ERROR CONTROL AND ADAPTIVITY FOR A VARIATIONAL MODEL PROBLEM DEFINED ON FUNCTIONS OF BOUNDED VARIATION

SÖREN BARTELS

Abstract. We derive a fully computable, optimal a posteriori error estimate for the finite element approximation of a total variation regularized model problem and devise an adaptive refinement strategy. Numerical experiments reveal a significant improvement over related approximations on uniformly refined triangulations.

1. Introduction

A simple model problem in the calculus of variations defined on the space of functions of bounded variation seeks a function $u : \Omega \to \mathbb{R}$ that minimizes the functional

$$E(u) = \int_\Omega |Du| + \frac{\alpha}{2}\|u - g\|^2_{L^2(\Omega)}$$

with a given function $g \in L^2(\Omega)$, a parameter $\alpha > 0$, and a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$. The first term is the total variation of the distributional derivative Du. It coincides with the semi-norm in $W^{1,1}(\Omega)$ if u belongs to this space and is finite if Du is a Radon measure with bounded total variation, i.e., $u \in BV(\Omega)$. The minimization of the functional E has been proposed in [ROF92] as a simple model in image processing, in which g is a given noisy image and the minimizer u serves as a smoother reconstruction. The function u may have discontinuities, e.g., if u is piecewise constant then $\int_\Omega |Du|$ coincides with the perimeter of the discontinuity set. Closely related energy functionals occur in the modeling of perfect plasticity [Suq78, BMR12] and the description of material damage [Tho11]. The methods discussed in this paper transfer to minimization problems that have the structure of the sum of the total variation norm plus a uniformly convex lower order term. Such problems often occur in the implicit time-discretization of gradient flows.

The finite element discretization of the minimization problem and the iterative solution of the resulting finite dimensional problems are now well understood: if for every $h > 0$ the set $\mathbb{V}_h \subset BV(\Omega)$ is a finite element space such that the spaces $\mathbb{V}_h \cap W^{1,1}(\Omega)$ define a dense family of subspaces in $W^{1,1}(\Omega)$ then the restriction of E to $\mathbb{V}_h$ leads to a Γ-convergent approximation. This is true if $\mathbb{V}_h$ contains piecewise affine, globally continuous finite element functions but not for the space of piecewise constant functions on a nested sequence of triangulations, cf. [Bar12] for details. The resulting discrete problems can be solved effectively with primal-dual methods recently developed and analyzed in [CP11, Bar12].

For a discretization with piecewise affine functions on a triangulation with maximal mesh-size $h > 0$ it can be shown that if $\Omega \subset \mathbb{R}^2$ is star-shaped and $g \in L^\infty(\Omega)$ then the exact and discrete minimizers u and u_h are related by $\|u - u_h\|_{L^2(\Omega)} \le ch^{1/4}$ so that a large number of degrees of freedom is required to guarantee a small error with respect to the norm in $L^2(\Omega)$. This may be suboptimal and the

Date: July 17, 2012.

1991 *Mathematics Subject Classification.* 65N30, 65N50.

Key words and phrases. Total variation, functions of bounded variation, finite elements, adaptivity, error estimation.

bound does not reflect the special structure of the exact solution which is often discontinuous but piecewise smooth with a simple jump set. Therefore, the simultaneous accurate resolution of the lower dimensional interface that seperates regions in which u is smooth and the approximation of the piecewise smooth functions can benefit from a treatment with different scales. While this is difficult to realize on the basis of a priori information, a posteriori error estimates that control the approximation error in terms of computable quantities can realize this goal automatically.

An abstract approach to the a posteriori error control for minimization problems with functionals J of the form

$$J(u) = F(u) + G(\Lambda u)$$

for a bounded linear operator $\Lambda : V \to Q$, and proper, convex, lower-semicontinuous functionals $F : V \to \overline{\mathbb{R}}$ and $G : Q \to \overline{\mathbb{R}}$ for Banach spaces V and Q and $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ has been developed in [Rep00]. It employs the dual formulation that consists in the minimization of the functional

$$J^*(q) = F^*(-\Lambda^* q) + G^*(q)$$

(or more precisely the maximization of $q \mapsto -J^*(q)$) with the Fenchel conjugates F^* and G^* of F and G. Provided that F or G has some coercivity properties the result controls the approximation error $u - u_h$ by the primal-dual gap with an arbitrary admissible function q, e.g., if F is quadratic then

$$2\gamma \|u - u_h\|^2 \le F(u_h) + F^*(\Lambda^* q) + G(\Lambda u_h) + G^*(-q) = J(u) + J^*(q). \tag{1.1}$$

The estimate can only be efficient if the primal and dual formulation satisfy a strong duality principle, i.e., the solutions u and p of the primal and dual problem satisfy $J(u) = -J^*(p)$. We will provide a refined, optimal version of the estimate (1.1). We remark that u_h and q in (1.1) are arbitrary admissible functions for the primal and dual problem, respectively. In particular, no exact solution of the discretized primal problem is required.

Owing to the lack of reflexivity of the Banach space $BV(\Omega)$ the dual of the minimization problem defined by the functional $E : BV(\Omega) \to \mathbb{R}$ is difficult to characterize. It has however been shown in [HK04] that the problem itself is the Fenchel-dual of the minimization problem defined by the functional

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

on the space $H_N(\operatorname{div};\Omega) = \{q \in L^2(\Omega;\mathbb{R}^d) : \operatorname{div} q \in L^2(\Omega),\ q \cdot \nu = 0 \text{ on } \partial\Omega\}$, where ν is the outer unit normal on $\partial\Omega$, and with the indicator functional $I_{K_1(0)}$ that vanishes for vector fields satisfying $|p| \le 1$ in Ω. This important observation implies that strong duality holds and thus that the error estimate (1.1) is sharp in the sense that the right-hand side vanishes if $u_h = u$ and $q = p$. Notice that minimizers of D are non-unique in general.

The abstract error estimate (1.1) is closely related to recovery error estimators for elliptic problems. In particular, for the Poisson problem it is known that a simple averaging of the discrete flux ∇u_h leads to reliable and efficient error control, i.e., up to higher order terms it is possible to show that,

$$\|\nabla(u - u_h)\|_{L^2(\Omega)} \le c\|\nabla u_h - \mathcal{A}_h \nabla u_h\|_{L^2(\Omega)}$$

with $c = 1$ in some situations and a converse estimate also applies, cf. [Bra07, CB02]. The estimate (1.1) actually implies this upper bound with $c = 1$ if $-\operatorname{div} \mathcal{A}_h \nabla u_h = f$ and thereby justifies a certain locality property of the right-hand side in (1.1). The proofs of the estimates for the Poisson problem in [Bra07, CB02] make essential use of the quadratic structure of the Dirichlet functional and cannot be transferred to the problem of minimizing the non-smooth energy functional E. Surprsingly, even for simple Helmholtz type problems that fit into the above framework simple averaging does not lead to efficient error control on unstructured triangulations.

To obtain a good bound on the approximation error, i.e., to find a good function q in (1.1), we will consistently discretize the dual formulation and solve it iteratively. This poses several difficulties. First, a subspace of $H_N(\mathrm{div};\Omega)$ has to be chosen in which the constraint $|p_h| \le 1$ can be imposed efficiently. Second, the discretization of the dual problem has to be done in such a way that it can be solved reliably with a computational effort comparable to the solution of the primal problem. We will show that lowest order H^1 conforming elements allow to establish both requirements.

The rest of this article is organized as follows. Notation and some auxiliary results are specified and stated in Section 2. In Section 3 we give a refined version of Repin's abstract a posteriori error estimate for convex optimization problems. The predual problem for the minimization problem defined by E and its discretization and iterative solution will be discussed in Section 4. Numerical experiments that illustrate the performance of the error estimate and the induced refinement indicators will be presented in Section 5.

2. Preliminaries

We include in this section some elementary facts about finite element spaces and a result on the approximation of the model problem.

2.1. **Function spaces.** We use standard notation for Lebesgue and Sobolev spaces and abbreviate the inner product and the norm in $L^2(\Omega)$ by $(\cdot,\cdot)$ and $\|\cdot\|$, respectively. The Banach space $BV(\Omega)$ consists of all functions $v \in L^1(\Omega)$ for which, with the space of compactly supported, continuously differentiable vector fields $C^1_c(\Omega;\mathbb{R}^d)$,

$$\int_\Omega |Dv| = \sup_{q\in C^1_c(\Omega;\mathbb{R}^d),\,|q|\le 1} -\int_\Omega v \,\mathrm{div}\, q \,\mathrm{d}x < \infty$$

and is equipped with the norm $\|v\|_{BV(\Omega)} = \|v\|_{L^1(\Omega)} + \int_\Omega |Dv|$. We let $H_N(\mathrm{div};\Omega)$ denote the space of all $q \in L^2(\Omega;\mathbb{R}^d)$ for which $\mathrm{div}\, q \in L^2(\Omega)$ with norm $\|q\|_{H_N(\mathrm{div};\Omega)} = (\|q\|^2 + \|\mathrm{div}\, q\|^2)^{1/2}$.

2.2. **Discrete time derivatives.** Given a time-step size $\tau > 0$ and a sequence of functions or real numbers $(v^n)_{n\in\mathbb{N}}$ in an inner product space X we define $d_t v^{n+1} = (v^{n+1} - v^n)/\tau$ and notice that for every $v \in X$ we have

$$d_t v^{n+1} \cdot (v^{n+1} - v) = \frac{d_t}{2}\|v - v^{n+1}\|^2 + \frac{\tau}{2}\|d_t v^{n+1}\|^2.$$

We also note that for sequences $(a^n)_{n\in\mathbb{N}}$ and $(b^n)_{n\in\mathbb{N}}$ we have the summation by parts formula $\tau\sum_{n=0}^N \big\{(d_t a^{n+1})\cdot b^{n+1} + a^n \cdot (d_t b^{n+1})\big\} = a^{N+1}\cdot b^{N+1} - a^0 \cdot b^0$.

