ESAIM: M2AN 54 (2020) 1635–1660
https://doi.org/10.1051/m2an/2019074

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

PRIMAL-DUAL GAP ESTIMATORS FOR *A POSTERIORI* ERROR ANALYSIS OF NONSMOOTH MINIMIZATION PROBLEMS

Sören Bartels* and Marijo Milicevic

Abstract. The primal-dual gap is a natural upper bound for the energy error and, for uniformly convex minimization problems, also for the error in the energy norm. This feature can be used to construct reliable primal-dual gap error estimators for which the constant in the reliability estimate equals one for the energy error and equals the uniform convexity constant for the error in the energy norm. In particular, it defines a reliable upper bound for any functions that are feasible for the primal and the associated dual problem. The abstract *a posteriori* error estimate based on the primal-dual gap is provided in this article, and the abstract theory is applied to the nonlinear Laplace problem and the Rudin–Osher–Fatemi image denoising problem. The discretization of the primal and dual problems with conforming, low-order finite element spaces is addressed. The primal-dual gap error estimator is used to define an adaptive finite element scheme and numerical experiments are presented, which illustrate the accurate, local mesh refinement in a neighborhood of the singularities, the reliability of the primal-dual gap error estimator and the moderate overestimation of the error.

Mathematics Subject Classification. 49M29, 65K15, 65N15, 65N50.

Received February 11, 2019. Accepted October 16, 2019.

1. Introduction

Many problems in various applications like partial differential equations, mechanics, imaging, and operations research can be formulated as convex minimization problems of the form

$$\inf_{u\in X} E(u) = \inf_{u\in X} F(Bu) + G(u)$$

with convex functionals F, G and a bounded linear operator B. Examples are the nonlinear Laplace equation, the Rudin–Osher–Fatemi model for image denoising, obstacle problems or convex programming. Depending on the data and the geometry of the problem a solution $u \in X$ of the above minimization problem may suffer from singularities which can harm the convergence rate as the mesh size $h > 0$ of a finite element method tends to zero. A well-known example for this phenomenon is the linear Laplace problem on the L-shaped domain. The geometry of the domain leads to a convergence rate of order $\mathcal{O}(h^\gamma)$ instead of $\mathcal{O}(h)$ in the energy norm, where $0 < \gamma < 1$ and γ depends on the angle at the reentrant corner. Singularities may also arise due to intrinsic

Keywords and phrases. Convex minimization, primal-dual gap, adaptive mesh refinement, nonlinear Laplace, image denoising.

Department of Applied Mathematics, Mathematical Institute, University of Freiburg, Hermann-Herder-Str. 9, 79104 Freiburg im Breisgau, Germany.
*Corresponding author: bartels@mathematik.uni-freiburg.de

properties of the functions in the underlying space X. An example is the space of functions with bounded variation $\mathrm{BV}(\Omega)$, which allows for jumps along interfaces, which is of interest, *e.g.*, in image processing to preserve sharp edges. Yet, these jumps cause problems in the finite element approximation of BV-functions.

One way to overcome these drawbacks is adaptive mesh refinement. The general procedure of adaptive routines is to compute an approximation of the minimizer in the discrete space with a given underlying triangulation, compute *a posteriori* error estimators on the basis of the computed approximation, refine the mesh locally where the error estimators are relatively large and to compute a new approximate solution corresponding to the new mesh. In this sense, adaptive methods are iterative numerical methods. The reader is referred to, *e.g.*, [1, 5, 40, 49, 52] to get an overview of adaptive finite element methods.

The design of *a posteriori* error estimators is fundamental to adaptive finite element methods. Particularly, it is crucial that the error estimators define upper (*reliability*) and lower (*efficiency*) bounds for an appropriate measure of the error and that the constant in the upper bound is small and known. We will consider primal-dual gap error estimators which can be derived using duality theory from convex analysis. In the contributions [8, 11, 42–47] these primal-dual gap error estimators have been introduced and used for various problems, *e.g.*, elasto-plasticity and optimal transport. In [45], the primal-dual gap error estimator has been analyzed for general convex minimization problems with uniformly convex functionals and the relation to other *a posteriori* error estimators based on, *e.g.*, residual and gradient recovery methods has been addressed. Yet, the numerical study of primal-dual gap error estimators has not been considered in any of those contributions. We will analyze primal-dual gap based error estimators for the nonlinear Laplace problem

$$E_{\Delta_\sigma}(u) = \frac{1}{\sigma}\int_\Omega |\nabla u|^\sigma \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x \quad \longrightarrow \quad \text{Min.!}$$

with $1 < \sigma < \infty$, which has also been addressed in [43] without a numerical study, and for the Rudin–Osher–Fatemi (ROF) model

$$E_{\mathrm{rof}}(u) = |Du|(\Omega) + \frac{\alpha}{2}\|u-g\|^2_{L^2(\Omega)} \quad \longrightarrow \quad \text{Min.!}$$

with $|Du|(\Omega)$ the total variation of u, which has been analyzed in, *e.g.*, [8].

The nonlinear Laplace problem serves as a model problem for degenerate nonlinear systems. Results concerning the regularity of solutions, their approximation by finite elements and *a priori* error estimates can be found, *e.g.*, [7, 18, 22, 23, 25, 26, 31, 34, 35]. An important observation in the *a priori* error analysis was that the energy norm is not well suited for the analysis since optimal convergence rates can only be guaranteed under restrictive assumptions on the regularity of the solution, *cf.* [7, 18, 31, 34, 35]. It turned out that a so-called *quasi-norm*, which is a weighted L^2-norm of the gradient with a weight depending on the gradient and which has been introduced in [7], is more appropriate for the analysis of the nonlinear Laplacian, *cf.* [20, 25]. Particularly, the optimal convergence rate $\mathcal{O}(h)$ for P1 finite elements can be proven under much less restrictive regularity assumptions on the solution, *cf.* [20, 25, 26]. In [36–38] residual-based *a posteriori* error estimators have been proposed and reliability and efficiency has been established with respect to the quasi-norm. However, the involved constants are not explicitly available. Residual-based quasi-norm error estimators yielding explicit constants in the reliability estimate have been discussed in [16] under the assumption that the modulus of the gradient is greater than zero almost everywhere in the domain whereas the reliability and efficiency of quasi-norm error estimators based on gradient recovery techniques has been established in [17]. The convergence of an adaptive scheme with residual-based *a posteriori* error estimators has been proven in [51]. In [13, 21] the linear convergence and optimality of an adaptive method driven by residual-based quasi-norm error estimators has been proven. The involved constants particularly for the upper bound depend on the nonlinearity of the problem. In [28, 29] the error is measured in a residual flux-based dual norm and the *a posteriori* error estimator consists of a residual term, a diffusive flux term and a linearization term. Flux reconstruction techniques are presented to compute the error estimator and reliability (with constant one) and efficiency (with a constant independent of the nonlinearity of the problem) are shown. Particular focus is on the balance of linearization and discretization errors.

The ROF model serves as a prototype for BV-regularized minimization problems with applications, *e.g.*, in image processing (*cf.* [4, 48]) and mechanics (*cf.* [50]). A primal-dual gap error estimator has been proposed to define an adaptive algorithm for the ROF problem in [8], which has proven to accurately detect the *a priori* unknown jump sets of the minimizer yielding locally refined meshes in a neighborhood of the jump sets. Therein, a finite element method has been proposed where the primal and dual problem have been discretized with continuous, elementwise affine finite elements. However, the approximation of the dual ROF problem by continuous finite elements is suboptimal since the dual ROF problem is posed on $H_N(\mathrm{div};\Omega)$. This is reflected in the experiments in [8] where oscillations of the approximations along the interface can be observed.

The advantage of primal-dual gap error estimators is that they are applicable to a large class of convex minimization problems and naturally yield upper bounds for the energy difference between the energy of an arbitrary admissible test function and the optimal energy with constant one. In case of F or G being strongly convex (or coercive) they also define upper bounds for some appropriate error measure with a constant depending on the coercivity constant. Particularly, they define reliable upper bounds independently of the iterative solver used to approximate discrete solutions to the primal and dual problem, *i.e.*, the primal-dual gap error estimator can be evaluated at any two feasible functions for the primal and the dual problem to obtain an upper bound for the error. Last but not least, the functionals F and G need not be assumed to be differentiable and there does not need to exist a variational formulation of the primal problem to establish the reliability of the primal-dual gap error estimators.

In this paper we will consider primal-dual gap error estimators for both the nonlinear Laplace problem and the ROF problem. While in [43] the primal-dual gap error estimator has been considered for the nonlinear Laplacian, the discretization and numerical implementation is missing. Furthermore, noting that the dual problem corresponding to the nonlinear Laplace problem is given by a smooth, linearly constrained optimization problem a modified error estimator, which is an upper bound for the primal-dual gap error estimator, is suggested in [43] allowing for dual test functions that do not satisfy the linear constraint. We will consider the "original" primal-dual gap error estimator to control the quasi-norm used in [7, 21]. In particular, the primal-dual gap error estimator η_{pd} can be used to improve the reliability estimate for the convergent, reliable and efficient residual-based error estimator η_{res} analyzed in [13, 21], *i.e.*, defining $\eta_{\mathrm{com}} = \min\{\eta_{\mathrm{pd}}, \eta_{\mathrm{res}}\}$ we obtain a reliable, robust, efficient and convergent error estimator. Continuous, piecewise affine finite elements are used for the discretization of the primal nonlinear Laplace problem and the ROF problem posed in $W^{1,\alpha}(\Omega)$ and $\mathrm{BV}(\Omega)\cap L^2(\Omega)$, respectively. The dual problems are posed in $W^{\beta}(\mathrm{div};\Omega)$, $\beta = \alpha/(\alpha-1)$, and $H_N(\mathrm{div};\Omega)$ in case of the nonlinear Laplacian and the ROF problem, respectively. In both cases we use the Brezzi–Douglas–Marini finite element (*cf.* [14]), which consists of discontinuous piecewise affine vector fields with continuous normal components across interelement sides, for the discretization. This is in contrast to the discretization in [8] where the dual ROF problem has been discretized with continuous, piecewise affine vector fields, which is known to be problematic in, *e.g.*, the discretization of the dual formulation of the linear Laplacian with mixed finite elements. Particularly, oscillations are observed in the approximation of u along the interface, *cf.* Section 6. The discrete optimization problems related to the primal and the dual problems are solved using the Variable-Alternating Direction Method of Multipliers (Variable-ADMM) proposed in [10] which is an operator splitting method with variable step sizes.

The paper is organized as follows. In Section 2 we introduce the notation, important function spaces and finite element spaces and state some approximation results. The abstract primal-dual gap error estimator and *a posteriori* error estimate are the subject of Section 3. In Sections 4 and 5 we state the nonlinear Laplace problem and the ROF problem, respectively, and the associated dual problems, summarize *a priori* and *a posteriori* error estimates and briefly address the numerical solution of the discrete primal and dual problems. Finally, we present in Section 6 our numerical results for both problems for examples for which the exact solutions are explicitly available.

Let us remark that this article is part of the thesis [39], in which certain arguments have been elaborated.

2. Preliminaries

2.1. Function spaces and convex analysis

We let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded, polygonal Lipschitz domain with Dirichlet boundary Γ_D and Neumann boundary Γ_N such that $\partial\Omega = \Gamma_D \cup \Gamma_N$. The L^2-norm on Ω is denoted by $\|\cdot\|$ and is induced by the scalar product

$$(v, w) := \int_\Omega v \cdot w \, \mathrm{d}x$$

for scalar functions or vector fields $v, w \in L^2(\Omega; \mathbb{R}^r)$, $r \in \{1, d\}$, and we write $|\cdot|$ for the Euclidean norm.

For $s \geq 0$ and $\sigma \geq 1$ we let $W^{s,\sigma}(\Omega; \mathbb{R}^r)$ be the standard Sobolev space with norm $\|\cdot\|_{W^{s,\sigma}(\Omega)}$ and seminorm $|\cdot|_{W^{s,\sigma}(\Omega)}$ with differentiability exponent s and integrability exponent σ. The subspace $W_D^{s,\sigma}(\Omega; \mathbb{R}^r)$ consists of all functions in $W^{s,\sigma}(\Omega; \mathbb{R}^r)$ that vanish on Γ_D for $s \geq 1$ in the sense of traces. If $s = 0$ we write $L^\sigma(\Omega; \mathbb{R}^r)$ instead of $W^{s,\sigma}(\Omega; \mathbb{R}^r)$.

Finally, for $\sigma' \geq 1$, we denote by $W^{\sigma'}(\mathrm{div}; \Omega)$ the function space consisting of all vector fields $p \in L^{\sigma'}(\Omega; \mathbb{R}^d)$ such that there exists a function $f \in L^{\sigma'}(\Omega)$ with

$$\int_\Omega p \cdot \nabla\varphi \, \mathrm{d}x = -\int_\Omega f\varphi \, \mathrm{d}x$$

for all continuously differentiable, compactly supported functions $\varphi \in C_c^1(\Omega)$. If such a function $f \in L^{\sigma'}(\Omega)$ exists, we write $\mathrm{div}\, p = f$. The space $W^{\sigma'}(\mathrm{div}; \Omega)$ is equipped with the norm

$$\|\cdot\|_{W^{\sigma'}(\mathrm{div};\Omega)} = \|\cdot\|_{L^{\sigma'}(\Omega)} + \|\mathrm{div}\,\cdot\|_{L^{\sigma'}(\Omega)}.$$

Furthermore, we denote by $W_N^{\sigma'}(\mathrm{div}; \Omega)$ all elements of $p \in W^{\sigma'}(\mathrm{div}; \Omega)$ with $p \cdot n = 0$ on Γ_N in distributional sense, *i.e.*,

$$\langle p \cdot n, u\rangle = \int_\Omega p \cdot \nabla u \, \mathrm{d}x + \int_\Omega u \, \mathrm{div}\, p \, \mathrm{d}x = 0$$

for all $u \in W_D^{1,\sigma}(\Omega)$, where $\sigma \geq 1$ is the dual exponent to $\sigma' \geq 1$, *i.e.*, $1/\sigma + 1/\sigma' = 1$. If $\sigma' = 2$ we write $H(\mathrm{div}; \Omega)$ instead of $W^2(\mathrm{div}; \Omega)$, and accordingly $H_N(\mathrm{div}; \Omega)$ instead of $W_N^2(\mathrm{div}; \Omega)$.

