NONCONFORMING DISCRETIZATIONS OF CONVEX MINIMIZATION PROBLEMS AND PRECISE RELATIONS TO MIXED METHODS

SÖREN BARTELS

ABSTRACT. This article discusses nonconforming finite element methods for convex minimization problems and systematically derives dual mixed formulations. Duality relations lead to simple error estimates that avoid an explicit treatment of nonconformity errors. A reconstruction formula provides the discrete solution of the dual problem via a simple postprocessing procedure which implies a strong duality relation and is of interest in a posteriori error estimation. The framework applies to differentiable and nonsmooth problems, examples include p-Laplace, total-variation regularized, and obstacle problems. Numerical experiments illustrate advantages of nonconforming over standard conforming methods.

1. INTRODUCTION

Mixed finite element methods as introduced in [44, 16] provide an attractive framework to approximate partial differential equations in divergence form since they lead to accurate approximations of fluxes. For the Poisson problem it is well understood that a close connection of mixed methods to nonconforming methods exists, cf. [40, 3]. This is of practical interest since mixed finite element methods require the solution of saddle-point problems while nonconforming methods lead to positive definite linear systems. Moreover, the nonconforming Crouzeix–Raviart element of [25] has proved to be particularly robust and flexible to provide accurate approximations for Stokes equations [33], for nearly incompressible Navier–Lamé equations [35], and for singular minimizers related to the Lavrentiev phenomenon in the calculus of variations [42]. Another useful feature is that the element is suitable to compute reliable lower bounds for eigenvalue problems [2, 20]. Further aspects of the Crouzeix–Raviart element are addressed in [18]. In this article we show that the relation to mixed methods applies to a large class of convex minimization problems provided an appropriate discretization is used. From a discrete duality relation we derive quasi-optimal error estimates for the modified discretizations, show that they apply to various nonlinear

Date: February 6, 2020.

2010 *Mathematics Subject Classification.* 65N12 65N15 65N30 49M25.

Key words and phrases. Nonconforming methods, mixed methods, nonsmooth problems, discrete duality, error estimates.

partial differential equations and variational inequalities, and illustrate the theoretical findings via simulations for certain singular limit settings. The results of this article are inspired by recent work on quasi-optimal convergence rates for nonconforming approximations of total-variation regularized problems in [23].

1.1. **Convex minimization.** To explain the main ideas we consider a convex variational problem defined via a minimization of the energy functional

$$I(u) = \int_\Omega \phi(\nabla u) \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x,$$

in a Sobolev space $W^{1,p}_D(\Omega)$, i.e., subject to homogeneous Dirichlet boundary conditions on a boundary part $\Gamma_D \subset \partial\Omega$; we set $\Gamma_N = \partial\Omega \setminus \Gamma_D$. The dual problem is obtained by using the relation $\phi^{**} = \phi$ with the convex conjugate

$$\phi^*(t) = \sup_{s\in\mathbb{R}^d} s \cdot t - \phi(s)$$

It consists in maximizing the functional

$$D(z) = -\int_\Omega \phi^*(z) \,\mathrm{d}x$$

in the space of vector fields $z \in L^{p'}(\Omega;\mathbb{R}^d)$ whose distributional divergence $\operatorname{div} z$ belongs to $L^{p'}(\Omega)$ with vanishing normal component on Γ_N and which satisfy the constraint

$$-\operatorname{div} z = f.$$

It turns out that solutions are related via

$$z = D\phi(\nabla u) \quad \Longleftrightarrow \quad \nabla u = D\phi^*(z),$$

and satisfy the Euler–Lagrange equation

$$-\operatorname{div} D\phi(\nabla u) = f$$

and the saddle-point system

$$D\phi^*(z) - \nabla\lambda = 0, \quad -\operatorname{div} z = f,$$

where λ is the Lagrange multiplier related to the divergence constraint. One directly verifies that $\lambda = u$.

1.2. **Mixed and nonconforming methods.** A low order finite element discretization of the dual problem uses the Raviart–Thomas finite element space $\mathcal{RT}^0_N(\mathcal{T}_h)$ that contains certain piecewise linear vector fields whose distributional divergence is given by a piecewise constant function and which have vanishing normal component on Γ_N. In the quadratic case with $\phi(s) = |s|^2/2$ and $\phi^*(t) = |t|^2/2$, corresponding to the Poisson problem, the numerical method determines a uniquely defined vector field $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and an elementwise constant function $\overline{u}_h \in \mathcal{L}^0(\mathcal{T}_h)$ that solve

$$(z_h, y_h) + (\overline{u}_h, \operatorname{div} y_h) = 0, \quad (\operatorname{div} z_h, \overline{v}_h) = -(f, \overline{v}_h) \tag{1}$$

for all $(y_h, \overline{v}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$, where $(\cdot,\cdot)$ denotes the L^2 inner product of functions or vector fields with associated norm $\|\cdot\|$. The low order nonconforming approximation of the primal problem uses the Crouzeix–Raviart finite element space $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ of piecewise linear functions that are continuous at midpoints of sides of elements and vanish at the midpoints of sides belonging to Γ_D. It provides a nonconforming approximation of the Sobolev space $W^{1,2}_D(\Omega)$. With the piecewise application of the gradient operator denoted by ∇_h we have that the discrete solution $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ satisfies

$$(\nabla_h u_h, \nabla_h v_h) = (f_h, v_h)$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. It has been shown in [40] that the solutions z_h and u_h are related via

$$z_h|_T(x) = \nabla_h u_h|_T - \frac{f_h|_T}{d}(x - x_T)$$

on every element $T \in \mathcal{T}_h$ with midpoint $x_T \in T$, provided that $f_h = \Pi_{h,0} f$ is the L^2 projection of f onto $\mathcal{L}^0(\mathcal{T}_h)$. Moreover, it follows that

$$\overline{u}_h|_T = u_h(x_T) + \frac{f_h|_T}{d^2|T|}\|x - x_T\|^2_{L^2(T)}.$$

Hence, the solution of the mixed finite element method can entirely be determined by the solution of the nonconforming discretization and vice versa. We show that the relations can be generalized and that a modification of the dual problem simplifies the second equation.

1.3. **Generalized reconstruction.** We consider the nonconforming discretization of the primal problem given by the minimization of

$$I_h(u_h) = \int_\Omega \phi(\nabla_h u_h)\,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x$$

in the set of all $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. Solutions satisfy

$$\big(D\phi(\nabla_h u_h), \nabla_h v_h\big) = (f_h, v_h)$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. The systematically obtained discretization of the dual problem consists in maximizing the discrete functional

$$D_h(z_h) = -\int_\Omega \phi^*(\Pi_{h,0} z_h)\,\mathrm{d}x$$

for $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ subject to the constraint

$$-\operatorname{div} z_h = f_h.$$

The existence of a solution z_h follows from surjectivity properties of the divergence operator restricted to $\mathcal{RT}^0_N(\mathcal{T}_h)$. In contrast to consistent discretizations of the dual problem, here the operator $\Pi_{h,0}$ is included in defining D_h leading to discrete duality relations. It does not limit the coercivity properties of the problem since for divergence-free vector fields in $\mathcal{RT}^0(\mathcal{T}_h)$ we have that $\Pi_{h,0} y_h = y_h$. In fact, including the operator $\Pi_{h,0}$ has the interpretation of using quadrature which makes the numerical realization substantially easier. By imposing the divergence constraint via a Lagrange multiplier $\overline{u}_h$ one finds that optimal pairs $(z_h, \overline{u}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ satisfy the mixed formulation of the dual problem

$$\begin{aligned} \big(D\phi^*(\Pi_{h,0} z_h), \Pi_{h,0} y_h\big) + (\overline{u}_h, \operatorname{div} y_h) &= 0, \\ (\operatorname{div} z_h, \overline{v}_h) &= -(f_h, \overline{v}_h), \end{aligned}$$

for all $(y_h, \overline{v}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$. We claim that we have

$$z_h|_T(x) = D\phi(\nabla_h u_h|_T) - \frac{f_h|_T}{d}(x - x_T)$$

and

$$\overline{u}_h|_T = u_h(x_T)$$

for all $T \in \mathcal{T}_h$. To see this, let $\widetilde{z}_h$ and $\widetilde{u}_h$ denote the right-hand sides of the asserted identities for z_h and $\overline{u}_h$. We have that $-\operatorname{div} \widetilde{z}_h|_T = f_h|_T$ for all $T \in \mathcal{T}_h$, and

$$\Pi_{h,0} \widetilde{z}_h = D\phi(\nabla_h u_h). \tag{2}$$

Hence, for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ we have

$$(\widetilde{z}_h, \nabla_h v_h) = \big(D\phi(\nabla_h u_h), \nabla_h v_h\big) = (f_h, v_h) = (z_h, \nabla_h v_h),$$

where we used an integration-by-parts formula for products of Raviart–Thomas vector fields and gradients of Crouzeix–Raviart functions. Since $\operatorname{div}(\widetilde{z}_h - z_h)|_T = 0$ for every $T \in \mathcal{T}_h$, this identity implies that $\widetilde{z}_h - z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and in particular that $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Using that $[D\phi^*]^{-1} = D\phi$ we find that

$$D\phi^*(\Pi_{h,0} \widetilde{z}_h) = \nabla_h u_h.$$

Since $\widetilde{u}_h$ coincides with the elementwise average of u_h this implies that

$$\big(D\phi^*(\Pi_{h,0} \widetilde{z}_h), \Pi_{h,0} y_h\big) + (\widetilde{u}_h, \operatorname{div} y_h) = 0$$

for all $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Hence, we see that $(\widetilde{z}_h, \widetilde{u}_h)$ solves the mixed finite element formulation and in case of uniqueness coincides with the pair $(z_h, \overline{u}_h)$. The crucial identity (2) also implies the important duality relation $I_h(u_h) = D_h(z_h)$. It is also possible to construct the solution u_h of the nonconforming discretization from the pair $(z_h, \overline{u}_h)$ solving the mixed formulation of the dual problem. One directly verifies that this is given by

$$u_h(x) = \overline{u}_h|_T + D\phi^*(\Pi_{h,0} z_h|_T) \cdot (x - x_T)$$

for every $T \in \mathcal{T}_h$ and all $x \in T$. The reconstruction formulas are related to discrete Lagrange functionals, e.g.,

$$L_h(u_h, z_h) = \int_\Omega \nabla_h u_h \cdot z_h - \phi^*(\Pi_{h,0} z_h) - f_h \Pi_{h,0} u_h \, \mathrm{d}x,$$

and imply weak and strong discrete duality principles. We note that related reconstructions in the case of the p-Laplace problem have been identified in [38].

1.4. **Error estimates.** The discrete duality relation $I_h(u_h) \geq D_h(z_h)$ provides a natural way to derive error estimates. With a coercivity functional σ_{I_h} that measures strong convexity properties of I_h, we have for a minimizing u_h that

$$\sigma_{I_h}^2(u_h, v_h) \leq I_h(v_h) - I_h(u_h)$$

for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$. Choosing $v_h = \mathcal{I}_{cr} u$ and using $I_h(u_h) \geq D_h(z_h) \geq D_h(\mathcal{I}_{\mathcal{RT}} z)$ leads to

$$\delta_h^2 = \sigma_{I_h}^2(u_h, \mathcal{I}_{cr} u) \leq \int_\Omega \phi(\nabla_h \mathcal{I}_{cr} u) - f_h \mathcal{I}_{cr} u + \phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x.$$

Noting that $-\operatorname{div} \mathcal{I}_{\mathcal{RT}} z = f_h$ and using an integration-by-parts formula show that

$$\delta_h^2 \leq \int_\Omega \phi(\nabla_h \mathcal{I}_{cr} u) - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z \cdot \nabla_h \mathcal{I}_{cr} u + \phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x.$$

Fenchel's inequality implies that the integrand is nonnegative and vanishes if $\nabla_h \mathcal{I}_{cr} u = D\phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z)$. The identity $\nabla_h \mathcal{I}_{cr} u = \Pi_{h,0} \nabla u$ in combination with Jensen's inequality, the duality relation $I(u) = D(z)$, and an integration by parts using $-\operatorname{div} z = f$ lead to

$$\begin{aligned} \delta_h^2 &\leq \int_\Omega \phi(\nabla u) - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z \cdot \nabla u + \phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x \\ &= \int_\Omega -\phi^*(z) + (z - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \cdot \nabla u + \phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x. \end{aligned}$$

Finally, using convexity of ϕ^*, i.e.,

$$\phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \leq \phi^*(z) - D\phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \cdot (z - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z),$$

and the relation $\nabla u = D\phi^*(z)$ lead to the general error estimate

$$\delta_h^2 \leq \int_\Omega \big(D\phi^*(z) - D\phi^*(\Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \big) \cdot (z - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x.$$

In case of a Lipschitz continuous mapping $D\phi^*$ and a regularity property $z \in W^{1,2}(\Omega; \mathbb{R}^d)$ we directly deduce a linear convergence rate for δ_h. The estimate and conceptual approach apply however to a significantly larger class of variational problems including nonsmooth problems. We remark that the same upper bound is obtained for the error in approximating the dual variable, i.e., for $\sigma_{D_h}^2(z_h, \mathcal{I}_{\mathcal{RT}} z)$. The error estimate can be improved by incorporating strong convexity properties of ϕ^*. For the Poisson problem the derivation then corresponds to the estimates

$$\|\nabla_h (u_h - \Pi_{h,0} \mathcal{I}_{cr} u)\| \leq \|\nabla_h \mathcal{I}_{cr} u - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z\| \leq \|z - \Pi_{h,0} \mathcal{I}_{\mathcal{RT}} z\|,$$

i.e., the discretization error related to the nonconforming discretization with the Crouzeix–Raviart element is controlled by the interpolation error for approximating the flux variable in the Raviart–Thomas finite element space.

By making use of interpolation estimates and the triangle inequality this estimate implies the well known error estimate

$$\|\nabla_h u_h - \nabla u\| \le ch\|D^2 u\|.$$

The derivation given here circumvents the use of a Strang lemma, cf. [19], or the decomposition of functions as in [34], to control nonconformity errors. Another application of duality relations arises in a posteriori error estimates for conforming discretizations [45, 17]. If $u_h^c \in W_D^{1,p}(\Omega)$ is a conforming approximation of the exact solution u then we have, assuming for simplicity that $f = f_h$ so that $I_h = I$ and $D_h = D$ on the discrete spaces, that for all $z_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ with $-\operatorname{div} z_h = f_h$ we have

$$\begin{aligned}\sigma_I^2(u, u_h^c) &\le I(u_h^c) - I(u) \le I(u_h^c) - D(z_h) \\ &= \int_\Omega \phi(\nabla u_h^c)\,\mathrm{d}x - z_h \cdot \nabla u_h^c + \phi^*(z_h)\,\mathrm{d}x =: \frac{1}{2}\eta^2(u_h^c, z_h)\end{aligned}$$

By Fenchel's inequality the integrand on the right-hand side is nonnegative and vanishes if the optimality condition $\nabla u_h^c = D\phi^*(z_h)$ holds which can in general not be satisfied on the discrete level. The optimal choice of z_h solves the discrete dual problem which by the arguments given above is obtained via solving the nonconforming discretization and using the reconstructed flux

$$z_h = D\phi(\nabla_h u_h) - (f_h/d)(\cdot - x_T).$$

For the Poisson problem we deduce the estimate

$$\|\nabla(u_h^c - u)\| \le \eta(u_h^c, z_h) = \|\nabla u_h^c - z_h\|,$$

and with the reconstruction relation $z_h = \nabla_h u_h - (f_h/d)(\cdot - x_T)$ in case that ∇u_h^c is elementwise constant,

$$\|\nabla(u_h^c - u)\| \le \|\nabla u_h^c - \nabla_h u_h\| + \|(f_h/d)(\cdot - x_T)\| = \widetilde{\eta}(u_h^c, u_h).$$

The error estimator $\widetilde{\eta}(u_h^c, u_h)$ is also efficient, which an application of the triangle inequality and the equivalence of the conforming and nonconforming method in case of the Poisson problem show, cf. [18].