2.3. **Finite element spaces.** For a sequence of regular triangulations $(\mathcal{T}_h)_{h>0}$ of Ω into triangles or tetrahedra with maximal diameters $h = \max_{T\in\mathcal{T}_h} \mathrm{diam}(T)$ we define the finite element spaces

$$\begin{aligned}
\mathcal{L}^0(\mathcal{T}_h) &= \{q_h \in L^1(\Omega) : q_h|_T \text{ is constant for each } T \in \mathcal{T}_h\},\\
\mathcal{S}^1(\mathcal{T}_h) &= \{v_h \in C(\overline{\Omega}) : v_h|_T \text{ is affine for each } T \in \mathcal{T}_h\}.
\end{aligned}$$

With the set $\mathcal{N}_h$ of vertices of triangles or tetrahedra the nodal basis of $\mathcal{S}^1(\mathcal{T}_h)$ is defined by the functions $(\varphi_z : z \in \mathcal{N}_h) \subset \mathcal{S}^1(\mathcal{T}_h)$ which satisfy $\varphi_z(y) = 0$ for distinct $y, z \in \mathcal{N}_h$ and $\varphi_z(z) = 1$ for all $z \in \mathcal{N}_h$. An elementwise inverse estimate shows that there exists $c > 0$ such that $\|\mathrm{div}\, q_h\| \le c h_{min}^{-1}\|q_h\|$ with $h_{min} = \min_{T\in\mathcal{T}_h} \mathrm{diam}(T)$ for all $q_h \in \mathcal{S}^1(\mathcal{T}_h)^d$. The nodal interpolant of a function $v \in C(\overline{\Omega})$ is defined by $\mathcal{I}_h v = \sum_{z\in\mathcal{N}_h} v(z)\varphi_z$. A discrete inner product that is equivalent to $(\cdot,\cdot)$ on $\mathcal{S}^1(\mathcal{T}_h)^d$ is for $q_h, r_h \in \mathcal{S}^1(\mathcal{T}_h)^d$ defined by

$$(q_h, r_h)_h = \int_\Omega \mathcal{I}_h[q_h \cdot r_h]\,\mathrm{d}x = \sum_{z\in\mathcal{N}_h} \beta_z q_h(z) \cdot r_h(z)$$

with $\beta_z = \int_\Omega \varphi_z \,\mathrm{d}x$ for all $z \in \mathcal{N}_h$. We let $\|q_h\|_h = (q_h, q_h)_h^{1/2}$ denote the corresponding norm. The L^2-projection $P_h : L^2(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)$ is characterized by $(P_h v - v, w_h) = 0$ for all $w_h \in \mathcal{S}^1(\mathcal{T}_h)$. We recall that the lowest order Raviart-Thomas finite element space $\mathcal{RT}^0(\mathcal{T}_h)$ consists of all $q_h \in H(\mathrm{div};\Omega)$ with $q_h|_T(x) = a + bx$ for $a \in \mathbb{R}^2$, $b \in \mathbb{R}$, and all $x \in T \in \mathcal{T}_h$. We define $\mathcal{RT}^0_N(\mathcal{T}_h) = \mathcal{RT}^0(\mathcal{T}_h) \cap H_N(\mathrm{div};\Omega)$.

2.4. **Duality and optimality.** As above we consider lower semicontinuous, proper, convex functionals $F : V \to \overline{\mathbb{R}}$ and $G : Q \to \overline{\mathbb{R}}$ for $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ on Banach spaces V and Q with duals V^* and Q^* and a bounded linear operator $\Lambda : V \to Q$. We assume that Q is reflexive and recall that the Fenchel conjugates $F^* : V^* \to \overline{\mathbb{R}}$ and $G^* : Q^* \to \overline{\mathbb{R}}$ are defined by

$$F^*(w) = \sup_{v\in V} \langle v, w\rangle - F(v), \quad G^*(q) = \sup_{r\in Q} \langle r, q\rangle - G(r).$$

With the adjoint operator $\Lambda^* : Q^* \to V^*$ that is defined by $\langle \Lambda^* q, v\rangle = \langle q, \Lambda v\rangle$ for $q \in Q^*$ and $v \in V$ and the identity $G^{**} = G$ we verify that, formally interchanging extrema, we have

$$\begin{aligned}\inf_{v\in V} F(v) + G(\Lambda v) &= \inf_{v\in V} \sup_{q\in Q^*} F(v) + \langle q, \Lambda v\rangle - G^*(q)\\ &= \sup_{q\in Q^*} \big(- \sup_{v\in V} \langle -\Lambda^* q, v\rangle - F(v)\big) - G^*(q) = \sup_{q\in Q^*} -F^*(-\Lambda^* q) - G^*(q).\end{aligned}$$

In general, we only have that the left-hand side is an upper bound for the right-hand side. Sufficient conditions for equality can be found in [ET99, Roc97]. The latter maximization problem defines the dual problem. The calculations show that u and p are optimal for the primal and dual formulation, reseptively, if and only if

$$\Lambda u \in \partial G^*(p), \quad -\Lambda^* p \in \partial F(u)$$

with the subdifferentials $\partial F(u)$ and $\partial G^*(p)$ defined by

$$\begin{aligned}\partial F(u) &= \{w \in V^* : \langle w, v - u\rangle \le F(v) - F(u) \text{ for all } v \in V\},\\ \partial G^*(p) &= \{q \in Q : \langle q, r - p\rangle \le G^*(r) - G^*(p) \text{ for all } r \in Q\}.\end{aligned}$$

We finally remark that the inclusions are equivalent to, cf., e.g., [Roc97, ET99],

$$p \in \partial G(\Lambda u), \quad u \in \partial F^*(-\Lambda^* p).$$

2.5. **Minimization of E.** A consistent discretization of the problem of minimizing E on $BV(\Omega)$ seeks a minimizer in $\mathcal{S}^1(\mathcal{T}_h)$. It can be shown that discrete minimizers converge with respect to intermediate convergence in $BV(\Omega)$ to the minimizer of E as $h \to 0$. The representation

$$E(v_h) = \sup_{q_h\in\mathcal{L}^0(\mathcal{T}_h)^d} \int_\Omega \nabla v_h \cdot q_h \,\mathrm{d}x + \frac{\alpha}{2}\|v_h - g\|^2 - I_{K_1(0)}(q_h),$$

for $v_h \in \mathcal{S}^1(\mathcal{T}_h)$ allows to formulate the minimization of E as a discrete saddle-point problem. This observation motivates the following algorithm that approximates the minimizer u_h of E restricted to $\mathcal{S}^1(\mathcal{T}_h)$ if $\tau \le ch_{min}$. We refer the reader to [Bar12] for details.

Algorithm (A). Let $(u_h^0, p_h^0) \in \mathcal{S}^1(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)^d$, set $d_t u_h^0 = 0$, and solve for $n = 0, 1, ...$ with $\widetilde{u}_h^{n+1} = u_h^n + \tau d_t u_h^n$ the equations

$$(-d_t p_h^{n+1} + \nabla \widetilde{u}_h^{n+1}, q_h - p_h^{n+1}) \le 0, \quad (d_t u_h^{n+1}, v_h) + (p_h^{n+1}, \nabla v_h^{n+1}) = -\alpha(u_h^{n+1} - g, v_h)$$

subject to $|p_h^{n+1}| \le 1$ in Ω for all $(v_h, q_h) \in \mathcal{S}^1(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)^d$ with $|q_h| \le 1$ in Ω.

Remarks 2.1. *(i) The algorithm computes a piecewise constant approximation of a (in general non-unique) solution of the dual problem, i.e., the iterates* $(p_h^n)_{n\geq 0}$ *converge to some* $p_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ *but in general this vector field does not belong to* $H_N(\mathrm{div};\Omega)$.
(ii) The equation and the variational inequality in the algorithm decouple and can be solved explicitly up to the inversion of a (lumped) mass matrix. In particular, we have for the piecewise constant vector field p_h^{n+1} *that*

$$p_h^{n+1} = \frac{p_h^n + \tau\nabla\widetilde{u}_h^{n+1}}{\max\{1, |p_h^n + \tau\nabla\widetilde{u}_h^{n+1}|\}}.$$

3. A sharp version of Repin's error estimate

For Banach spaces V and Q with duals V^* and Q^*, a bounded linear operator $\Lambda : V \to Q$, and convex, lower-semicontinuous, proper functionals $F : V \to \overline{\mathbb{R}}$ and $G : Q \to \overline{\mathbb{R}}$ we consider the problem of finding $u \in V$ with

$$J(u) = \inf_{v\in V} J(v), \quad J(v) = F(v) + G(\Lambda v).$$

The associated dual problem consists in finding $p \in Q^*$ with

$$-J^*(p) = \sup_{q\in Q^*} -J^*(q), \quad J^*(q) = F^*(-\Lambda^* q) + G^*(q)$$

We let Φ_G and Φ_F be non-negative functionals such that for all $q_1, q_2 \in Q$ and $v_1, v_2 \in V$ we have

$$G\big((q_1+q_2)/2\big) + \Phi_G(q_2 - q_1) \leq \big(G(q_1) + G(q_2)\big)/2,$$
$$F\big((v_1+v_2)/2\big) + \Phi_F(v_2 - v_1) \leq \big(F(v_1) + F(v_2)\big)/2.$$

By convexity we have that, e.g., $\Phi_G = \Phi_F \equiv 0$ satisfy the estimates. The primal and dual optimization problems are related by a weak complementarity principle which states that

$$J(u) = \inf_{v\in V} J(v) \geq \sup_{q\in Q^*} -J^*(q) = -J^*(p).$$

We say that strong duality applies if equality holds. Our final ingredient for the error estimate is a characterization of the optimality of the solution of the primal problem. For all $w \in \partial J(u)$ we have with a non-negative functional Ψ_J that

$$\langle w, v-u\rangle + \Psi_J(v-u) \leq J(v) - J(u)$$

and u is optimal if and only if $0 \in \partial J(u)$.