For the general, abstract *a posteriori* error estimate we will work with two reflexive Banach spaces X and Y equipped with the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. We denote their duals by X' and Y' and the corresponding duality pairings by $\langle\cdot,\cdot\rangle_{X',X}$ and $\langle\cdot,\cdot\rangle_{Y',Y}$, respectively. The double duals X'' and Y'' are identified with X and Y, respectively. If X is a Hilbert space with inner product $(\cdot,\cdot)_X$, we identify the dual X' with X. Given a bounded linear operator $B : X \to Y$ we denote by $B' : Y' \to X'$ its adjoint. For proper, convex and lower-semicontinuous functionals $F : Y \to \mathbb{R} \cup \{\infty\}$ and $G : X \to \mathbb{R} \cup \{\infty\}$ the subdifferentials $\partial G(u) \subset X'$ at $u \in X$ and $\partial F(p) \subset Y'$ at $p \in Y$ are defined by

$$\begin{aligned} \partial G(u) &= \{w \in X' : \ \langle w, v - u\rangle_{X',X} + G(u) \leq G(v) \quad \text{for all } v \in X\}, \\ \partial F(p) &= \{\lambda \in Y' : \ \langle \lambda, q - p\rangle_{Y',Y} + F(p) \leq F(q) \quad \text{for all } q \in Y\}. \end{aligned}$$

Possible coercivity of the functionals F and G is characterized by non-negative mappings $\varrho_F : Y \times Y \to \mathbb{R}_+$ and $\varrho_G : X \times X \to \mathbb{R}_+$ such that for $w \in \partial G(u)$ and $\lambda \in \partial F(p)$ we have

$$\begin{aligned} \langle w, v - u\rangle_{X',X} + G(u) + \varrho_G(v, u) &\leq G(v) \quad \text{for all } v \in X, \\ \langle \lambda, q - p\rangle_{Y',Y} + F(p) + \varrho_F(q, p) &\leq F(q) \quad \text{for all } q \in Y. \end{aligned} \tag{2.1}$$

This can be regarded as a generalization of the notion of uniform convexity and strong convexity. The existence of non-trivial ϱ_G or ϱ_F will induce an error measure for which we establish primal-dual gap error estimates.

Possible constraints will be encoded using indicator functionals I_K with $K \subset X$ defining the constraint and I_K being defined as

$$I_K(v) = \begin{cases} 0, & \text{if } v \in K, \\ +\infty, & \text{if } v \notin K. \end{cases}$$

For the *a posteriori* error analysis we will need the Fenchel conjugates F^* and G^*, which are defined by

$$F^*(q) = \sup_{p \in Y} \langle q, p \rangle_{Y',Y} - F(p), \quad G^*(v) = \sup_{u \in X} \langle v, u \rangle_{X',X} - G(u).$$

These are used to convert the primal problems into dual problems.

2.2. Finite element spaces

We let $(\mathcal{T}_h)_{h>0}$ be a family of regular triangulations of Ω. The set $\mathcal{S}_h$ consists of all edges ($d = 2$) or faces ($d = 3$) of elements of $\mathcal{T}_h$ and $\mathcal{N}_h$ denotes the set of nodes of $\mathcal{T}_h$. The elementwise constant mesh size function $h_{\mathcal{T}} \in \mathcal{L}^\infty(\Omega)$ is defined by

$$h_{\mathcal{T}}|_T = h_T = \operatorname{diam}(T)$$

for all $T \in \mathcal{T}_h$. In the context of locally refined meshes we employ the average mesh size

$$\overline{h} = |\mathcal{N}_h|^{-1/d}$$

defined with the cardinality $|\mathcal{N}_h|$ of $\mathcal{N}_h$. Throughout the paper c will denote a generic, positive and mesh-independent constant.

For an integer $k \geq 0$ and a triangle $T \in \mathcal{T}_h$ let $P_k(T)$ be the space of polynomials on T with total degree at most k. We then consider for $r \in \{1, d\}$ the finite element spaces

$$\mathcal{S}^k(\mathcal{T}_h)^r := \{v_h \in C(\overline{\Omega}; \mathbb{R}^r) : \ v_h|_T \in P_k(T)^r \text{ for all } T \in \mathcal{T}_h\}$$

and

$$\mathcal{L}^k(\mathcal{T}_h)^r := \{q_h \in L^1(\Omega; \mathbb{R}^r) : \ q_h|_T \in P_k(T)^r \text{ for all } T \in \mathcal{T}_h\}.$$

For an elementwise continuous function $v \in C(\mathcal{T}_h)$ the operator

$$\widehat{\mathcal{I}}_h : C(\mathcal{T}_h) \to \mathcal{L}^1(\mathcal{T}_h)$$

is defined by the elementwise application of the standard nodal interpolation operator $\mathcal{I}_h : C(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$, *i.e.*, the function $\widehat{\mathcal{I}}_h v \in \mathcal{L}^1(\mathcal{T}_h)$ is the piecewise affine function uniquely defined by

$$\widehat{\mathcal{I}}_h v|_T(z_i) = v|_T(z_i)$$

for all $T \in \mathcal{T}_h$ and its vertices z_i, $i = 1, \ldots, d+1$. Note that $\widehat{\mathcal{I}}_h|_{C(\overline{\Omega})} = \mathcal{I}_h$. With the nodal basis $\{\varphi_z : \ z \in \mathcal{N}_h\} \subset \mathcal{S}^1(\mathcal{T}_h)$ the bilinear form

$$(v, w)_h := \int_\Omega \widehat{\mathcal{I}}_h(vw) \, \mathrm{d}x = \sum_{T \in \mathcal{T}_h} \sum_{z \in \mathcal{N}_h \cap T} \beta_z^T v|_T(z) w|_T(z)$$

for $v, w \in \mathcal{L}^1(\mathcal{T}_h)$, where $\beta_z = \int_T \varphi_z \, \mathrm{d}x$, defines an inner product on $\mathcal{L}^1(\mathcal{T}_h)$. This mass lumping will allow for the nodewise solution of certain nonlinearities. We have the relation

$$\|v_h\| \leq \|v_h\|_h \leq (d+2)^{1/2} \|v_h\|$$

for all $v_h \in \mathcal{L}^1(\mathcal{T}_h)$, *cf.* Lemma 3.9 of [9].

For completeness we provide the next lemma which states that $\bigcup_{h>0} \mathcal{S}^1(\mathcal{T}_h)^d$ is dense in $W^{\sigma'}(\operatorname{div}; \Omega)$ for $1 < \sigma' < \infty$.

Lemma 2.1. *Let $1 < \sigma' < \infty$ and $p \in W^{\sigma'}(\mathrm{div};\Omega)$. For every $\varepsilon > 0$ there exists $h(\varepsilon) > 0$ such that for all $h \le h(\varepsilon)$ there exists a function $q_h \in \mathcal{S}^1(\mathcal{T}_h)^d$ with*

$$\|p - q_h\|_{W^{\beta}(\mathrm{div};\Omega)} < \varepsilon.$$

Proof. Since $C^\infty(\overline{\Omega};\mathbb{R}^d)$ is dense in $W^{\sigma'}(\mathrm{div};\Omega)$, there exists for given $\varepsilon > 0$ a function $q \in C^\infty(\overline{\Omega};\mathbb{R}^d)$ with

$$\|p - q\|_{H(\mathrm{div};\Omega)} < \varepsilon/2.$$

Standard nodal interpolation estimates yield

$$\|q - \mathcal{I}_h q\|_{W^{\sigma'}(\mathrm{div};\Omega)} \le \|q - \mathcal{I}_h q\|_{W^{1,\sigma'}(\Omega;\mathbb{R}^d)} \le ch|q|_{W^{2,\infty}(\Omega;\mathbb{R}^d)}.$$

Now let h be such that

$$\|q - \mathcal{I}_h q\|_{W^{\sigma'}(\mathrm{div};\Omega)} < \varepsilon/2.$$

Choosing $q_h = \mathcal{I}_h q$ and using the triangle inequality yields the assertion. □

For an element $T \in \mathcal{T}_h$ and $p_h \in P_k(T)^r$ we have by an inverse estimate

$$\|p_h\|^2_{L^2(T)} \le c h_T^{2\min\{0,d/2-d/\sigma\}} \|p_h\|^2_{L^\sigma(T)},$$

cf. [15]. Hence, we may introduce for $1 \le \sigma < 2$ the weighted L^2-inner product

$$(p_h, q_h)_{w_\sigma} = (h_{\mathcal{T}}^{d(2/\sigma-1)} p_h, q_h)$$

for $p_h, q_h \in \mathcal{L}^k(\mathcal{T}_h)$. Its induced norm then has the property $\|\cdot\|_{w_\sigma} \le c\|\cdot\|_{L^\sigma(\Omega)}$ on $\mathcal{L}^k(\mathcal{T}_h)$.

Let us finally introduce the so called *Brezzi–Douglas–Marini (BDM)* finite element space which is given by

$$\mathcal{BDM}(\Omega) = \mathcal{L}^1(\mathcal{T}_h)^d \cap H(\mathrm{div};\Omega) \subset H(\mathrm{div};\Omega),$$

cf. [14]. For an element $T \in \mathcal{T}_h$ we can define a local interpolation operator $\Pi_{h,T} : H^1(T)^d \to P_1(T)^d$ by

$$\int_S q \cdot n\psi \,\mathrm{d}s = \int_S \Pi_{h,T} q \cdot n\psi \,\mathrm{d}s$$

for all sides $S \in \mathcal{S}_h \cap T$ of the element T and all affine functions $\psi \in P_1(S)$ on S. Note that the interpolation operator is well-defined also for less regular functions, *e.g.*, for $q \in H(\mathrm{div};T) \cap L^\gamma(T;\mathbb{R}^d)$ with $\gamma > 2$, *cf.* [14]. The global interpolation operator $\Pi_h : H^1(\Omega)^d \to \mathcal{BDM}(\Omega)$ is then defined by

$$(\Pi_h q)|_T = \Pi_{h,T}(q|_T)$$

and, in particular, $\Pi_h q \in \mathcal{BDM}(\Omega) \subset H(\mathrm{div};\Omega)$. For more details on $H(\mathrm{div};\Omega)$-conforming finite element spaces we refer the reader to [14].

3. Abstract error estimate

In the following we recap the existing results on abstract *a posteriori* error estimation for convex minimization problems and refer to [8, 42, 43, 45] for further details.

Let $F : Y \to \mathbb{R} \cup \{\infty\}$ and $G : X \to \mathbb{R} \cup \{\infty\}$ be proper, convex and lower-semicontinuous functionals and $B : X \to Y$ be bounded and linear. Under these hypothesis there holds $F = (F^*)^*$ and we obtain

$$\begin{aligned}
\inf_{u\in X} E(u) &= \inf_{u\in X} F(Bu) + G(u) \\
&= \inf_{u\in X} \sup_{p\in Y'} \langle p, Bu\rangle_{Y',Y} - F^*(p) + G(u) \\
&\geq \sup_{p\in Y'} \inf_{u\in X} -F^*(p) + \langle p, Bu\rangle_{Y',Y} + G(u) \\
&= \sup_{p\in Y'} - \sup_{u\in X} F^*(p) + \langle -B'p, u\rangle_{X',X} - G(u) \\
&= \sup_{p\in Y'} -F^*(p) - G^*(-B'p) \\
&=: \sup_{p\in Y'} D(p).
\end{aligned}$$

Hence, the dual formulation seeks a maximizer $p \in Y'$ for D. Particularly, we have the weak duality relation

$$E(v) \geq D(q) \tag{3.1}$$

for all $v \in X$ and $q \in Y'$. If $u \in X$ is a minimizer for E, the necessary optimality condition reads

$$0 \in \partial E(u).$$

With a nonnegative coercivity functional $\varrho_E : X \times X \to [0, \infty)$ this is equivalent to

$$\varrho_E(v, u) + E(u) \leq E(v) \tag{3.2}$$

for all $v \in X$. A combination of (3.1) and (3.2) yields the following abstract *a posteriori* error estimate.

