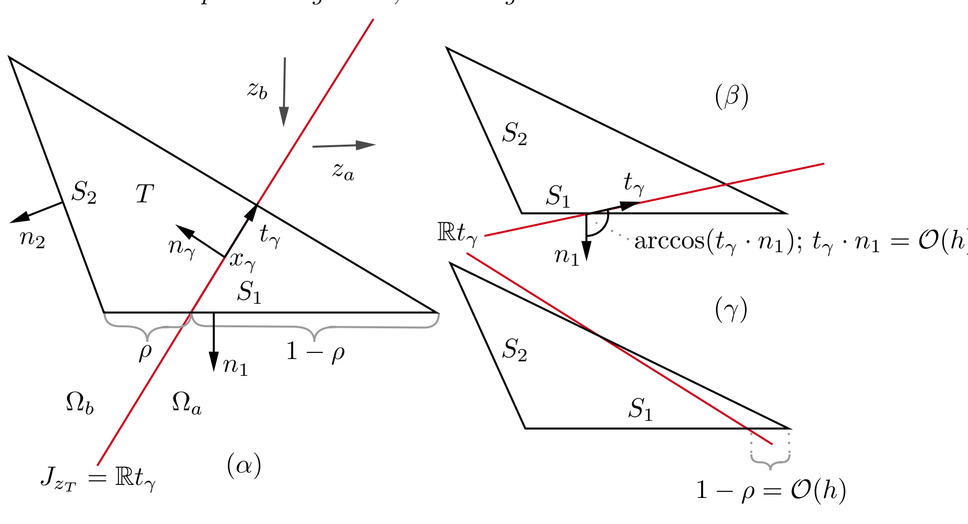

Figure 1: Sketch of the construction as described in Remark 4.8 with a discontinuity set J_{z_T} intersecting an element T. Part *(α)* depicts the setting of Remark 4.8, while part *(β)* and part *(γ)* illustrate the cases *(ii.a)* and *(ii.b)* in Remark 4.8.

5. Numerical experiments

In this section, we verify the theoretical findings of Section 4 via numerical experiments. To compare approximations to an exact solution, we impose Dirichlet boundary conditions on $\Gamma_D = \partial\Omega$, though an existence theory is difficult to establish, in general. However, the error estimates derived in Section 4 carry over verbatimly with $\Gamma_N = \emptyset$ provided that a minimizer exists.

All experiments were conducted using the finite element software FEniCS, cf. [29]. All graphics are generated using the Matplotlib library, cf. [27].

5.1 Experimental convergence rates

All computations are based on using the regularized discrete ROF functional, i.e., for $\varepsilon > 0$ and $g \in L^2(\Omega)$, the functional $I_h^\varepsilon : \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, defined by

$$I_h^\varepsilon(v_h) := \||\nabla v_h|_\varepsilon\|_{L^1(\Omega)} + \frac{\alpha}{2}\|\Pi_h(v_h - g)\|_{L^2(\Omega)}^2 \tag{5.1}$$

for all $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, where $|\cdot|_\varepsilon \in C^1(\mathbb{R}^d)$ is the regularized modulus, defined by $|a|_\varepsilon := (|a|^2+\varepsilon^2)^{\frac{1}{2}}$ for all $a \in \mathbb{R}^d$ and $\varepsilon > 0$. On the basis of $0 \le |a|_\varepsilon - |a| \le \varepsilon$ for all $a \in \mathbb{R}^d$ and $\varepsilon > 0$, for the minima $u_h, u_h^\varepsilon \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ of $I_h, I_h^\varepsilon : \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, resp., there holds

$$\frac{\alpha}{2}\|\Pi_h(u_h - u_h^\varepsilon)\|_{L^2(\Omega)}^2 \le \varepsilon|\Omega|.$$

Thus, in order to bound the error $\|u - \Pi_h u_h\|_{L^2(\Omega)}$, it suffices to determine the error $\|u - \Pi_h u_h^\varepsilon\|_{L^2(\Omega)}$, e.g., for $\varepsilon = h$. The iterative minimization of $I_h^h : \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, i.e., for $\varepsilon = h$, is realized using the unconditionally strongly stable semi-implicit discretized L^2–gradient flow from [10], see also [8, Section 5].