1.5. **Outline.** The article is organized as follows. We collect various relevant facts about Crouzeix–Raviart and Raviart–Thomas finite element spaces in Section 2. In Section 3 we present a general theory leading to an error estimate for differentiable convex minimization problems and a general flux reconstruction formula. Nonsmooth problems including a quadratic obstacle problem, a total-variation regularized problem, and an infinity Laplace problem require certain modifications and are discussed in Section 4. In preparation of numerical experiments we devise iterative algorithms for the practical realization in Section 5. The results of various numerical experiments that reveal certain advantages of nonconforming methods are presented in Section 6.

2. Finite element spaces

Throughout what follows we let $(\mathcal{T}_h)_{h>0}$ be a sequence of regular triangulations of the bounded polyhedral Lipschitz domain $\Omega \subset \mathbb{R}^d$ into triangles or tetrahedra for $d = 2$ and $d = 3$, respectively. We let $P_k(T)$ denote the set of polynomials of maximal total degree k on $T \in \mathcal{T}_h$ and define the set of discontinuous, elementwise polynomial functions or vector fields

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

The parameter $h > 0$ refers to the maximal mesh-size of the triangulation $\mathcal{T}_h$. The set of sides of elements is denoted by $\mathcal{S}_h$. We let x_S and x_T denote the midpoints (barycenters) of sides and elements, respectively. The L^2 projection onto piecewise constant functions or vector fields is denoted by

$$\Pi_{h,0} : L^1(\Omega;\mathbb{R}^\ell) \to \mathcal{L}^0(\mathcal{T}_h)^\ell.$$

For an elementwise affine function it corresponds to the evaluation at element midpoints. Standard notation is used for Sobolev spaces, in particular

$$\begin{aligned} W^{1,p}_D(\Omega) &= \{v \in W^{1,p}(\Omega) : v|_{\Gamma_D} = 0\}, \\ W^q_N(\mathrm{div};\Omega) &= \{y \in L^q(\Omega;\mathbb{R}^d) : \operatorname{div} y \in L^q(\Omega),\, y \cdot n = 0 \text{ on } \Gamma_N\}. \end{aligned}$$

We let $BV(\Omega)$ denote space of functions in $L^1(\Omega)$ with finite total variation denoted $|\mathrm{D}u|(\Omega)$. Most estimates derived below follow from the boundedness of the trace operator

$$\mathrm{tr} : W^{1,p}(\Omega;\mathbb{R}^\ell) \to L^p(\partial\Omega;\mathbb{R}^\ell), \quad v \mapsto v|_{\partial\Omega},$$

and the Poincaré inequality

$$\|v - \overline{v}\|_{L^p(\omega)} \le c_{p,\omega} \operatorname{diam}(\omega) \|\nabla v\|_{L^p(\omega)}, \quad \overline{v} = |\omega|^{-1} \int_\omega v \,\mathrm{d}x,$$

for Lipschitz domains $\omega \subset \Omega$, functions $v \in W^{1,p}(\Omega;\mathbb{R}^\ell)$ with mean integral $\overline{v}$ on ω, and $1 \le p \le \infty$. We occasionally make use of indicator functionals, which are for sets $K \subset X$ are defined by

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

for every $s \in X$. For details on the properties of finite element methods listed below we refer the reader to [24, 16, 19, 29, 11].

2.1. Crouzeix–Raviart finite elements.

The Crouzeix–Raviart finite element space of lowest order consists of piecewise affine functions that are continuous at the midpoints of sides of elements, i.e.,

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

The space provides nonconforming approximations of Sobolev spaces $W^{1,p}(\Omega)$. The elementwise application of the gradient operator to a function $v_h \in$

$\mathcal{S}^{1,cr}(\mathcal{T}_h)$ defines an elementwise constant vector field $\nabla_h v_h$ via

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

for all $T \in \mathcal{T}_h$. For weakly differentiable functions $v \in W^{1,p}(\Omega)$ we have $\nabla_h v = \nabla v$. The subset of functions vanishing at midpoints of boundary sides on Γ_D is denoted by

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

We note that the jump of a function $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ over an inner element side $S \in \mathcal{S}_h$ with neighboring elements $T_-, T_+ \in \mathcal{T}_h$, defined by

$$[v_h](x) = v_h|_{T_+}(x) - v_h|_{T_-}(x),$$

has vanishing integral mean over S. Similarly, if $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ then the integral of $v_h|_S$ vanishes on every boundary side $S \in \mathcal{S}_h \cap \Gamma_D$. A basis of the space $\mathcal{S}^{1,cr}(\mathcal{T}_h)$ is given by the functions $\varphi_S \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying

$$\varphi_S(x_{S'}) = \delta_{S,S'}$$

for all $S, S' \in \mathcal{S}_h$. The function φ_S vanishes on elements that do not contain the side S and is continuous with value 1 on S. A quasi-interpolation operator is for $v \in W^{1,p}(\Omega)$ defined via

$$\mathcal{I}_{cr} v = \sum_{S \in \mathcal{S}_h} v_S \varphi_S, \quad v_S = |S|^{-1} \int_S v \,\mathrm{d}s,$$

Since $\mathcal{I}_{cr}$ is bounded and preserves affine functions and averages of gradients, i.e., $\nabla_h \mathcal{I}_{cr} v = \Pi_{h,0} \nabla v$, we have the estimates

$$\|v - \mathcal{I}_{cr} v\|_{L^p(\Omega)} \le c_{cr} h \|\nabla v\|_{L^p(\Omega)}, \quad \|\nabla_h \mathcal{I}_{cr} v\|_{L^p(\Omega)} \le \|\nabla v\|_{L^p(\Omega)}$$

for all $v \in W^{1,p}(\Omega)$, $1 \le p \le \infty$. Moreover, we have $\|\mathcal{I}_{cr} v\|_{L^\infty(\Omega)} \le c_d \|v\|_{L^\infty(\Omega)}$ with $c_d = (d-1)(d+1)$. For $v \in W^{2,p}(\Omega)$ with $1 \le p \le \infty$ we also have the interpolation estimates

$$\|v - \mathcal{I}_{cr} v\|_{L^p(\Omega)} + h \|\nabla_h \mathcal{I}_{cr} v - \nabla v\|_{L^p(\Omega)} \le c'_{cr} h^2 \|D^2 v\|_{L^p(\Omega)}.$$

Finally, we note that there exists a linear enriching operator

$$E^{cr}_h : \mathcal{S}^{1,cr}_D(\mathcal{T}_h) \to W^{1,p}_D(\Omega)$$

such that

$$\|\nabla E^{cr}_h v_h\|_{L^p(\Omega)} + h^{-1} \|E^{cr}_h v_h - v_h\|_{L^p(\Omega)} \le c_E \|\nabla_h v_h\|_{L^p(\Omega)}$$

for $1 \le p < \infty$, cf. [18] in case $p = 2$ and Appendix A.1 for $p \ne 2$.

2.2. **Raviart–Thomas finite elements.** The lowest order Raviart–Thomas finite element space is defined as

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

Vector fields in $\mathcal{RT}^0(\mathcal{T}_h)$ have continuous constant normal components on element sides. The subset of vector fields with vanishing normal component on the Neumann boundary Γ_N is defined as

$$\mathcal{RT}^0_N(\mathcal{T}_h) = \{y_h \in \mathcal{RT}^0(\mathcal{T}_h) : y_h \cdot n = 0 \text{ on } \Gamma_N\},$$

where n denotes the outer unit normal on $\partial\Omega$. A basis of the space $\mathcal{RT}^0(\mathcal{T}_h)$ is given by vector fields ψ_S, $S \in \mathcal{S}_h$, supported on adjacent elements with

$$\psi_S(x) = \pm \frac{|S|}{d!|T_\pm|}(z_{S,T_\pm} - x) \tag{3}$$

for $x \in T_\pm$ with opposite vertex $z_{S,T_\pm}$ to $S \subset \partial T_\pm$. We have that $\psi_S|_{S'} \cdot n_{S'} = 0$ for all sides $S' \neq S$ with unit normal vector $n_{S'}$. If n_S is the unit normal vector on S and points from T_- into T_+ then we have $\psi_S|_S \cdot n_S = 1$. A quasi-interpolation operator is for vector fields $z \in W^{1,1}(\Omega;\mathbb{R}^d)$ given by

$$\mathcal{I}_{\mathcal{RT}} z = \sum_{S \in \mathcal{S}_h} z_S \psi_S, \quad z_S = |S|^{-1} \int_S z \cdot n_S \,\mathrm{d}s.$$

The operator $\mathcal{I}_{\mathcal{RT}}$ is bounded on $C^0(\overline{\Omega};\mathbb{R}^d)$ and we have

$$\|z - \mathcal{I}_{\mathcal{RT}} z\|_{L^p(\Omega)} \le c_{\mathcal{RT}} h \|\nabla z\|_{L^p(\Omega)}$$

and $\operatorname{div} \mathcal{I}_{\mathcal{RT}} z = \Pi_{h,0} \operatorname{div} z$ for all $z \in W^{1,p}(\Omega;\mathbb{R}^d)$. The latter property implies that the divergence operator defines a surjection from $\mathcal{RT}^0_N(\mathcal{T}_h)$ into $\mathcal{L}^0(\mathcal{T}_h)$, provided that constants are eliminated from $\mathcal{L}^0(\mathcal{T}_h)$ if $\Gamma_D = \emptyset$.

2.3. **Orthogonality relations.** An elementwise integration by parts implies that for $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}^0(\mathcal{T}_h)$ we have the integration-by-parts formula

$$\int_\Omega \nabla_h v_h \cdot y_h \,\mathrm{d}x + \int_\Omega v_h \operatorname{div} y_h \,\mathrm{d}x = \int_{\partial\Omega} v_h\, y_h \cdot n \,\mathrm{d}s. \tag{4}$$

Here we used that y_h has continuous constant normal components on inner element sides and that jumps of v_h have vanishing integral mean. If an elementwise constant vector field $w_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ satisfies

$$\int_\Omega w_h \cdot \nabla_h v_h \,\mathrm{d}x = 0$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ then its normal components are continuous on inner element sides and vanish on the Γ_N, so that it belongs to $\mathcal{RT}^0_N(\mathcal{T}_h)$. The following elementary identity is used repeatedly.

Lemma 2.1 (Exchange of projections). *For $z \in W^p_N(\mathrm{div};\Omega) \cap W^{1,1}(\Omega;\mathbb{R}^d)$ and $u \in W^{1,p}_D(\Omega)$ and their interpolants $\widetilde{z} = \mathcal{I}_{\mathcal{RT}} z \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and $\widetilde{u}_h = \mathcal{I}_{cr} u \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ we have*

$$\int_\Omega \operatorname{div} z (u - \Pi_{h,0}\widetilde{u}_h)\,\mathrm{d}x + \int_\Omega \nabla u \cdot (z - \Pi_{h,0}\widetilde{z}_h)\,\mathrm{d}x = 0.$$

Proof. Since $\operatorname{div} \widetilde{z}_h = \Pi_{h,0} \operatorname{div} z$ and $\nabla_h \widetilde{u}_h = \Pi_{h,0} \nabla u$, we verify that

$$\begin{aligned}\int_\Omega \operatorname{div} z (u - \Pi_{h,0}\widetilde{u}_h)\,\mathrm{d}x &= -\int_\Omega z \cdot \nabla u + \operatorname{div} \widetilde{z}_h \widetilde{u}_h \,\mathrm{d}x \\ &= -\int_\Omega \nabla u \cdot (z - \Pi_{h,0}\widetilde{z}_h)\,\mathrm{d}x,\end{aligned}$$

which proves the asserted equality. □

2.4. **Convex conjugates.** Given a proper, convex, and lower semicontinuous functional $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ the convex conjugate $\phi^* : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is defined via

$$\phi^*(t) = \sup_{s \in \mathbb{R}^d} t \cdot s - \phi(s).$$

The function ϕ^* is proper, convex, and lower semicontinuous and we have the relations

$$\phi^{**} = \phi, \quad s = D\phi^*\big(D\phi(s)\big),$$

where the second identy can be generalized to subdifferentials. We refer the reader to [46] for details and note the Fenchel–Young inequality which states that for $s, t \in \mathbb{R}^d$ we have

$$t \cdot s \le \phi(s) + \phi^*(t)$$

with equality if and only if $t = D\phi(s)$. Certain duality relations can be transferred to discretizations of variational problems. We provide a modified version and a different proof of an important formula identified in [23].

Proposition 2.2 (Discrete duality). *Given $\overline{u}_h \in \Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \subset \mathcal{L}^0(\mathcal{T}_h)$ we have*

$$\begin{aligned}&\inf\Big\{ \int_\Omega \phi\big(\nabla_h u_h\big)\,\mathrm{d}x : u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h),\, \Pi_{h,0} u_h = \overline{u}_h \Big\} \\ &\qquad \ge \sup\Big\{ -\int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) + \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x : z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)\Big\}.\end{aligned}$$

If $\phi \in C^1(\mathbb{R}^d)$ then equality holds.