Proposition 3.1 (Primal-dual gap estimates)**.** *Let $X_h \subset X$ and $Y_h \subset Y'$ and $u \in X$ and $u_h \in X_h$ be minimial for E in X and X_h, respectively. We then have the a priori error estimate*

$$\varrho_E(u, u_h) \leq E(u_h) - E(u) \leq \inf_{v_h\in X_h} E(v_h) - E(u).$$

For any $w_h \in X_h$ and $q_h \in Y_h$ we have with $\eta(w_h, q_h) := (E(w_h) - D(q_h))^{1/2}$ the a posteriori error estimate

$$\varrho_E(u, u_h) \leq \eta^2(w_h, q_h).$$

Proof. The *a priori* error estimate is a direct consequence of (3.2). Using the optimality (3.2) of $u \in X$, the weak duality (3.1) and $Y_h \subset Y'$ we then obtain

$$\varrho_E(u, u_h) \leq E(w_h) - E(u) \leq E(u_h) - \sup_{p\in Y'} D(p) \leq E(w_h) - D(q_h),$$

which concludes the proof. □

Remark 3.2. (1) Note that in case of strong duality, *i.e.*, there holds equality in (3.1), the *a posteriori* error estimate stated in Proposition 3.1 is sharp in the sense that if we use $w_h = u$ and $q_h = p$ in η with $u \in X$ and $p \in Y'$ being solutions to the primal and the dual problem, respectively, we have

$$\eta^2(u, p) = E(u) - D(p) = \inf_{v\in X} E(v) - \sup_{q\in Y'} D(q) = 0.$$

Sufficient for strong duality is that there exists $w \in X$ with $F(Bw) < \infty$, $G(w) < \infty$ and F being continuous at Bw. In this case the solutions are related by the inclusions

$$-B'p \in \partial G(u), \quad p \in \partial F(Bu),$$

cf. [27], which are equivalent to the variational inequalities

$$\begin{aligned}\langle -B'p, v-u\rangle_{X',X} + \varrho_G(v,u) + G(u) &\le G(v),\\ \langle p, Bv - Bu\rangle_{Y',Y} + \varrho_F(Bv,Bu) + F(Bu) &\le F(Bv).\end{aligned}$$

Adding both inequalities gives (3.2) with

$$\varrho_E(v,u) = \varrho_F(Bv,Bu) + \varrho_G(v,u),$$

which serves as an error measure.

(2) Let us emphasize that for the derivation of the reliability estimate for the primal-dual gap error estimator η we did not need to make any assumptions on the differentiability of the functionals F and G.

(3) One is free in the construction of feasible functions $w_h \in X_h$ and $q_h \in Y_h$ to define the error estimator $\eta(w_h, q_h)$. We will use the Variable-ADMM introduced in [10] to approximately solve the primal and the dual problem for the nonlinear Laplace problem and the ROF problem. However, feasible functions, *e.g.*, for the dual problem, may be constructed using other techniques like gradient recovery or flux reconstruction techniques, if they are applicable for the specific problem. The relation between primal-dual gap error estimators and other error estimators is discussed in [45] for a certain class of convex minimization problems.

4. Nonlinear Laplace equation

4.1. Primal and dual formulation

The nonlinear Laplace problem seeks for $\sigma \in (1,\infty)$, $\sigma' = \sigma/(\sigma-1)$, $f \in L^{\sigma'}(\Omega)$, $g \in L^{\sigma'}(\Gamma_N)$, $\widetilde{u}_D \in W^{1,\sigma}(\Omega)$ and $u_D = \widetilde{u}_D|_{\Gamma_D}$ a minimizer $u \in W^{1,\sigma}(\Omega)$ for

$$E_{\Delta_\sigma}(u) = \frac{1}{\sigma}\int_\Omega |\nabla u|^\sigma \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x - \int_{\Gamma_N} g u \,\mathrm{d}s + I_{u_D}(u|_{\Gamma_D}).$$

The indicator functional I_{u_D} encodes the boundary condition $u|_{\Gamma_D} = u_D$ on $\Gamma_D = \partial\Omega \setminus \Gamma_N$. The minimization problem admits a unique minimizer, *cf.* [31]. Minimization problems of the above structure arise in various areas of interest, *e.g.*, nonlinear diffusion [41], nonlinear elasticity [2], and fluid mechanics [3,6].

Let us make the following assumption that will simplify the presentation.

Assumption 4.1. *For ease of presentation we restrict to the case $g = 0$ and $u_D = 0$ in what follows. We then omit the indicator functional $I_{u_D}(u|_{\Gamma_D})$ in the definition of E_{Δ_σ} and seek for a minimizer $u \in W^{1,\sigma}_D(\Omega)$ instead.*

The dual nonlinear Laplace problem seeks $p \in W^{\sigma'}_N(\mathrm{div};\Omega)$ that maximizes the functional

$$D_{\Delta_\sigma}(p) := -\frac{1}{\sigma'}\int_\Omega |p|^{\sigma'}\,\mathrm{d}x - I_{\{f\}}(-\operatorname{div} p).$$

The following result (*cf.* [43], Thm. 1) shows that the dual nonlinear Laplace problem is in fact the dual problem to the primal nonlinear Laplace problem in the sense of Fenchel duality. It further ensures the strong duality between the primal and the dual nonlinear Laplace problem.

Theorem 4.2 (Strong duality)**.** *There exists a unique minimizer* $u \in W^{1,\sigma}_D(\Omega)$ *for* E_{Δ_σ} *and a unique maximizer* $p \in W^{\sigma'}_N(\operatorname{div};\Omega)$ *for* D_{Δ_σ}*. The functions* u *and* p *are related by* $\operatorname{div} p = -f$, $p = |\nabla u|^{\sigma-2}\nabla u$ *(or, equivalently,* $\nabla u = |p|^{\sigma'-2}p$*) and*

$$E_{\Delta_\sigma}(u) = D_{\Delta_\sigma}(p).$$

Proof. The assertion follows from standard arguments in duality theory, *cf.* [27,43]. □

The unique minimizer $u \in W^{1,\sigma}_D(\Omega)$ satisfies the variational equality

$$\int_\Omega |\nabla u|^{\sigma-2}\nabla u \cdot \nabla v \,\mathrm{d}x = \int_\Omega f v \,\mathrm{d}x$$

for all $v \in W^{1,\sigma}_D(\Omega)$, *cf.* [19,31].

Next, we introduce suitable finite element spaces for the primal and dual nonlinear Laplace problem.

4.2. Finite element spaces and *a priori* estimates

To make use of the primal-dual gap estimator we need to choose conforming finite element spaces $X_h \subset W^{1,\sigma}_D(\Omega)$ and $Y_h \subset W^{\sigma'}_N(\operatorname{div};\Omega)$. We let

$$X_h = \mathcal{S}^1(\mathcal{T}_h) \cap W^{1,\sigma}_D(\Omega), \quad Y_h = \mathcal{BDM}(\Omega) \cap W^{\sigma'}_N(\operatorname{div};\Omega).$$

For $f_h \in \mathcal{L}^0(\mathcal{T}_h)$ being the elementwise L^2-projection of f we set

$$\begin{aligned} E^h_{\Delta_\sigma}(u_h) &= \frac{1}{\sigma}\int_\Omega |\nabla u_h|^\sigma \,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x, \\ D^h_{\Delta_\sigma}(p_h) &= -\frac{1}{\sigma'}\int_\Omega |p_h|^{\sigma'} \,\mathrm{d}x - I_{\{f_h\}}(-\operatorname{div} p_h), \\ \widehat{D}^h_{\Delta_\sigma}(p_h) &= -\frac{1}{\sigma'}\int_\Omega \widehat{\mathcal{I}}_h |p_h|^{\sigma'} \,\mathrm{d}x - I_{\{f_h\}}(-\operatorname{div} p_h). \end{aligned}$$

As it has been found out in earlier contributions, the energy norm $\|\nabla \cdot\|_{L^\sigma(\Omega)}$ is not well suited for the *a priori* and *a posteriori* error analysis for the nonlinear Laplacian, since one obtains convergence rates that are not optimal for a discretization with linear finite elements, *cf.* [7]. Instead, for a fixed function $v \in W^{1,\sigma}(\Omega)$, a so called *quasi-norm* defined by

$$\|\nabla w\|^2_{(v,\sigma)} := \int_\Omega (|\nabla v| + |\nabla w|)^{\sigma-2}|\nabla w|^2 \,\mathrm{d}x$$

has been introduced and widely used in the literature, *cf.* [7,13,17,21,25,36–38]. Defining

$$V(\nabla v) := |\nabla v|^{\frac{\sigma-2}{2}}\nabla v$$

it has been shown in [20,21] that there exist constants $c, C > 0$ with

$$c\|\nabla v - \nabla w\|^2_{(v,\sigma)} \le \|V(\nabla v) - V(\nabla w)\|^2 \le C\|\nabla v - \nabla w\|^2_{(v,\sigma)}.$$

The following *a priori* estimate for the quasi-norm has been shown in Lemma 5.2 of [20].

Proposition 4.3 (*A priori* estimate)**.** *Let* u *and* u_h *be the minimizers for* E_{Δ_σ} *in* $W^{1,\sigma}_D(\Omega)$ *and in* X_h*, respectively. Then we have*

$$\|V(\nabla u) - V(\nabla u_h)\| \le c \inf_{v_h \in X_h} \|V(\nabla u) - V(\nabla v_h)\|.$$

If the minimizer u *additionally satisfies* $V(\nabla u) \in W^{1,2}(\Omega;\mathbb{R}^d)$*, there holds*

$$\|V(\nabla u) - V(\nabla u_h)\| \le c \inf_{v_h \in X_h} \|V(\nabla u) - V(\nabla v_h)\| \le ch\|\nabla V(\nabla u)\|.$$

Proof. A complete proof is given in [20]. □

Remark 4.4. Under certain regularity assumptions on the data f and the boundary $\partial\Omega$ one can prove $V(\nabla u) \in W^{1,2}(\Omega;\mathbb{R}^d)$, *cf.* [24, 26]. In general, one may only expect $V(\nabla u) \in W^{1,\gamma}(\Omega;\mathbb{R}^d)$ for some $\gamma > 1$ and, in this case,

$$\|V(\nabla u) - V(\nabla u_h)\| \le ch^s$$

with $s = \min\{1, 2 - 2/\gamma\}$, *cf.* Remark 5.2 of [13].

To obtain an *a posteriori* error estimate in the style of Proposition 3.1 we need to bound the error in the quasi-norm by the energy difference. This is established in Lemma 16 of [21] for the difference between two finite element solutions of the nonlinear Laplace problem on nested finite element spaces.

Proposition 4.5 ([21], Lem. 16)**.** *Let $u \in W^{1,\sigma}_D(\Omega)$ be the unique minimizer of E_{Δ_σ} and $v_h \in X_h$ be arbitrary. Then we have*

$$c\|V(\nabla u) - V(\nabla v_h)\|^2 \le E_{\Delta_\sigma}(v_h) - E_{\Delta_\sigma}(u).$$

Proof. A proof is presented in Lemma 16 of [21], where the error between two minimizers $u_h \in X_h$ and $u_{h'} \in X_{h'}$ of E_{Δ_σ} in nested spaces $X_h \subset X_{h'} \subset W^{1,\sigma}_D(\Omega)$ is considered. However, the minimality property of u_h is not used so that we may replace it by any test function $v_h \in X_h$, see also Lemma 3.2, Remark 3.3 of [13]. We refer the reader to Lemma 16 of [21] for details. □

The previous proposition enables us to follow the arguments for the *a posteriori* error analysis presented in the abstract setting.

4.3. *A posteriori* estimate and error estimator

By Proposition 4.5 and the strong duality ensured by Theorem 4.2 we obtain an *a posteriori* error estimate and an error estimator in the fashion of Proposition 3.1, which can be used for adaptive local mesh refinement. The next result is a special case of Proposition 3.1 for the nonlinear Laplace problem, where also the data approximation error is taken into account.

Proposition 4.6 (*A posteriori* estimate)**.** *Let u and u_h be the unique minimizers for E_{Δ_σ} in $W^{1,\sigma}_D(\Omega)$ and for $E^h_{\Delta_\sigma}$ in X_h, respectively, and let p_h be the unique maximizer for $D^h_{\Delta_\sigma}$ in Y_h. Then we have for any $v_h \in X_h$ and $q_h \in Y_h$ with $\operatorname{div} q_h = -f_h$ that*

$$c\|V(\nabla u) - V(\nabla u_h)\|^2 \le \eta^h_{\Delta_\sigma}(v_h, q_h)^2 + c\|p_h\|^{1/(\sigma-1)}_{L^{\sigma'}(\Omega)}\|f - f_h\|_{L^{\sigma'}(\Omega)},$$

with $\eta^h_{\Delta_\sigma}(v_h, q_h)^2 = E^h_{\Delta_\sigma}(v_h) - D^h_{\Delta_\sigma}(q_h)$.