Algorithm 5.1 (Semi-implicit discretized L^2–gradient flow)**.** *Let $g_h \in \mathcal{L}^0(\mathcal{T}_h)$ and choose $\tau, \varepsilon_{stop} > 0$. Moreover, let $u_h^0 \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and set $k = 1$. Then, for $k \ge 1$:*

(i) Compute $u_h^k \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ such that for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, there holds

$$\big(d_t u_h^k, v_h\big)_{L^2(\Omega)} + \left(\frac{\nabla_h u_h^k}{\big|\nabla_h u_h^{k-1}\big|_h}, \nabla_h v_h\right)_{L^2(\Omega;\mathbb{R}^d)} + \alpha\big(\Pi_h u_h^k - g_h, \Pi_h v_h\big)_{L^2(\Omega)} = 0,$$

where $d_t u_h^k := \frac{1}{\tau}(u_h^k - u_h^{k-1})$ denotes the backward difference quotient.

(ii) Stop if $\big\|d_t u_h^k\big\|_{L^2(\Omega)} \le \varepsilon_{stop}$; otherwise, increase $k \to k+1$ and continue with (i).

It is shown in [10, Proposition 3.4] and [8, Proposition 5.3], that Algorithm 5.1 is unconditionally strongly stable, energy decreasing as well as converging, i.e., stops after finitely many iteration steps. To be more specific, for arbitrary $l \in \mathbb{N}$, one has the discrete energy estimate

$$I_h^h\big(u_h^l\big) + \tau\sum_{k=1}^{l}\big\|d_t u_h^k\big\|_{L^2(\Omega)}^2 + \frac{\tau^2}{2}\sum_{k=1}^{l}\int_\Omega \frac{\big|d_t\nabla u_h^k\big|^2 + \big(d_t\big|\nabla u_h^k\big|_h\big)^2}{\big|\nabla u_h^{k-1}\big|_h}\,\mathrm{d}x \le I_h^h\big(u_h^0\big),$$

which mainly results from $d_t\big|\nabla u_h^k\big|_h = \frac{1}{2}\big|\nabla u_h^{k-1}\big|_h^{-1}\big(d_t\big|\nabla u_h^k\big|^2 - \big(d_t\big|\nabla u_h^k\big|_h\big)^2\big)$.

We will always employ the h–independent step-size $\tau = 1$ but the h–dependent stopping criteria $\big\|d_t u_h^k\big\|_{L^2(\Omega)} \le \varepsilon_{stop}^h := \frac{h}{20}$, i.e., $\big\|u_h^k - u_h^{k-1}\big\|_{L^2(\Omega)} \le \frac{h}{20}$ as $\tau = 1$.

Example 5.2 (Two disks problem). *Let* $\Omega=(-1,1)^2\subseteq\mathbb{R}^2$, $r=0.4$, $\alpha=10$, *and* $\tilde{g}:=g\circ\Phi\in BV(\Omega)\cap L^\infty(\Omega)$, *where* $g:=\chi_{B_r^2(re_1)}-\chi_{B_r^2(-re_1)}\in BV(\Omega)\cap L^\infty(\Omega)$ *and for some angle* $\phi\in[0,2\pi]$ *and some vector* $b_\gamma=(b_1,b_2)^\top\in\mathbb{R}^2$,

$$\Phi(x):=\begin{bmatrix}\cos(\phi)(x_1-b_1)+\sin(\phi)(x_2-b_2)\\ \cos(\phi)(x_2-b_2)-\sin(\phi)(x_1-b_1)\end{bmatrix}\tag{5.2}$$

for all $x=(x_1,x_2)^\top\in\mathbb{R}^2$, *i.e.,* $\Phi:\mathbb{R}^2\to\mathbb{R}^2$ *performs a rotation by* ϕ *and a shift by* b. *The same argumentation as in the proof of Proposition 2.1 demonstrates that the corresponding primal solution is given via* $\tilde{u}:=u\circ\Phi=(1-\frac{2}{\alpha r})\tilde{g}\in BV(\Omega)\cap L^\infty(\Omega)$, *where* $u:=(1-\frac{2}{\alpha r})g\in BV(\Omega)\cap L^\infty(\Omega)$, *cf. Proposition 2.1.*

For $z\in W^\infty(\mathrm{div};\Omega)$ *defined as in Proposition 2.3, we define the vector field* $\tilde{z}:=\det(D\Phi)(D\Phi)^{-1}z\circ\Phi=(D\Phi)^{-1}z\circ\Phi\in W^\infty(\mathrm{div};\Omega)$. *Then, resorting to properties of the contra-variant Piola transform, cf. [13, (2.1.71)], we find that*

$$\mathrm{div}(\tilde{z})=\det(D\Phi)\mathrm{div}(z)\circ\Phi=\mathrm{div}(z)\circ\Phi=\alpha(\tilde{u}-\tilde{g})\quad\text{ in }\Omega.\tag{5.3}$$