Proof. We let $L(\overline{u}_h)$ and $R(\overline{u}_h)$ denote the terms on the left- and right-hand side of the asserted inequality and show that $R(\overline{u}_h) \le L(\overline{u}_h)$. For this, let $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ with $\Pi_{h,0} u_h = \overline{u}_h$. Given any $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ we then have that

$$-\int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) + \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x = -\int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) - \nabla_h u_h \cdot \Pi_{h,0} z_h \,\mathrm{d}x.$$

Hence, only the midpoint values of z_h matter and the supremum is larger if it is taken over elementwise constant vector fields $p_h \in \mathcal{L}^0(\mathcal{T}_h)^d$. This corresponds to computing elementwise the values $\phi(\nabla_h u_h) = \phi^{**}(\nabla_h u_h)$. Since u_h is arbitrary with $\Pi_{h,0} u_h = \overline{u}_h$ we deduce that $R(\overline{u}_h) \le L(\overline{u}_h)$. If $\phi \in C^1(\mathbb{R}^d)$ then an optimal $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ for $L(\overline{u}_h)$ satisfies

$$\int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h v_h \, \mathrm{d}x + \int_\Omega \mu_h \Pi_{h,0} v_h \, \mathrm{d}x = 0,$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, where $\mu_h \in \Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \subset \mathcal{L}^0(\mathcal{T}_h)$ is a Lagrange multiplier related to the constraint $\Pi_{h,0} u_h = \overline{u}_h$. For $T \in \mathcal{T}_h$ and $x \in T$ we define

$$z_h(x) = D\phi(\nabla_h u_h|_T) + \frac{\mu_h|_T}{d}(x - x_T)$$

and note that $\operatorname{div} z_h|_T = \mu_h|_T$. We choose an arbitrary element $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ with $\operatorname{div} \widetilde{z}_h = \mu_h$ and verify that the elementwise constant vector field $z_h - \widetilde{z}_h$ satisfies

$$\int_\Omega (z_h - \widetilde{z}_h) \cdot \nabla_h v_h \, \mathrm{d}x = \int_\Omega \big(D\phi(\nabla u_h) - \widetilde{z}_h\big) \cdot \nabla_h v_h \, \mathrm{d}x = 0$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, i.e., $z_h - \widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and in particular $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. The identity $\Pi_{h,0} z_h = D\phi(\nabla_h u_h)$ implies that

$$\phi^*(\Pi_{h,0} z_h) = \Pi_{h,0} z_h \cdot \nabla_h u_h - \phi(\nabla_h u_h).$$

An integration over Ω and the integration-by-parts formula (4) lead to

$$\int_\Omega \phi(\nabla_h u_h) \, \mathrm{d}x = -\int_\Omega \phi^*(\Pi_{h,0} z_h) + \overline{u}_h \operatorname{div} z_h \, \mathrm{d}x,$$

which implies that $L(\overline{u}_h) = R(\overline{u}_h)$. □

Remark 2.3. *The condition $\phi \in C^1(\mathbb{R}^d)$ can be avoided provided there exists a sequence of regularizations ϕ_ε of ϕ such that ϕ_ε and ϕ^*_ε converge uniformly to ϕ and ϕ^* on their domains. This applies, e.g., to the truncated regularization $\phi_\varepsilon(s) = \min\{|s| - \varepsilon/2, |s|^2/(2\varepsilon)\}$ of the modulus for which we have $\phi^*_\varepsilon(t) = I_{K_1(0)}(t) + t^2/(2\varepsilon)$, where $K_1(0) = \{t \in \mathbb{R}^d : |t| \le 1\}$.*

3. General results

We consider the minimization of the abstract functional

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

in a Sobolev space $W^{1,p}_D(\Omega)$ for $1 < p < \infty$ and $f \in L^{p'}(\Omega)$. We assume that the convex and measurable integrands

$$\phi : \mathbb{R}^d \to \mathbb{R}, \quad \psi : \Omega \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$$

are such that I is bounded from below, coercive, not identical to $+\infty$, and weakly lower semicontinuous so that the direct method in the calculus of variations implies the existence of a solution $u \in W^{1,p}_D(\Omega)$. The dual problem consists in maximizing the functional

$$D(z) = -\int_\Omega \phi^*(z)\,\mathrm{d}x - \int_\Omega \psi^*(x, \operatorname{div} z)\,\mathrm{d}x$$

in the space $W^{p'}_N(\Omega; \operatorname{div})$ with $p' = p/(p-1)$ and we assume that a solution exists. We also assume the strong duality relation

$$\inf_{u \in W^{1,p}_D(\Omega)} I(u) = \sup_{z \in W^{p'}_N(\Omega;\operatorname{div})} D(z)$$

to hold and refer the reader to, e.g., [5, 46], for conditions leading to this equality. We recall that in this case we have the relations

$$z = D\phi(\nabla u), \quad \operatorname{div} z = D\psi(u)$$

for solutions u and z, where $D\psi$ stands for the derivative of ψ with respect to the second argument. The derivatives can be replaced by subdifferentials.

3.1. **Discrete duality.** The discrete primal problem is defined by minimizing the functional

$$I_h(u_h) = \int_\Omega \phi(\nabla_h u_h)\,\mathrm{d}x + \int_\Omega \psi_h(x, \Pi_{h,0} u_h)\,\mathrm{d}x$$

in the nonconforming finite element space $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ with suitable convex approximations ψ_h of ψ that are elementwise constant with respect to the first argument. The corresponding discrete dual problem consists in maximizing the functional

$$D_h(z_h) = -\int_\Omega \phi^*(\Pi_{h,0} z_h)\,\mathrm{d}x - \int_\Omega \psi_h^*(x, \operatorname{div} z_h)\,\mathrm{d}x$$

in the set $\mathcal{RT}^0_N(\mathcal{T}_h)$.

Proposition 3.1 (Duality relations)**.** *The discrete primal and dual problems satisfy the duality relation*

$$\inf_{u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)} I_h(u_h) \ge \sup_{z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)} D_h(z_h).$$

If ϕ and ψ are differentiable then solutions u_h and z_h are related via

$$z_h(x) = D\phi\big(\nabla_h u_h|_T\big) + d^{-1} D\psi_h\big(x, u_h(x_T)\big)(x - x_T)$$

for every $T \in \mathcal{T}_h$ and $x \in T$. The pair $(z_h, \overline{u}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ with $\overline{u}_h|_T = u_h(x_T)$ for all $T \in \mathcal{T}_h$ solves the corresponding saddle-point problem

$$\begin{aligned} \big(D\phi^*(\Pi_{h,0} z_h), \Pi_{h,0} y_h\big) + (\overline{u}_h, \operatorname{div} y_h) &= 0, \\ (\operatorname{div} z_h, \overline{v}_h) - \big(D\psi_h(\overline{u}_h), \overline{v}_h\big) &= 0, \end{aligned}$$

for all $(y_h, v_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$. Moreover, in this case strong duality applies, i.e., $I_h(u_h) = D_h(z_h)$.

Proof. We use the duality formula of Proposition 2.2 and exchange extrema, to verify that, indicating by u_h, z_h arbitrary functions from the spaces $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $\mathcal{RT}^0_N(\mathcal{T}_h)$, and abbrevating $\mathcal{P}_h = \Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \subset \mathcal{L}^0(\mathcal{T}_h)$,

$$\begin{aligned}
\inf_{u_h} I_h(u_h) &= \inf_{\overline{u}_h \in \mathcal{P}_h} \inf_{u_h :\, \Pi_{h,0} u_h = \overline{u}_h} \int_\Omega \phi(\nabla_h u_h) \,\mathrm{d}x + \int_\Omega \psi_h(\Pi_{h,0} u_h) \,\mathrm{d}x \\
&\geq \inf_{\overline{u}_h \in \mathcal{P}_h} \sup_{z_h} - \int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) + \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x + \int_\Omega \psi_h(\overline{u}_h) \,\mathrm{d}x \\
&\geq \inf_{\overline{u}_h \in \mathcal{L}^0(\mathcal{T}_h)} \sup_{z_h} - \int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) + \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x + \int_\Omega \psi_h(\overline{u}_h) \,\mathrm{d}x \\
&\geq \sup_{z_h} \inf_{\overline{u}_h \in \mathcal{L}^0(\mathcal{T}_h)} - \int_\Omega \phi^*\big(\Pi_{h,0} z_h\big) + \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x + \int_\Omega \psi_h(\overline{u}_h) \,\mathrm{d}x.
\end{aligned}$$

The infimum is eliminated by using the convex conjugate of ψ_h, i.e., by noting that

$$\psi_h^*(x,t) = \sup_{s\in\mathbb{R}} t\, s - \psi_h(x,s) = - \inf_{s\in\mathbb{R}} -t\, s + \psi_h(x,s),$$

we find that

$$\inf_{\overline{u}_h \in \mathcal{L}^0(\mathcal{T}_h)} - \int_\Omega \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x + \int_\Omega \psi_h(\overline{u}_h) \,\mathrm{d}x = - \int_\Omega \psi_h^*(\operatorname{div} z_h) \,\mathrm{d}x.$$

This implies that we have

$$\inf_{u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)} I_h(u_h) \geq \sup_{z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)} D_h(z_h).$$

Assume that ϕ and ψ are differentiable, let u_h be a solution of the primal problem, and let z_h be defined as in the proposition. Furthermore, let $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ be such that $\operatorname{div} \widetilde{z}_h = D\psi_h(\Pi_{h,0} u_h)$. Since $\operatorname{div} z_h|_T = D\psi_h(u_h(x_T))$ for all $T \in \mathcal{T}_h$ it follows that $z_h - \widetilde{z}_h \in \mathcal{L}^0(\mathcal{T}_h)^d$. Using the discrete Euler–Lagrange equations

$$\int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h v_h \,\mathrm{d}x + \int_\Omega D\psi_h(\Pi_{h,0} u_h) v_h \,\mathrm{d}x = 0$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, we find that

$$\int_\Omega (z_h - \widetilde{z}_h) \cdot \nabla_h v_h \,\mathrm{d}x = \int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h v_h \,\mathrm{d}x + \int_\Omega \operatorname{div} \widetilde{z}_h v_h \,\mathrm{d}x = 0.$$

Hence $z_h - \widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and in particular $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Since $\Pi_{h,0} z_h = D\phi(\nabla_h u_h)$ and $\operatorname{div} z_h = D\psi_h(\Pi_{h,0} u_h)$ we verify that

$$\begin{aligned}
\phi^*(\Pi_{h,0} z_h) &= \Pi_{h,0} z_h \cdot \nabla_h u_h - \phi(\nabla_h u_h), \\
\psi_h^*(\operatorname{div} z_h) &= \operatorname{div} z_h \Pi_{h,0} u_h - \psi_h(\Pi_{h,0} u_h).
\end{aligned}$$

Adding these identities and incorporating the integration-by-parts formula (4) implies that

$$- \int_\Omega \phi^*(\Pi_{h,0} z_h) \,\mathrm{d}x - \int_\Omega \psi_h^*(\operatorname{div} z_h) \,\mathrm{d}x = \int_\Omega \phi(\nabla_h u_h) \,\mathrm{d}x + \int_\Omega \psi_h(\Pi_{h,0} u_h) \,\mathrm{d}x,$$

i.e., that $I_h(u_h) = D_h(z_h)$. Noting that $D\phi^*(\Pi_{h,0} z_h) = \nabla_h u_h$ we find that the pair $(z_h, \overline{u}_h)$ solves the saddle-point problem. □

Remark 3.2. *In general, the inclusion* $\Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \subset \mathcal{L}^0(\mathcal{T}_h)$ *is strict, e.g., for* $\mathcal{T}_h = \{T\}$, $\overline{\Omega} = T$, *and* $\Gamma_D = \partial\Omega$. *The implicit treatment of Dirichlet boundary conditions in the dual formulation implies that strong duality still applies.*

3.2. **Γ-convergence of** I_h**.** A general justification of the discrete problems I_h as correct discretizations of the functional I is established via a Γ-convergence result. For this, we extend the functionals I and I_h to functionals $\widetilde{I}$ and $\widetilde{I}_h$ on $L^p(\Omega)$ by formally assigning the value $+\infty$ to arguments not belonging to $W^{1,p}_D(\Omega)$ and $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, respectively.

Proposition 3.3 (Γ-convergence)**.** *Assume that* $\phi : \mathbb{R}^d \to \mathbb{R}$ *satisfies*

$$\big|\phi(s) - \phi(r)\big| \le c_5\big(1 + |r| + |s|\big)^{p-1}|r - s|$$

for all $s, r \in \mathbb{R}^d$ *and that* $\psi_h : \Omega\times\mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ *and* $\psi : \Omega\times\mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ *are related via*

$$\psi_h(\cdot, v_h) \to \psi(\cdot, v)$$

in $L^1(\Omega)$ *as* $h \to 0$ *whenever* $\psi_h(\cdot, v_h) \in L^1(\Omega)$ *for all* $h > 0$ *and* $v_h \to v$ *in* $L^p(\Omega)$. *Then, the extended functionals* $\widetilde{I}_h : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ *are* Γ*-convergent to* $\widetilde{I} : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ *as* $h \to 0$ *with respect to strong convergence in* $L^p(\Omega)$.

Proof. Let $(u_h)_{h>0}$ be such that $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $\liminf_{h\to 0} \widetilde{I}_h(u_h) < \infty$. The assumed coercivity property implies that for a subsequence we have $\|\nabla_h u_h\|_{L^p(\Omega)} \le c$. Incorporating the enriching operator E^{cr}_h shows that there exists $u \in W^{1,p}_D(\Omega)$ such that $u_h \to u$ in $L^p(\Omega)$ and $\nabla_h u_h \rightharpoonup \nabla u$ for an appropriate subsequence as $h \to 0$. Convexity of ϕ and the assumption on ψ then imply that

$$\liminf_{h\to 0} I_h(u_h) \le I(u).$$

Given $u \in W^{1,p}_D(\Omega)$ we use regularizations of u and the interpolation operator $\mathcal{I}_{cr}$ to construct a seqence $(u_h)_{h>0}$ with $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ such that

$$\|u - u_h\|_{L^p(\Omega)} + \|\nabla_h u_h - \nabla u\|_{L^p(\Omega)} \to 0,$$

as $h \to 0$. Noting that $\Pi_{h,0} u_h \to u$ in $L^p(\Omega)$ and using the assumed local Lipschitz continuity of ϕ and the approximability condition on ψ_h and ψ we deduce that

$$\begin{aligned}|I_h(u_h) - I(u)| \le c_5 \int_\Omega \big(1 + |\nabla u| + |\nabla_h u_h|\big)^{p-1} |\nabla u - \nabla_h u_h| \,\mathrm{d}x \\ + \|\psi(u) - \psi_h(\Pi_{h,0} u_h)\|_{L^1(\Omega)}.\end{aligned}$$

This implies that $I_h(u_h) \to I(u)$ as $h \to 0$. □

3.3. **Error estimate.** We next derive an abstract error estimate for the approximation of I with the nonconforming discretization I_h. We assume that the functionals I_h provide a uniform strong coercivity property, i.e., with the variational derivative δI_h, that

$$I_h(v_h) + \delta I_h(v_h)[w_h - v_h] + \sigma^2_{I_h}(v_h, w_h) \le I_h(w_h)$$

for all $v_h, w_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ with a definite functional σ_{I_h}. This implies that minimizers for I_h are unique.