Proof. By Proposition 4.5, the strong duality given by Theorem 4.2 and the optimality of u_h and p_h in X_h and Y_h, respectively, we have

$$\begin{aligned} c\|V(\nabla u) - V(\nabla u_h)\|^2 &\le E_{\Delta_\sigma}(u_h) - E_{\Delta_\sigma}(u) \\ &= E_{\Delta_\sigma}(u_h) - D_{\Delta_\sigma}(p) \\ &= E^h_{\Delta_\sigma}(u_h) - D^h_{\Delta_\sigma}(p_h) \\ &\quad + E_{\Delta_\sigma}(u_h) - E^h_{\Delta_\sigma}(u_h) + D^h_{\Delta_\sigma}(p_h) - D_{\Delta_\sigma}(p) \\ &\le \eta^h_{\Delta_\sigma}(v_h, q_h)^2 \\ &\quad + E_{\Delta_\sigma}(u_h) - E^h_{\Delta_\sigma}(u_h) + D^h_{\Delta_\sigma}(p_h) - D_{\Delta_\sigma}(p). \end{aligned}$$

Using the Poincaré inequality and choosing $v_h = u_h$ in the variational equality satisfied by u_h we obtain $\|u_h\|_{L^\sigma(\Omega)} \le c\|\nabla u_h\|_{L^\sigma(\Omega)} \le c\|f_h\|_{L^{\sigma'}(\Omega)}^{1/(\sigma-1)}$. Hence, the first data approximation error can be estimated by

$$E_{\Delta_\sigma}(u_h) - E^h_{\Delta_\sigma}(u_h) = \int_\Omega u_h(f - f_h)\,\mathrm{d}x \le c\|f_h\|_{L^{\sigma'}(\Omega)}^{1/(\sigma-1)}\|f - f_h\|_{L^{\sigma'}(\Omega)}.$$

To estimate the second error involving the discretization of the dual functional we will construct a function $\widetilde{p}_h \in W_N^{\sigma'}(\mathrm{div};\Omega)$ for which D_{Δ_σ} is finite, *i.e.*, $\operatorname{div}\widetilde{p}_h = -f$, and which relates p and p_h. Let $w^{(h)} \in W^{1,\sigma}(\Omega)$ be the unique weak solution with vanishing mean of

$$-\operatorname{div}(|\nabla w^{(h)}|^{\sigma-2}\nabla w^{(h)}) = f - f_h, \quad |\nabla w^{(h)}|^{\sigma-2}\nabla w^{(h)} \cdot n = 0 \text{ on } \partial\Omega.$$

The existence of $w^{(h)}$ follows with standard arguments from the theory of monotone operators, see, *e.g.*, [19,31], using the closed subspace $\widetilde{W}^{1,\sigma}(\Omega) \subset W^{1,\sigma}(\Omega)$ consisting of all functions in $W^{1,\sigma}(\Omega)$ with vanishing mean. Set $p^{(h)} = |\nabla w^{(h)}|^{\sigma-2}\nabla w^{(h)}$. Then we have $p^{(h)} \in W_N^{\sigma'}(\mathrm{div};\Omega)$ with

$$\|p^{(h)}\|_{L^{\sigma'}(\Omega)} \le c\|f - f_h\|_{L^{\sigma'}(\Omega)}.$$

For $\widetilde{p}_h = p_h + p^{(h)}$ there holds $-\operatorname{div}\widetilde{p}_h = f$, *i.e.*, $D_{\Delta_\sigma}(\widetilde{p}_h) < \infty$. With the optimality of p, the feasibility of $\widetilde{p}_h$ with respect to D_{Δ_σ} and the monotonicity

$$|a|^{\sigma'} - |b|^{\sigma'} \le \sigma'|a|^{\sigma'-2}a \cdot (a - b)$$

for $a, b \in \mathbb{R}^d$ we can then bound the error $D^h_{\Delta_\sigma}(p_h) - D_{\Delta_\sigma}(p)$ by

$$\begin{aligned} D^h_{\Delta_\sigma}(p_h) - D_{\Delta_\sigma}(p) &\le D^h_{\Delta_\sigma}(p_h) - D_{\Delta_\sigma}(\widetilde{p}_h) \\ &\le \int_\Omega |p_h|^{\sigma'-2}p_h \cdot (p_h - \widetilde{p}_h)\,\mathrm{d}x \\ &\le \||p_h|^{\sigma'-1}\|_{L^\sigma(\Omega)}\|p^{(h)}\|_{L^{\sigma'}(\Omega)} \\ &\le c\|p_h\|_{L^{\sigma'}(\Omega)}^{1/(\sigma-1)}\|f - f_h\|_{L^{\sigma'}(\Omega)}, \end{aligned}$$

which completes the proof. □

Remark 4.7. (1) In our numerical experiments below the sequence of discrete solutions to the dual nonlinear Laplace problem $(p_h)_{h>0}$ remained bounded in $L^{\sigma'}(\Omega)$. Unfortunately, we were not able to prove this theoretically in general.
(2) In Proposition 4.9 we prove that the error indicator is nonnegative.

Remark 4.8. (1) Note that the (discrete) primal-dual gap error estimator $\eta^h_{\Delta_\sigma}$ defines for arbitrary $v_h \in X_h$ and $q_h \in Y_h$ with $\operatorname{div} q_h = -f_h$ a reliable upper bound (up to data oscillations) for the error in the quasi-norm, *i.e.*, we do not need to compute exact discrete solutions u_h and p_h of the primal and dual nonlinear Laplace problem, respectively.
(2) The proof of the reliability of the primal-dual gap error estimator did not require any differentiability assumptions on E_{Δ_σ} or a variational formulation of the primal nonlinear Laplace problem.
(3) Using integration by parts and $\operatorname{div} q_h = -f_h$ we obtain the expression

$$\eta^h_{\Delta_\sigma}(v_h, q_h)^2 = \int_\Omega \frac{1}{\sigma}|\nabla v_h|^\sigma + \frac{1}{\sigma'}|q_h|^{\sigma'} - q_h \cdot \nabla v_h\,\mathrm{d}x.$$

(4) In our numerical experiments we will use the computable (lumped) discrete primal-dual gap error estimator

$$\widehat{\eta}^h_{\Delta_\sigma}(v_h, q_h)^2 = E^h_{\Delta_\sigma}(v_h) - \widehat{D}^h_{\Delta_\sigma}(q_h)$$

so that subproblems appearing in the iterative solution of the dual nonlinear Laplace problem can be efficiently solved nodewise. As before, integration by parts and the relation $\operatorname{div} q_h = -f_h$ yield

$$\widehat{\eta}^h_{\Delta_\sigma}(v_h, q_h)^2 = \int_\Omega \frac{1}{\sigma}|\nabla v_h|^\sigma + \frac{1}{\sigma'}\widehat{\mathcal{I}}_h|q_h|^{\sigma'} - q_h \cdot \nabla v_h \,\mathrm{d}x.$$

For $T \in \mathcal{T}_h$ the local error indicator is given by restriction of the global error estimator to the element T. We have the following nonnegativity result.

Proposition 4.9. *Let for any $T \in \mathcal{T}_h$ the local error indicator be defined by*

$$\eta^{h,T}_{\Delta_\sigma}(v_h, q_h)^2 = \int_T \frac{1}{\sigma}|\nabla v_h|^\sigma + \frac{1}{\sigma'}|q_h|^{\sigma'} - q_h \cdot \nabla v_h \,\mathrm{d}x,$$
$$\widehat{\eta}^{h,T}_{\Delta_\sigma}(v_h, q_h)^2 = \int_T \frac{1}{\sigma}|\nabla v_h|^\sigma + \frac{1}{\sigma'}\widehat{\mathcal{I}}_h|q_h|^{\sigma'} - q_h \cdot \nabla v_h \,\mathrm{d}x.$$

Then we have for any $v_h \in X_h$ and $q_h \in Y_h$

$$\widehat{\eta}^{h,T}_{\Delta_\sigma}(v_h, q_h) \geq \eta^{h,T}_{\Delta_\sigma}(v_h, q_h) \geq 0.$$

Proof. Using that for an element $T \in \mathcal{T}_h$ and $x \in T$ the mapping $x \mapsto |q_h(x)|^{\sigma'}$ is convex we conclude that $\widehat{\mathcal{I}}_h|q_h|^{\sigma'} \geq |q_h|^{\sigma'}$ on T since $q_h|_T$ is affine, and, therefore,

$$\widehat{\eta}^{h,T}_{\Delta_\sigma}(v_h, q_h)^2 \geq \eta^{h,T}_{\Delta_\sigma}(v_h, q_h)^2.$$

Note that the integrand in the definition of $\eta^{h,T}_{\Delta_\sigma}$ is nonnegative, because for arbitrary $b \in \mathbb{R}^d$ we have by Young's inequality

$$(1/\sigma')|b|^{\sigma'} = \sup_{a \in \mathbb{R}^d} a \cdot b - (1/\sigma)|a|^\sigma.$$

Particularly, we have

$$\eta^{h,T}_{\Delta_\sigma}(v_h, q_h)^2 = \int_T \frac{1}{\sigma}|\nabla v_h|^\sigma + \frac{1}{\sigma'}|q_h|^{\sigma'} - q_h \cdot \nabla v_h \,\mathrm{d}x \geq 0$$

for every element $T \in \mathcal{T}_h$. Hence, putting everything together, we arrive at

$$\widehat{\eta}^{h,T}_{\Delta_\sigma}(v_h, q_h) \geq \eta^{h,T}_{\Delta_\sigma}(v_h, q_h) \geq 0$$

for any $T \in \mathcal{T}_h$. □

In the sequel we briefly discuss the explicit computation of the primal-dual gap error estimator.

4.4. Iterative solution

As we have pointed out in Remark 4.8 the quantity $\eta^h_{\Delta_\sigma}(v_h, q_h)$, and therefore also $\widehat{\eta}^h_{\Delta_\sigma}(v_h, q_h)$ by Proposition 4.9, defines a reliable upper bound for any feasible functions $v_h \in X_h$ and $q_h \in Y_h$. Since the minimizer u_h of $E^h_{\Delta_\sigma}$ in X_h and the maximizer p_h of $D^h_{\Delta_\sigma}$ in Y_h are not directly available, a reasonable choice of functions v_h and q_h with $\operatorname{div} q_h = -f_h$ are approximate discrete solutions of the primal and dual nonlinear Laplace problem.

These will be computed using splitting methods based on augmented Lagrange functionals, which have been introduced in [30, 31]. For the primal problem we define

$$L_\tau^E(u_h, r_h; \lambda_h) = \frac{1}{\sigma}\int_\Omega |r_h|^\sigma \,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x \\ + (\lambda_h, \nabla u_h - r_h)_{w_\sigma} + \frac{\tau}{2}\|\nabla u_h - r_h\|_{w_\sigma}^2$$

for $u_h \in X_h$ and $r_h, \lambda_h \in \mathcal{L}^0(\mathcal{T}_h)^d$. For the dual problem we consider

$$L_\tau^D(p_h, q_h; \mu_h) = \frac{1}{\sigma'}\int_\Omega \widehat{\mathcal{I}}_h |q_h|^{\sigma'} \,\mathrm{d}x + I_{\{-f_h\}}(\operatorname{div} p_h) \\ + (\mu_h, p_h - q_h)_{h,w_{\sigma'}} + \frac{\tau}{2}\|p_h - q_h\|_{h,w_{\sigma'}}^2$$

for $q_h, \mu_h \in \mathcal{L}^1(\mathcal{T}_h)^d$ and $p_h \in Y_h$. The minimization of $E_{\Delta_\sigma}^h$ and $-\widehat{D}_{\Delta_\sigma}^h$ is equivalent to seeking a saddle point for L_τ^E and L_τ^D, respectively, *i.e.*,

$$\min_{u_h \in X_h} E_{\Delta_\sigma}^h(u_h) = \min_{(u_h, r_h) \in X_h \times \mathcal{L}^0(\mathcal{T}_h)^d} \max_{\lambda_h \in \mathcal{L}^0(\mathcal{T}_h)^d} L_\tau^E(u_h, r_h; \lambda_h),$$
$$\min_{p_h \in Y_h} -\widehat{D}_{\Delta_\sigma}^h(p_h) = \min_{(p_h, q_h) \in Y_h \times \mathcal{L}^1(\mathcal{T}_h)^d} \max_{\mu_h \in \mathcal{L}^1(\mathcal{T}_h)^d} L_\tau^D(p_h, q_h; \mu_h).$$

The associated saddle-point problems are then solved using the Variable-ADMM, which alternately optimizes L_τ^E and L_τ^D with repsect to u_h and r_h, and p_h and q_h, respectively, followed by an update of the Lagrange multipliers, *cf.* [10] for details on the Variable-ADMM. The optimization problems related to L_τ^E boil down to elementwise optimization problems, which is due to the choice of the piecewise constant finite element space $\mathcal{L}^0(\mathcal{T}_h)^d$, whereas the optimization problems related to L_τ^D are given by nodewise optimization problems, since mass lumping and the piecewise affine finite element space $\mathcal{L}^1(\mathcal{T}_h)^d$ are used. The elementwise and nodewise optimization problems are solved using Newton's method, see [39] for details.

5. Rudin–Osher–Fatemi image denoising

5.1. Primal and dual formulation

In this section we consider a variant of the nonlinear Laplacian with limit exponent $\sigma = 1$. For a given function $g \in L^2(\Omega)$ and a fidelity parameter $\alpha > 0$ we seek a minimizer $u \in \mathrm{BV}(\Omega) \cap L^2(\Omega)$ of the functional

$$E_{\mathrm{rof}}(u) = \int_\Omega |Du| + \frac{\alpha}{2}\|u - g\|^2.$$

This particular minimization problem has been proposed in image processing for denoising a given noisy image g and is known as the Rudin–Osher–Fatemi (ROF) image denoising problem [48]. It also serves as a model problem for general BV-regularized minimization problems and evolutions, *cf.*, *e.g.*, [50]. The (pre-)dual problem is given by the maximization of the functional

$$D_{\mathrm{rof}}(p) = -\frac{1}{2\alpha}\|\operatorname{div} p + \alpha g\|^2 + \frac{\alpha}{2}\|g\|^2 - I_{K_1(0)}(p)$$

in the set of vector fields $p \in H_N(\mathrm{div};\Omega)$ with square integrable distributional divergence and vanishing normal component on $\partial\Omega$, *cf.* [33]. The indicator functional $I_{K_1(0)}$ of the set of vector fields $q \in L^2(\Omega;\mathbb{R}^d)$ which satisfy $|q| \leq 1$ in Ω introduces a pointwise constraint. Note that a maximizer of D_{rof} may not be unique. The primal and the dual ROF problem are in strong duality and the unique minimizer $u \in \mathrm{BV}(\Omega) \cap L^2(\Omega)$ of E_{rof} and any maximizer $p \in H_N(\mathrm{div};\Omega)$ of D_{rof} are related by

$$\operatorname{div} p = \alpha(u - g), \quad -(u, \operatorname{div}(q - p)) \leq 0$$

for all $q \in H_N(\mathrm{div};\Omega) \cap K_1(0)$, *cf.* [33].