We define the decomposition $\Omega_\Phi^+:=\Omega\cap\Phi(\mathbb{R}_{>0}\times\mathbb{R})$ *and* $\Omega_\Phi^-:=\Omega\cap\Phi(\mathbb{R}_{<0}\times\mathbb{R})$. *Then, using that* $\tilde{u}=0$ *continuously on* $b_\gamma+\mathbb{R}t_\gamma\cap\Omega$, *where* $t_\gamma=(-\sin(\phi),\cos(\phi))^\top$, *and the transformation theorem, we further obtain that*

$$\begin{aligned}|D\tilde{u}|(\Omega)&=|D\tilde{u}|(\Omega_\Phi^+)+|D\tilde{u}|(\Omega_\Phi^-)\\&=|Du|(\Omega^+)+|Du|(\Omega^-)\\&=(u,\mathrm{div}(z))_{L^2(\Omega^+)}+(u,\mathrm{div}(z))_{L^2(\Omega^-)}\\&=(u\circ\Phi,\mathrm{div}(z)\circ\Phi)_{L^2(\Phi^{-1}(\Omega))}=(\tilde{u},\mathrm{div}(\tilde{z}))_{L^2(\Omega)},\end{aligned}\tag{5.4}$$

where we used in the last equality sign that $\mathrm{supp}(\tilde{u})\subseteq\Phi^{-1}(\Omega)\cap\Omega$. *Consequently, if we combine (5.3) and (5.4) and refer to the optimality conditions (2.6), then we find that* $\tilde{z}\in W^\infty(\mathrm{div};\Omega)$ *is a dual solution to* $\tilde{u}\in BV(\Omega)\cap L^\infty(\Omega)$. *Apparently,* $\tilde{z}\in W^\infty(\mathrm{div};\Omega)$ *is piece-wise Lipschitz continuous in the sense of Remark 4.8 and its jump set is given via* $J_{\tilde{z}}=b_\gamma+\mathbb{R}t_\gamma$. *As a consequence, if for every* $T\in\mathcal{T}_h$, *either of the cases (ii.a)–(ii.d) in Remark 4.8 is satisfied, then the quasi-optimal rate* $\mathcal{O}(h^{\frac{1}{2}})$ *is guaranteed by Remark 4.8, Lemma 4.7, Lemma 4.1 and Theorem 4.2.*

Example 5.3 (Four disks problem). *Let* $\Omega=(-1,1)^2\subseteq\mathbb{R}^2$, $r=0.4$, $\alpha=10$, *and*

$$g:=\chi_{B_r^2(r,r)}+\chi_{B_r^2(-r,-r)}-\chi_{B_r^2(r,-r)}-\chi_{B_r^2(-r,r)}\in BV(\Omega)\cap L^\infty(\Omega).$$

The same argumentation as for the proof of Proposition 2.1 shows that a minimum of (2.3) is given via $u:=(1-\frac{2}{r\alpha})g\in BV(\Omega)\cap L^\infty(\Omega)$. *A straightforward adaption of the proof of Corollary 2.2 implies that any dual solution* $z\in W^\infty(\mathrm{div};\Omega)$ *is not* θ*–Hölder continuous at* $x=\pm re_1$ *and* $x=\pm re_2$ *if* $\theta>\frac{1}{2}$. *Apart from that, arguing as in the proof of Proposition 2.3, we find that an example of a dual solution* $\overline{z}\in W^\infty(\mathrm{div};\Omega)$ *is given via* $\overline{z}(x):=\pm z(x\mp re_2)$ *if* $\pm x_2\geq 0$ *for all* $x=(x_1,x_2)^\top\in\Omega$, *where* $z\in W^\infty(\mathrm{div};\Omega)$ *is defined as in Proposition 2.3. In addition, if* $\Phi:\mathbb{R}^2\to\mathbb{R}^2$ *is defined as in Example 5.2, then for* $\tilde{g}:=g\circ\Phi\in BV(\Omega)\cap L^\infty(\Omega)$, *the primal solution is given via* $\tilde{u}:=u\circ\Phi\in BV(\Omega)\cap L^\infty(\Omega)$ *and a dual solution is given via* $\tilde{z}:=(D\Phi)^{-1}\overline{z}\circ\Phi\in W^\infty(\mathrm{div};\Omega)$. *Apparently,* $\tilde{z}\in W^\infty(\mathrm{div};\Omega)$ *is piece-wise Lipschitz continuous in the sense of Remark 4.8 and its jump set is given via* $J_{\tilde{z}}=b_\gamma+\mathbb{R}t_\gamma+\mathbb{R}n_\gamma$, *where* $t_\gamma=(-\sin(\phi),\cos(\phi))^\top$ *and* $n_\gamma=(\cos(\phi),\sin(\phi))^\top$. *As a consequence, if for every* $T\in\mathcal{T}_h$, *either of the cases (ii.a)–(ii.d) in Remark 4.8 is satisfied, then the quasi-optimal rate* $\mathcal{O}(h^{\frac{1}{2}})$ *is guaranteed by Remark 4.8, Lemma 4.7, Lemma 4.1 and Theorem 4.2.*