Theorem 3.4 (Discretization error)**.** *For minimizers $u \in W^{1,p}_D(\Omega)$ and $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ of the functionals I and I_h and a dual solution $z \in W^1_N(\mathrm{div};\Omega) \cap W^{1,1}(\Omega;\mathbb{R}^d)$ we have that*

$$\begin{aligned}\sigma^2_{I_h}(u_h, \mathcal{I}_{cr} u) \le &\int_\Omega \big(D\phi^*(z) - D\phi^*(\Pi_{h,0}\mathcal{I}_{\mathcal{RT}} z)\big) \cdot (z - \Pi_{h,0}\mathcal{I}_{\mathcal{RT}} z) \,\mathrm{d}x \\ &+ \int_\Omega \big(D\psi(u) - D\psi_h(\Pi_{h,0}\mathcal{I}_{cr} u)\big) \cdot (u - \Pi_{h,0}\mathcal{I}_{cr} u)\,\mathrm{d}x \\ &+ \int_\Omega \psi_h(u) - \psi(u)\,\mathrm{d}x + \int_\Omega \psi_h^*(\Pi_{h,0}\operatorname{div} z) - \psi^*(\operatorname{div} z)\,\mathrm{d}x.\end{aligned}$$

Proof. The interpolants $\mathcal{I}_{\mathcal{RT}} z$ and $\mathcal{I}_{cr} u$ are well defined and we abbreviate

$$\widetilde{z}_h = \mathcal{I}_{\mathcal{RT}} z, \quad \widetilde{u}_h = \mathcal{I}_{cr} u.$$

By minimality of u_h and the duality relation $I_h(u_h) \ge D_h(\mathcal{I}_{\mathcal{RT}} z)$, we have

$$\begin{aligned}\sigma^2_{I_h}(u_h, \widetilde{u}_h) &\le I_h(\widetilde{u}_h) - I_h(u_h) \le I_h(\widetilde{u}_h) - D_h(\widetilde{z}_h) \\ &= \int_\Omega \phi(\nabla_h \widetilde{u}_h) + \psi_h(\Pi_{h,0}\widetilde{u}_h) + \phi^*(\Pi_{h,0}\widetilde{z}_h) + \psi_h^*(\operatorname{div}\widetilde{z}_h)\,\mathrm{d}x.\end{aligned}$$

The identity $\Pi_{h,0}\nabla u = \nabla_h \widetilde{u}_h$ in combination with convexity of ϕ and Jensen's inequality on every element leads to

$$\begin{aligned}\sigma^2_{I_h}(u_h, \widetilde{u}_h) &\le \int_\Omega \phi(\nabla u) + \psi_h(\Pi_{h,0}\widetilde{u}_h) + \phi^*(\Pi_{h,0}\widetilde{z}_h) + \psi_h^*(\operatorname{div}\widetilde{z}_h)\,\mathrm{d}x \\ &= \int_\Omega \phi(\nabla u) + \psi_h(\Pi_{h,0}\widetilde{u}_h) + \phi^*(\Pi_{h,0}\widetilde{z}_h) + \psi^*(\operatorname{div} z)\,\mathrm{d}x \\ &\quad + \int_\Omega \psi_h^*(\operatorname{div}\widetilde{z}_h) - \psi^*(\operatorname{div} z)\,\mathrm{d}x.\end{aligned}$$

The duality relation $I(u) = D(z)$ allows us to replace the sum of the first and the last term in the first integral and implies that we have

$$\begin{aligned}\sigma^2_{I_h}(u_h, \widetilde{u}_h) \le &\int_\Omega -\psi(u) + \psi_h(\Pi_{h,0}\widetilde{u}_h) + \phi^*(\Pi_{h,0}\widetilde{z}_h) - \phi^*(z)\,\mathrm{d}x \\ &+ \int_\Omega \psi_h^*(\operatorname{div}\widetilde{z}_h) - \psi^*(\operatorname{div} z)\,\mathrm{d}x.\end{aligned}$$

We next use convexity of ϕ^* and ψ_h at $\Pi_{h,0}\widetilde{z}_h$ and $\Pi_{h,0}\widetilde{u}_h$, respectively, to deduce that

$$\sigma^2_{I_h}(u_h,\widetilde{u}_h) \le -\int_\Omega D\psi_h(\Pi_{h,0}\widetilde{u}_h)\cdot(u-\Pi_{h,0}\widetilde{u}_h)\,\mathrm{d}x - \int_\Omega \psi(u)-\psi_h(u)\,\mathrm{d}x$$
$$-\int_\Omega D\phi^*(\Pi_{h,0}\widetilde{z}_h)\cdot(z-\Pi_{h,0}\widetilde{z}_h)\,\mathrm{d}x + \int_\Omega \psi_h^*(\operatorname{div}\widetilde{z}_h)-\psi^*(\operatorname{div} z)\,\mathrm{d}x.$$

Using the relations $D\phi^*(z)=\nabla u$ and $\operatorname{div} z = D\psi(u)$ in Lemma 2.1 implies that

$$\int_\Omega D\phi^*(z)\cdot(z-\Pi_{h,0}\widetilde{z}_h)\,\mathrm{d}x = -\int_\Omega (u-\Pi_{h,0}\widetilde{u}_h)D\psi(u)\,\mathrm{d}x.$$

In combination with the previous estimate we deduce the error bound. □

Remark 3.5. *If ψ is independent of $x\in\Omega$ and $\psi_h=\psi$ then the last two integrals on the right-hand side of the error estimate are nonpositive by Jensen's inequality.*

3.4. **Examples.** The abstract error estimates applies to various partial differential equations. For linear and quadratic low order terms ψ the corresponding error terms simplify, provided the approximations ψ_h are suitably chosen.

Proposition 3.6 (Low order terms)**.** *(i) Assume that for $f\in L^q(\Omega)$ and $f_h=\Pi_{h,0}f$ we have*

$$\psi(x,s)=-f(x)s,\quad \psi_h(x,s)=-f_h(x)s.$$

Then the error estimate of Theorem 3.4 reduces to

$$\sigma^2_{I_h}(\widetilde{u}_h,u_h)\le \int_\Omega \big(D\phi^*(z)-D\phi^*(\Pi_{h,0}\mathcal{I}_{\mathcal{RT}}z)\big)\cdot(z-\Pi_{h,0}\mathcal{I}_{\mathcal{RT}}z)\,\mathrm{d}x.$$

(ii) Assume that for $g\in L^2(\Omega)$ and $g_h=\Pi_{h,0}g$ we have

$$\psi(x,s)=(g(x)-s)^2/2,\quad \psi_h(x,s)=(g_h(x)-s)^2/2.$$

Then the error estimate of Theorem 3.4 reduces to

$$\sigma^2_{I_h}(\widetilde{u}_h,u_h)\le \int_\Omega \big(D\phi^*(z)-D\phi^*(\Pi_{h,0}\mathcal{I}_{\mathcal{RT}}z)\big)\cdot(z-\Pi_{h,0}\mathcal{I}_{\mathcal{RT}}z)\,\mathrm{d}x$$
$$+\|u-\Pi_{h,0}\mathcal{I}_{cr}u\|^2.$$

Proof. As above we abbreviate $\widetilde{u}_h=\mathcal{I}_{cr}u$ and $\widetilde{z}_h=\mathcal{I}_{\mathcal{RT}}z$. In the first case we have

$$\psi^*(x,t)=I_{\{-f(x)\}}(t),\quad \psi_h^*(x,t)=I_{\{-f_h(x)\}}(t).$$

Hence, the last three integrals in the error estimate of Theorem 3.4 become

$$E_\psi = -\int_\Omega (f-f_h)(u-\Pi_{h,0}\widetilde{u}_h)\,\mathrm{d}x - \int_\Omega f_h u - f u\,\mathrm{d}x$$
$$+\int_\Omega I_{\{-f_h\}}(\operatorname{div}\widetilde{z}_h)-I_{\{-f\}}(\operatorname{div} z)\,\mathrm{d}x,$$

so that $E_\psi = 0$ since $f - f_h$ is orthogonal to $\Pi_{h,0}\widetilde{u}_h$ and $\operatorname{div} z = -f$ and $\operatorname{div}\widetilde{z}_h = -f_h$. In the second case we have

$$\psi^*(x,t) = \frac{1}{2}(t+g(x))^2 - \frac{1}{2}g(x)^2, \quad \psi_h^*(x,t) = \frac{1}{2}(t+g_h(x))^2 - \frac{1}{2}g_h(x)^2.$$

The corresponding error terms are given by

$$\begin{aligned} E_\psi = &\int_\Omega \big((u-g) - (\Pi_{h,0}\widetilde{u}_h - g_h)\big)(u - \Pi_{h,0}\widetilde{u}_h)\,\mathrm{d}x + \frac{1}{2}\int_\Omega (g_h - u)^2 - (g-u)^2\,\mathrm{d}x \\ &+ \frac{1}{2}\int_\Omega (\operatorname{div}\widetilde{z}_h + g_h)^2 - g_h^2 - (\operatorname{div} z + g)^2 + g^2\,\mathrm{d}x. \end{aligned}$$

The relation $\operatorname{div}\widetilde{z}_h + g_h = \Pi_{h,0}(\operatorname{div} z + g)$ in combination with Jensen's inequality and elementary calculations imply that

$$E_\psi \le \int_\Omega (u - \Pi_{h,0}\widetilde{u}_h)^2\,\mathrm{d}x.$$

This proves the simplified error estimate. □

Remark 3.7. *Using the strong convexity of ψ in case (ii) of Proposition 3.6 the factor* 1 *in front of the term* $\|u - \Pi_{h,0}\mathcal{I}_{cr}u\|^2$ *can be replaced by 1/2.*

Typical choices for the function ϕ correspond to p-Laplace equations.

Example 3.8 (p-Dirichlet problems)**.** *For $1 < p < \infty$ let $\phi(s) = |s|^p/p$ and $\psi(x,s) = -f(x)s$ for $f \in L^q(\Omega)$ with $q = p' = p/(p-1)$. Noting that $\phi^*(t) = |t|^q/q$ we define*

$$F(a) = |a|^{(p-2)/2}a, \quad \widetilde{S}(v) = D\phi^*(v) = |v|^{q-2}v, \quad \widetilde{F}(v) = |v|^{(q-2)/2}v.$$

We abbreviate $\widetilde{z}_h = \mathcal{I}_{\mathcal{R}T}z$ and use inequalities from [26] *which are explained in Appendix A.3 to verify that the error estimate of Theorem 3.4 becomes*

$$\begin{aligned} c_p\big\|F(\nabla_h\mathcal{I}_{cr}u) - F(\nabla_h u_h)\big\|^2 &\le \int_\Omega \big(\widetilde{S}(z) - \widetilde{S}(\Pi_{h,0}\widetilde{z}_h)\big)\cdot(z - \Pi_{h,0}\widetilde{z}_h)\,\mathrm{d}x \\ &\le c_p'\|\widetilde{F}(z) - \widetilde{F}(\Pi_{h,0}\widetilde{z}_h)\|^2. \end{aligned}$$

The right-hand side can be bounded using techniques from [28] *provided $\widetilde{F}(z) \in W^{1,2}(\Omega;\mathbb{R}^d)$. The results provided there also imply that $\|F(\nabla_h\mathcal{I}_{cr}u) - F(\nabla u)\| \le ch\|\nabla F(\nabla u)\|$. The estimate confirms error estimates from* [6, 39, 28]*.*

4. Nonsmooth problems

We discuss in this section necessary adjustments of the general theory to apply it to nondifferentiable problems, where, e.g., well-posedness and admissibility of modified interpolants has to be ensured.

4.1. **Obstacle problem.** We consider a prototypical obstacle problem defined by minimizing

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x + I_+(u)$$

in the set $W^{1,2}_D(\Omega)$, where I_+ is the indicator functional of functions that are nonnegative almost everywhere. With the

$$\psi(x,s) = -f(x)u + I_+(s)$$

we have

$$\psi^*(x,t) = I_-(t + f(x)).$$

The dual problem thus determines a maximizing $z \in L^2(\Omega;\mathbb{R}^d)$ for

$$D(z) = -\frac{1}{2}\int_\Omega |z|^2 \,\mathrm{d}x - I_-(\operatorname{div} z + f),$$

where the indicator functional I_- is finite if $\operatorname{div} z + f$ is nonpositive as a functional on $W^{1,2}_D(\Omega)$. We have $z = \nabla u$ and a complementarity principle implies that $\operatorname{div} z + f = 0$ whenever $u > 0$. We remark that general obstacles $\chi \in H^1_D(\Omega)$ can be treated via a substitution $u = \widetilde{u} + \chi$ which leads to a modified function f provided that $\Delta\chi \in L^2(\Omega)$.

Discretization. The discrete primal problem imposes the obstacle constraint at midpoints of elements, i.e., we consider

$$I_h(u_h) = \frac{1}{2}\int_\Omega |\nabla_h u_h|^2 \,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x + I_+(\Pi_{h,0} u_h),$$

where $f_h = \Pi_{h,0} f$. Proposition 3.1 shows that the discrete dual problem consists in determining a maximizing vector field $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ for

$$D_h(z_h) = -\frac{1}{2}\int_\Omega |\Pi_{h,0} z_h|^2 \,\mathrm{d}x - I_-(\operatorname{div} z_h + f_h).$$

Adopting the ideas of the general error analysis leads to a quasi-optimal error estimate. We note that imposing the obstacle condition at midpoints of elements instead of midpoints of element sides as in the two-dimensional setting considered in [21] simplifies the error analysis.

Proposition 4.1 (Error estimate)**.** *Let $u \in H^1_D(\Omega)$ and $u_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ are the solutions of the primal and discrete primal problem, respectively. If the solution $z \in L^2(\Omega;\mathbb{R}^d)$ of the dual problem satisfies $z \in H^1(\Omega;\mathbb{R}^d)$ then we have*

$$\|\nabla_h(u_h - u)\| \le ch\big(\|D^2 u\| + \|f + \operatorname{div} z\|\big).$$

Proof. Throughout this proof we abbreviate $\widetilde{u}_h = \mathcal{I}_{cr} u$ and $\widetilde{z}_h = \mathcal{I}_{\mathcal{RT}} z$. Minimality of u_h, strong convexity of I_h, and discrete duality imply that

$$\delta_h^2 = \frac{1}{2}\|\nabla_h(u_h - \widetilde{u}_h)\|^2 \le I_h(\widetilde{u}_h) - D_h(\widetilde{z}_h).$$

The relation $\nabla\widetilde{u}_h = \Pi_{h,0}\nabla u$ in combination with Jensen's inequality and the identity $I(u) = D(z)$ show that

$$\delta_h^2 \le -\frac{1}{2}\int_\Omega |z|^2 \,\mathrm{d}x + \int_\Omega fu - f_h\Pi_{h,0}\widetilde{u}_h \,\mathrm{d}x + \frac{1}{2}\int_\Omega |\Pi_{h,0}\widetilde{z}_h|^2 \,\mathrm{d}x.$$

Using Lemma 2.1 with $\nabla u = z$ and noting $f_h = \Pi_{h,0} f$ leads to

$$\delta_h^2 \le \int_\Omega (f + \operatorname{div} z)(u - \Pi_{h,0}\widetilde{u}_h) \,\mathrm{d}x + \frac{1}{2}\int_\Omega |z - \Pi_{h,0}\widetilde{z}_h|^2 \,\mathrm{d}x.$$

We abbreviate $\mu = f + \operatorname{div} z \in L^2(\Omega)$ and insert $\widetilde{u}_h = \mathcal{I}_{cr} u$ to rewrite the first term on the right-hand side as

$$\int_\Omega \mu(u - \Pi_{h,0}\widetilde{u}_h) \,\mathrm{d}x = \int_\Omega \mu(u - \widetilde{u}_h) \,\mathrm{d}x + \int_\Omega \mu(\widetilde{u}_h - \Pi_{h,0}\widetilde{u}_h) \,\mathrm{d}x.$$

To deduce the error estimate it remains to bound the second term on the right-hand side. For $T \in \mathcal{T}_h$ let $\mathcal{C}_T = \{x \in T : u(x) = 0\}$ and note that $\lambda|_{T\setminus\mathcal{C}_T} = 0$. Since $\nabla u = 0$ almost everywhere on $\mathcal{C}_T$ and since $\Pi_{h,0}\widetilde{u}_h|_T = \widetilde{u}_h(x_T)$ it follows from $\widetilde{u}_h(x) = \widetilde{u}_h(x_T) + \nabla_h\widetilde{u}_h|_T \cdot (x - x_T)$ that

$$\begin{aligned}\int_T \mu(\widetilde{u}_h - \Pi_{h,0}\widetilde{u}_h) \,\mathrm{d}x &= \int_{\mathcal{C}_T} \mu\, \nabla_h(\widetilde{u}_h - u) \cdot (x - x_T) \,\mathrm{d}x \\ &\le h_T \|\mu\|_{L^2(T)} \|\nabla_h(\widetilde{u}_h - u)\|_{L^2(T)}.\end{aligned}$$

We thus deduce that

$$\delta_h^2 \le \|\mu\|\big(\|u - \widetilde{u}_h\| + h\|\nabla_h(u - \widetilde{u}_h)\|\big) + \frac{1}{2}\|z - \Pi_{h,0}\widetilde{z}_h\|^2,$$

which implies the error estimate. □

Flux reconstruction. The discrete flux z_h can be constructed if a discrete Lagrange multiplier $\mu_h \in \Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ is given, i.e., $\mu_h \le 0$ is such that

$$(\mu_h, v_h) = (f_h, v_h) - (\nabla_h u_h, \nabla_h v_h)$$

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. We then have that

$$z_h(x) = \nabla_h u_h|_T - \frac{(f_h - \mu_h)|_T}{d}(x - x_T)$$

for all $T \in \mathcal{T}_h$ and $x \in T$.