Theorem 3.1 ([Rep00]). *For the solution* $u \in V$ *of the primal problem and arbitrary* $v \in V$ *and* $q \in Q^*$ *we have*

$$\Phi_G\big(\Lambda(u-v)\big) + \Phi_F(u-v) + \Psi_J\big((u-v)/2\big) \leq (1/2)\big[J(v) + J^*(q)\big].$$

Proof. The convexity estimates imply that

$$\Phi_G\big(\Lambda(u-v)\big) + \Phi_F\big(u-v\big) \leq (1/2)\big[F(v) + G(\Lambda v) + F(u) + G(\Lambda u)\big] - F\big((v+u)/2\big) - G\big(\Lambda(v+u)/2\big).$$

The optimality of u shows that we have

$$F(u) + G(\Lambda u) + \Psi_J\big(u - (u+v)/2\big) \leq F\big((u+v)/2\big) + G\big(\Lambda(u+v)/2\big).$$

It follows that

$$\Phi_G\big(\Lambda(u-v)\big) + \Phi_F\big(u-v\big) \leq (1/2)\big[F(v) + G(\Lambda v) - F(u) - G(\Lambda u)\big] - \Psi\big((u-v)/2\big).$$

The weak complementarity principle

$$F(u) + G(\Lambda u) = J(u) \geq -J^*(q) = -F^*(\Lambda^* q) - G^*(-q)$$

yields the asserted estimate. $\square$

Remarks 3.1. *(i) Notice that the only relation between the problems defined by the functionals J and J^* needed in the proof is the weak duality $J(u) = \inf_{v\in V} J(v) \ge \sup_{q\in Q^*} -J^*(q)$, in particular, $-J^*(q)$ can be replaced by any quantity that is a lower bound for $J(u)$.*
(ii) If the primal and dual problem are related by a strong complementarity property then the estimate of the theorem is sharp in the sense that the right-hand side vanishes if $v = u$ and q solves the dual problem.
(iii) The estimate differs from the one given in [Rep00] *by the term Ψ_J which is necessary to obtain optimal constants, cf. Example 3.1 below.*
(iv) Notice that the coercivity of a convex functional ϕ is often defined by, cf., e.g., [NSV00],

$$\sigma(w, v) = \phi(v) - \phi(w) - \sup_{q\in\partial\phi(w)} \langle q, v - w\rangle.$$

The map σ is also known as the Brègman distance defined by ϕ, cf. [Brè67, MO08].
(v) Assume that $\partial\phi$ is single-valued. If

$$\langle D\phi(u), v - u\rangle + \Phi_\phi(v - u) \le \phi(v) - \phi(u)$$

then with $2\Psi_\phi(w) = \Phi_\phi(w) + \Phi_\phi(-w)$ we have

$$\phi\big((v_1 + v_2)/2\big) + \Psi_\phi(v_2 - v_1) \le \big(\phi(v_1) + \phi(v_2)\big)/2.$$

Example 3.1. *For the Poisson problem $-\Delta u = f$ in Ω, $u|_{\partial\Omega} = 0$, we have $V = H^1_0(\Omega)$, $Y = L^2(\Omega;\mathbb{R}^d)$, $\Lambda = \nabla$, $G(\Lambda v) = (1/2)\int_\Omega |\nabla v|^2 \,\mathrm{d}x$, and $F(v) = -\int_\Omega f v \,\mathrm{d}x$. Since $F^*(w) = I_{\{-f\}}(w)$, $G^*(q) = (1/2)\int_\Omega |q|^2 \,\mathrm{d}x$, $\Lambda^* = -\operatorname{div} : L^2(\Omega;\mathbb{R}^d) \to H^1_0(\Omega)^*$, we find that the right-hand side $\eta^2(v, q)$ of the estimate of Theorem 3.1 is given by*

$$\begin{aligned}\eta^2(v, q) &= (1/2)\Big[-\int_\Omega f v \,\mathrm{d}x + I_{\{-f\}}(\operatorname{div} q) + (1/2)\int_\Omega |\nabla v|^2 \,\mathrm{d}x + (1/2)\int_\Omega |q|^2 \,\mathrm{d}x\Big] \\ &= (1/2)\Big[\int_\Omega (\operatorname{div} q) v \,\mathrm{d}x + (1/2)\|\nabla v\|^2 + (1/2)\|q\|^2\Big] = (1/4)\|\nabla v - q\|^2\end{aligned}$$

provided that $-\operatorname{div} q = f$. We also have that

$$(1/2)\big((q_1 + q_2)/2\big)^2 - (1/4)(q_1^2 + q_2^2) = (1/8)(q_1^2 + 2q_1q_2 + q_2^2 - 2q_1^2 - 2q_2^2) = -(1/8)(q_1 - q_2)^2$$

so that $\Phi_G(q) = (1/8)\|q\|^2$ and

$$(1/2)q_1^2 - (1/2)q_2^2 - q_1(q_1 - q_2) = -(1/2)q_1^2 - (1/2)q_2^2 + q_1q_2 = -(1/2)(q_1 - q_2)^2,$$

i.e., $\Psi_J\big((v - u)/2\big) = (1/8)\|\nabla(v - u)\|^2$. With $\Phi_F \equiv 0$ Theorem 3.1 implies the estimate

$$\|\nabla(u - v)\| \le \inf_{-\operatorname{div} q = f} \|\nabla v - q\|$$

and equality occurs for $q = \nabla u$.

Example 3.2. *For the Helmholtz type problem defined by the minimization of*

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x + \frac{\alpha}{2}\|v - g\|^2$$

for $v \in V = H^1(\Omega)$ and with $\nabla : H^1(\Omega) \to H_N(\operatorname{div};\Omega)^*$ given by $\langle\nabla v, q\rangle = (v, -\operatorname{div} q)$ for $v \in H^1(\Omega)$ and $q \in Q = H_N(\operatorname{div};\Omega)^*$ the dual problem is for $q \in H_N(\operatorname{div};\Omega)$ defined by

$$J^*(q) = \frac{1}{2}\int_\Omega |q|^2 \,\mathrm{d}x + \frac{1}{2\alpha}\|\operatorname{div} q + \alpha g\|^2 - \frac{\alpha}{2}\|g\|^2.$$

We may choose $\Phi_F(v) = (\alpha/8)\|v\|^2$, $\Phi_G(q) = (1/8)\|q\|^2$, and $\Psi_J(v) = (1/2)\|\nabla v\|^2 + (\alpha/2)\|v\|^2$. Theorem 3.1 leads to the error estimate

$$\begin{aligned}\|\nabla(u-v)\|^2 + \alpha\|u-v\|^2 &\le \alpha\|v-g\|^2 + \frac{1}{\alpha}\|\operatorname{div} q + \alpha g\|^2 - \alpha\|g\|^2 + \int_\Omega |\nabla v|^2 \,\mathrm{d}x + \int_\Omega |q|^2 \,\mathrm{d}x \\ &= \|\nabla v - q\|^2 - 2(v, \operatorname{div} q) + \frac{1}{\alpha}\|\operatorname{div} q + \alpha g\|^2 + \alpha\|v-g\|^2 - \alpha\|g\|^2 \\ &= \|\nabla v - q\|^2 + \frac{1}{\alpha}\|\operatorname{div} q - \alpha(v-g)\|^2.\end{aligned}$$

Here equality occurs for $q = \nabla u$ since $-\Delta u + \alpha(u-g) = 0$. One may also regard the minimization of $J^*(q)$ as the primal problem in which $\Lambda = \operatorname{div} : H_N(\operatorname{div};\Omega) \to L^2(\Omega)$ and then identify J as the corresponding dual.

4. Approximation of the Fenchel predual

The identification of the dual problem defined by the problem of minimizing $E(v)$ among $v \in BV(\Omega)$ is difficult since $BV(\Omega)$ is not reflexive and its dual is difficult to characterize. It turns out that the predual can be described very efficiently, i.e., a minimization problem whose dual consists in the minimization of E on $BV(\Omega)$. We recall in this section the result on Fenchel duality of [HK04] and discuss the discretization and iterative solution of the predual formulation.

4.1. Fenchel predual.

For $q \in H_N(\operatorname{div};\Omega)$ let

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

where $I_{K_1(0)}(q) = 0$ if $|q| \le 1$ almost everywhere in Ω and $I_{K_1(0)}(q) = \infty$ otherwise. The essential link between the functionals E and D is an equivalent characterization of the total variation norm of $v \in BV(\Omega)$, i.e., we have

$$\int_\Omega |Dv| = \sup_{q \in H_N(\operatorname{div};\Omega)} -\int_\Omega v \operatorname{div} q \,\mathrm{d}x - I_{K_1(0)}(q).$$

It has been verified in [HK04] that E is the dual functional related to D and that Fenchel duality theory in the sense of [ET99] applies, i.e., that strong duality holds. This allows us to deduce the following result.

Theorem 4.1. *Let $u \in BV(\Omega) \cap L^2(\Omega)$ be minimal for E. Then for every $u_h \in BV(\Omega) \cap L^2(\Omega)$ and $q \in H_N(\operatorname{div};\Omega)$ we have*

$$\frac{\alpha}{2}\|u - u_h\|^2 = \int_\Omega |Du_h| + \frac{\alpha}{2}\|u_h - g\|^2 + \frac{1}{2\alpha}\|\operatorname{div} q + \alpha g\|^2 - \frac{\alpha}{2}\|g\|^2 + I_{K_1(0)}(q).$$

Proof. Owing to the results of [HK04] we have $E(u) \ge -D(q)$ for every $q \in H_N(\operatorname{div};\Omega)$. The theorem is therefore a consequence of Theorem 3.1 upon noting that we may choose $\Phi_F(v) = (\alpha/8)\|v\|^2$, $\Phi_G = 0$, and $\Psi_J = (\alpha/2)\|v\|^2$. □

The remainder of this section is devoted to the computation of a discrete element $q \in H_N(\operatorname{div};\Omega)$ that leads to a finite and nearly optimal upper error bound.