The experimental convergence rates in Figure 2 are obtained on k–times red-refined triangulations $\mathcal{T}_{h_k}$, $k=1,\dots,10$, of an initial triangulation $\mathcal{T}_{h_0}$ with two elements, i.e., $h_k=h_0 2^{-k}$ and $\varepsilon_{stop}^{h_k}=\frac{h_k}{20}$ for every $k=1,\dots,10$. In addition, for a simple implementation, we employ $\tilde{g}_{h_k}\in\mathcal{L}^0(\mathcal{T}_{h_k})$, defined by $\tilde{g}_{h_k}:=\tilde{g}(x_{\mathcal{T}_{h_k}})$, where $x_{\mathcal{T}_{h_k}}|_T:=x_T$ for all $T\in\mathcal{T}_{h_k}$, instead of $g_{h_k}:=\Pi_{h_k}\tilde{g}\in\mathcal{L}^0(\mathcal{T}_{h_k})$. However, since for each input data $\tilde{g}\in BV(\Omega)\cap L^\infty(\Omega)$ considered in this section, it holds $\|\tilde{g}-\tilde{g}_{h_k}\|_{L^1(\Omega)}\le ch_k|\partial B_r^2(0)|$, the error estimate remains valid. To be more specific, Figure 2 contains logarithmic plots for the experimental convergence rates of the error quantities

$$\big\|u(x_{\mathcal{T}_{h_k}})-\Pi_{h_k}u_{h_k}\big\|^2_{L^2(\Omega)},\quad k=3,\dots,10, \tag{5.5}$$

versus the total number of vertices $N_k=(2^k+1)^2\sim h_k^{-2}$ for $k=3,\dots,10$. In it, we find that the L^2–errors (5.5) converge at the quasi-optimal convergence rate $\mathcal{O}(h^{\frac{1}{2}})$. This behavior is reported for both examples, i.e., Example 5.2 and Example 5.3, for $\phi=0.0$ and $b_\gamma=(0.0,0.0)^\top$ as well as for $\phi=\frac{7\pi}{18}$ and $b_\gamma=(0.1,0.0)^\top$. Recall that for $\phi=0.0$ and $b_\gamma=(0.0,0.0)^\top$, the quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ is analytically guaranteed in both examples (cf. Example 5.2 and Example 5.3). Apart from that, we also could report the quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ for $d=3$, a uniform triangulation of $\Omega=(-1,1)^3$ and $g\in L^\infty(\Omega)\cap BV(\Omega)$ given via two or four touching balls, with several rotations and shifts, for which no Lipschitz continuous dual solution exists.

In Figure 3, the numerical solution $u_{h_5}\in\mathcal{S}^{1,cr}(\mathcal{T}_{h_5})$ obtained in Example 5.3 and its L^2–projection $\Pi_{h_5}u_{h_5}\in\mathcal{L}^0(\mathcal{T}_{h_5})$ are displayed for $\phi=0.0$ and $b_\gamma=(0.0,0.0)^\top$. Large gradients occur near the contact points of the disks, the midpoint values do not, however, show artifacts. In Figure 4, the L^2–projection $\Pi_{h_5}z_{h_5}\in\mathcal{L}^0(\mathcal{T}_{h_5})^2$ of the discrete dual solution $z_{h_5}:=\nabla_{h_5}u_{h_5}|\nabla_{h_5}u_{h_5}|^{-1}_{h_5}+\frac{\alpha}{2}\Pi_{h_5}(u_{h_5}-g)(\mathrm{id}_{\mathbb{R}^2}-\Pi_{h_5}\mathrm{id}_{\mathbb{R}^2})$ $\in\mathcal{RT}^0(\mathcal{T}_{h_5})$ with respect to the regularized ROF functional (5.1) (cf. [12, Section 5]) is displayed for $\phi=0.0$ and $b_\gamma=(0.0,0.0)^\top$.