4.2. **Total variation minimization.** Given a function $g \in L^2(\Omega)$ we consider the primal problem that consists in determining a function $u \in BV(\Omega) \cap L^2(\Omega)$ which is minimial for the functional

$$I(u) = |\mathrm{D}u|(\Omega) + \frac{1}{2}\|u - g\|^2.$$

The corresponding dual problem determines a maximizing vector field $z \in W^2_N(\mathrm{div};\Omega)$ with $\Gamma_N = \partial\Omega$ for the functional

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

subject to the pointwise constraint $|z| \le 1$ in Ω. From the characterization

$$|\mathrm{D}u|(\Omega) = \sup\Big\{ -\int_\Omega u \operatorname{div} z \,\mathrm{d}x : z \in W^2_N(\operatorname{div};\Omega),\ |z| \le 1 \text{ in } \Omega\Big\},$$

we obtain the strong duality relation

$$I(u) = D(z)$$

for solutions u and z of the primal and dual problems, where u and z are related via $\operatorname{div} z = u - g$ and the subdifferential inclusion $z \in \partial|\nabla u|$, cf., e.g., [37].

Discretization. With $g_h = \Pi_{h,0} g$ the discrete minimization problem is defined as the minimization of

$$I_h(u_h) = \int_\Omega |\nabla_h u_h| \,\mathrm{d}x + \frac{1}{2}\|\Pi_{h,0} u_h - g_h\|^2$$

in the set of all $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$. The discrete dual formulation consists in a maximization of

$$D_h(z_h) = -\frac{1}{2}\| \operatorname{div} z_h + g_h\|^2 + \frac{1}{2}\|g_h\|^2$$

in the set $\mathcal{RT}^0_N(\mathcal{T}_h)$ subject to the constraints $|z_h(x_T)| \le 1$ for all $T \in \mathcal{T}_h$. Related discretizations have been used in [36]. The discretization used here is obtained from Proposition 2.2 which shows that for every $\overline{u}_h \in \Pi_{h,0}\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ we have

$$\begin{aligned} &\inf\Big\{\int_\Omega |\nabla_h u_h| \,\mathrm{d}x :\ u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h), \Pi_{h,0} u_h = \overline{u}_h\Big\} \\ &\ge \sup\Big\{ -\int_\Omega \overline{u}_h \operatorname{div} z_h \,\mathrm{d}x :\ z_h \in \mathcal{RT}^0_N(\mathcal{T}_h),\ |z_h(x_T)| \le 1 \text{ for all } T \in \mathcal{T}_h\Big\}, \end{aligned}$$

and by using the relation $\operatorname{div} z_h = \Pi_{h,0} u_h - g_h$ and arguing as in Proposition 3.1 we obtain the discrete duality relation

$$I_h(u_h) \ge D_h(z_h)$$

for optimal elements u_h and z_h, respectively. The following quasi-optimal error estimate is obtained via constructing appropriate comparison functions. It confirms an estimate from [23] in which a discretization using piecewise constant functions and implicitly incorporating Crouzeix–Raviart elements has been considered. We closely follow the arguments used therein. It is remarkable that the data approximation error $g - g_h$ does not occur explicitly which avoids imposing restrictive conditions on g.

Proposition 4.2 (Error estimate)**.** *Let $u \in BV(\Omega) \cap L^2(\Omega)$ and $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ be optimal for I and I_h, respectively. Assume that $g \in L^\infty(\Omega)$ and there exists an optimal $z \in W^2_N(\operatorname{div};\Omega)$ for D with $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$. We then have that*

$$\|u - u_h\| \le ch^{1/2}\big(\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega) + \|g\|\|\nabla z\|_{L^\infty(\Omega)}\| \operatorname{div} z\|\big)^{1/2},$$

where $\|u\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}$.

Proof. The strong convexity properties of I_h and the discrete duality relation yield that

$$\frac{1}{2}\|\Pi_{h,0}(u_h - \widetilde{u}_h)\|^2 \le I_h(\widetilde{u}_h) - I_h(u_h) \le I_h(\widetilde{u}_h) - D_h(\widetilde{z}_h)$$

for every $\widetilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ with $|\widetilde{z}_h(x_T)| \le 1$. Since $g \in L^\infty(\Omega)$ we have that $u \in L^\infty(\Omega)$ with $\|u\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}$ and Lemma 4.3 below yields the existence of $\widetilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ with

$$I_h(\widetilde{u}_h) \le I(u) + ch - \frac{1}{2}\|g - g_h\|^2$$

and

$$\|u - \widetilde{u}_h\|_{L^1(\Omega)} \le ch, \quad \|\widetilde{u}_h\|_{L^\infty(\Omega)} \le c.$$

Letting $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ be the function constructed in Lemma 4.4 below we find that

$$D_h(\widetilde{z}_h) \ge D(z) - ch - \frac{1}{2}\|g - g_h\|^2.$$

On combining the previous estimates, and noting that $I(u) = D(z)$, we deduce that

$$\frac{1}{2}\|\Pi_{h,0}(u_h - \widetilde{u}_h)\|^2 \le ch.$$

We incorporate the estimates

$$\|u - \widetilde{u}_h\| \le \|u - \widetilde{u}_h\|^{1/2}_{L^1(\Omega)}\|u - \widetilde{u}_h\|_{L^\infty(\Omega)} \le ch^{1/2}$$

and

$$\|v_h - \Pi_{h,0}v_h\|_{L^2(\Omega)} \le ch\|\nabla_h v_h\|_{L^1(\Omega)}\|v\|_{L^\infty(\Omega)}$$

for every $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ to deduce the error bound. □

Modified interpolants. The following lemma provides the primal comparison function with explicit constants.

Lemma 4.3 (Primal quasi-interpolant)**.** *Given any $u \in BV(\Omega) \cap L^\infty(\Omega)$ there exists $\widetilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ such that*

$$I_h(\widetilde{u}_h) \le I(u) + 2c_d c_{cr} h\|u\|_{L^\infty(\Omega)}|\mathrm{D}u|(\Omega) - \frac{1}{2}\|g - g_h\|^2.$$

Proof. We choose a sequence $(u_\varepsilon)_{\varepsilon>0} \in C^\infty(\overline{\Omega}) \cap BV(\Omega)$ such that

$$\|u - u_\varepsilon\|_{L^1(\Omega)} \to 0, \quad \|\nabla u_\varepsilon\|_{L^1(\Omega)} \to |Du|(\Omega), \quad \|u_\varepsilon\|_{L^\infty(\Omega)} \to \|u\|_{L^\infty(\Omega)},$$

cf. [1, 13]. We then define $\widetilde{u}^\varepsilon_h = \mathcal{I}_{cr} u_\varepsilon$ and note that

$$\|\nabla_h \widetilde{u}^\varepsilon_h\|_{L^1(\Omega)} \le \|\nabla u_\varepsilon\|_{L^1(\Omega)}$$

We pass to an accumulation point $\widetilde{u}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ as $\varepsilon \to 0$ for which we have that

$$\begin{aligned} \|\nabla_h \widetilde{u}_h\|_{L^1(\Omega)} &\le |\mathrm{D}u|(\Omega), \\ \|\widetilde{u}_h\|_{L^\infty(\Omega)} &\le 2c_d\|u\|_{L^\infty(\Omega)}, \\ \|\widetilde{u}_h - u\|_{L^1(\Omega)} &\le c_{cr} h|\mathrm{D}u|(\Omega). \end{aligned}$$

For ease of notation we abbreviate $\overline{u}_h = \Pi_{h,0}\widetilde{u}_h$ and $g_h = \Pi_{h,0}g$. We have that

$$\|\overline{u}_h - g_h\|^2 = \|\overline{u}_h - g\|^2 - \|g - g_h\|^2$$

and

$$\|\overline{u}_h - g\|^2 = \|u - g\|^2 + \int_\Omega (\overline{u}_h - u)(\overline{u}_h + u - 2g)\,\mathrm{d}x.$$

These identities imply that we have

$$\begin{aligned} I_h(\widetilde{u}_h) &= \|\nabla \widetilde{u}_h\|_{L^1(\Omega)} + \frac{1}{2}\|\overline{u}_h - g_h\|^2 \\ &\le I(u) + \frac{1}{2}\|\overline{u}_h - u\|_{L^1(\Omega)}\|\overline{u}_h + u - 2g\|_{L^\infty(\Omega)} - \frac{1}{2}\|g - g_h\|^2. \end{aligned}$$

This implies the assertion. □

A comparison function for the discrete dual problem is constructed in the following lemma.

Lemma 4.4 (Dual quasi-interpolant)**.** *Given any $z \in W^2_N(\mathrm{div};\Omega)$ with $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$ there exists $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathrm{div};\Omega)$ with $|\widetilde{z}_h(x_T)| \le 1$ for all $T \in \mathcal{T}_h$ and*

$$D_h(\widetilde{z}_h) \ge D(z) - c_{\mathcal{RT}} hL\|g\|\|\operatorname{div} z\| - \frac{1}{2}\|g - g_h\|^2,$$

with the Lipschitz constant L of z.

Proof. The interpolant $\mathcal{I}_{\mathcal{RT}} z$ satisfies $\operatorname{div} \mathcal{I}_{\mathcal{RT}} z = \Pi_{h,0} \operatorname{div} z$ and we have with the constant function $\overline{z}|_T = z(x_T)$ that

$$|\mathcal{I}_{\mathcal{RT}} z(x_T)| \le \|\mathcal{I}_{\mathcal{RT}}(z - \overline{z})\|_{L^\infty(T)} + |\overline{z}| \le c_{\mathcal{RT}} hL + 1 = \gamma_h.$$

Hence, for $\widetilde{z}_h = \gamma_h^{-1}\mathcal{I}_{\mathcal{RT}} z = \mathcal{I}_{\mathcal{RT}}\widetilde{z}$ with $\widetilde{z} = \gamma_h^{-1} z$ we have $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and $|\widetilde{z}_h(x_T)| \le 1$ for all $T \in \mathcal{T}_h$. Noting that

$$\operatorname{div} \widetilde{z}_h + g_h = \Pi_{0,h}(\operatorname{div} \widetilde{z} + g)$$

and $\|g\|^2 - \|g_h\|^2 = \|g - g_h\|^2$ we deduce that

$$\begin{aligned} D_h(\widetilde{z}_h) &= -\frac{1}{2}\int_\Omega (\operatorname{div} \widetilde{z}_h + g_h)^2 - g_h^2\,\mathrm{d}x \\ &\ge -\frac{1}{2}\int_\Omega (\operatorname{div} \widetilde{z} + g)^2 - g^2\,\mathrm{d}x - \frac{1}{2}\|g - g_h\|^2. \end{aligned}$$

Hence, we have that

$$\begin{aligned} D_h(\widetilde{z}_h) &\ge -\frac{1}{2}\int_\Omega (\operatorname{div} \widetilde{z})^2 + 2g \operatorname{div} \widetilde{z}\,\mathrm{d}x - \frac{1}{2}\|g - g_h\|^2 \\ &= -\frac{1}{2}\gamma_h^{-2}\int_\Omega (\operatorname{div} z)^2\,\mathrm{d}x - \gamma_h^{-1}\int_\Omega g \operatorname{div} z\,\mathrm{d}x - \frac{1}{2}\|g - g_h\|^2 \\ &\ge -\frac{1}{2}\int_\Omega (\operatorname{div} z)^2 + 2g \operatorname{div} z\,\mathrm{d}x - (1 - \gamma_h^{-1})\|g\|\|\operatorname{div} z\| - \frac{1}{2}\|g - g_h\|^2, \end{aligned}$$

where we also used that $\gamma_h^{-2} \le 1$. The estimate $1 - \gamma_h^{-1} \le c_{\mathcal{RT}} hL$ implies the assertion. □

Remark 4.5. *In the absence of the regularity condition $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$ one can establish Γ-convergence $I_h \to I$ in $L^1(\Omega)$. Alternatively, one may choose a regularization z_ε of z so that Lemma 4.4 holds with $L_\varepsilon = c\varepsilon^{-1}$. An approximability condition on g then implies $\|\operatorname{div} z - \operatorname{div} z_\varepsilon\| \le \varepsilon$ and leads to the convergence rate $\mathcal{O}(h^{1/4})$, cf.* [23]. *This rate has also been obtained in* [9, 10] *for conforming approximations and was improved in* [14] *in the case of certain anisotropic functionals.*

Flux reconstruction. The ideas that lead to the reconstruction of the solution of the dual problem can be transferred to the nonsmooth situation if a regularization of the modulus function is used to approximate the discrete primal functional I_h, i.e., if $|\cdot|_\varepsilon : \mathbb{R}^d \to \mathbb{R}$ is a differentiable approximation of euclidean length, then the discrete primal and dual problems correspond to the Lagrange functional

$$L_{h,\varepsilon}(u_h, z_h) = -\int_\Omega u_h \operatorname{div} z_h + |\Pi_{h,0} z_h|_\varepsilon^* \,\mathrm{d}x + \frac{1}{2}\|\Pi_{h,0} u_h - g_h\|^2.$$

and the relations

$$\operatorname{div} z_h = \Pi_{h,0} u_h - g_h, \quad \nabla_h u_h \in D|\Pi_{h,0} z_h|_\varepsilon^*,$$

where the second identity is equivalent to

$$\Pi_{h,0} z_h = D|\nabla_h u_h|_\varepsilon.$$

If, e.g., $|s| = (|s|^2 + \varepsilon^2)^{1/2}$ then we obtain on every $T \in \mathcal{T}_h$

$$z_h = \frac{\nabla_h u_h}{|\nabla_h u_h|_\varepsilon} + \frac{\Pi_{h,0} u_{h,\varepsilon} - g_h}{d}(\cdot - x_T).$$

4.3. **Infinity Laplacian.** A variant of the p-Laplace problem with $p \to \infty$ arises in problems of optimal transportation and leads to a minimization of

$$I(u) = I_{K_1(0)}(\nabla u) - \int_\Omega f u \,\mathrm{d}x$$

in the space $W_D^{1,\infty}(\Omega)$ for a given function $f \in L^1(\Omega)$. The dual problem consists in maximizing the functional

$$D(z) = -\int_\Omega |z| \,\mathrm{d}x - I_{\{-f\}}(\operatorname{div} z)$$

in the space $W_N^1(\operatorname{div};\Omega)$. We refer the reader to [30] for existence and strong duality results.