5.2. Finite element spaces

As for the nonlinear Laplace equation we let

$$X_h = \mathcal{S}^1(\mathcal{T}_h) \subset \mathrm{BV}(\Omega) \cap L^2(\Omega).$$

The discrete space Y_h is chosen to consist of continuous or discontinuous, elementwise affine vector fields

$$Y_h^C = \mathcal{S}^1(\mathcal{T}_h)^d \cap H_N(\mathrm{div};\Omega), \quad \text{or} \quad Y_h^{dC} = \mathcal{L}^1(\mathcal{T}_h)^d \cap H_N(\mathrm{div};\Omega).$$

We have the consistency relation $Y_h^C \subset Y_h^{dC} \subset H_N(\mathrm{div};\Omega)$ and denote by Y_h either of the two spaces. Let $g_h \in \mathcal{L}^0(\mathcal{T}_h)$ be the elementwise L^2-projection of g. The discretized functionals are then defined by

$$E_{\mathrm{rof}}^h(u_h) = \int_\Omega |\nabla u_h| \,\mathrm{d}x + \frac{\alpha}{2}\|u_h - g_h\|^2,$$
$$D_{\mathrm{rof}}^h(p_h) = -\frac{1}{2\alpha}\|\operatorname{div} p_h + \alpha g_h\|^2 - I_{K_1(0)}(p_h) + \frac{\alpha}{2}\|g_h\|^2.$$

Remark 5.1. The discretization of the dual ROF problem with the lowest order Raviart–Thomas finite element is not suitable since it does not include nodal degrees of freedom which is required to ensure the pointwise constraint $|p_h| \le 1$ which in turn is mandatory to derive a meaningful and useful *a posteriori* error estimate.

Let u and u_h be the unique minimizers of E_{rof} in $\mathrm{BV}(\Omega) \cap L^2(\Omega)$ and X_h, respectively. The strong convexity of E_{rof} can be used to derive the *a priori* error estimate

$$\frac{\alpha}{2}\|u - u_h\|^2 \le ch^{1/2}$$

if $u \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$, *cf.* [9,12]. The optimal convergence rate for the approximation with continuous, piecewise linear functions is, however, given by

$$\min_{v_h \in \mathcal{S}^1(\mathcal{T}_h)} \|u - v_h\|^2 \le ch,$$

which cannot be improved in general, *cf.* [9,12].

Motivated by the relation $\operatorname{div} p = \alpha(u - g)$ we also consider for any discrete maximizer $p_h \in Y_h$ of D_{rof}^h the approximation

$$\overline{u}_h = \frac{1}{\alpha}\operatorname{div} p_h + g_h \in \mathcal{L}^0(\mathcal{T}_h)$$

of u, for which the following convergence result can be proven.

Proposition 5.2. *Let for any $h > 0$ the function p_h be a discrete maximizer of D_{rof}^h in Y_h and let $\overline{u}_h = (1/\alpha)\operatorname{div} p_h + g_h$. If $g_h \to g$ in $L^2(\Omega)$, we have*

$$\|u - \overline{u}_h\| \to 0$$

as $h \to 0$.

Proof. The sequence $(g_h)_h \subset L^2(\Omega)$ is uniformly bounded since $g_h \to g$ in $L^2(\Omega)$. Using that p_h is a minimizer for $-D_{\mathrm{rof}}^h$ in Y_h we can bound

$$\frac{1}{2\alpha}\|\operatorname{div} p_h + \alpha g_h\|^2 - \frac{\alpha}{2}\|g_h\|^2 = -D_{\mathrm{rof}}^h(p_h) \le -D_{\mathrm{rof}}^h(0) = 0,$$

i.e.,

$$\frac{1}{2\alpha}\|\operatorname{div} p_h + \alpha g_h\|^2 \le \frac{\alpha}{2}\|g_h\|^2.$$

Thus, the sequence $(p_h)_{h>0}$ is uniformly bounded in $H_N(\operatorname{div};\Omega)$. Hence, we can choose a subsequence $(p_{h'})_{h'>0}$ with $p_{h'} \rightharpoonup p$ for a function $p \in H_N(\operatorname{div};\Omega)$. On the other hand there exists for any $q \in H_N(\operatorname{div};\Omega)$ a sequence $(q_h)_{h>0} \subset Y_h^C$ with $|q_h| \le 1$ for all $h > 0$ and $q_h \to q$ in $H_N(\operatorname{div};\Omega)$. Indeed, for given $q \in H_N(\operatorname{div};\Omega)$ one can construct a smooth function $\widetilde{q} \in C_c^\infty(\Omega;\mathbb{R}^d)$ *via* convolution of q with a nonnegative convolution kernel noting that this process does not increase the L^∞-norm. One then procedes as in the proof of Lemma 2.1 noting again that neither the nodal interpolation operator increases the L^∞-norm. The weak lower-semicontinuity of $-D_{\text{rof}}$ and the optimality of each $p_{h'}$ yield

$$\begin{aligned}
-D_{\text{rof}}(p) &\le \liminf_{h'\to 0} -D_{\text{rof}}(p_{h'}) \\
&\le \limsup_{h'\to 0} -D_{\text{rof}}^{h'}(p_{h'}) + D_{\text{rof}}^{h'}(p_{h'}) - D_{\text{rof}}(p_{h'}) \\
&\le \limsup_{h'\to 0} -D_{\text{rof}}^{h'}(p_{h'}) + c\|g - g_{h'}\| \\
&\le \limsup_{h'\to 0} -D_{\text{rof}}^{h'}(q_{h'}) \\
&= \limsup_{h'\to 0} -D_{\text{rof}}(q_{h'}) + D_{\text{rof}}(q_{h'}) - D_{\text{rof}}^{h'}(q_{h'}) \\
&\le \limsup_{h'\to 0} -D_{\text{rof}}(q_{h'}) + c\|g - g_{h'}\| = -D_{\text{rof}}(q).
\end{aligned}$$

Hence, p is a minimizer of $-D_{\text{rof}}$. By choosing $q = p$ and a (possibly different) sequence $(q_h)_{h>0} \subset Y_h^C$ with $q_h \to p$ in $H_N(\operatorname{div};\Omega)$ we particularly find that

$$-D_{\text{rof}}(p) = \lim_{h'\to 0} -D_{\text{rof}}(p_{h'}).$$

Since $g_{h'} \to g$ and $\operatorname{div} p_{h'} \rightharpoonup \operatorname{div} p$, this implies that

$$\|\operatorname{div} p_{h'}\| \to \|\operatorname{div} p\|.$$

Altogether, we have that $\operatorname{div} p_{h'} \to \operatorname{div} p$ in $H_N(\operatorname{div};\Omega)$. By strong duality of the primal and dual ROF problem we have

$$u = \frac{1}{\alpha}\operatorname{div} p + g.$$

With $\operatorname{div} p_{h'} \to \operatorname{div} p$ and $g_{h'} \to g$ it follows that

$$u - \overline{u}_{h'} = \frac{1}{\alpha}\operatorname{div} p + g - \frac{1}{\alpha}\operatorname{div} p_{h'} - g_{h'} \to 0.$$

Thus, every convergent subsequence of $(\overline{u}_h)_{h>0}$ converges to u. Therefore, the whole sequence converges to u. □

Using the strong convexity of the functional E_{rof}, *i.e.*, there holds

$$\frac{\alpha}{2}\|u - v_h\|^2 \le E_{\text{rof}}(v_h) - E_{\text{rof}}(u) \tag{5.1}$$

for any $v_h \in \mathcal{S}^1(\mathcal{T}_h)$, we can carry out the *a posteriori* error analysis.

5.3. *A posteriori* estimate and error estimator

By the strong convexity (5.1) and the strong duality of the primal and dual ROF problem we can establish an *a posteriori* error estimate and an error estimator in the fashion of Proposition 3.1, which can be used for adaptive mesh refinement. The following reliability result is a special case of Proposition 3.1 for the ROF problem, where also the data approximation error is taken into account.

Proposition 5.3. *Let u and u_h be the unique minimizers for $E_{\rm rof}$ in $\mathrm{BV}(\Omega)\cap L^2(\Omega)$ and $E_{\rm rof}^h$ in X_h, respectively, and let p_h be a maximizer for $D_{\rm rof}^h$ in Y_h. Then we have for any $v_h \in X_h$ and $q_h \in Y_h$ with $|q_h| \le 1$ that*

$$\frac{\alpha}{2}\|u-u_h\|^2 \le \eta_{\rm rof}^h(v_h,q_h)^2 + c\|g-g_h\|$$

with $\eta_{\rm rof}^h(v_h,q_h)^2 = E_{\rm rof}^h(v_h) - D_{\rm rof}^h(q_h)$ and c depending on $\|g\|$.

Proof. Let $p \in H_N(\mathrm{div};\Omega)$ be a maximizer of $D_{\rm rof}$. Taking $v = u_h$ in (5.1) and using the strong duality $E_{\rm rof}(u) = D_{\rm rof}(p)$, the optimality of p in $H_N(\mathrm{div};\Omega)$, the optimality of p_h in $Y_h \subset H_N(\mathrm{div};\Omega)$ and the optimality of u_h in X_h we have

$$\begin{aligned}\frac{\alpha}{2}\|u-u_h\|^2 \le{}& E_{\rm rof}(u_h) - E_{\rm rof}(u)\\ ={}& E_{\rm rof}(u_h) - D_{\rm rof}(p)\\ \le{}& E_{\rm rof}(u_h) - D_{\rm rof}(p_h)\\ ={}& \eta_{\rm rof}^h(u_h,p_h)^2 + E_{\rm rof}(u_h) - E_{\rm rof}^h(u_h)\\ &+ D_{\rm rof}^h(p_h) - D_{\rm rof}(p_h)\\ \le{}& \eta_{\rm rof}^h(v_h,q_h)^2 + E_{\rm rof}(u_h) - E_{\rm rof}^h(u_h)\\ &+ D_{\rm rof}^h(p_h) - D_{\rm rof}(p_h).\end{aligned}$$

The first data approximation error can be bounded by

$$E_{\rm rof}(u_h) - E_{\rm rof}^h(u_h) = \frac{\alpha}{2}\int_\Omega (g_h - g)(2u_h - g - g_h)\,\mathrm{d}x \le c\|g-g_h\|,$$

where we used that $\|u_h\| \le c\|g_h\|$ and $\|g_h\| \le c\|g\|$. The second data approximation error can be analogously estimated by

$$\begin{aligned}D_{\rm rof}^h(p_h) - D_{\rm rof}(p_h) &= \frac{1}{2}\left[\int_\Omega (g_h-g)(g_h+g)\,\mathrm{d}x + \int_\Omega (g-g_h)(2\,\mathrm{div}\,p_h + \alpha(g+g_h))\,\mathrm{d}x\right]\\ &\le c\|g-g_h\|\end{aligned}$$

using that $\|\,\mathrm{div}\,p_h\| \le c\|g_h\| \le c\|g\|$, which completes the proof. $\square$

Remark 5.4. (1) Note that, as for the nonlinear Laplace problem, the (discrete) primal-dual gap error estimator $\eta_{\rm rof}^h$ defines for arbitrary $v_h \in X_h$ and $q_h \in Y_h$ with $|q_h| \le 1$ a reliable upper bound (up to data oscillations) for the error. Particularly, the exact discrete solutions u_h and p_h of the primal and dual ROF problem, respectively, need not to be computed exactly to estimate the error.

(2) Using binomial formulas and integration by parts we obtain the representation

$$\eta_{\rm rof}^h(v_h,q_h)^2 = \int_\Omega |\nabla v_h| - \nabla v_h \cdot q_h\,\mathrm{d}x + \frac{1}{2\alpha}\|\,\mathrm{div}\,q_h - \alpha(v_h - g_h)\|^2$$

for $v_h \in X_h$ and $q_h \in Y_h$ with $|q_h| \le 1$.

As for the nonlinear Laplace problem, for $T \in \mathcal{T}_h$ the local error indicators are defined *via* restricting the global error estimator to the simplex T. The local error indicators are non-negative due to the condition $|q_h| \le 1$ as the next proposition shows.

Proposition 5.5. *Let for any $T \in \mathcal{T}_h$ the local error indicator be defined by*

$$\eta_{\rm rof}^{h,T}(v_h,q_h)^2 = \int_T |\nabla v_h| - \nabla v_h \cdot q_h\,\mathrm{d}x + \frac{1}{2\alpha}\|\,\mathrm{div}\,q_h - \alpha(v_h - g)\|_{L^2(T)}^2.$$

Then we have for any $v_h \in X_h$ and $q_h \in Y_h$ with $|q_h| \le 1$ that

$$\eta_{\text{rof}}^{h,T}(v_h, q_h) \ge 0.$$

Proof. The non-negativity immediately follows from $|q_h| \le 1$ and the Cauchy–Schwarz inequality. □

To obtain a computable *a posteriori* error estimator we iteratively solve the primal and dual ROF problem.

5.4. Iterative solution

We approximate discrete minimizers u_h and p_h of E_{rof}^h and $-D_{\text{rof}}^h$ as in the case of the nonlinear Laplacian *via* an augmented Lagrangian approach. To this end, we introduce for the primal problem

$$L_\tau^E(u_h, r_h; \lambda_h) = \int_\Omega |r_h| \, \mathrm{d}x + \frac{\alpha}{2}\|u_h - g_h\|^2 + (\lambda_h, \nabla u_h - r_h)_w + \frac{\tau}{2}\|\nabla u_h - r_h\|_w^2$$

for $u_h \in X_h$ and $r_h, \lambda_h \in \mathcal{L}^0(\mathcal{T}_h)^d$, and, for the dual problem,

$$L_\tau^D(p_h, q_h; \mu_h) = \frac{1}{2\alpha}\| \operatorname{div} p_h + \alpha g_h\|^2 - \frac{\alpha}{2}\|g_h\|^2 + I_{K_1(0)}(q_h) + (\mu_h, p_h - q_h)_h + \frac{\tau}{2}\|p_h - q_h\|_h^2$$

for $q_h, \mu_h \in \mathcal{L}^1(\mathcal{T}_h)^d$ and $p_h \in Y_h$. The corresponding saddle-point problems are again solved using the Variable-ADMM presented in [10]. The elementwise optimization problem with respect to r_h appearing in the Variable-ADMM is solved using for given u_h and λ_h the explicit formula

$$r_h = \max\left\{|\lambda_h + \tau\nabla u_h| - h^{-d}\tau^{-1}, 0\right\} \frac{\lambda_h + \tau\nabla u_h}{|\lambda_h + \tau\nabla u_h|}.$$

For given p_h and μ_h the optimization problem with respect to q_h is a nodewise optimization problem due to mass lumping and is solved using the shrinkage operator

$$q_h = \frac{\mu_h/\tau + p_h}{\max\{1, |\mu_h/\tau + p_h|\}},$$

see [39] for details.