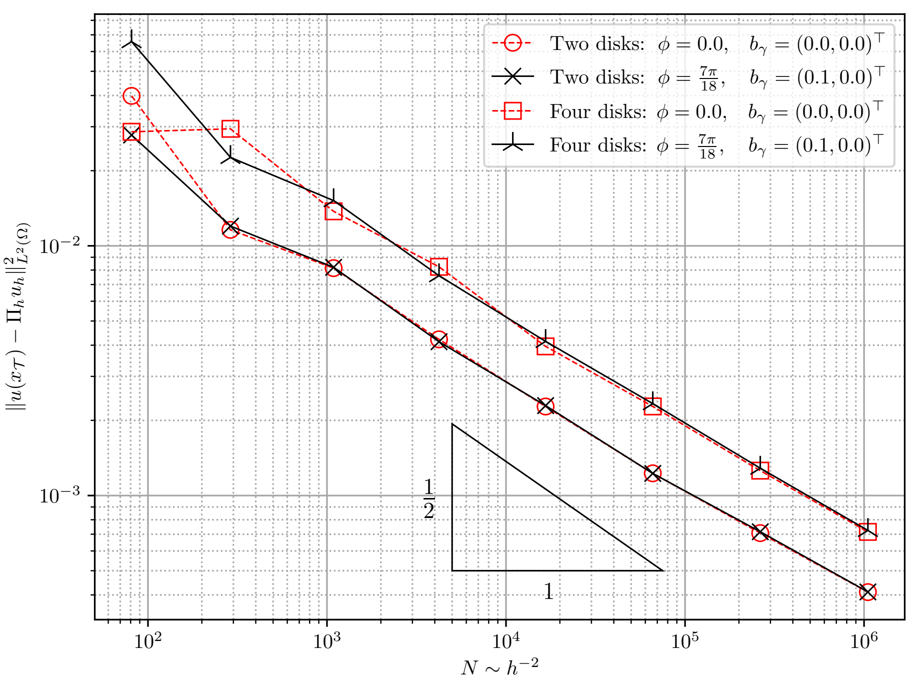

Figure 2: Logarithmic plots for the experimental convergence rates of the error quantities (5.5) in Example 5.2 and in Example 5.3. The rate $\mathcal{O}(h^{\frac{1}{2}})$ is observed.

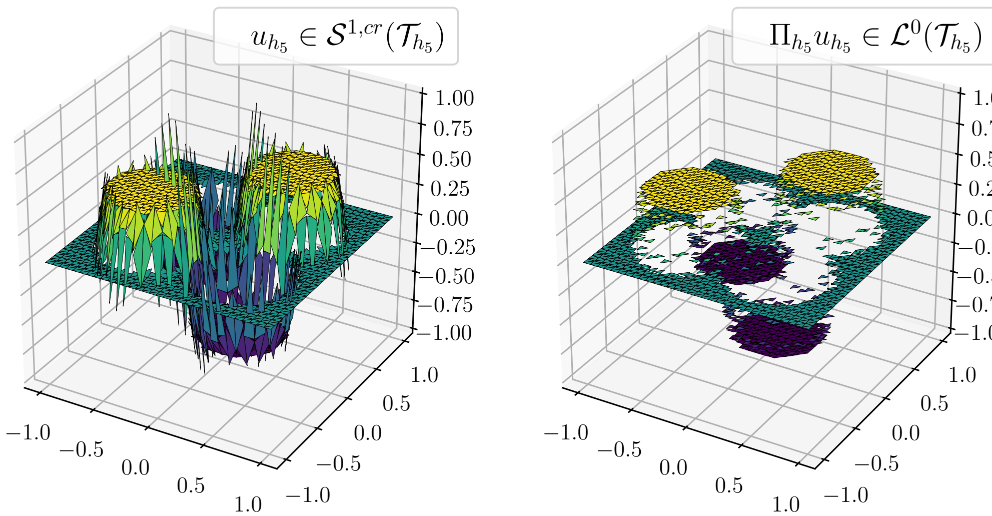

Figure 3: Numerical solution $u_{h_5} \in \mathcal{S}^{1,cr}(\mathcal{T}_{h_5})$ in Example 5.3 displayed as piecewise affine function (left) and via its L^2–projection $\Pi_{h_5} u_{h_5} \in \mathcal{L}^0(\mathcal{T}_{h_5})$ (right) for $r = 0.4$, $\alpha = 10$, $\phi = 0.0$ and $b_\gamma = (0.0, 0.0)^\top$. Large discrete gradients occur near $\pm re_1$ and $\pm re_2$, where no dual solution is θ–Hölder continuous for $\theta > \frac{1}{2}$.