Discretization. We define a discrete approximation of I via

$$I_h(u_h) = I_{K_1(0)}(\nabla_h u_h) - \int_\Omega f_h u_h \,\mathrm{d}x$$

on the set $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ using $f_h = \Pi_{h,0} f$. Proposition 3.1 implies that the discrete dual problem consists in maximizing the functional

$$D_h(z_h) = -\int_\Omega |\Pi_{h,0} z_h| \,\mathrm{d}x - I_{\{-f_h\}}(\operatorname{div} z_h)$$

in the set of all discrete vector fields $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Other discretizations are addressed in [7, 41, 8, 15, 43]. We have the following approximation result.

Proposition 4.6 (Approximation). *If a solution $z \in W^1_N(\mathrm{div};\Omega)$ of the dual problem with $z \in W^{1,1}(\Omega;\mathbb{R}^d)$ exists and if u and u_h solves the primal and discrete primal problem, respectively, then we have*

$$\big|I_h(u_h) - I(u)\big| \le ch\big(\|f\|_{L^1(\Omega)} + \|\nabla z\|_{L^1(\Omega)}\big).$$

Proof. Establishing the existence of a discrete solution $u_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ is straightforward by continuity of the discrete problem and boundedness of the admissible set. Abbreviating $\widetilde{u}_h = \mathcal{I}_{cr}u$ and $\widetilde{z}_h = \mathcal{I}_{\mathcal{RT}}z$ we note that $|\nabla_h \widetilde{u}_h| \le 1$ and hence

$$\begin{aligned}
0 \le I_h(\widetilde{u}_h) - I_h(u_h) &\le I_h(\widetilde{u}_h) - D_h(\widetilde{z}_h) \\
&= -\int_\Omega f_h \widetilde{u}_h \,\mathrm{d}x + \int_\Omega |\Pi_{h,0}\widetilde{z}_h| \,\mathrm{d}x \\
&= -\int_\Omega \Pi_{h,0}\widetilde{z}_h \cdot \nabla_h \widetilde{u}_h \,\mathrm{d}x + \int_\Omega |\Pi_{h,0}\widetilde{z}_h| \,\mathrm{d}x \\
&= -\int_\Omega \Pi_{h,0}\widetilde{z}_h \cdot \nabla u \,\mathrm{d}x + \int_\Omega |\Pi_{h,0}\widetilde{z}_h| \,\mathrm{d}x.
\end{aligned}$$

The duality relation $I(u) = D(z)$ shows that

$$-\int_\Omega |z| \,\mathrm{d}x = -\int_\Omega fu \,\mathrm{d}x = \int_\Omega z \cdot \nabla u \,\mathrm{d}x.$$

This leads to

$$\begin{aligned}
0 &\le I_h(\widetilde{u}_h) - I_h(u_h) \\
&\le \int_\Omega (z - \Pi_{h,0}\widetilde{z}_h) \cdot \nabla u \,\mathrm{d}x + \int_\Omega |\Pi_{h,0}\widetilde{z}_h| - |z| \,\mathrm{d}x \\
&\le 2\|z - \Pi_{h,0}\widetilde{z}\|_{L^1(\Omega)} \le ch\|\nabla z\|_{L^1(\Omega)}.
\end{aligned}$$

Finally, we verify that

$$I_h(\widetilde{u}_h) - I(u) = \int_\Omega f_h \Pi_{h,0}\widetilde{u}_h - fu \,\mathrm{d}x = -\int_\Omega f(u - \Pi_{h,0}\widetilde{u}_h) \,\mathrm{d}x,$$

and deduce the asserted estimate. □

Remark 4.7. *On right-angled triangulations the conforming $P1$ finite element method leads to a similar estimate since we have that*

$$\|\nabla \mathcal{I}_{p1} u\|_{L^\infty(\Omega)} \le \|\nabla u\|_{L^\infty(\Omega)}$$

for every $u \in W^{1,\infty}(\Omega)$ and hence if $u_h^c \in \mathcal{S}^{1,0}_D(\mathcal{T}_h) \subset W^{1,\infty}_D(\Omega)$ is minimal for I_h in this set then

$$\begin{aligned}0 \le I(u_h^c) - I(u) &= I(u_h^c) - I_h(u_h^c) + I_h(u_h^c) - I_h(\mathcal{I}_{p1}u) + I_h(\mathcal{I}_{p1}u)\\ &\quad - I(\mathcal{I}_{p1}u) + I(\mathcal{I}_{p1}u) - I(u)\\ &\le \|f - f_h\|_{L^1(\Omega)}\big(\|u_h^c - \Pi_{h,0}u_h^c\|_{L^\infty(\Omega)} + \|\mathcal{I}_{p1}u - \Pi_{h,0}\mathcal{I}_{p1}u\|_{L^\infty(\Omega)}\big)\\ &\quad + \|f\|_{L^1(\Omega)}\|u - \mathcal{I}_{p1}u\|_{L^\infty(\Omega)},\end{aligned}$$

where we used that $f_h = \Pi_{h,0}f$ and $I_h(u_h) \le I_h(\mathcal{I}_{p1}u)$. Hence, without additional regularity assumptions we have that $|I(u_h^c) - I(u)| \le ch$; if $f \in W^{1,1}(\Omega)$ and $u \in W^{2,\infty}(\Omega)$ then this can be improved to $\mathcal{O}(h^2)$. A realistic regularity property is $u \in W^{4/3,\infty}(\Omega)$, cf. [4].

Flux reconstruction. To construct the discrete flux z_h from the solution u_h of the nonconforming method for the primal problem we consider a regularization $|\cdot|_\varepsilon$ of the euclidean length which defines regularizations $|\cdot|_\varepsilon^*$ of $I_{K_1(0)}$. We then find that on every $T \in \mathcal{T}_h$ we have

$$z_h = D|\nabla_h u_h|_\varepsilon^* - (f_h/d)(\cdot - x_T),$$

where z_h and u_h are the solutions of the regularized problems.

5. Iterative solution

To solve the discrete problems we devise iterative algorithms for problems with sub- and superquadratic growth properties that result from semi-implicit discretizations of appropriate gradient flows for the primal and dual problem, respectively. A gradient flow for the primal minimization problem determines a family $(u(t))_{t\ge 0} \subset W^{1,p}_D(\Omega)$ of functions for an initial $u^0 \in W^{1,p}_D(\Omega)$ via $u(0) = u^0$ and

$$(\partial_t u, v)_* = -\int_\Omega D\phi(\nabla u)\cdot\nabla v \,\mathrm{d}x - \int_\Omega D\psi(u)v \,\mathrm{d}x$$

for all $v \in W^{1,p}_D(\Omega)$ and all $t > 0$. To avoid solving nonlinear systems of equations a semi-implicit discretization in time is used. We consider the case that ϕ only depends on the length of its argument, i.e., $\phi(s) = \varphi(|s|)$ with a convex function $\varphi \in C^1(\mathbb{R}_{\ge 0})$. In this case we have

$$D\phi(s) = \frac{\varphi'(|s|)s}{|s|},$$

which naturally leads to a semi-implicit treatment. To discretize the time derivative we use the the backward difference quotient operator

$$d_t u^k = \tau^{-1}(u^k - u^{k-1})$$

for a sequence $(u^k)_{k=0,1,\dots}$ and a step-size $\tau > 0$.

Algorithm 5.1 (Subquadratic case, primal iteration)**.** *Let* $u^0 \in W^{1,p}_D(\Omega)$ *and choose* $\tau, \varepsilon_{\rm stop} > 0$*, set* $k = 0$*.*
(1) Compute $u^k \in W^{1,p}_D(\Omega)$ *such that*

$$(d_t u^k, v)_* + \int_\Omega \frac{\varphi'(|\nabla u^{k-1}|)}{|u^{k-1}|} \nabla u^k \cdot \nabla v \,\mathrm{d}x + \int_\Omega D\psi(u) v \,\mathrm{d}x = 0$$

for all $v \in W^{1,p}_D(\Omega)$*.*
(2) Stop if $\|d_t u^k\|_* \le \varepsilon_{\rm stop}$*; otherwise increase* $k \to k+1$ *and continue with (1).*

It is shown below that the iteration is unconditionally energy decreasing and convergent if φ has subquadratic growth. If this is not the case then we expect the dual problem to have this property and consider a gradient descent for $-D$, i.e., we determine a family $(z(t))_{t\ge 0}$ satisfying $z(0) = z^0$ and the constrained evolution equation

$$(\partial_t z, y)_\dagger = -\int_\Omega D\phi^*(z) \cdot y \,\mathrm{d}x - \int_\Omega D\psi^*(\operatorname{div} z) \operatorname{div} y \,\mathrm{d}x$$

for all $y \in W^{p'}_N(\operatorname{div};\Omega)$. In case of a linear functional ψ, the differential $D\psi^*$ becomes a subdifferential and the equation a variational inequality or constrained equation. Similarly to the gradient flow for the primal problem we assume that the integrand is isotropic, i.e., $\phi^*(r) = \varphi(|r|)$ with a convex function $\varphi \in C^1(\mathbb{R}_{\ge 0})$. In this case we have

$$D\phi^*(r) = \frac{\varphi'(|r|) r}{|r|}$$

and the semi-implicit iteration is similar to that of Algorithm 5.1.

Algorithm 5.2 (Superquadratic case, dual iteration)**.** *Let* $z^0 \in W^{p'}_N(\operatorname{div};\Omega)$ *and choose* $\tau, \varepsilon_{\rm stop} > 0$*, set* $k = 0$*.*
(1) Compute $z^k \in W^{p'}_N(\operatorname{div};\Omega)$ *such that*

$$(d_t z^k, y)_\dagger + \int_\Omega \frac{\varphi'(|z^{k-1}|)}{|z^{k-1}|} z^k \cdot y \,\mathrm{d}x + \int_\Omega D\psi^*(\operatorname{div} z^k) \operatorname{div} y \,\mathrm{d}x = 0,$$

for all $y \in W^{p'}_N(\operatorname{div};\Omega)$*.*
(2) Stop if $\|d_t z^k\|_\dagger \le \varepsilon_{\rm stop}$*; otherwise increase* $k \to k+1$ *and continue with (1).*

If $\psi(x,s) = -f(x)s$ then the system in Step (1) includes the constraints $-\operatorname{div} z^k = f$ and $\operatorname{div} y = 0$ instead of the integral involving $D\psi^*$. The algorithms converge for subquadratic growth of ϕ and ϕ^*, respectively. We adopt arguments from [12].

Proposition 5.3 (Unconditional convergence)**.** *Assume that* $r \mapsto \varphi'(r)/r$ *is positive, non-increasing, and continuous on* $\mathbb{R}_{\ge 0}$*. If* $\phi(s) = \varphi(|s|)$ *for all*

$s \in \mathbb{R}^d$ then the iteration of Algorithm 5.1 is well-posed, convergent, and monotone with

$$I(u^\ell) + \tau \sum_{k=1}^{\ell} \|d_t u^k\|_*^2 \le I(u^0).$$

If $\phi^*(t) = \varphi(|t|)$ then the iteration of Algorithm 5.2 is well-posed, convergent, and monotone with

$$-D(z^\ell) + \tau \sum_{k=1}^{\ell} \|d_t z^k\|_\dagger^2 \le -D(z^0).$$

Proof. (i) The conditions on φ imply that the iteration is well posed and that we have

$$\frac{\varphi'(|a|)}{|a|} b \cdot (b-a) \ge \varphi(|b|) - \varphi(|a|) + \frac{1}{2}\frac{\varphi'(|a|)}{|a|}|b-a|^2 \tag{5}$$

for all $a, b \in \mathbb{R}^d$, cf. Appendix A.2 for a proof of (5). Hence, by choosing $v = d_t u^k$ in Algorithm 5.1 we find that

$$\|d_t u^k\|_*^2 + \int_\Omega \frac{\varphi'(|\nabla u^{k-1}|)}{|u^{k-1}|} \nabla u^k \cdot \nabla d_t u^k \,\mathrm{d}x + \int_\Omega D\psi(u^k) d_t u^k \,\mathrm{d}x = 0.$$

Using $a = \nabla u^{k-1}$ and $b = \nabla u^k$ in (5) shows that

$$\frac{\varphi'(|\nabla u^{k-1}|)}{|\nabla u^{k-1}|} \nabla u^k \cdot \nabla(u^k - u^{k-1}) \ge \varphi(|\nabla u^k|) - \varphi(|\nabla u^{k-1}|).$$

By combining the last two equations, using convexity of ψ, and summing over $k = 1, 2, \dots, \ell$ we deduce the asserted estimate.
(ii) If the conditions on ϕ^* are satisfied then the arguments used to show (i) apply to Algorithm 5.2 and we deduce the estimate. □

Example 5.4. *The conditions of the proposition apply to typical regularized p-Dirichlet energies $\phi(s) = |s|_\varepsilon^p$ for $\varepsilon > 0$, cf. [12]. Algorithm 5.1 converges if $1 \le p \le 2$ while Algorithm 5.2 converges if $2 \le p < \infty$.*

Remark 5.5. *Note that owing to the semi-implicit discretization the functions $d_t u^k$ and $d_t z^k$ are not residuals. If, e.g., $\widetilde{u} = u^k$ for some $k \ge 0$ and the residual r is defined via*

$$(D\phi_\varepsilon(\nabla \widetilde{u}), \nabla v) + (D\psi(\widetilde{u}), v) = (r, v)_*$$

for all $v \in W^{1,p}_D(\Omega)$, then by convexity of I_ε we have

$$I_\varepsilon(\widetilde{u}) + (r, v - \widetilde{u})_* + \sigma_I^2(\widetilde{u}, v) \le I_\varepsilon(v)$$

for all $v \in W^{1,p}_D(\Omega)$, where we assume that coercivity holds uniformly with respect to $\varepsilon \ge 0$. In case of the L^2 scalar product $(\cdot,\cdot)_* = (\cdot,\cdot)$, and if, e.g., $\sigma_I^2(\widetilde{u}, v) \ge (\alpha_I/2)\|\widetilde{u} - v\|^2$, we deduce that

$$I_\varepsilon(\widetilde{u}) + \frac{\alpha_I}{4}\|v - \widetilde{u}\|^2 \le I_\varepsilon(v) + \frac{1}{\alpha_I}\|r\|^2.$$

With the minimizing u_ε for I_ε we deduce $\|u_\varepsilon - \widetilde{u}\| \le (2/\alpha_I)\|r\|$.

Two alternative approaches to the iterative solution of the discrete problems are described in the following remarks.

Remarks 5.6. *(i) The ADMM iteration (alternating direction of multiplier method) as in* [31] *decouples the gradient operator from ϕ by introducing $q = \nabla u$ via a Lagrange multiplier λ. With the augmented Lagrange functional*

$$L_\tau(u,q,\lambda) = \int_\Omega \phi(q)\,\mathrm{d}x + \int_\Omega \psi(u)\,\mathrm{d}x + (\lambda, \nabla u - q)_H + \frac{\tau}{2}\|\nabla u - q\|_H^2,$$

with a suitable Hilbert space norm H and a stabilization parameter $\tau > 0$, the algorithm successively minimizes L_τ with respect to u and q, and then performs an ascent step with respect to λ.
(ii) Primal-dual methods as investigated in [22] *alternatingly update the variable u and z in the Lagrange functional*

$$L(u,z) = \int_\Omega -u \operatorname{div} z - \phi^*(z) + \psi(u)\,\mathrm{d}x.$$

via discretizations of $\partial_t z = \delta_z L(u,z)$ and $\partial_t u = -\delta_u L(u,z)$ using an extrapolated quantity to decouple the equations. The application to Raviart–Thomas methods is not straightforward due to their nonlocal character.