6. Numerical experiments

In this section we present our numerical results for the approximation of solutions for the nonlinear Laplace equation and the ROF problem using mesh adaptivity which is based on the primal-dual gap estimators $\eta(u_h, p_h)$. The refinement of a given triangulation $\mathcal{T}_h$ relies on the Dörfler marking and consists in the bisection of elements $T \in \mathcal{M}_h$ of a minimal set $\mathcal{M}_h \subset \mathcal{T}_h$ for which

$$\left[\sum_{T\in\mathcal{M}_h} \eta^T(u_h, p_h)^2\right]^{1/2} \ge 1/2 \left[\sum_{T\in\mathcal{T}_h} \eta^T(u_h, p_h)\right]^{1/2}$$

holds. Additional elements then are refined to avoid hanging nodes. The numerical approximations u_h and p_h for the primal and dual problem, respectively, are obtained using the corresponding saddle-point formulations and the Variable-ADMM presented in [10].

Before we report the performance of the adaptive algorithm for the nonlinear Laplace equation and the ROF problem in this section, we will first briefly comment on the hybrid realization of the Brezzi–Douglas–Marini finite element space.

6.1. Hybrid implementation of $\mathcal{BDM}(\Omega)$

We first of all define the space

$$Z_h = \left\{ r_h \in L^\infty(\cup \mathcal{S}_h) : r_h|_S \text{ affine for all } S \in \mathcal{S}_h \right\},$$

i.e., Z_h contains all functions r_h that are piecewise affine, discontinuous functions on the skeleton $\mathcal{S}_h$ of the triangulation $\mathcal{T}_h$. The space $\mathcal{BDM}(\Omega)$ consists of all elementwise affine vector fields q_h for which the normal component is continuous across interelement sides $S \in \mathcal{S}_h$, *i.e.*,

$$[\![q_h \cdot n_S]\!]|_S(x) = \lim_{\varepsilon \to 0} \big(q_h(x + \varepsilon n_S) - q_h(x - \varepsilon n_S) \big) \cdot n_S = 0$$

for all $x \in S$ with a unit normal n_S on S. If $\mathcal{BDM}(\Omega)$ is defined to be a subspace of $H_N(\mathrm{div};\Omega)$, the normal component on Γ_N vanishes, *i.e.*,

$$[\![q_h \cdot n_S]\!]|_S(x) = q_h(x) \cdot n_S = 0$$

for all boundary sides $S \in \mathcal{S}_h \cap \Gamma_N$ and $x \in S$. This means that $q_h \in \mathcal{BDM}(\Omega)$, if and only if $q_h \in \mathcal{L}^1(\mathcal{T}_h)^d$ and

$$\int_{\cup(\mathcal{S}_h \setminus (\mathcal{S}_h \cap \Gamma_D))} [\![q_h \cdot n_S]\!] r_h \, \mathrm{d}s = 0$$

for all $r_h \in Z_h$.

6.2. Nonlinear Laplace equation

We consider the nonlinear Laplace problem with inhomogeneous Dirichlet data on the L-shaped domain and let $\Omega = (-1,1)^2 \setminus ([0,1] \times [-1,0])$, $\Gamma_D = \partial\Omega$ and $g = 0$, and define the Dirichlet data $u_D = u|_{\partial\Omega}$ through restriction of the exact solution given in polar coordinates by

$$u(r,\theta) = r^\delta \sin(\delta\theta)$$

to the boundary. The choice of δ will be specified later in dependence of the choice of σ. The nonsmooth source term f is then given in polar coordinates by

$$f(r,\theta) = -(2-\sigma)\delta^{\sigma-1}(1-\delta) r^{(\delta-1)(\sigma-1)-1} \sin(\delta\theta).$$

We let $\delta = (6/5)(1 - 1/\sigma)$. Then we have that $u \in W^{1,\sigma}(\Omega)$ but $u \notin W^{2,\sigma}(\Omega)$. In what follows $u_h \in X_h$ and $p_h \in Y_h$ denote approximate solutions to the primal and dual nonlinear Laplace problem obtained with the iterative scheme Variable-ADMM (*cf.* [10]).

In Figure 1 the error estimator $\widehat{\eta}^h_{\Delta_\sigma}(u_h, p_h)$ and the error in the quasi-norm on the left-hand side of the estimate in Proposition 4.6

$$\varrho^{1/2}_{\Delta_\sigma} = \|V(\nabla u) - V(\nabla u_h)\|$$

are plotted against the number of degrees of freedom $N = |\mathcal{N}_h|$ in a loglog-plot. One can clearly observe that mesh adaptivity yields the quasi-optimal convergence rate $\overline{h} \sim N^{-1/2}$. Particularly, the primal-dual gap error estimator $\widehat{\eta}^h_{\Delta_\sigma}(u_h, p_h)$ defines a reliable upper bound for the error in the quasi-norm. On the right-hand side of Figure 1 we displayed the energy curves for the primal and dual energy $E^h_{\Delta_\sigma}(u_h)$ and $\widehat{D}^h_{\Delta_\sigma}(p_h)$, respectively. The primal and dual energy converge to the optimal value and the primal-dual gap $E^h_{\Delta_\sigma}(u_h) - \widehat{D}^h_{\Delta_\sigma}(p_h)$ converges to zero as $N \to \infty$ and at a higher rate, when local mesh refinement is used. In Figure 2 three snapshots of the refined mesh are displayed, which show that the primal-dual gap error estimator yields triangulations that are locally refined in the neighborhood of the singularity. The high resolution is even more localized for $\sigma \to 1$, since the singularity at the reentrant corner increases.

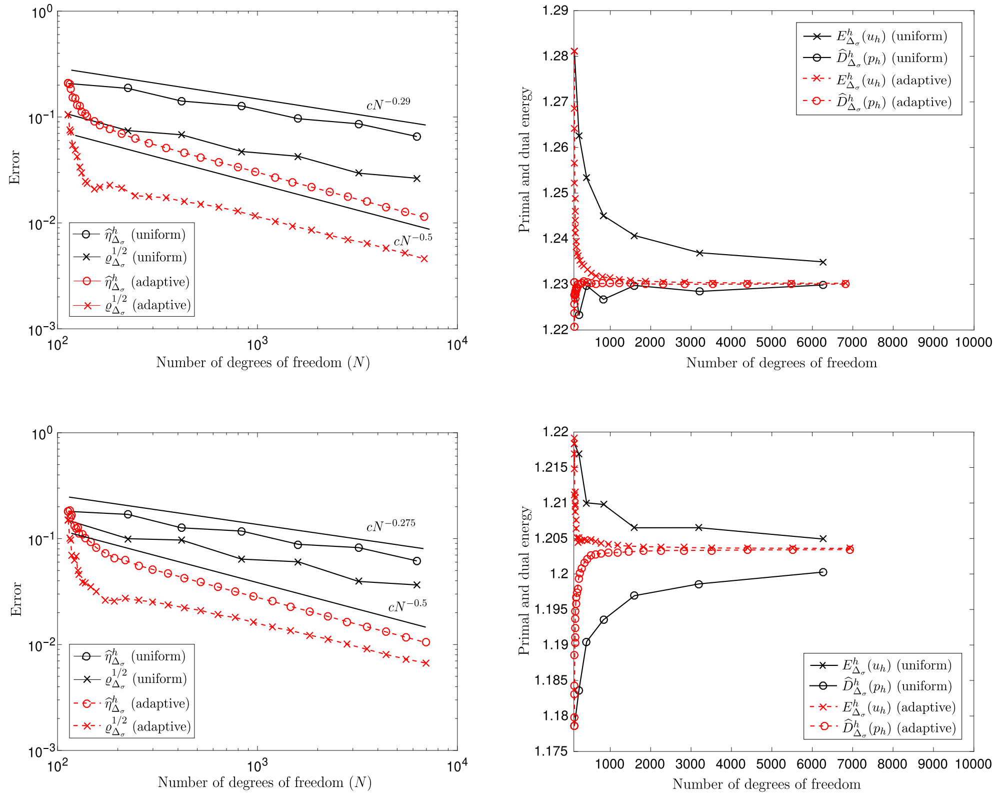

FIGURE 1. Primal-dual gap error estimators $\widehat{\eta}^h_{\Delta_\sigma}$ and error $\varrho^{1/2}_{\Delta_\sigma} = \|V(\nabla u) - V(\nabla u_h)\|$ (*left*) and primal and dual energy $E^h_{\Delta_\sigma}(u_h)$ and $\widehat{D}^h_{\Delta_\sigma}(p_h)$ (*right*) for uniform and adaptive mesh refinement. *Top*: nonlinear Laplace problem with $\sigma = 1.6$. *Bottom*: primal-dual gap estimator Laplace problem with $\sigma = 1.2$.

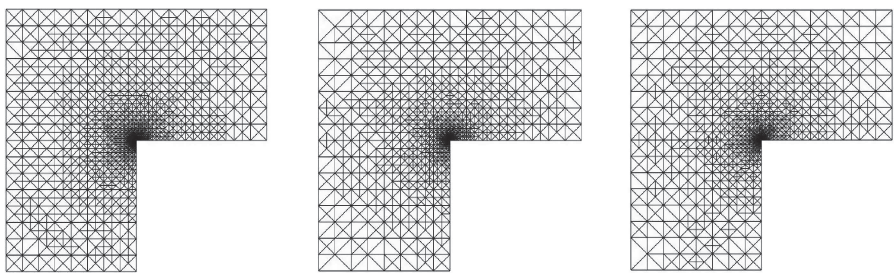

FIGURE 2. Snapshots of refined meshes for nonlinear Laplace problem with $\sigma = 1.6$ (*left*), $\sigma = 1.2$ (*middle*) and $\sigma = 1.05$ (*right*). The mesh is locally refined in a neighborhood of the reentrant corner. The resolution at the reentrant corner increases as $\sigma \to 1$.

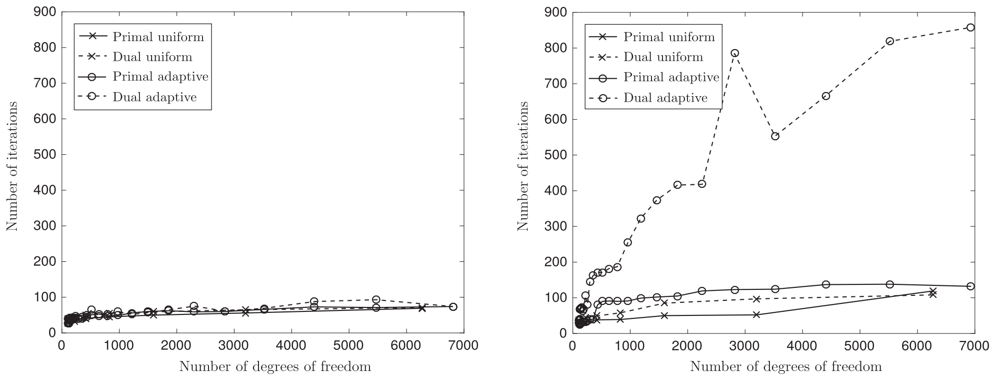

FIGURE 3. Iterations numbers for Variable-ADMM for the minimization of $E^h_{\Delta_\alpha}$ $\widehat{D}^h_{\Delta_\alpha}$ for both uniform and adaptive refinement. *Left*: $\sigma = 1.6$; *right*: $\sigma = 1.2$.

In Figure 3 the iteration numbers for the Variable-ADMM for the primal and dual problem are plotted *versus* the number of degrees of freedom for both uniform and adaptive mesh refinement and for parameters $\sigma = 1.6$ and $\sigma = 1.2$. The error tolerance for the residual in the Variable-ADMM was of order $\mathcal{O}(h^2)$. One can observe that the iteration numbers for the dual problem critically increase as σ is decreased.

Let us finally consider the residual-based error estimator

$$\eta^h_{\mathrm{res}}(u_h)^2 = \sum_{T\in\mathcal{T}_h} \eta^{h,T}_{\mathrm{res}}(u_h)^2$$

from [13, 21, 36–38] with

$$\eta^{h,T}_{\mathrm{res}}(u_h)^2 = \eta^{h,T}_{E}(u_h)^2 + \sum_{S\in\mathcal{S}_h\setminus\partial\Omega, S\subset\partial T} \eta^{h,S}_{J}(u_h)^2$$

and

$$\eta^{h,T}_{E}(u_h)^2 = \int_T (|\nabla u_h|^{\sigma-1} + h_T|f_h|)^{\sigma'-2} h_T^2 |f_h|^2 \,\mathrm{d}x,$$
$$\eta^{h,S}_{J}(u_h)^2 = \int_{\omega_S} (|\nabla u_h| + |[\![\nabla u_h]\!]_S|)^{\sigma-2} |[\![\nabla u_h]\!]_S|^2 \,\mathrm{d}x,$$

where $\omega_S = \bigcup\{T_1, T_2 \in \mathcal{T}_h : \ S = T_1 \cap T_2\}$ for $S \in \mathcal{S}_h \setminus \partial\Omega$ and u_h is the unique discrete minimizer of $E^h_{\Delta_\sigma}$. The expression $[\![\nabla u_h]\!]_S$ denotes the jump of ∇u_h across an inner side $S \in \mathcal{S}_h$ defined by

$$[\![\nabla u_h]\!]_S = \nabla u_h|_{T_1} - \nabla u_h|_{T_2}$$

for $S = T_1 \cap T_2$. The error estimator $\eta^h_{\mathrm{res}}(u_h)$ has been extensively studied in [13, 21, 36–38], where the efficiency and reliability of the estimator has been proven and the linear convergence as well as the optimality of the corresponding adaptive finite element scheme have been shown.