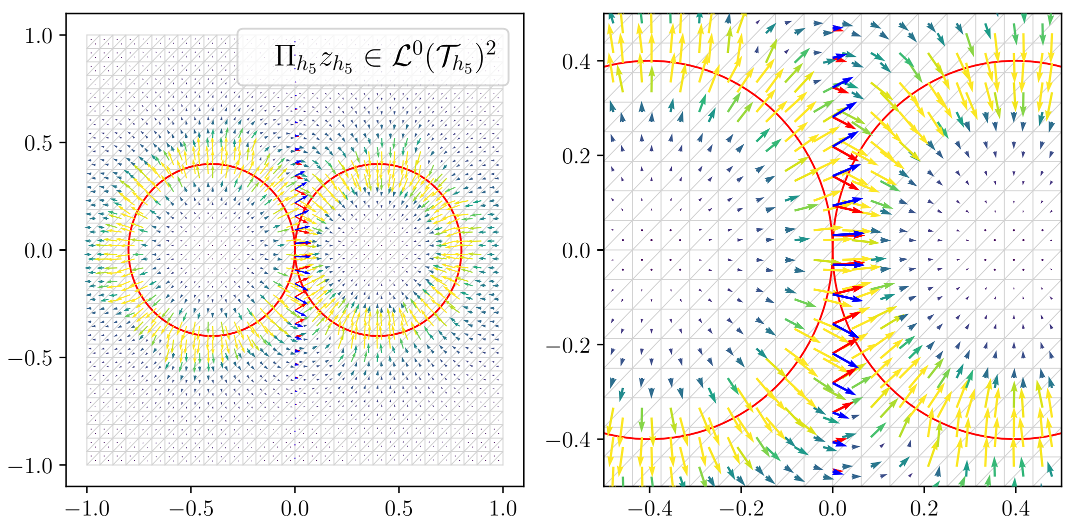

Figure 4: L^2–projection $\Pi_{h_5} z_{h_5} \in \mathcal{L}^0(\mathcal{T}_{h_5})^2$ of the discrete dual solution $z_{h_5} \in \mathcal{R}T^0(\mathcal{T}_{h_5})$ with respect to the regularized ROF functional (5.1) (cf. [12, Section 5]) displayed for $\phi = 0.0$ and $b_\gamma = (0.0, 0.0)^\top$. The red and blue arrows represent the values of $z_{h_5} \in \mathcal{R}T^0(\mathcal{T}_{h_5})$ at the midpoints of element sides along the $\mathbb{R}e_2$–axis, i.e., $\lim_{\varepsilon \to 0} z_{h_5}(x_S - \varepsilon e_1)$ (blue arrows) and $\lim_{\varepsilon \to 0} z_{h_5}(x_S + \varepsilon e_1)$ (red arrows). Here, the different orientations of the arrows indicate that $z_{h_5} \in \mathcal{R}T^0(\mathcal{T}_{h_5})$ approximates a discontinuous vector field – empirically $z \in W^\infty(\mathrm{div}; \Omega)$ defined in Proposition 2.3. Moreover, the red circles display the discontinuity set J_u of the minimizer $u \in BV(\Omega) \cap L^\infty(\Omega)$ defined in Proposition 2.1.

5.2 Experimental verification of condition (4.4)

In this section, we examine whether the dual solutions given in Example 5.2 for every $\phi \in [0, 2\pi]$ and $b_\gamma \in \mathbb{R}^2$ comply with condition (4.4) in Lemma 4.7, which, in view of Lemma 4.1 and Theorem 4.2 yields a guarantee for the quasi-optimal convergence rate $\mathcal{O}(h^{\frac{1}{2}})$. If we compute the quantities

$$\|\Pi_{h_k} I_{\mathcal{RT}} \tilde{z}\|_{L^\infty(\Omega;\mathbb{R}^d)}, \quad k = 1, \dots, 8, \tag{5.6}$$

where $\tilde{z} \in W^\infty(\mathrm{div};\Omega)$ is defined as in Example 5.2, then we find that for $\phi = 0.0$ and $b_\gamma = (0.0, 0.0)^\top$, $\phi = \frac{\pi}{2}$ and $b_\gamma = (0.0, 0.1)^\top$, $\phi = 0.0$ and $b_\gamma = (0.1, 0.0)^\top$, and $\phi = -\frac{\pi}{4}$ and $b_\gamma = (0.0, 0.0)^\top$, there exists a constant $c_z > 0$ – presumably, one has that $c_z = 1$ – such that for $k = 1, \dots, 8$, there holds

$$\|\Pi_{h_k} I_{\mathcal{RT}} \tilde{z}\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1 + c_z h_k. \tag{5.7}$$