4.2. **Discretization of D.** For a regular triangulation $\mathcal{T}_h$ of Ω we set

$$\mathcal{S}^1_N(\mathcal{T}_h)^d = \{q_h \in \mathcal{S}^1(\mathcal{T}_h)^d : q_h \cdot \nu = 0 \text{ on } \partial\Omega\}$$

and note that $\mathcal{S}^1_N(\mathcal{T}_h)^d \subset H_N(\operatorname{div};\Omega)$. We have that a vector field $q_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ satisfies $|q_h| \le 1$ in Ω if and only if $|q_h(z)| \le 1$ for all $z \in \mathcal{N}_h$. We then consider the minimization problem defined by the functional

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

with the L^2 projection $P_h : L^2(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)$. We formally extend D_h to vector fields in $H_N(\operatorname{div};\Omega)$ by setting $D_h(q) = \infty$ if $q \in H_N(\operatorname{div};\Omega) \setminus \mathcal{S}^1_N(\mathcal{T}_h)^d$.

Theorem 4.2. *For a sequence of triangulations $(\mathcal{T}_h)_{h>0}$ with maximal mesh-size $h \to 0$ the functionals D_h converge in the sense of Γ-convergence to the functional D with respect to strong convergence in $H_N(\operatorname{div};\Omega)$, i.e., (i) for every sequence $(q_h)_{h>0} \subset H_N(\operatorname{div};\Omega)$ with $q_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ for all $h > 0$ and $\lim_{h\to 0} q_h = q$ in $H_N(\operatorname{div};\Omega)$ we have*

$$D(q) \le \liminf_{h\to 0} D_h(q_h)$$

and conversely, (ii) for every $q \in H_N(\operatorname{div};\Omega)$ there exists a sequence $(q_h)_{h>0} \subset H_N(\operatorname{div};\Omega)$ with $q_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ for every $h > 0$ and $\lim_{h\to 0} q_h = q$ such that

$$D(q) = \lim_{h\to 0} D_h(q_h).$$

Proof. To show the first statement let $(q_h)_{h>0}$ and q be as in the theorem. By lower semicontinuity of D it suffices to show that

$$\|P_h(\operatorname{div} q_h + \alpha g)\| \to \|\operatorname{div} q_h + \alpha g\|$$

as $h \to 0$. Letting $\psi = \operatorname{div} q + \alpha g$ and $\psi_h = \operatorname{div} q_h + \alpha g$ and noting that P_h is a bounded operator on $L^2(\Omega)$ with operator norm $\|P_h\| = 1$ we have

$$\begin{aligned}\|P_h\psi_h - \psi\| &\le \|(1-P_h)(\psi_h - \psi)\| + \|\psi - \psi_h\| + \|\psi - \psi_h\| \\ &\le 3\|\psi - \psi_h\| + \|\psi - P_h\psi\| \to 0\end{aligned}$$

as $h \to 0$ by density of $\mathcal{S}^1(\mathcal{T}_h)$ in $L^2(\Omega)$ and properties of the projection. To prove the second statement we may assume by density of compactly supported smooth vector fields in $H_N(\operatorname{div};\Omega)$ that $q \in C^1_c(\Omega;\mathbb{R}^d) \cap H^2(\Omega;\mathbb{R}^d)$ and employ the nodal interpolant $q_h = \mathcal{I}_h q \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ for every $h > 0$. The convergence then follows from standard results on nodal interpolation. $\square$

Remarks 4.1. *(i) For vector fields in the lowest order Raviart-Thomas finite element space $\mathcal{RT}_N(\mathcal{T}_h) \subset H_N(\operatorname{div};\Omega)$ the condition $|q_h| \le 1$ is difficult to formulate in terms of the natural degrees of freedom, i.e., the normal components on edges or faces.*
(ii) The reason for the inconsistent discretization of D, i.e., for incorporating the projection operator P_h is that this allows for a reliable iterative solution of the discrete problems.

4.3. **Discrete duality.** The following lemma shows that the discretization of the predual may be regarded as the discrete (pre-)dual of a discretization of the functional E.

Lemma 4.1. *Let $\widetilde{\nabla}_{h,N} : \mathcal{S}^1(\mathcal{T}_h) \to \mathcal{S}^1_N(\mathcal{T}_h)^d$ be for $v_h \in \mathcal{S}^1(\mathcal{T}_h)$ defined by*

$$(\widetilde{\nabla}_{h,N} v_h, q_h)_h = -(v_h, P_h \operatorname{div} q_h)$$

for all $q_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$. Then the Fenchel dual of the minimzation of D_h is defined through the functional

$$E_h(u_h) = \int_\Omega \mathcal{I}_h|\widetilde{\nabla}_{h,N} u_h| \,\mathrm{d}x + \frac{\alpha}{2}\|u_h - P_h g\|^2$$

Proof. We identify the spaces $\mathcal{S}^1(\mathcal{T}_h)$ and $\mathcal{S}^1_N(\mathcal{T}_h)^d$ with their duals via the L^2 inner product $(\cdot,\cdot)$ and the discrete L^2 inner product $(\cdot,\cdot)_h$, respectively, so that in particular $-P_h \operatorname{div} = \widetilde{\nabla}^*_{h,N} = \Lambda^*_h$. For $F_h(v_h) = (\alpha/2)\|v_h - P_h g\|^2$ we have $F^*_h(w_h) = (1/(2\alpha))\|w_h + \alpha P_h g\|^2 - (\alpha/2)\|P_h g\|^2$. The functional $G_h(q_h) = \int_\Omega \mathcal{I}_h |q_h| \,\mathrm{d}x$ can be written in the form

$$G_h(r_h) = \sum_{z\in\mathcal{N}_h} \beta_z |r_h(z)| = \sup_{q_h\in\mathcal{S}^1_N(\mathcal{T}_h)^d:\, |q_h|\le 1} \beta_z r_h(z)\cdot q_h(z) = \sup_{q_h\in\mathcal{S}^1_N(\mathcal{T}_h)^d} (r_h,q_h)_h - I_{K_1(0)}(q_h),$$

i.e., $G^*_h = I_{K_1(0)}$. Since the discrete spaces are reflexive the statement follows from the fact that $G^{**}_h = G_h$, $F^{**}_h = F_h$, and $\Lambda^{**}_h = \Lambda_h$. □

Remark 4.1. *It can be shown that the discrete functionals E_h are Γ-convergent to the functional E with respect to intermediate convergence in $BV(\Omega)$. Numerical experiments reported below show however that minimizers of E_h develop oscillations at discontinuities.*

To characterize solutions of the discrete formulations more precisely, we define the discrete subdifferential $\partial_h I_{K_1(0)}(p_h)$ at $p_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ with $|p_h| \le 1$ in Ω as the set of all elements $\xi_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ with

$$(\xi_h, q_h - p_h)_h \le 0$$

for all $q_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ with $|q_h| \le 1$ in Ω.

Lemma 4.2. *Minimizers $u_h \in \mathcal{S}^1(\mathcal{T}_h)$ and $p_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ of the discrete functionals E_h and D_h, respectively, are saddle points of the functional*

$$S_h(v_h,q_h) = \frac{\alpha}{2}\|v_h - P_h g\|^2 - (v_h, P_h \operatorname{div} q_h) - I_{K_1(0)}(q_h),$$

in particular, they are solutions if and only if

$$\alpha(u_h - P_h g) - P_h \operatorname{div} p_h = 0, \quad \widetilde{\nabla}_{h,N} u_h \in \partial_h I_{K_1(0)}(p_h).$$

Proof. The first statement follows from noting that we have, as in the proof of Lemma 4.1,

$$\sup_{q_h\in\mathcal{S}^1_N(\mathcal{T}_h)^d} -(v_h, \operatorname{div} q_h) - I_{K_1(0)}(q_h) = \sup_{q_h\in\mathcal{S}^1_N(\mathcal{T}_h)^d} (\widetilde{\nabla}_{h,N} v_h, q_h)_h - I_{K_1(0)}(q_h) = \int_\Omega \mathcal{I}_h |\widetilde{\nabla}_{h,N} v_h| \,\mathrm{d}x$$

and

$$\inf_{v_h\in\mathcal{S}^1(\mathcal{T}_h)} \frac{\alpha}{2}\|v_h - P_h g\|^2 - (v_h, P_h \operatorname{div} q_h) = \frac{1}{\alpha}\|P_h(\operatorname{div} q_h + \alpha g)\|^2 - \frac{\alpha}{2}\|P_h g\|^2.$$

The optimal q_h and v_h satisfy the equations stated in the lemma. □

4.4. **Iterative solution.** We approximately solve the saddle-point formulation of Lemma 4.2 by a simultaneous gradient flow for u_h and p_h in descent and ascent directions, respectively, i.e., we consider temporal discretizations of the system of ordinary differential equations and inclusions

$$\partial_t u_h = -\alpha(u_h - P_h g) + P_h \operatorname{div} q_h, \quad -\partial_t p_h + \widetilde{\nabla}_{h,N} u_h \in \partial_h I_{K_1(0)}(p_h).$$

The following algorithm defines a time-stepping scheme and states the equations and inclusions in variational form.

Algorithm (A'). Let $\tau > 0$ and $(u^0_h, p^0_h) \in \mathcal{S}^1(\mathcal{T}_h) \times \mathcal{S}^1_N(\mathcal{T}_h)^d$ with $|p^0_h(z)| \le 1$ for all $z \in \mathcal{N}_h$, set $d_t u^0_h = 0$, and solve for $n = 0, 1, ...$ with $\widetilde{u}^{n+1}_h = u^n_h + \tau d_t u^n_h$ the equations

$$(-d_t p^{n+1}_h + \widetilde{\nabla}_{h,N} \widetilde{u}^{n+1}_h, q_h - p^{n+1}_h)_h \le 0, \quad (d_t u^{n+1}_h, v_h) - (\operatorname{div} p^{n+1}_h, v^{n+1}_h) = -\alpha(u^{n+1}_h - P_h g, v_h)$$

subject to $|p^{n+1}_h(z)| \le 1$ for all $z \in \mathcal{N}_h$ and for all $(v_h, q_h) \in \mathcal{S}^1(\mathcal{T}_h) \times \mathcal{S}^1_N(\mathcal{T}_h)^d$ with $|q_h(z)| \le 1$ for all $z \in \mathcal{N}_h$.