In Figure 4 we compare the primal-dual gap error estimator $\widehat{\eta}^h_{\Delta_\sigma}(u_h, p_h)$ with the residual error estimator $\eta^h_{\mathrm{res}}(u_h)$ for the nonlinear Laplace problem with inhomogeneous Dirichlet data on the L-shaped domain for $\sigma = 1.6$ and $\sigma = 1.2$ as before. One can observe that both estimators decay at the same rate $\mathcal{O}(N^{-1/2})$ on a sequence of locally refined meshes driven by an element marking strategy based on $\widehat{\eta}^h_{\Delta_\sigma}(u_h, p_h)$.

However, the overestimation of the primal-dual gap estimator $\widehat{\eta}^h_{\Delta_\sigma}(u_h, p_h)$ is moderate compared to the residual-based error estimator $\eta^h_{\mathrm{res}}(u_h)$. While the overestimation of $\eta^h_{\mathrm{res}}(u_h)$ for $\sigma = 1.6$ and $\sigma = 1.2$ do not

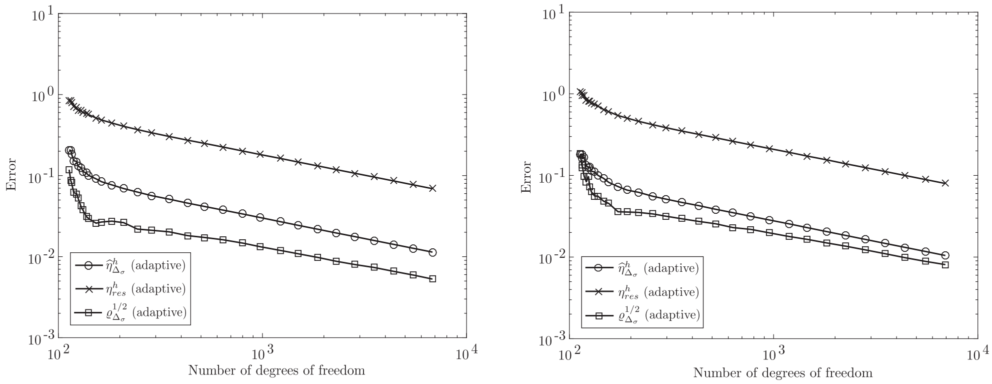

FIGURE 4. Primal-dual gap estimator $\widehat{\eta}^h_{\Delta_\sigma}$, residual-based estimator η^h_{res} and error $\varrho^{1/2}_{\Delta_\sigma} = \|V(\nabla u) - V(\nabla u_h)\|$ for a sequence of adaptively refined meshes driven by $\widehat{\eta}^h_{\Delta_\sigma}$. *Left*: nonlinear laplace problem with $\sigma = 1.6$. *Right*: nonlinear laplace problem with $\sigma = 1.2$.

TABLE 1. Effectivity indices $\widehat{e}^h_{\Delta_\sigma}$ and e^h_{res} for primal-dual gap estimator and residual-based estimator, respectively, for a sequence of adaptively refined meshes with number of degrees of freedom $|\mathcal{N}_h|$.

Effectivity indices for nonlinear Laplace problem					
$\sigma = 1.6$			$\sigma = 1.2$		
$\lvert\mathcal{N}_h\rvert$	$\widehat{\eta}^h_{\Delta_\sigma}$	η^h_{res}	$\lvert\mathcal{N}_h\rvert$	$\widehat{\eta}^h_{\Delta_\sigma}$	η^h_{res}
209	2.64	15.45	218	1.76	13.08
1221	2.26	13.66	1186	1.43	10.75
1863	2.23	13.48	1824	1.38	10.42
2306	2.23	13.52	2254	1.35	10.22
2840	2.21	13.38	2815	1.34	10.18
3524	2.15	13.06	3519	1.34	10.17
4398	2.15	13.05	4401	1.34	10.16
5468	2.14	12.97	5524	1.32	10.07
6808	2.15	13.05	6934	1.29	9.86

differ significantly, the gap between the primal-dual gap error estimator and the error diminishes for $\sigma = 1.2$. Let us also remark that in the proofs of the reliability and the efficiency of the residual-based error estimator $\eta^h_{\mathrm{res}}(u_h)$ it is crucial that u_h is the unique solution to the primal nonlinear Laplace problem in X_h, *cf.* [21]. Its robustness regarding inexact iterative solutions is not addressed in the aforementioned articles.

Finally, in Table 1 the effectivity indices

$$\widehat{e}^h_{\Delta_\sigma} = \frac{\widehat{\eta}^h_{\Delta_\sigma}}{\varrho^{1/2}_{\Delta_\sigma}}, \quad e^h_{\mathrm{res}} = \frac{\eta^h_{\mathrm{res}}}{\varrho^{1/2}_{\Delta_\sigma}}$$

corresponding to the primal-dual gap error estimator and the residual-based estimator, respectively, are shown depending on the number of degrees of freedom $|\mathcal{N}_h|$. It has to be taken into account that the uniform convexity constant of the functional E_{Δ_σ} enters the *a posteriori* error estimate in Proposition 4.6.

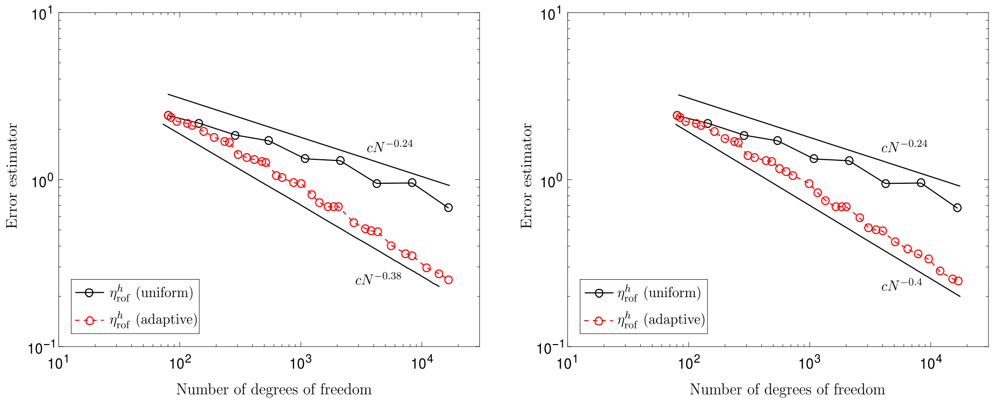

FIGURE 5. Error estimator η^h_{rof} for Example 6.1 with discretization of the dual problem with continuous finite element space Y_h^C (*left*) and $H(\mathrm{div};\Omega)$-conforming finite element space Y_h^{dC} (*right*) for uniform and adaptive mesh refinement.

6.3. Rudin–Osher–Fatemi image denoising

We let $\Omega = (-1,1)^2$ and consider two examples with the given function g being the characteristic function of a set, the first one with homogeneous Neumann boundary conditions and the second one with homogeneous Dirichlet boundary conditions, for which we have an explicit solution at hand. In the case of Dirichlet boundary conditions the dual energy functional D_{rof} is maximized over $H(\mathrm{div};\Omega)$ instead of $H_N(\mathrm{div};\Omega)$. The calculations remain valid, but in general it is nontrivial to guarantee the existence of solutions for Dirichlet boundary conditions. Since our principal motivation for considering the ROF problem is the application to total variation regularized damage evolution models from continuum mechanics, *cf.* [50], rather than image processing we do not include experiments with g being a real image. For adaptive mesh refinement techniques applied to real images we refer the reader, *e.g.*, to [32].

Example 6.1. We set $\Gamma_D = \emptyset$, $\Gamma_N = \partial\Omega$, $\alpha = 100$, and $g = \chi_{B^\infty_{1/2}(0)}$ the characteristic function of $B^\infty_{1/2}(0) = \{(x_1,x_2) \in \mathbb{R}^2 : \max\{|x_1|,|x_2|\} \le 1/2\}$.

In Figure 5 the error estimator η^h_{rof} is plotted against the number of degrees of freedom $N = |\mathcal{N}_h|$ using a logarithmic scaling on both axes both for uniform and adaptive mesh refinement and with the dual problem discretized with the continuous finite element space $Y_h^C = \mathcal{S}^1(\mathcal{T}_h)^d$ and the $H(\mathrm{div};\Omega)$-conforming finite element space $Y_h^{dC} = \mathcal{L}^1(\mathcal{T}_h)^d \cap H_N(\mathrm{div};\Omega)$. Again, one can observe that using locally refined meshes with Y_h^C as the discrete space for the dual problem yields a better convergence rate $\overline{h}^{0.76} \sim N^{-0.38}$ as compared to uniform refinement with an experimental convergence rate of $\overline{h}^{0.47}$. For the choice Y_h^{dC} we record the rates $\overline{h}^{0.81} \sim N^{-0.4}$ (adaptive) and $\overline{h}^{0.47} \sim N^{-0.24}$ (uniform). The choice of the finite element space for the discretization of the dual problem does not significantly affect the rate of convergence of the primal-dual gap error estimator η^h_{rof}.

Example 6.2. We set $\Gamma_D = \partial\Omega$, $\Gamma_N = \emptyset$, $\alpha = 10$ and $g = \chi_{B^2_{1/2}(0)}$ with $B^2_{1/2}(0) = \{x \in \mathbb{R}^2 : |x| \le 1/2\}$.

In this case the exact solution is given by $u = (3/5)\chi_{B^2_{1/2}(0)}$, *cf.* [9].

In Figure 6 the error estimator η^h_{rof} and the L^2-error

$$\varrho^{1/2}_{\mathrm{rof}} = (\alpha/2)^{1/2}\|u-u_h\|$$

are plotted against the number of degrees of freedom in a loglog-plot and again, as before, both for uniform and adaptive mesh refinement and for the discretization of the dual problem with Y_h^C (left) and Y_h^{dC} (right).

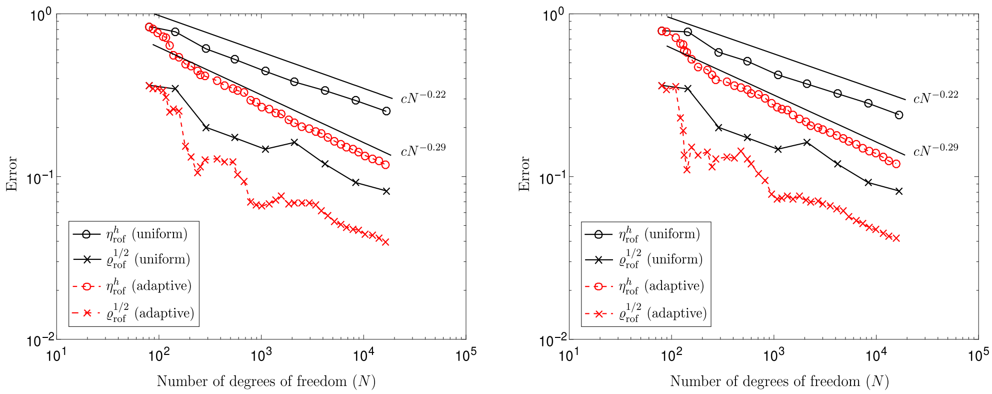

FIGURE 6. Primal-dual gap error estimator η_{rof}^h and L^2-error $\varrho_{\mathrm{rof}}^{1/2} = (\alpha/2)^{1/2}\|u - u_h\|$ for Example 6.2 with discretization of the dual problem with continuous finite element space Y_h^C (*left*) and $H(\mathrm{div};\Omega)$-conforming finite element space Y_h^{dC} (*right*) for uniform and adaptive mesh refinement.

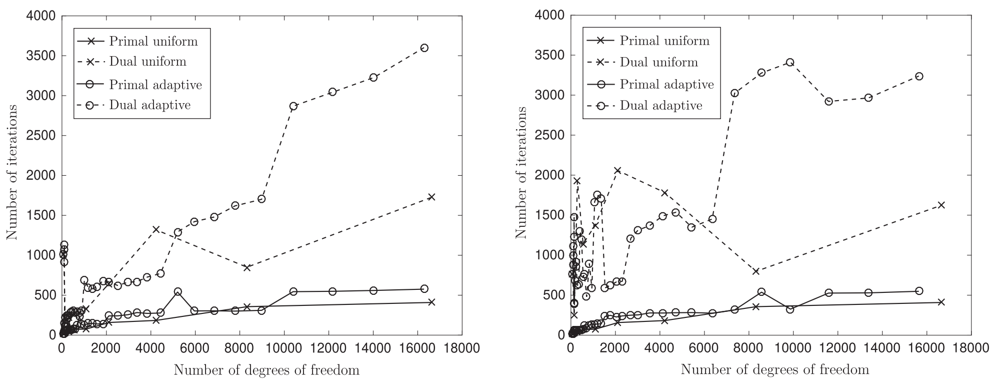

FIGURE 7. Iterations numbers for Variable-ADMM for the minimization of E_{rof}^h and $-D_{\mathrm{rof}}^h$ for both uniform and adaptive refinement. *Left*: $Y_h = Y_h^C$; *right*: $Y_h = Y_h^{dC}$.