These results confirm the findings in Remark 4.8 as they fall within one of the cases *(ii.a)–(ii.d)*. Apart from that, for $\phi = \frac{\pi}{4}$ and $b_\gamma = (0.0, 0.0)^\top$ as well as for $\phi = \frac{7\pi}{18}$ and $b_\gamma = (0.0, 0.0)^\top$, we cannot report the existence of a constant $c_z > 0$ such that (5.7) holds. This behavior can also be easily predicted analytically by resorting to the formula (4.8). All results can be found in Figure 5, which displays the quantities (5.6) versus the total number of vertices $N_k = (2k+1)^2 \sim h_k^{-2}$ for $k = 1, \dots, 8$.

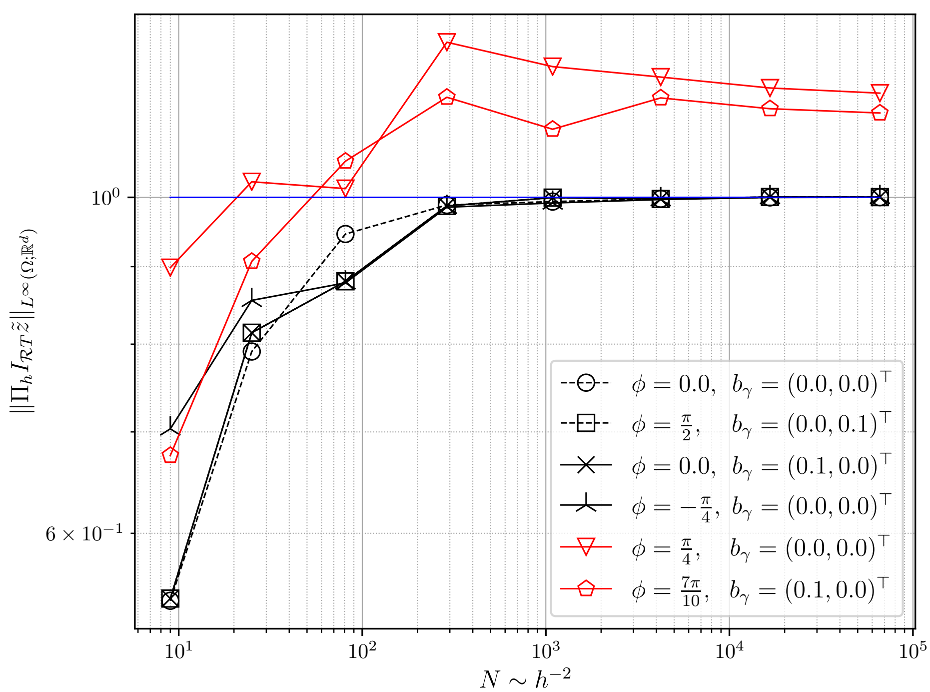

Figure 5: Logarithmic plots of the quantities (5.6) in Example 5.2.

Possible explanations for the observed quasi-optimal rate $\mathcal{O}(h^{\frac{1}{2}})$ for $\phi = \frac{\pi}{4}$ and $b_\gamma = (0.0, 0.0)^\top$ as well as for $\phi = \frac{7\pi}{18}$ and $b_\gamma = (0.0, 0.0)^\top$, even though (5.7) could not be reported, might be that this violation is merely pre-asymptotic or occurs only along the interface $(b_\gamma + \mathbb{R}t_\gamma) \cap \Omega$ (the latter, we observed experimentally), that the proofs presented are still sub-optimal, or that there exists an alternative dual solution for which (5.7) can be reported.

References

[1] E. Acerbi and N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Rational Mech. Anal.* **86** (1984), 125–145. doi:10.1007/BF00275731.

[2] E. Acerbi and N. Fusco, A regularity theorem for minimizers of quasiconvex integrals, *Arch. Rational Mech. Anal.* **99** (1987), 261–281. doi:10.1007/BF00284509.

[3] E. Acerbi and N. Fusco, An approximation lemma for $W^{1,p}$ functions, in *Material instabilities in continuum mechanics (Edinburgh, 1985–1986)*, *Oxford Sci. Publ.*, Oxford Univ. Press, New York, 1988, pp. 1–5.

[4] 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.