Remark 4.2. *Notice that p_h^{n+1} is the unique minimizer of*

$$q_h \mapsto \frac{1}{2\tau}\|q_h - p_h\|_h^2 - (q_h, \widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1})_h + I_{K_1(0)}(q_h)$$

and is for every $z \in \mathcal{N}_h$ given by

$$p_h^{n+1}(z) = \frac{p_h^n(z) + \tau\widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1}(z)}{\max\{1, |p_h^n(z) + \tau\widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1}(z)|\}}.$$

The iterates of Algorithm (A') converge to a stationary point, e.g., if $\tau \le ch_{min}$. We denote $\|\widetilde{\nabla}_{h,N}\| = \sup_{0\neq v_h \in \mathcal{S}^1(\mathcal{T}_h)} \|\widetilde{\nabla}_{h,N} v_h\|_h/\|v_h\|$ and owing to an inverse estimate on $\mathcal{S}_N^1(\mathcal{T}_h)$ we have $\|\widetilde{\nabla}_{h,N}\| \le ch_{min}^{-1}$.

Proposition 4.1. *Let $u_h \in \mathcal{S}^1(\mathcal{T}_h)$ be the unique minimizer for E_h in $\mathcal{S}^1(\mathcal{T}_h)$. If $\theta = \tau^2\|\widetilde{\nabla}_{h,N}\|^2 < 1$ then the iterates of Algorithm (A') satisfy for every $N \ge 1$*

$$\tau\sum_{n=0}^{N}\Big((1-\theta)\frac{\tau}{2}\big(\|d_t u_h^{n+1}\|^2 + \|d_t p_h^{n+1}\|_h^2\big) + \alpha\|u_h - u_h^{n+1}\|^2\Big) \le C.$$

In particular, we have that $u_h^n \to u_h$ and $p_h^n \to \widetilde{p}_h$ for a minimizer $\widetilde{p}_h \in \mathcal{S}_N^1(\mathcal{T}_h)^d$ of D_h as $n \to \infty$.

Proof. Let $p_h \in \mathcal{S}_N^1(\mathcal{T}_h)^d$ be as in Lemma 4.2. Upon choosing $v_h = u_h - u_h^{n+1}$ and $q_h = p_h$ in Algorithm (A') and $q_h = p_h^{n+1}$ in the variational inclusion of Lemma 4.2 and using

$$(u_h^{n+1} - P_h g, u_h - u_h^{n+1}) + \|u_h - u_h^{n+1}\|^2 = (u_h - P_h g, u_h - u_h^{n+1})$$

we find that

$$\begin{aligned}
&\frac{d_t}{2}\big(\|u_h - u_h^{n+1}\|^2 + \|p_h - p_h^{n+1}\|_h^2\big) + \frac{\tau}{2}\big(\|d_t u_h^{n+1}\|^2 + \|d_t p_h^{n+1}\|_h^2\big) + \alpha\|u_h - u_h^{n+1}\|^2\\
&= -(d_t u_h^{n+1}, u_h - u_h^{n+1}) - (d_t p_h^{n+1}, p_h - p_h^{n+1})_h + \alpha\|u_h - u_h^{n+1}\|^2\\
&\le (p_h^{n+1}, \widetilde{\nabla}_{h,N}(u_h - u_h^{n+1}))_h + \alpha(u_h^{n+1} - P_h g, u_h - u_h^{n+1}) - (p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1})_h + \alpha\|u_h - u_h^{n+1}\|^2\\
&= (p_h^{n+1}, \widetilde{\nabla}_{h,N}(u_h - u_h^{n+1})) - (p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1})_h + \alpha(u_h - P_h g, u_h - u_h^{n+1})\\
&= (p_h^{n+1}, \widetilde{\nabla}_{h,N}(u_h - u_h^{n+1}))_h - (p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N}\widetilde{u}_h^{n+1})_h - (p_h, \widetilde{\nabla}_{h,N}(u_h - u_h^{n+1}))_h\\
&= (p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N}(u_h^{n+1} - \widetilde{u}_h^{n+1}))_h + (p_h^{n+1} - p_h, \widetilde{\nabla}_{h,N} u_h)_h\\
&\le (p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N}(u_h^{n+1} - \widetilde{u}_h^{n+1}))_h = \tau^2(p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N} d_t^2 u_h^{n+1})_h,
\end{aligned}$$

where we used $u_h^{n+1} - \widetilde{u}_h^{n+1} = \tau^2 d_t^2 u_h^{n+1}$. Multiplication by τ, summation over $n = 0, ..., N$, discrete integration by parts, Young's inequality, and $d_t u_h^0 = 0$ show that for the right-hand side we have

$$\begin{aligned}
\tau^3\sum_{n=0}^{N}(p_h - p_h^{n+1}, \widetilde{\nabla}_{h,N} d_t^2 u_h^{n+1})_h &= \tau^3\sum_{n=0}^{N}(d_t p_h^{n+1}, \widetilde{\nabla}_{h,N} d_t u_h^n)_h + \tau^2(p_h - p_h^n, \widetilde{\nabla}_{h,N} d_t u_h^n)_h\Big|_{n=0}^{N+1}\\
&\le \frac{\tau^2}{2}\Big(\sum_{n=0}^{N}\tau^2\|\widetilde{\nabla}_{h,N} d_t u_h^n\|^2 + \|d_t p_h^{n+1}\|_h^2\Big) + \frac{1}{2}\|p_h - p_h^{N+1}\|_h^2 + \frac{\tau^4}{2}\|\widetilde{\nabla}_{h,N} d_t u_h^{N+1}\|^2.
\end{aligned}$$

A combination of the estimates proves the asserted bound. The bound implies that $u_h^n \to u_h$ and $p_h^n \to \widetilde{p}_h$ for some $\widetilde{p}_h \in \mathcal{S}_N^1(\mathcal{T}_h)^d$ as $n \to \infty$. Since $d_t u_h^{n+1} \to 0$ and $d_t p_h^{n+1} \to 0$ we have that the pair $(u_h, \widetilde{p}_h)$ is a saddle point as in Lemma 4.2. $\square$

5. Numerical experiments

We discuss in this section the practical performance of the proposed error estimate. To illustrate some of its features in a well understood setting, we first report our experience in the case of elliptic model problems. The theory is then applied to the non-smooth functional E from Section 4 defined on $BV(\Omega)$.

5.1. **Elliptic problems.** We consider two elliptic problems defined on $H^1(\Omega)$ that reveal some fundamental properties of the error estimate provided by Theorem 3.1.

Example 5.1. *For $\Omega = (-1,1)^2 \setminus ([0,1) \times (-1,0])$ and $f \equiv 1$ consider*

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x - \int_\Omega f v \,\mathrm{d}x$$

for $v \in V = H^1_0(\Omega)$. For the unique minimizer $u \in H^1_0(\Omega)$ and a Galerkin approximation $u_h \in \mathcal{S}^1_0(\mathcal{T}_h)$ we have according to Example 3.1

$$\|\nabla(u - u_h)\| \le \|\nabla u_h - q_h\|$$

for every $q_h \in H(\mathrm{div};\Omega)$ with $-\operatorname{div} q_h = f$ in Ω. The employed initial triangulation was obtained from two uniform refinements of a coarse triangulation of Ω consisting of 6 triangles with diameters $\sqrt{2}$ that are halved squares along the diagonal $(1,1)$.

Our first choice of $q_h \in H(\mathrm{div};\Omega)$ is obtained by an averaging of the gradient of the approximate solution u_h, i.e., we employ $q_h = \mathcal{A}_h(\nabla u_h) \in \mathcal{S}^1(\mathcal{T}_h)^2$ defined by $q_h = \sum_{z\in\mathcal{N}_h} q_z \varphi_z$ with

$$q_z = \frac{1}{|\omega_z|}\int_{\omega_z} \nabla u_h \,\mathrm{d}x$$

for every $z \in \mathcal{N}_h$ and $\omega_z = \operatorname{supp}\varphi_z$. In general, the vector field q_h does not satisfy $-\operatorname{div} q_h = f$. The second choice results from the solution of a discretization of the dual problem with the lowest order Raviart-Thomas finite element space, i.e., we compute $(p_h, \overline{u}_h) \in \mathcal{RT}^0(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ with

$$\begin{aligned} (p_h, q_h) \quad &+ (\operatorname{div} q_h, \overline{u}_h) = 0, \\ (\operatorname{div} p_h, v_h) \quad & \qquad\qquad = -(f, v_h) \end{aligned}$$

for all $(q_h, v_h) \in \mathcal{RT}^0(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$. The corresponding error estimators are given by

$$\eta_{\mathcal{A}} = \|\nabla u_h - \mathcal{A}_h(\nabla u_h)\|, \qquad \eta_{DP} = \|\nabla u_h - p_h\|.$$

We associate to η_{DP} the elementwise refinement indicator specified by $\eta_{DP}(T) = \|\nabla u_h - p_h\|_{L^2(T)}$ for $T \in \mathcal{T}_h$. Notice that η_{DP} is a reliable error estimator owing to the results of Section 3 and the fact that f is elementwise constant. The reliability of $\eta_{\mathcal{A}}$ does not follow from the arguments above but can be proved up to a generic constant with different arguments, cf. [CB02].