The plot underlines that the quantity η_{rof}^h defines a reliable estimator for the L^2-error $\varrho_{\mathrm{rof}}^{1/2}$ as predicted by Proposition 5.3. One can, once again, observe that adaptive mesh refinement leads to an improvement of the convergence rate from $\overline{h}^{0.44} \sim N^{-0.22}$ to $\overline{h}^{0.62} \sim N^{-0.31}$ for both discretization methods for the dual problem. In Figure 7 the iteration numbers for the Variable-ADMM for the primal and dual problem are plotted against the number of degrees of freedom for both uniform and adaptive mesh refinement and for discretizations of the dual problem with $Y_h = Y_h^C$ and $Y_h = Y_h^{dC}$. The error tolerance for the residual in the Variable-ADMM was of order $\mathcal{O}(h)$. The iteration numbers for $Y_h = Y_h^C$ and $Y_h = Y_h^{dC}$ do not differ significantly. However, one can observe that the iteration numbers of the Variable-ADMM as a function of the degrees of freedom grow significantly faster for the dual problem compared to the primal problem reflecting the weaker coercivity property.

In Table 2 the effectivity index

$$e_{\mathrm{rof}}^h = \frac{\eta_{\mathrm{rof}}^h}{\varrho_{\mathrm{rof}}^{1/2}}$$

TABLE 2. Effectivity index e^h_{rof} for primal-dual gap estimator for a sequence of adaptively refined meshes with number of degrees of freedom $|\mathcal{N}_h|$.

Effectivity index for ROF problem								
$\lvert\mathcal{N}_h\rvert$	1531	3542	5432	7356	9868	11 579	13 373	15 638
e^h_{rof}	3.27	2.82	2.91	2.91	2.95	2.92	2.90	2.86

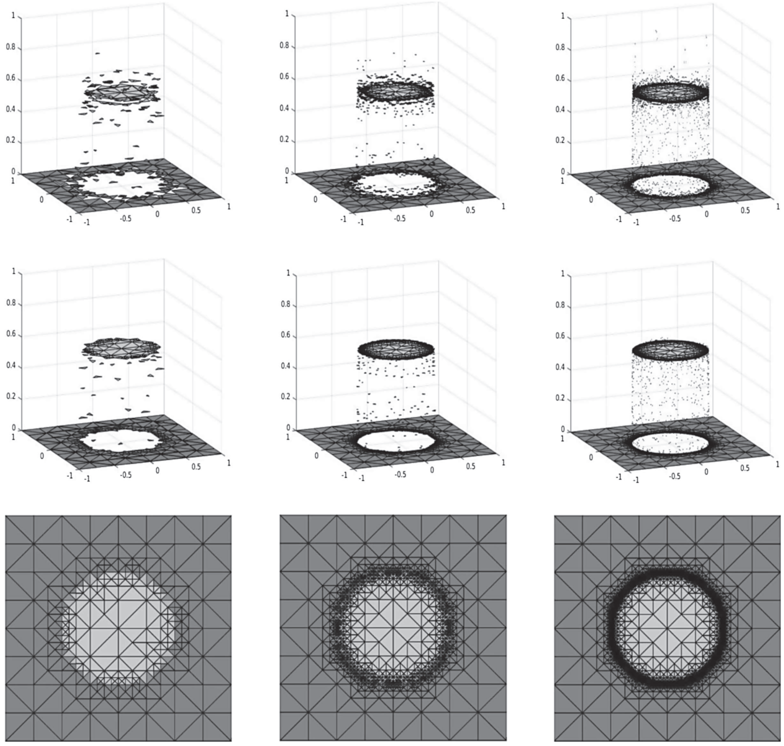

FIGURE 8. Piecewise constant approximations $\overline{u}_h = (1/\alpha)\,\mathrm{div}\,p_h + g_h$ for a sequence of adaptively refined triangulations for Example 6.2. *Top*: dual variable is approximated in $Y_h^C = \mathcal{S}^1(\mathcal{T}_h)^d$. *Middle*: dual variable is approximated in $Y_h^{dC} = \mathcal{L}^1(\mathcal{T}_h)^d \cap H_N(\mathrm{div};\Omega)$. One can observe oscillations of $\overline{u}_h$ along the jump set for the discretization of the dual ROF problem with Y_h^C. *Bottom*: bird's eye view of the middle row. The mesh is locally refined in a neighborhood of the circular jump set.

depending on the number of degrees of freedom $|\mathcal{N}_h|$ is presented and shows that the overestimation of the error is moderate.

In Figure 8 we depicted for a sequence of adaptively refined triangulations the piecewise constant approximations $\overline{u}_h = (1/\alpha)\,\mathrm{div}\,p_h + g_h$ with $p_h \in Y_h^C$ (top) and $p_h \in Y_h^{dC}$ (bottom), *cf.* Proposition 5.2. Although the different discretization methods for the dual problem do not affect the convergence rates in the presented experiments, the discretization of the dual problem with the continuous finite element space Y_h^C causes oscillations in $\overline{u}_h$ along the jump set.

7. Conclusion

We have seen that the primal-dual gap error estimator defines a reliable upper bound with constant one for the error in the energy for convex minimization problems. For uniformly convex minimization problems it also controls the error with respect to a distance induced by the uniform convexity. The primal-dual gap error estimator has been introduced in [45] in an abstract setting and has been applied to several minimization problems in an infinite-dimensional framework. We extended the theory to general finite discretizations of convex minimization problems and applied the theory to the nonlinear Laplace problem and the ROF problem, which serve as model problems for a wide class of convex minimization problems. The theoretical results, especially the reliability of the primal-dual gap error estimator, has been confirmed in several numerical experiments. In order to compute the estimator we approximately solved the primal and dual problems using the Variable-ADMM provided in [10]. Yet, it seems necessary to consider more efficient strategies to construct feasible functions especially for the dual problems.

Acknowledgements. The authors acknowledge support by the DFG via the priority programme SPP 1748 *Reliable Simulation Techniques in Solid Mechanics: Development of Non-standard Discretization Methods, Mechanical and Mathematical Analysis.*

References

[1] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Eng.* **142** (1997) 1–88.
[2] C. Atkinson and C.R. Champion, Some boundary-value problems for the equation $\nabla.(|\nabla\phi|^N\nabla\phi) = 0$. *Q. J. Mech. Appl. Math.* **37** (1984) 401–419.
[3] C. Atkinson and C.W. Jones, Similarity solutions in some nonlinear diffusion problems and in boundary-layer flow of a pseudo plastic fluid. *Q. J. Mech. Appl. Math.* **27** (1974) 193–211.
[4] G. Aubert and P. Kornprobst, Mathematical problems in image processing, 2nd edition. In: Vol. 147 of *Applied Mathematical Sciences.* Springer (2006).
[5] I. Babuška and W.C. Rheinboldt, Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.* **15** (1978) 736–754.
[6] J. Baranger and K. Najib, Numerical analysis of quasi-newtonian flow obeying the power low or the carreau flow. *Numer. Math.* **58** (1990) 35–49.
[7] J.W. Barrett and W.B. Liu, Finite element approximation of the p-Laplacian. *Math. Comput.* **61** (1993) 523–537.
[8] S. Bartels, Error control and adaptivity for a variational model problem defined on functions of bounded variation. *Math. Comput.* **84** (2015) 1217–1240.
[9] S. Bartels, Numerical methods for nonlinear partial differential equations. In: Vol. 47 of *Springer Series in Computational Mathematics.* Springer (2015).
[10] S. Bartels and M. Milicevic, Alternating direction method of multipliers with variable step sizes. Preprint arXiv:1704.06069 (2017).
[11] S. Bartels and P. Schön, Adaptive approximation of the monge-kantorovich problem *via* primal-dual gap estimates. *ESAIM: M2AN* **51** (2017) 2237–2261.
[12] S. Bartels, R.H. Nochetto and A.J. Salgado, Discrete total variation flows without regularization. *SIAM J. Numer. Anal.* **52** (2014) 363–385.
[13] L. Belenki, L. Diening and C. Kreuzer, Optimality of an adaptive finite element method for the p-Laplacian equation. *IMA J. Numer. Anal.* **32** (2012) 484–510.
[14] D. Boffi, F. Brezzi and M. Fortin. Mixed finite element methods and applications. In: Vol. 44 of *Springer Series in Computational Mathematics.* Springer (2013).
[15] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods, 3rd edition. In: Vol. 15 of *Texts in Applied Mathematics.* Springer (2008).
[16] C. Carstensen and R. Klose, A posteriori finite element error control for the p-Laplace problem. *SIAM J. Sci. Comput.* **25** (2003) 792–814.
[17] C. Carstensen, W.B. Liu and N.N. Yan, A posteriori FE error control for p-Laplacian by gradient recovery in quasi-norm. *Math. Comput.* **75** (2006) 1599–1616.
[18] S.-S. Chow, Finite element error estimates for non-linear elliptic equations of monotone type. *Numer. Math.* **54** (1989) 373–393.
[19] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. Society for Industrial and Applied Mathematics (2002).
[20] L. Diening and M.Růžička, Interpolation operators in orlicz–sobolev spaces. *Numer. Math.* **107** (2007) 107–129.
[21] L. Diening and C. Kreuzer, Linear convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM J. Numer. Anal.* **46** (2008) 614–638.

[22] C. Ebmeyer, Nonlinear elliptic problems with p-structure under mixed boundary value conditions in polyhedral domains. *Adv. Differ. Equ.* **6** (2001) 873–895.

[23] C. Ebmeyer, Mixed boundary value problems for nonlinear elliptic systems with p-structure in polyhedral domains. *Math. Nachr.* **236** (2002) 91–108.

[24] C. Ebmeyer, Global regularity in Sobolev spaces for elliptic problems with p-structure on bounded domains, edited by J.F. Rodrigues, G. Seregin and J.M. Urbano. In: Trends in Partial Differential Equations of Mathematical Physics. Birkhäuser Basel, Basel (2005) 81–89.

[25] C. Ebmeyer and W.B. Liu, Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-Laplacian problems. *Numer. Math.* **100** (2005) 233–258.

[26] C. Ebmeyer, W.B. Liu and M. Steinhauer, Global regularity in fractional order sobolev spaces for the p-Laplace equation on polyhedral domains. *Z. Anal. Anwend.* **24** (2005) 353–374.

[27] I. Ekeland and R. Témam, Convex Analysis and Variational Problems. Society for Industrial and Applied Mathematics (1999).

[28] L. El Alaoui, A. Ern and M. Vohralík, Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. Special Issue on Modeling Error Estimation and Adaptive Modeling. *Comput. Methods Appl. Mech. Eng.* **200** (2011) 2782–2795.

[29] A. Ern and M. Vohralík, Adaptive inexact newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.* **35** (2013) A1761–A1791.

[30] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Comput. Math. App.* **2** (1976) 17–40.

[31] R. Glowinski and A. Marroco, Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *RAIRO Anal. Numer.* **9** (1975) 41–76.

[32] M. Hintermüller and M. Rincon-Camacho, An adaptive finite element method in L^2-TV-based image denoising. *Inverse Prob. Imaging* **8** (2014) 685–711.

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

[34] W.B. Liu and J.W. Barrett, A further remark on the regularity of the solutions of the p-Laplacian and its applications to their finite element approximation. *Nonlinear Anal. Theory Methods Appl.* **21** (1993) 379–387.

[35] W.B. Liu and J.W. Barrett, A remark on the regularity of the solutions of the p-Laplacian and its application to their finite element approximation. *J. Math. Anal. App.* **178** (1993) 470–487.

[36] W.B. Liu and N.N. Yan, Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian. *Numer. Math.* **89** (2001) 341–378.

[37] W.B. Liu and N.N. Yan, Quasi-norm local error estimators for p-Laplacian. *SIAM J. Numer. Anal.* **39** (2001) 100–127.

[38] W.B. Liu and N.N. Yan, On quasi-norm interpolation error estimation and a posteriori error estimates for p-Laplacian. *SIAM J. Numer. Anal.* **40** (2002) 1870–1895.

[39] M. Milicevic, *Finite element discretization and iterative solution of total variation regularized minimization problems and application to the simulation of rate-independent damage evolutions*. Ph.D. thesis, Albert-Ludwigs-Universität Freiburg (2019).

[40] R.H. Nochetto, K.G. Siebert and A. Veeser, Theory of adaptive finite element methods: an introduction, edited by R. DeVore and A. Kunoth. In: Multiscale, Nonlinear and Adaptive Approximation. Springer Berlin Heidelberg, Berlin, Heidelberg (2009) 409–542.

[41] J.R. Philip, n-diffusion. *Aust. J. Phys.* **14** (1961) 1–13.

[42] S.I. Repin, A posteriori error estimates for approximate solutions to variational problems with strongly convex functionals. *J. Math. Sci.* **97** (1999) 4311–4328.

[43] S.I. Repin, A posteriori error estimates for approximate solutions of variational problems with functionals of power growth. *J. Math. Sci.* **101** (2000) 3531–3538.

[44] S.I. Repin, A posteriori error estimation for nonlinear variational problems by duality theory. *J. Math. Sci.* **99** (2000) 927–935.

[45] S.I. Repin, A posteriori error estimation for variational problems with uniformly convex functionals. *Math. Comput.* **69** (2000) 481–500.

[46] S.I. Repin and L.S. Xanthis, A posteriori error estimation for elasto-plastic problems based on duality theory. *Comput. Methods Appl. Mech. Eng.* **138** (1996) 317–339.

[47] S.I. Repin and L.S. Xanthis, A posteriori error estimation for nonlinear variational problems. *C.R. Acad. Sci. – Ser. I Math.* **324** (1997) 1169–1174.

[48] L.I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. D: Nonlinear Phenom.* **60** (1992) 259–268.

[49] R. Stevenson, Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7** (2007) 245–269.

[50] M. Thomas, Quasistatic damage evolution with spatial BV-regularization. *Discrete Contin. Dyn. Syst. – S* **6** (2013) 235–255.

[51] A. Veeser, Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.* **92** (2002) 743–770.

[52] R. Verfürth, A Posteriori Error Estimation Techniques for Finite Element Methods. Oxford University Press (2013).