[5] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, second ed., *MOS-SIAM Series on Optimization* **17**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2014, Applications to PDEs and optimization. doi:10.1137/1.9781611973488.

[6] S. Bartels, Total variation minimization with finite elements: convergence and iterative solution, *SIAM J. Numer. Anal.* **50** (2012), 1162–1180. doi:10.1137/11083277X.

[7] S. Bartels, *Numerical methods for nonlinear partial differential equations*, *Springer Series in Computational Mathematics* **47**, Springer, Cham, 2015.doi:10.1007/978-3-319-13797-1.

[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. doi:10.1090/mcom/3656.

[9] S. Bartels, Nonconforming discretizations of convex minimization problems and precise relations to mixed methods, *Comput. Math. Appl.* **93** (2021), 214–229.doi:10.1016/j.camwa.2021.04.014.

[10] 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. doi:10.1137/17M1159166.

[11] S. Bartels, R. H. Nochetto, and A. J. Salgado, A total variation diminishing interpolation operator and applications, *Math. Comp.* **84** (2015), 2569–2587. doi:10.1090/mcom/2942.

[12] S. Bartels, R. Tovey, and F. Wassmer, Singular solutions, graded meshes, and adaptivity for total-variation regularized minimization problems, 2021.

[13] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, *Springer Series in Computational Mathematics* **44**, Springer, Heidelberg, 2013.doi:10.1007/978-3-642-36519-5.

[14] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, third ed., *Texts in Applied Mathematics* **15**, Springer, New York, 2008. doi:10.1007/978-0-387-75934-0.

[15] H. Brézis, *Function Analysis, Sobolev Spaces and Partial Differential Equations*, 01 2010. doi:10.1007/978-0-387-70914-7.

[16] 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. doi:10.1137/070683003.

[17] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, T. Pock, *Theoretical foundations and numerical methods for sparse recovery*, *Radon Series on Computational and Applied Mathematics* **9**, Walter de Gruyter GmbH & Co. KG, Berlin, 2010, Papers from the Summer School held in Linz, August 31–September 4, 2009. doi:10.1515/9783110226157.

[18] 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. doi:10.1137/090752754.

[19] A. Chambolle and P.-L. Lions, Image recovery via total variation minimization and related problems, *Numer. Math.* **76** (1997), 167–188. doi:10.1007/s002110050258.

[20] A. Chambolle and T. Pock, Crouzeix-Raviart approximation of the total variation on simplicial meshes, *J. Math. Imaging Vision* **62** (2020), 872–899. doi:10.1007/s10851-019-00939-3.

[21] M. Crouzeix and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I, *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge* **7** (1973), 33–75.

[22] L. Diening, J. Málek, and M. Steinhauer, On Lipschitz truncations of Sobolev functions (with variable exponent) and their selected applications, *ESAIM: Control, Optimisation and Calculus of Variations* **14** (2008), 211–232. doi:10.1051/cocv:2007049.

[23] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, *Applied Mathematical Sciences* **159**, Springer-Verlag, New York, 2004. doi:10.1007/978-1-4757-4355-5.

[24] A. Ern and J. L. Guermond, *Finite Elements I: Approximation and Interpolation*, *Texts in Applied Mathematics* no. 1, Springer International Publishing, 2021. doi:10.1007/978-3-030-56341-7.

[25] 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. doi:10.1007/s10851-018-0852-7.

[26] M. Hintermüller and K. Kunisch, Total bounded variation regularization as a bilaterally constrained optimization problem, *SIAM J. Appl. Math.* **64** (2004), 1311–1333. doi:10.1137/S0036139903422784.

[27] J. D. Hunter, Matplotlib: A 2d graphics environment, *Computing in Science & Engineering* **9** (2007), 90–95. doi:10.1109/MCSE.2007.55.

[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. doi:10.1137/110854539.

[29] A. Logg and G. N. Wells, Dolfin: Automated finite element computing, *ACM Transactions on Mathematical Software* **37** (2010). doi:10.1145/1731022.1731030.

[30] J. Malý and W. P. Ziemer, *Fine regularity of solutions of elliptic partial differential equations*, *Mathematical Surveys and Monographs* **51**, American Mathematical Society, Providence, RI, 1997. doi:10.1090/surv/051.

[31] P.-A. Raviart and J. M. Thomas, A mixed finite element method for 2nd order elliptic problems, in *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*, 1977, pp. 292–315. Lecture Notes in Math., Vol. 606.

[32] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, **60**, 1992, Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991), pp. 259–268. doi:10.1016/0167-2789(92)90242-F.

[33] 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. doi:10.1137/090769594.