Figure 1 shows the error estimators and the error for sequences of uniformly and adaptively refined triangulations. The adaptively refined meshes were obtained through a red-green-blue refinement strategy and a set of marked elements given by $\mathcal{M} = \{T \in \mathcal{T}_h : \eta_{DP}(T) \ge (1/2)\max_{T'\in\mathcal{T}_h} \eta_{DP}(T')\}$. The error $\delta = \|\nabla(u - u_h)\| = (\|\nabla u\|^2 - \|\nabla u_h\|^2)^{1/2}$ was computed with an approximation of $\|\nabla u\|$ obtained by an extrapolation of corresponding values for finite element approximations on a sequence of uniform triangulations. We see that the estimator $\eta_{\mathcal{A}}$ provides an accurate approximation of the error δ but is not a reliable upper bound. This is satisfied for the estimator η_{DP} which leads to some overestimation but defines a guaranteed upper bound for the error. The adaptive strategy improves the suboptimal experimental convergence rate $\delta \sim N^{-1/3}$ of uniform mesh-refinement to the quasi-optimal rate $\delta \sim N^{-1/2}$, where $N = \#\mathcal{N}_h$ is the number of nodes that define the approximation u_h on a triangulation $\mathcal{T}_h$.

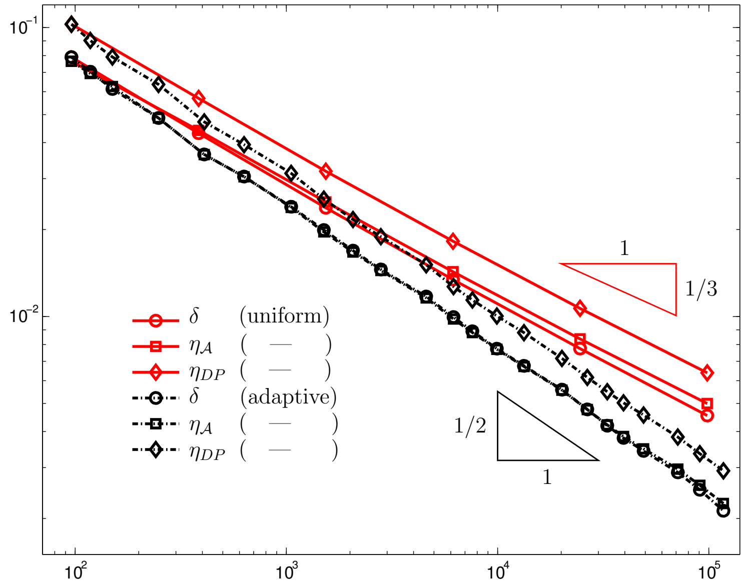

FIGURE 1. Error $\delta = \|\nabla(u - u_h)\|$ and error estimators $\eta_{\mathcal{A}}$ and η_{DP} versus degrees of freedom N for a Poisson problem on an L-shaped domain defined in Example 5.1 for uniformly and adaptively refined triangulations. All quantities decay with the same rates $N^{-1/3}$ and $N^{-1/2}$ for uniform and adaptive mesh-refinement, respectively.

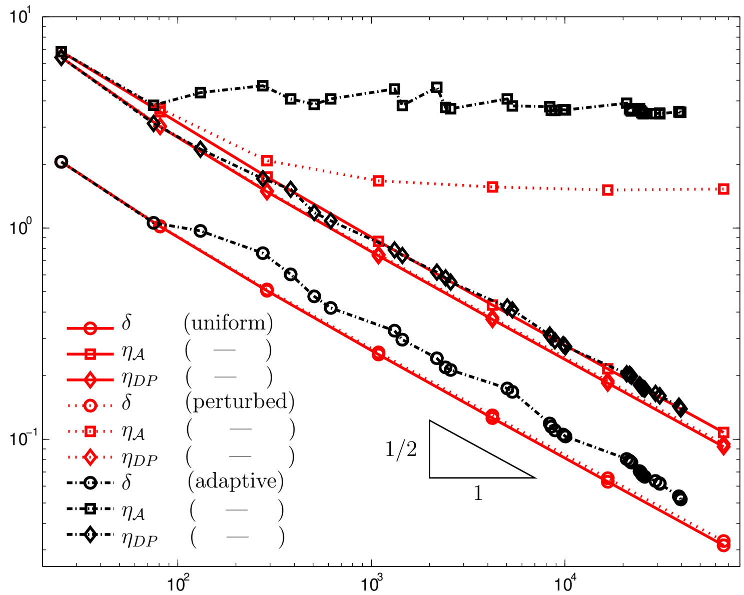

FIGURE 2. Error δ and error estimators $\eta_{\mathcal{A}}$ and η_{DP} versus degrees of freedom N for a Helmholtz type problem on a square defined in Example 5.2 for uniformly, perturbed uniformly, and adaptively refined triangulations. The reliable estimator η_{DP} decays at the same optimal rate as the error on all sequences of triangulations while the reliable estimator $\eta_{\mathcal{A}}$ is efficient only on uniformly refined, highly symmetric triangulations.

Example 5.2. *For* $\Omega = (-1,1)^2$ *and* $g(x_1,x_2) = (1+\pi^2/\alpha)\cos(\pi x_1)$ *consider*

$$J(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 \,\mathrm{d}x + \frac{\alpha}{2}\int_\Omega |v-g|^2 \,\mathrm{d}x$$

for $v \in V = H^1(\Omega)$. *For the unique minimizer* $u(x_1,x_2) = \cos(\pi x_1)$ *and a Galerkin approximation* $u_h \in \mathcal{S}^1(\mathcal{T}_h)$ *we have according to Example 3.2*

$$\|\nabla(u-u_h)\|^2 + \alpha\|u-u_h\|^2 \le \|\nabla u_h - q_h\|^2 + (1/\alpha)\|\operatorname{div} q_h - \alpha(u_h - g)\|^2$$

for every $q_h \in H_N(\mathrm{div};\Omega)$. *We start with a triangulation that is obtained from two uniform refinements of a coarse triangulation of* Ω *consisting of 2 triangles with diameters* $2\sqrt{2}$ *that are halved squares along the diagonal* $(1,1)$.

Again, we consider estimators obtained by simple averaging and by a numerical solution of the dual problem. To satisfy the condition $q_h \cdot \nu = 0$ on $\partial\Omega$, we modify the coefficients in the definition of $\mathcal{A}_h(\nabla u_h)$ above by projecting q_z onto the orthogonal complement of the space spanned by the normals at a boundary node z, i.e.,

$$\widetilde{q}_z = \Pi_{N_z^\perp} q_z,$$

where $N_z = \operatorname{span}\{\nu_E : E \in \mathcal{E} \cap \partial\Omega,\, z \in E\}$ is the span of all normals of boundary edges E of the triangulation that have the node z as an endpoint. We then set

$$\widetilde{\mathcal{A}}_h(\nabla u_h) = \sum_{z\in\mathcal{N}_h} \widetilde{q}_z \varphi_z \in \mathcal{S}^1_N(\mathcal{T}_h)^2 \subset H_N(\mathrm{div};\Omega).$$

In contrast to Example 5.1 the vector field $q_h = \widetilde{\mathcal{A}}_h(\nabla u_h)$ leads to a finite guranteed upper bound for the error. Our employed numerical solution of the discretized dual problem is the unique pair $(p_h, \overline{u}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ with

$$\begin{aligned}
(p_h, q_h) \qquad &+ (\operatorname{div} q_h, \overline{u}_h) = 0,\\
(\operatorname{div} p_h, v_h) - \alpha(\overline{u}_h, v_h) \quad &= -\,\alpha(g_h, v_h)
\end{aligned}$$

for all $(q_h, v_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$. For the Helmholtz problem defined in Example 5.2 we have that both estimators

$$\begin{aligned}
\eta^2_{\mathcal{A}} &= \|\nabla u_h - \widetilde{\mathcal{A}}_h(\nabla u_h)\|^2 + (1/\alpha)\|\operatorname{div}\widetilde{\mathcal{A}}_h(\nabla u_h) - \alpha(u_h - P_h g)\|^2,\\
\eta^2_{DP} &= \|\nabla u_h - p_h\|^2 + (1/\alpha)\|\operatorname{div} p_h - \alpha(u_h - P_h g)\|^2
\end{aligned}$$

are according to Example 3.2 reliable upper bounds for the error

$$\delta^2 = \|\nabla(u-u_h)\|^2 + \alpha\|u-u_h\|^2$$

if $g = P_h g$. Data oscillation terms that are related to an approximation error $g \neq P_h g$ are neglected in the following.

The numerical results shown in Figure 2 confirm that the estimators $\eta_{\mathcal{A}}$ and η_{DP} serve as guaranteed upper bounds for the error. In case of uniformly refined triangulations they converge with the same optimal rate as the error δ. The error estimator η_{DP} obtained through the solution of the discrete dual problem is also efficient on adaptively refined and perturbed uniform triangulations where inner nodes z are randomly perturbed by a vector ξ_z with $|\xi_z| \le h/10$, i.e., it decays at the same rate as the error δ, and is nearly insensitive to mesh perturbations. This is not the case for the estimator $\eta_{\mathcal{A}}$ which remains almost constant for adaptively and perturbed uniformly refined triangulations with more than 200 nodes. A more careful investigation of the contributions to the estimators show that the failure of $\eta_{\mathcal{A}}$ on non-symmetric triangulations is due to the term $\|\operatorname{div}\widetilde{\mathcal{A}}_h(\nabla u_h) - \alpha(u_h - g_h)\|$.

Remark 5.1. *We also tried an estimator defined through averaging in the Raviart-Thomas space $\mathcal{RT}^0(\mathcal{T}_h)$ defined through*

$$\mathcal{A}_h^{\mathcal{RT}}(\nabla u_h) = \sum_{E\in\mathcal{E}} \left\{\nabla u_h\right\}_E \cdot \nu_E \psi_E,$$

where $\{\nabla u_h\}_E$ is the average of ∇u_h on an edge E, ν_E a fixed unit normal for every $E \in \mathcal{E}$, and ψ_E the basis of $\mathcal{RT}^0(\mathcal{T}_h)$ satisfying $\psi_E \cdot \nu_{E'} = \delta_{EE'}$ for all $E, E' \in \mathcal{E}$. The corresponding error estimator led to similar results as the estimator $\eta_\mathcal{A}$. In particular, the failure of estimation by averaging cannot be attributed solely to the limited approximation properties of the H^1 conforming space $\mathcal{S}^1(\mathcal{T}_h)^d$ in $H(\operatorname{div};\Omega)$.