6. Numerical experiments

In this section we verify the theoretical findings via numerical experiments and illustrate advantages of nonconforming and mixed methods over standard conforming methods.

6.1. Total variation minimization.

We consider the numerical approximation of the functional

$$I(u) = |\mathrm{D}u|(\Omega) + \frac{\alpha}{2}\|u-g\|^2.$$

To compare approximations to an exact solution we impose Dirichlet boundary conditions on $\Gamma_D = \partial\Omega$. Although it is difficult to establish a general existence theory, the error estimates of Section 4.2 carry over verbatimly with $\Gamma_N = \emptyset$ provided a minimizer exists. This is the case in the setting of the following example.

Example 6.1. *For $\Omega \subset \mathbb{R}^d$ and $r > 0$ with $B_r(0) \subset \Omega$, and $g = \chi_{B_r(0)}$ the unique minimizer for I subject to Dirichlet boundary conditions is given by*

$$u = \max\{0, 1 - d/(\alpha r)\}\chi_{B_r(0)}.$$

If $d \le \alpha r$ then the Lipschitz continuous vector field

$$z(x) = \begin{cases} -r^{-1}x & \text{for } |x| \le r, \\ -rx/|x|^2 & \text{for } |x| \ge r, \end{cases}$$

solves the dual problem, cf., e.g., [10]*. We use $d = 2$, $\Omega = (-1,1)^2$, $r = 1/2$, and $\alpha = 10$.*

Iterative solution. For the practical solution of the minimization problem we use a regularization defined with the regularized euclidean length

$$|s|_\varepsilon = (|s|^2 + \varepsilon^2)^{1/2}$$

for $\varepsilon > 0$ and $s \in \mathbb{R}^d$. The uniform approximation property $0 \le |s|_\varepsilon - |s| \le \varepsilon$ for all $s \in \mathbb{R}^d$ implies that with the regularized functional

$$I_\varepsilon(u) = \int_\Omega |\nabla u|_\varepsilon \, \mathrm{d}x + \frac{\alpha}{2}\|u - g\|^2,$$

we have for minimizers u of I and u_ε of I_ε that

$$\frac{\alpha}{2}\|u - u_\varepsilon\|^2 \le I(u_\varepsilon) - I(u) \le \varepsilon.$$

This justifies using the regularized functional with $\varepsilon = h$ to compute approximations for minimizers of I. We use Algorithm 5.1 to decrease the energy and stop the iteration when $\|d_t u^k\| \le \varepsilon_{\mathrm{stop}} = h/20$. We always use the L^2 inner product and the step size $\tau = 1$.

Experimental results. For triangulations $\mathcal{T}_\ell$ of $\Omega = (-1,1)^2$ resulting from $\ell \ge 0$ uniform refinements of a coarse triangulation of Ω into two triangles we have that the maximal mesh-size of $\mathcal{T}_\ell$ is proportional to $h_\ell = 2^{-\ell}$. For a simple implementation we use the function $\widetilde{g}_h \in \mathcal{L}^0(\mathcal{T}_h)$ via

$$\widetilde{g}_h(x_T) = g(x_T) = \begin{cases} 1 & |x_T| < 1/2, \\ 0 & |x_T| \ge 1/2, \end{cases}$$

instead of the L^2 projection $g_h = \Pi_{h,0} g$. Since for $g = \chi_{B_r(0)}$ we have $\|g - \widetilde{g}_h\|_{L^1(\Omega)} \le ch|\partial B_r(0)|$, the error estimate remains valid. The top row in Figure 1 shows the numerical solutions obtained for the discretizations using a standard $P1$ method and the Crouzeix–Raviart method on the triangulation $\mathcal{T}_5$. At first glance the $P1$ approximation appears superior as, e.g., the Crouzeix–Raviart approximation does not satisfy a discrete maximum principle. The projections of the approximations onto piecewise constant functions are shown in the bottom row of Figure 1 and lead to a different interpretation. The circular discontinuity set is better resolved by the discontinuous method and we observe a more localized approximation of the jump set. Figure 2 supports the latter interpretation via logarithmic plots for the experimental convergence rates of the error quantity

$$\|e_h\|^2 = \|\Pi_{h,0} u_h - u(x_\mathcal{T})\|^2,$$

where $x_\mathcal{T}|_T = x_T$ for every $T \in \mathcal{T}_\ell$, versus the number of vertices $N_\ell \sim h_\ell^{-2}$. We observe that the L^2 error for the Crouzeix–Raviart method converges at the quasi-optimal rate $\mathcal{O}(h^{1/2})$ while the $P1$ error is larger and decays at a lower rate. The approximations were computed on the triangulations $\mathcal{T}_\ell$ for $\ell = 3, 4, \dots, 9$ with $N_\ell = (2^\ell + 1)^2 = 81, 289, \dots, 66049, 263169$ vertices.

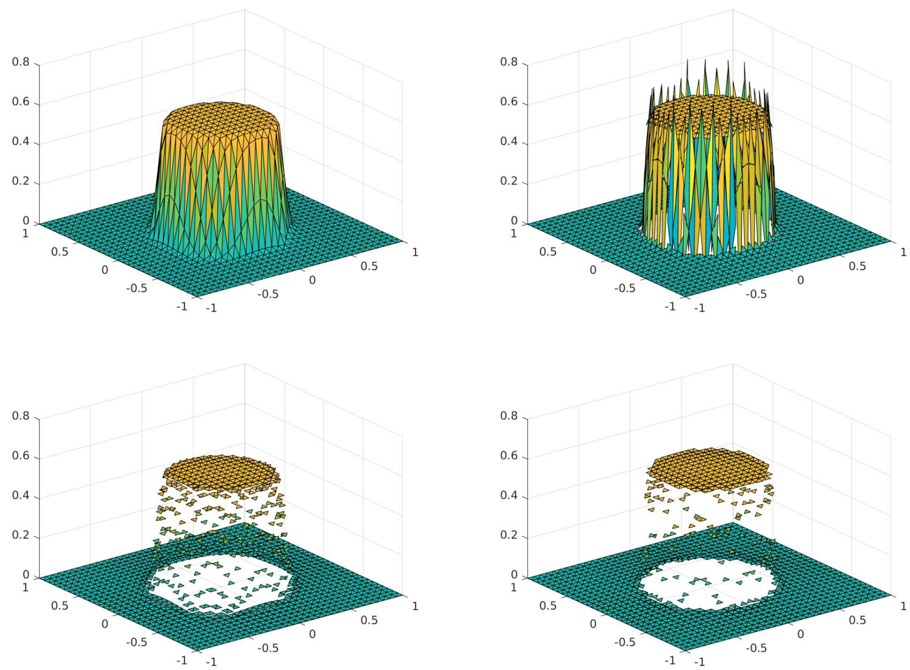

FIGURE 1. Continuous $P1$ (left) and Crouzeix–Raviart approximations (right) in Example 6.1 displayed as piecewise affine functions (top) and via their projections onto piecewise constant functions (bottom).

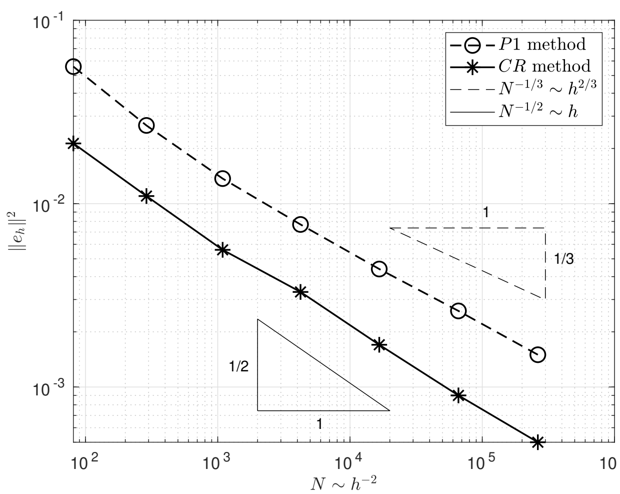

FIGURE 2. Squared L^2 errors in Example 6.1 for $P1$ and Crouzeix-Raviart approximations. The predicted rate $\mathcal{O}(h^{1/2})$ is observed for the Crouzeix–Raviart method while the $P1$ method leads to larger errors and a reduced rate.

6.2. **Effect of modification.** The operator $\Pi_{h,0}$ that occurs in the discrete dual problem via the term $\phi^*(\Pi_{h,0} z_h)$ is crucial for the discrete duality theory and in fact simplifies the realization of the method as quadrature becomes trivial. This does not affect the discrete flux variable z_h but leads to a modified discrete Langrange multiplier $\overline{u}_h$. To illustrated this effect we consider the standard dual mixed formulation (1) of the Poisson problem and the modified version which seeks $(z_h, \overline{u}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ satisfying

$$(\Pi_{h,0} z_h, y_h) + (\overline{u}_h, \operatorname{div} z_h) = 0, \quad (\overline{v}_h, \operatorname{div} y_h) = -(f_h, v_h),$$

for all $(y_h, \overline{v}_h) \in \mathcal{RT}^0_N(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$. We specify the problem as follows.

Example 6.2 (Poisson problem)**.** *Let $d = 2$, $\Omega = (-1,1)^2$, $\Gamma_D = \partial\Omega$, and $f(x,y) = 2(1-x^2) + 2(1-y^2)$. The solution of the dual mixed formulation of the Poisson problem is given by $u(x,y) = (1-x^2)(1-y^2)$ and $z = \nabla u$.*

Figure 3 shows the L^2 errors

$$\|e_h\| = \|\overline{u}_h - u(x_{\mathcal{T}})\|,$$

versus numbers of vertices in $\mathcal{T}_\ell$ with a logarithmic scaling on both axes. The L^2 error for the modified treatment is larger than that for the exact treatment but converges at the same quadratic rate. This rate is higher than the expected linear convergence rate for the difference $\|\overline{u}_h - u\|$. An explanation is provided by the relation $\overline{u}_h = u_h(x_T)$ to solutions u_h of the Crouzeix–Raviart discretization for which we have $\|u_h - u\|_{L^\infty(\Omega)} = \mathcal{O}(h^2 \log(h))$, cf. [32].

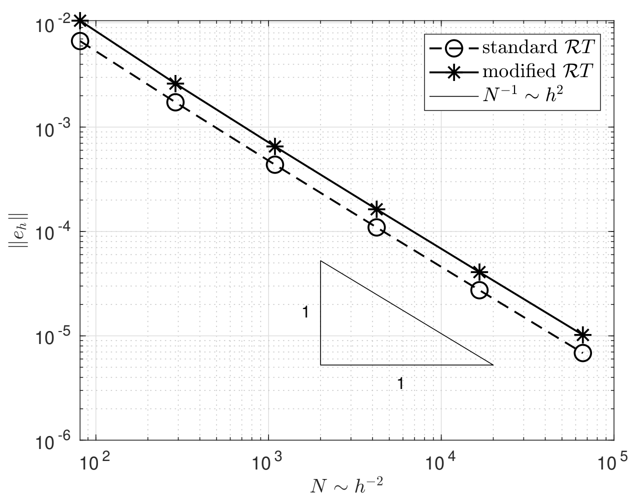

FIGURE 3. L^2 errors for the standard and modified Raviart–Thomas approximations of the Poisson problem of Example 6.2. An increased L^2 error is observed for the modified treatment but both approximations converge with nearly qudratic rate.

6.3. **Infinity Laplacian.** We define an infinity Laplace problem via the primal functional

$$I(u) = -\int_\Omega f u \,\mathrm{d}x, \quad |\nabla u| \le 1,$$

on the set $W^{1,\infty}_D(\Omega)$ for a given function $f \in L^1(\Omega)$. We approximate solutions by determining nearly maximizing discrete vector fields for the regularized dual functional

$$D_\varepsilon(z) = -\int_\Omega |z|_\varepsilon \,\mathrm{d}x, \quad -\operatorname{div} z = f,$$

with $|s|_\varepsilon = (|s|^2 + \varepsilon^2)^{1/2}$. We consider the following specification that leads to a Lipschitz continuous solution.

Example 6.3 (Infinity Laplacian)**.** *Let $d = 2$, $\Omega = (-1,1)^2$, $\Gamma_D = \partial\Omega$, and $f(x,y) = 1$. Then the solution of the primal problem is given by $u(x,y) = 1 - \max\{|x|,|y|\}$.*

We use Algorithm 5.2 with the L^2 scalar product and $\tau = 1$ to iteratively determine discrete minimizers for D_ε using $\varepsilon = h$. We also compute conforming approximations u_h^c for the primal problem using a conforming $P1$ finite element method and the ADMM iteration described in Remarks 5.6. Figure 5 displays the resulting approximation errors

$$|D(z) - D_{\varepsilon,h}(z_h)|, \quad |I(u) - I(u_h^c)|,$$

obtained using the Raviart-Thomas method for the dual problem and a standard conforming $P1$ method for the primal problem. We observe that on right-angled triangulations the $P1$ method leads to an almost quadratic convergence rate which is slightly better than the experimental convergence rate $\mathcal{O}(h^{5/3})$ observed for the Raviart–Thomas method. Surprisingly, the nearly quadratic convergence behavior is also observed for $P1$ finite element approximations on perturbed triangulations. We note however that in this case the admissibility of the nodal interpolant is not true in general, cf. Remark 4.7.

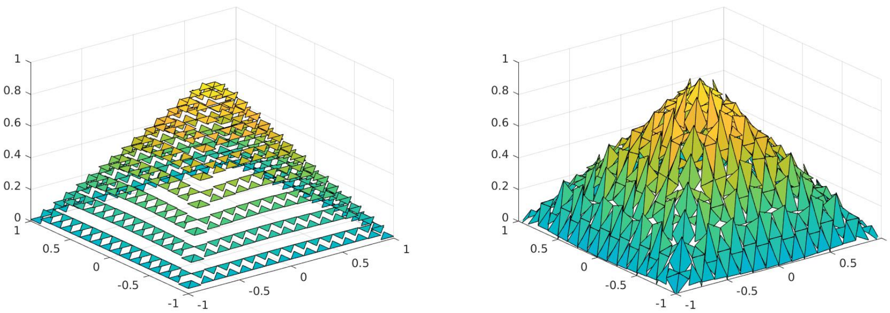

FIGURE 4. Piecewise constant approximation of the solution of the infinity Laplacian defined in Example 6.3 obtained with the Raviart–Thomas discretization of the regularized dual problem (left) and the corresponding reconstructed Crouzeix–Raviart approximation (right).