5.2. **Total variation regularization.** The experiments for elliptic problems show that simple averaging may not lead to efficient error control. Therefore, we solve the dual problem of the non-smooth model problem as discussed in Section 4 to define a reliable error estimator.

Example 5.3. *Set $\Omega = (-1,1)^2$, $\alpha = 100$, $g(x) = \chi_{B^\infty_{1/2}}(x)$, where $B^\infty_{1/2} = \{(x_1,x_2) \in \mathbb{R}^2 : \max\{|x_1|,|x_2|\} \le 1/2\}$, and consider the minimization problem defined through the functional*

$$E(v) = \int_\Omega |Dv| + \frac{\alpha}{2}\|v-g\|^2$$

for $v \in BV(\Omega)\cap L^2(\Omega)$. For the minimal $u \in BV(\Omega)\cap L^2(\Omega)$, arbitrary $q_h \in H_N(\operatorname{div};\Omega)$ with $|q_h| \le 1$ in Ω, and an arbitrary appoximation $u_h \in W^{1,1}(\Omega)\cap L^2(\Omega)$ we have according to Theorem 4.1

$$\frac{\alpha}{2}\|u-u_h\|^2 \le \int_\Omega |\nabla u_h|\,\mathrm{d}x + \frac{\alpha}{2}\int_\Omega |u_h-g|^2\,\mathrm{d}x + \frac{1}{2\alpha}\int_\Omega |\operatorname{div} q_h + \alpha g|^2\,\mathrm{d}x - \frac{\alpha}{2}\int_\Omega |g|^2\,\mathrm{d}x.$$

The initial triangulation was obtained from two uniform refinements of a coarse triangulation with 2 triangles that are halved squares with diameter $\sqrt{2}2$.

An approximation $p_h \in \mathcal{S}^1_N(\mathcal{T}_h)^d$ of a solution of the dual problem that satisfies $|p_h(z)| \le 1$ for all nodes $z \in \mathcal{N}_h$ and hence $|p_h| \le 1$ in Ω is computed with Algorithm (A'). For this and the approximate solution of the discretized primal problem with Algorithm (A) we always used the step-size $\tau = h_{min}/10$ with the minimial mesh-size h_{min} of a triangulation $\mathcal{T}_h$. The approximations u_h and p_h define the error estimator

$$\eta_{DP}^2 = \int_\Omega |\nabla u_h|\,\mathrm{d}x + \frac{\alpha}{2}\int_\Omega |u_h-g|^2\,\mathrm{d}x + \frac{1}{2\alpha}\int_\Omega |\operatorname{div} p_h + \alpha g|^2\,\mathrm{d}x - \frac{\alpha}{2}\int_\Omega |g|^2\,\mathrm{d}x.$$

To the reliable estimator η_{DP} we associate the elementwise defined refinement indicators

$$\eta_{DP}(T) = \Big|\int_T |\nabla u_h|\,\mathrm{d}x + \frac{\alpha}{2}\int_T |u_h-g|^2\,\mathrm{d}x + \frac{1}{2\alpha}\int_T |\operatorname{div} p_h + \alpha g|^2\,\mathrm{d}x - \frac{\alpha}{2}\int_T |g|^2\,\mathrm{d}x\Big|^{1/2}$$

that steer the adaptive algorithm.
Figure 3 shows the behaviour of the estimator η_{DP} for a sequence of uniformly and adaptively refined triangulations. On the uniform triangulations we observe that

$$\eta_{DP} \sim h^{1/4}$$

which coincides with the theoretically predicted convergence rate. The adaptive strategy leads to a smaller error estimator and an improved experimental rate of convergence $\eta_{DP} \sim h^{1/2}$ which is the optimal rate for the approximation of a function $u \in BV(\Omega)\cap L^\infty(\Omega)$ on a sequence of uniform triangulations. The refinement strategy refines the grid in a neighbourhood of the discontinuity set of the data where a discontinuity of the exact solution is expected. This is illustrated in the sequence of adaptively generated triangulations shown in Figure 4.

The refinement indicators $(\eta_{DP}(T) : T \in \mathcal{T}_h)$ can also be used for de-refinement of a triangulation, i.e., for mesh coarsening. This is of particular importance when a fine initial grid is required

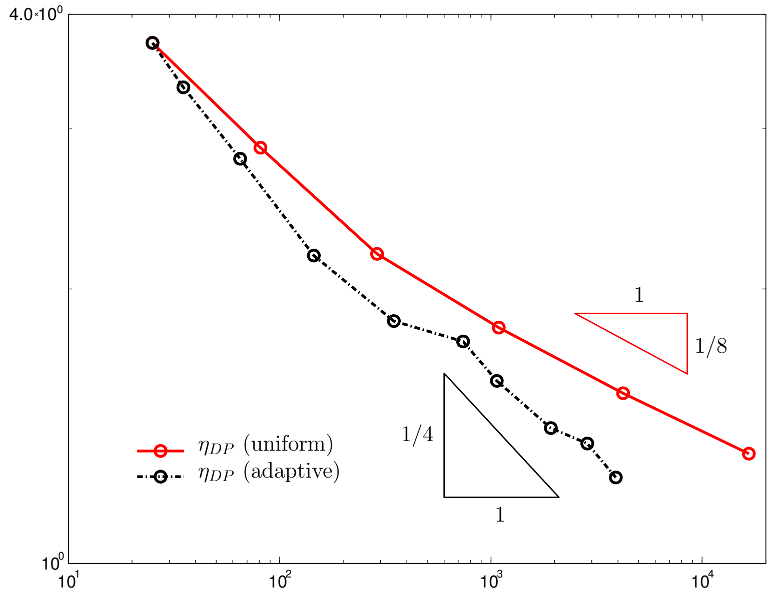

FIGURE 3. Error estimator η_{DP} versus degrees of freedom N for the total variation minimization problem defined in Example 5.3. The estimator decays at a rate $N^{-1/8}$ for uniform refinement. The numbers are reduced on adaptively refined meshes with comparable numbers of degrees of freedom and the rate is improved to $\eta_{DP} \sim N^{-1/4}$.

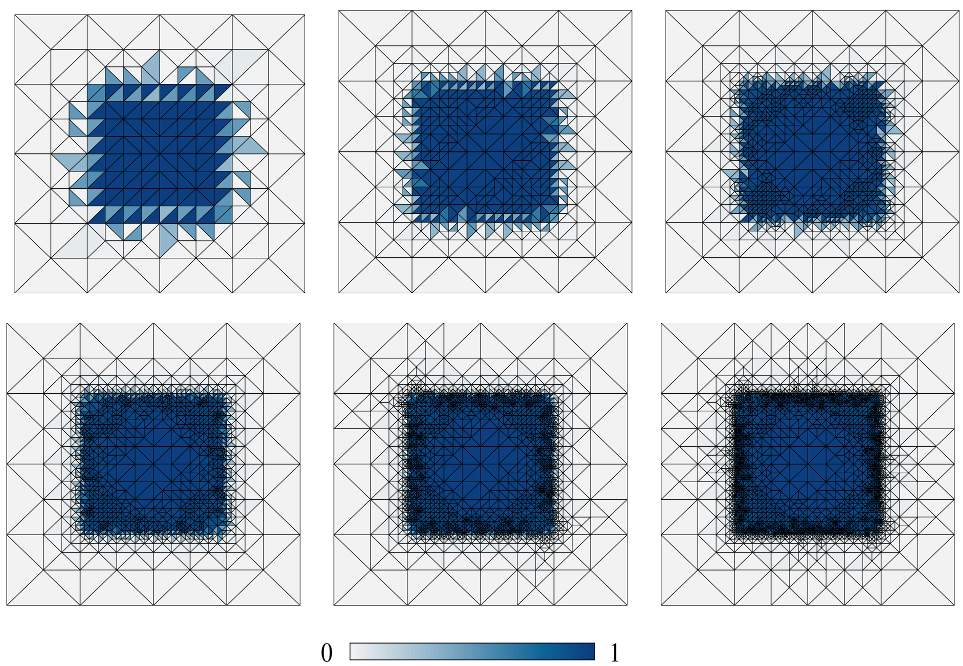

FIGURE 4. Adaptively generated grids and numerical solutions u_h in Example 5.3. The adaptive strategy steered by the refinement indicators $\eta_{DP}(T)$ automatically refines a neighbourhood of the discontinuity set of the data function g which is expected to coincide with the singuarity set of the exact solution u.