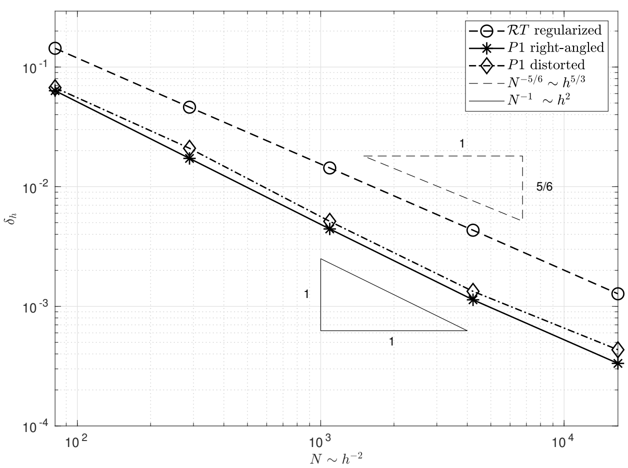

FIGURE 5. Experimental convergence rate for the approximation of the value $D(z)$ using the Raviart–Thomas discretization $D_{h,\varepsilon}$ and of $I(u)$ using a conforming $P1$ method in the case of the inifinity Laplace problem of Example 6.3.

Appendix A. Auxiliary results

A.1. **Enriching operator.** To prove the estimates for the enriching operator E_h^{cr} we follow the ideas described in [18] and define for given $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ a continuous, piecewise quadratic function $\alpha_h = E_h^{cr} v_h \in \mathcal{S}^{2,0}_D(\mathcal{T}_h) \subset W^{1,p}_D(\Omega)$ by defining the nodal values α_z associated with vertices $z \in \mathcal{N}_h$ and α_S associated with sides $S \in \mathcal{S}_h$ via

$$\alpha_S = v_h(x_S)$$

for all $S \in \mathcal{S}_h$, and in case of vertices $\alpha_z = 0$ if $z \in \mathcal{N}_h \cap \Gamma_D$ and otherwise

$$\alpha_z = n_z^{-1} \sum_{T \in \mathcal{T}_h : z \in T} v_h|_T(z),$$

where n_z is the number of elements in $\mathcal{T}_h$ that contain z, so that constant functions are reproduced. With a nodal basis $(\widetilde{\varphi}_{\widetilde{z}})_{\widetilde{z} \in \mathcal{N}_h \cup \mathcal{S}_h}$ for $\mathcal{S}^{2,0}(\mathcal{T}_h)$ which defines a partition of unity with supports $\omega_{\widetilde{z}}$ of diameters $h_{\widetilde{z}}$ we have

$$\begin{aligned} \|\nabla \alpha_h\|_{L^p(\Omega)}^p &\le \sum_{\widetilde{z} \in \mathcal{N}_h \cup \mathcal{S}_h} \int_{\omega_{\widetilde{z}}} |\alpha_{\widetilde{z}} - v_h| |\nabla \widetilde{\varphi}_{\widetilde{z}}| |\nabla \alpha_h|^{p-1} \,\mathrm{d}x \\ &\le c \sum_{\widetilde{z} \in \mathcal{N}_h \cup \mathcal{S}_h} \|\alpha_{\widetilde{z}} - v_h\|_{L^p(\omega_{\widetilde{z}})} h_{\widetilde{z}}^{-1} \|\nabla \alpha_h\|_{L^p(\omega_z)}^{p-1}. \end{aligned}$$

To show that $\|\nabla \alpha_h\|_{L^p(\Omega)} \le c \|\nabla_h v_h\|_{L^p(\Omega)}$ it suffices to prove that $\|\alpha_{\widetilde{z}} - v_h\|_{L^p(\omega_{\widetilde{z}})} \le c h_{\widetilde{z}} \|\nabla_h v_h\|_{L^p(\omega_{\widetilde{z}})}$. If $\widetilde{z} = S$ then the piecewise affine function $w_h = \alpha_{\widetilde{z}} - v_h$ vanishes at x_S and the estimate follows. If $\widetilde{z} = z \in \Gamma_D$ then there exist $S \in \mathcal{S}_h \cap \Gamma_D$ with $z \in S$ and we have $\alpha_{\widetilde{z}} = 0$ and $v_h(x_S) = 0$ and again the estimate follows. Finally, for $\widetilde{z} = z \in \mathcal{N}_h \setminus \Gamma_D$ we choose a side $S \in \mathcal{S}_h$ with $z \in S$ and replace v_h by $v_h - v_h(x_S)$ which corresponds to replacing α_z by $\alpha_z - v_h(x_S)$. In particular, the difference $\alpha_z - v_h$ remains unchanged. Hence, we assume $v_h(x_S) = 0$ and estimate, using $\|v_h\|_{L^\infty(T)} \le c h_z^{-d/p} \|v_h\|_{L^p(T)}$,

$$\|\alpha_z\|_{L^p(\omega_z)} \le |\alpha_z| |\omega_z|^p \le c h_z^{d/p} n_z^{-1} \sum_{T \in \mathcal{T}_h,\, z \in T} \big| v_h|_T(z) \big| \le c \|v_h\|_{L^p(\omega_z)}.$$

Using that $v_h(x_S) = 0$ this implies that

$$\|\alpha_z - v_h\|_{L^p(\omega_z)} \le c \|v_h\|_{L^p(\omega_z)} \le c h_z \|\nabla_h v_h\|_{L^p(\omega_z)}.$$

The estimate $\|E_h^{cr} v_h - v_h\|_{L^p(\Omega)} \le c h \|\nabla_h v_h\|_{L^p(\Omega)}$ follows from noting that for every $\widetilde{z} \in \mathcal{N}_h \cup \mathcal{S}_h$ we may choose $S \in \mathcal{S}_h$ belonging to $\omega_{\widetilde{z}}$ and hence

$$\begin{aligned} \|E_h^{cr} v_h - v_h\|_{L^p(\omega_{\widetilde{z}})} &\le \|E_h^{cr} v_h - v_h(x_S)\|_{L^p(\omega_{\widetilde{z}})} + \|v_h - v_h(x_S)\|_{L^p(\omega_{\widetilde{z}})} \\ &\le c h_{\widetilde{z}} \big(\|\nabla E_h^{cr} v_h\|_{L^p(\omega_{\widetilde{z}})} + \|\nabla_h v_h\|_{L^p(\omega_{\widetilde{z}})} \big), \end{aligned}$$

since $E_h^{cr} v_h(x_S) = v_h(x_S)$.

A.2. **Proof of inequality** (5)**.** We assume that $\varphi \in C^1(\mathbb{R}_{\geq 0})$ is convex and that $r \mapsto \varphi'(r)/r$ is positive, nonincreasing, and continuous on $\mathbb{R}_{\geq 0}$ and follow [12]. For $a, b \in \mathbb{R}^d$ the identity $2b\cdot(b-a) = |b|^2 - |a|^2 + |b-a|^2$ yields that

$$\frac{\varphi'(|a|)}{|a|} b\cdot(b-a) = \frac{1}{2}\frac{\varphi'(|a|)}{|a|}\big(|b|^2 - |a|^2\big) + \frac{1}{2}\frac{\varphi'(|a|)}{|a|}|b-a|^2.$$

Since $r \mapsto \varphi'(r)/r$ is nonincreasing, the function $\widetilde{\varphi}(y) = \varphi(y^{1/2})$ is concave on $\mathbb{R}_{\geq 0}$, so that we have

$$\widetilde{\varphi}'(y)(z-y) \geq \widetilde{\varphi}(z) - \widetilde{\varphi}(y),$$

for all $y, z \geq 0$. With $y = |a|^2$ and $z = |b|^2$ we deduce that

$$\frac{1}{2}\frac{\varphi'(|a|)}{|a|}\big(|b|^2 - |a|^2\big) \geq \varphi(|b|) - \varphi(|a|).$$

Combining these inequalities implies the asserted inequality

$$\frac{\varphi'(|a|)}{|a|} b\cdot(b-a) \geq \varphi(|b|) - \varphi(|a|) + \frac{1}{2}\frac{\varphi'(|a|)}{|a|}|b-a|^2.$$

A.3. p**-Dirichlet energies.** For the p-Dirichlet energy defined via $\phi(a) = |a|^p/p$ it is shown in [27] that if $u \in W^{1,p}_D(\Omega)$ is minimal then we have with $F(a) = |a|^{p/2-1}a$

$$c_p\|F(\nabla u) - F(\nabla v)\|^2 \leq I(v) - I(u).$$

The estimate carries over to the discretized functional I_h using the Crouzeix–Raviart method. It is shown in [26] via Taylor approximations that with $S(a) = D\phi(a) = |a|^{p-2}a$ and $\varphi_{|a|}(|c|) = (|a| + |c|)^{p-2}|c|^2$ we have

$$(S(a) - S(b))\cdot(a-b) \approx |F(a) - F(b)|^2 \approx \varphi_{|a|}(|a-b|).$$

The relations hold also for the functionals $\widetilde{S}$ and $\widetilde{F}$ which are obtained by replacing p by $p' = p/(p-1)$. The article [28] implies the estimate

$$\|F(\nabla u) - F(\nabla_h \mathcal{I}_{cr} u)\| \leq ch\|\nabla F(\nabla u)\|,$$

provided that $F(\nabla u) \in W^{1,2}(\Omega; \mathbb{R}^d)$.

Acknowledgments. (i) Some time before his death John W. Barrett told me that he was working on a generalization of Marini's formula, which unfortunately he was not able to complete. His comment was an important inspiration for the results of this article. I believe that many of the ideas presented here were known to him.
(ii) The author acknowledges support by the DFG via the priority programmes SPP 1748 *Reliable Simulation Techniques in Solid Mechanics: Development of Non-standard Discretization Methods, Mechanical and Mathematical Analysis* and SPP 1962 *Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization.*

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[2] M. G. Armentano and R. G. Durán. Asymptotic lower bounds for eigenvalues by nonconforming finite element methods. *Electron. Trans. Numer. Anal.*, 17:93–101, 2004.

[3] D. N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.*, 19(1):7–32, 1985.

[4] G. Aronsson. On the partial differential equation ${u_x}^2 u_{xx} + 2u_x u_y u_{xy} + {u_y}^2 u_{yy} = 0$. *Ark. Mat.*, 7:395–425 (1968), 1968.

[5] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006. Applications to PDEs and optimization.

[6] J. W. Barrett and W. B. Liu. Finite element approximation of the p-Laplacian. *Math. Comp.*, 61(204):523–537, 1993.

[7] J. W. Barrett and L. Prigozhin. A mixed formulation of the Monge-Kantorovich equations. *M2AN Math. Model. Numer. Anal.*, 41(6):1041–1060, 2007.

[8] J. W. Barrett and L. Prigozhin. Sandpiles and superconductors: nonconforming linear finite element approximations for mixed formulations of quasi-variational inequalities. *IMA J. Numer. Anal.*, 35(1):1–38, 2015.

[9] S. Bartels. Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.*, 50(3):1162–1180, 2012.

[10] S. Bartels. *Numerical methods for nonlinear partial differential equations*, volume 47 of *Springer Series in Computational Mathematics*. Springer, Cham, 2015.

[11] S. Bartels. *Numerical approximation of partial differential equations*, volume 64 of *Texts in Applied Mathematics*. Springer, [Cham], 2016.

[12] S. Bartels, L. Diening, and R. H. Nochetto. Unconditional stability of semi-implicit discretizations of singular flows. *SIAM J. Numer. Anal.*, 56(3):1896–1914, 2018.

[13] S. Bartels, R. H. Nochetto, and A. J. Salgado. Discrete total variation flows without regularization. *SIAM J. Numer. Anal.*, 52(1):363–385, 2014.

[14] S. Bartels, R. H. Nochetto, and A. J. Salgado. A total variation diminishing interpolation operator and applications. *Math. Comp.*, 84(296):2569–2587, 2015.

[15] S. Bartels and P. Schön. Adaptive approximation of the Monge-Kantorovich problem *via* primal-dual gap estimates. *ESAIM Math. Model. Numer. Anal.*, 51(6):2237–2261, 2017.

[16] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[17] D. Braess. An a posteriori error estimate and a comparison theorem for the nonconforming P_1 element. *Calcolo*, 46(2):149–155, 2009.

[18] S. C. Brenner. Forty years of the Crouzeix-Raviart element. *Numer. Methods Partial Differential Equations*, 31(2):367–396, 2015.

[19] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[20] C. Carstensen and J. Gedicke. Guaranteed lower bounds for eigenvalues. *Math. Comp.*, 83(290):2605–2629, 2014.

[21] C. Carstensen and K. Köhler. Nonconforming FEM for the obstacle problem. *IMA J. Numer. Anal.*, 37(1):64–93, 2017.

[22] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40(1):120–145, 2011.

[23] A. Chambolle and T. Pock. Crouzeix-Raviart approximation of the total variation on simplicial meshes. working paper or preprint, June 2019.

[24] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[25] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(no. , no. R-3):33–75, 1973.

[26] L. Diening, C. Ebmeyer, and M. Růžička. Optimal convergence for the implicit space-time discretization of parabolic systems with p-structure. *SIAM J. Numer. Anal.*, 45(2):457–472, 2007.

[27] L. Diening and C. Kreuzer. Linear convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM J. Numer. Anal.*, 46(2):614–638, 2008.

[28] L. Diening and M. Růžička. Interpolation operators in Orlicz-Sobolev spaces. *Numer. Math.*, 107(1):107–129, 2007.

[29] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[30] L. C. Evans. Partial differential equations and Monge-Kantorovich mass transfer. In *Current developments in mathematics, 1997 (Cambridge, MA)*, pages 65–126. Int. Press, Boston, MA, 1999.

[31] M. Fortin and R. Glowinski. *Augmented Lagrangian methods*, volume 15 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.

[32] L. Gastaldi and R. Nochetto. Optimal L^∞-error estimates for nonconforming and mixed finite element methods of lowest order. *Numer. Math.*, 50(5):587–611, 1987.

[33] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[34] T. Gudi. A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.*, 79(272):2169–2189, 2010.

[35] P. Hansbo and M. G. Larson. Discontinuous Galerkin and the Crouzeix-Raviart element: application to elasticity. *M2AN Math. Model. Numer. Anal.*, 37(1):63–72, 2003.

[36] M. Herrmann, R. Herzog, S. Schmidt, J. Vidal-Núñez, and G. Wachsmuth. Discrete total variation with finite elements and applications to imaging. *J. Math. Imaging Vision*, 61(4):411–431, 2019.

[37] M. Hintermüller and K. Kunisch. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM J. Appl. Math.*, 64(4):1311–1333, 2004.

[38] D. J. Liu, A. Q. Li, and Z. R. Chen. Nonconforming FEMs for the p-Laplace problem. *Adv. Appl. Math. Mech.*, 10(6):1365–1383, 2018.

[39] W. Liu and N. Yan. Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian. *Numer. Math.*, 89(2):341–378, 2001.

[40] L. D. Marini. An inexpensive method for the evaluation of the solution of the lowest order Raviart-Thomas mixed method. *SIAM J. Numer. Anal.*, 22(3):493–496, 1985.

[41] A. M. Oberman. Finite difference methods for the infinity Laplace and p-Laplace equations. *J. Comput. Appl. Math.*, 254:65–80, 2013.

[42] C. Ortner. Nonconforming finite-element discretization of convex variational problems. *IMA J. Numer. Anal.*, 31(3):847–864, 2011.

[43] T. Pryer. On the finite-element approximation of ∞-harmonic functions. *Proc. Roy. Soc. Edinburgh Sect. A*, 148(4):819–834, 2018.

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

[45] S. I. Repin. A posteriori error estimation for variational problems with uniformly convex functionals. *Math. Comp.*, 69(230):481–500, 2000.

[46] R. T. Rockafellar. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

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`