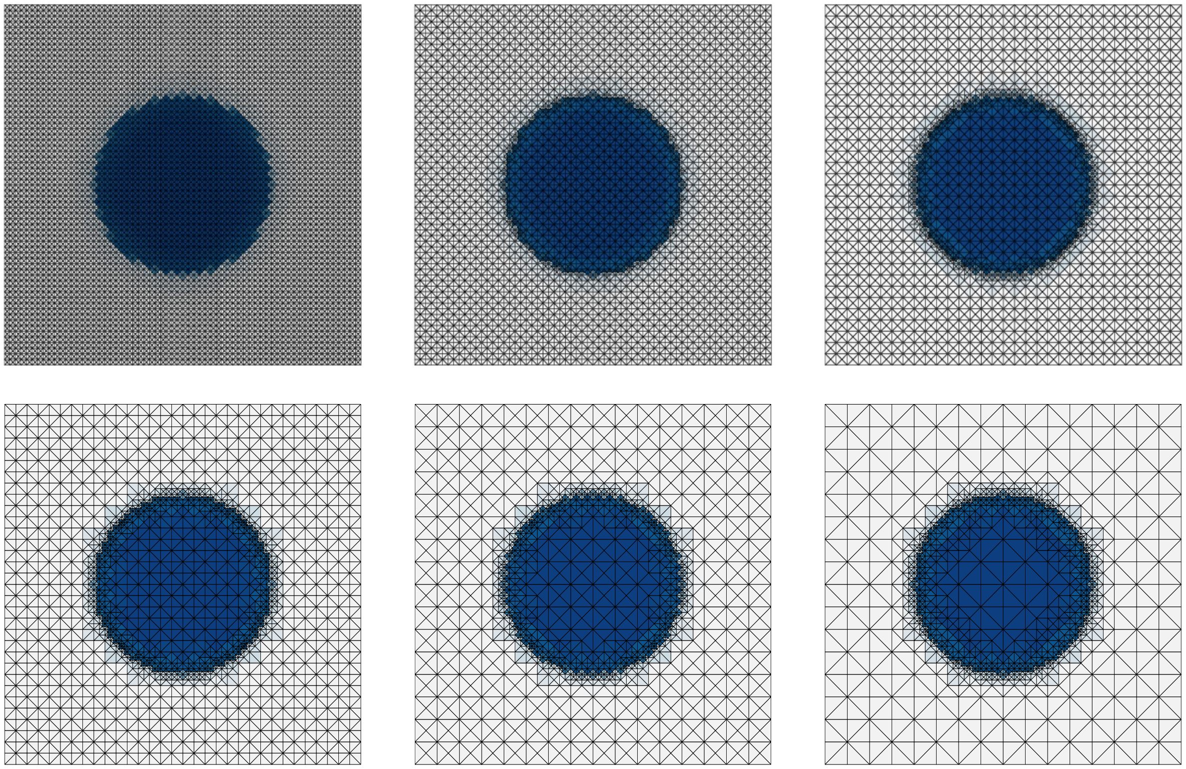

FIGURE 5. Adaptively refined and coarsened grids and numerical solutions u_h in Example 5.4. The mesh is significantly coarsened away from the circular discontinuity set and refined therein.

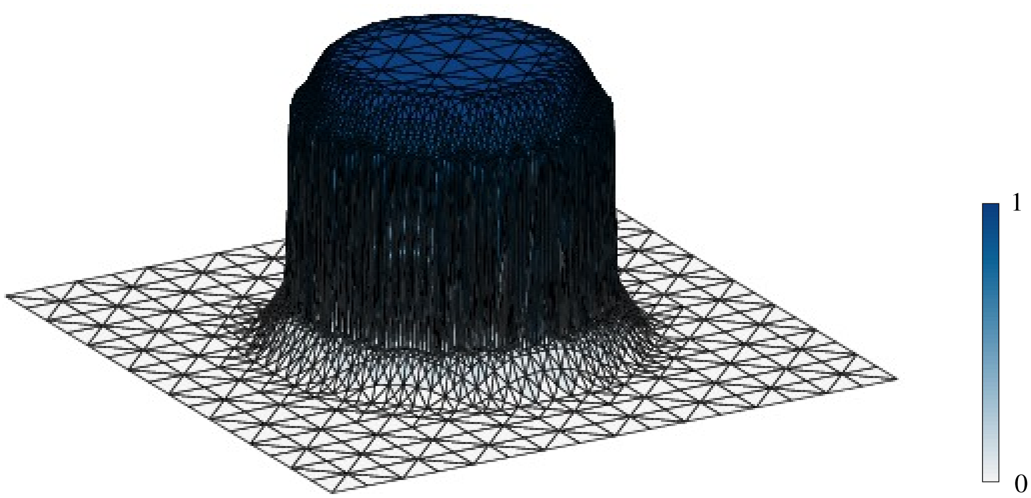

FIGURE 6. Numerical solution obtained through adaptive refinement and coarsening based on the indicators $(\eta_{DP}(T) : T \in \mathcal{T}_h)$ in Example 5.4. The algorithm automatically coarsens the triangulation away from the discontinuity set and refines the mesh in a neighbourhood of it leading to an efficient approximation scheme.

to resolve given data g. To illustrate this we consider the following experiment in which the discontinuity set of g is circular and not exactly resolved on the employed triangulations.

Example 5.4. *Let* $\Omega = (-1,1)^2$, $\alpha = 100$, *and* $g(x) = \chi_{B^2_{1/2}}(x)$ *for* $B^2_{1/2} = \{x \in \mathbb{R}^2 : (x_1^2 + x_2^2)^{1/2} \le 1/2\}$. *The initial triangulation was obtained from 13 global bisection steps of a coarse triangulation with 2 triangles and whose refinement edges are chosen as longest edges.*

Our adaptive mesh-refinement and coarsening strategy consisted in marking elements in a triangulation $\mathcal{T}_h$ for coarsening if $\eta_{DP}(T) \le (1/2)\max_{T' \in \mathcal{T}_h} \eta_{DP}(T')$. Subsequently, the mesh was refined according the rule $\widetilde{\eta}_{DP}(T) \ge (1/2)\max_{T' \in \mathcal{T}_h} \widetilde{\eta}_{DP}(T')$ using the elementwise function $\widetilde{\eta}_{DP}$ obtained from a restriction of η_{DP} onto the coarsened mesh. The mesh coarsening was based on the algorithm proposed and analyzed in [BS12].
The sequence of meshes shown in Figures 5 and 6 show that adaptive mesh coarsening can substantially improve the efficiency of the image regularization and it can be regarded as an adaptive image compression technique. Since we use an approximation $g_h \in \mathcal{L}^0(\mathcal{T}_h)$ of g on the initial mesh and prolongate and restrict this function to coarser and finer meshes we cannot expect convergence to a circular interface. The role of data oscillation and numerical integration is not investigated in this article. Figure 7 reveals the limited applicability of Algorithm (A') to compute solutions of the primal problem. Artificial oscillations occur at the discontinuity set and a discrete maximum principle is violated.

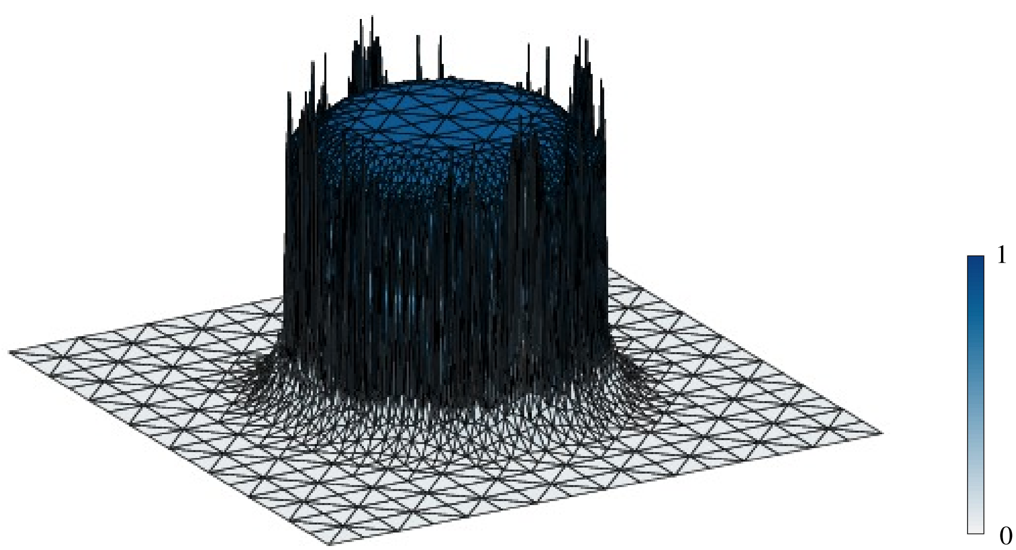

FIGURE 7. Numerical approximation u'_h obtained with Algorithm (A') in Example 5.4. In contrast to the numerical solution computed with Algorithm (A) oscillations occur at the discontinuity set. The purpose of Algorithm (A') is to compute a conforming solution of the dual problem.

REFERENCES

[Bar12] Sören Bartels, *Total variation minimization with finite elements: Convergence and iterative solution*, SIAM J. Numer. Anal. **50** (2012), no. 3, 1162–1180.

[BMR12] Sören Bartels, Alexander Mielke, and Tomáš Roubíček, *Quasi-static small-strain plasticity in the limit of vanishing hardening and its numerical approximation*, SIAM J. Numer. Anal. **50** (2012), no. 2, 951–976.

[Bra07] Dietrich Braess, *Finite elements*, third ed., Cambridge University Press, Cambridge, 2007.

[Brè67] Lev M. Brègman, *A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 620–631.

[BS12] Sören Bartels and Patrick Schreier, *Local coarsening of simplicial finite element meshes generated by bisections*, BIT Numerical Mathematics (2012), 1–11, 10.1007/s10543-012-0378-0.

[CB02] Carsten Carstensen and Sören Bartels, *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I. Low order conforming, nonconforming, and mixed FEM*, Math. Comp. **71** (2002), no. 239, 945–969 (electronic).

[CP11] Antonin Chambolle and Thomas Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision **40** (2011), no. 1, 120–145.

[ET99] Ivar Ekeland and Roger Témam, *Convex analysis and variational problems*, english ed., Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[HK04] Michael Hintermüller and Karl Kunisch, *Total bounded variation regularization as a bilaterally constrained optimization problem*, SIAM J. Appl. Math. **64** (2004), no. 4, 1311–1333 (electronic).

[MO08] Antonio Marquina and Stanley J. Osher, *Image super-resolution by TV-regularization and Bregman iteration*, J. Sci. Comput. **37** (2008), no. 3, 367–382.

[NSV00] Ricardo H. Nochetto, Giuseppe Savaré, and Claudio Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math. **53** (2000), no. 5, 525–589.

[Rep00] Sergey I. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals*, Math. Comp. **69** (2000), no. 230, 481–500.

[Roc97] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Reprint of the 1970 original, Princeton Paperbacks.

[ROF92] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena **60** (1992), no. 1-4, 259 – 268.

[Suq78] Pierre-Marie Suquet, *Existence et régularité des solutions des équations de la plasticité*, C. R. Acad. Sci. Paris Sér. A-B **286** (1978), no. 24, A1201–A1204.

[Tho11] Marita Thomas, *Quasistatic damage evolution with spatial bv-regularization*, Weierstraß Institute Berlin, Preprint No. 1638, 2011.

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

E-mail address: `bartels@mathematik.uni-freiburg.de`