ERROR ESTIMATES FOR A CLASS OF DISCONTINUOUS GALERKIN METHODS FOR NONSMOOTH PROBLEMS VIA CONVEX DUALITY RELATIONS

SÖREN BARTELS

Abstract. We devise and analyze a class of interior penalty discontinuous Galerkin methods for nonlinear and nonsmooth variational problems. Discrete duality relations are derived that lead to optimal error estimates in the case of total-variation regularized minimization or obstacle problems. The analysis provides explicit estimates that precisely determine the role of stabilization parameters. Numerical experiments suppport the optimality of the estimates.

1. Introduction

Total-variation minimization. As a particular example of a nonsmooth convex variational problem we consider the total variation regularized optimization problem that determines a function $u \in BV(\Omega) \cap L^2(\Omega)$ via a minimization of

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

where $|Du|(\Omega)$ is the total variation of $u \in L^2(\Omega)$, which coincides with $\|\nabla u\|_{L^1(\Omega)}$ if $u \in W^{1,1}(\Omega)$, while $\alpha > 0$ and $g \in L^2(\Omega)$ are given data, cf. [1, 4, 16] for analytical features and numerical methods. Since discontinuous solutions are expected and since continuous methods are known to provide suboptimal results [8, 10], it is attractive to discretize the minimization problem by a discontinuous finite element method, e.g., via determining an elementwise affine, possibly discontinuous function $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ on a triangulation $\mathcal{T}_h$ as a minimizer of the functional

$$I_h(u_h) = \int_\Omega |\nabla_h u_h| \,\mathrm{d}x + \frac{1}{r}\int_{\mathcal{S}_h\setminus\partial\Omega} h_{\mathcal{S}}^{-\gamma r} |[\![u_h]\!]_h|^r \,\mathrm{d}s + \frac{\alpha}{2}\|\Pi_h(u_h-g)\|^2.$$

Here, ∇_h denotes the elementwise application of the gradient, $\mathcal{S}_h$ stands for the union of element sides in $\mathcal{T}_h$, the operator Π_h is the projection onto piecewise constant functions or vector fields on $\mathcal{T}_h$, the function $h_{\mathcal{S}} : \mathcal{S}_h \to \mathbb{R}_{>0}$ is a mesh-size function, and $[\![\cdot]\!]$ and $[\![\cdot]\!]_h$ denote the jump and the mean of a jump of a piecewise polynomial function. Since u_h is piecewise affine

---

Date: April 20, 2020.

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

Key words and phrases. Nonsmooth problems, discontinuous Galerkin method, error estimates, total variation, variational inequalities.

we have that $\Pi_h u_h|_T = u_h(x_T)$ for every element $T \in \mathcal{T}_h$ with midpoint (barycenter) x_T and

$$[\![u_h]\!]_h|_S = [\![u_h]\!](x_S) = \lim_{\varepsilon \to 0} u_h(x + \varepsilon n_S) - u_h(x_S - \varepsilon n_S)$$

for every side $S \in \mathcal{S}_h$ with midpoint (barycenter) x_S and unit normal n_S. The second term in the discrete energy functional I_h thus penalizes averages of jumps across interelement sides. The use of the mean has the alternative interpretation of using quadrature which makes the scheme practical. The choice of the parameters r and γ is crucial for an accurate approximation of the exact solution u. We remark that our approach is motivated and inspired by recent results in [18, 10] on discretizations of nonsmooth problems using Crouzeix–Raviart elements.

To quantify the accuracy we define a discrete dual problem. We show that a naturally associated maximization problem is defined for a discrete vector field $z_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ by the functional

$$\begin{aligned} D_h(z_h) = &- I_{K_1(0)}(\Pi_{0,h} z_h) - \frac{1}{r'} \int_{\mathcal{S}_h \setminus \Gamma_N} h_{\mathcal{S}}^{\gamma r'} |\{z_h \cdot n_{\mathcal{S}}\}|^{r'} \,\mathrm{d}s \\ &- \frac{1}{2\alpha} \| \operatorname{div} z_h + \alpha g_h\|^2 + \frac{\alpha}{2} \|g_h\|^2, \end{aligned}$$

where $r' = r/(r-1)$. The Raviart–Thomas finite element space $\mathcal{RT}^0_N(\mathcal{T}_h)$ consists of certain elementwise affine vector fields whose weak divergence is a function and whose normal component vanishes on $\Gamma_N = \partial\Omega$. In particular, the normal components $z_h \cdot n_{\mathcal{S}}$ are continuous and constant on element sides, so that their averages $\{z_h \cdot n_{\mathcal{S}}\}$ coincide with the values $z_h \cdot n_{\mathcal{S}}$ on every side. It turns out that the penalty terms in the discrete primal problem I_h are related to stabilizing terms on element sides in the discrete dual functional D_h. The indicator functional $I_{K_1(0)}$ enforces the length of the vector field z_h to be bounded by 1 at element midpoints. Thus, in the discrete duality relation, jumps of functions in the primal problem lead to averages in the dual formulation. On the continuous level the dual formulation consists in maximizing the functional

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

in the set of vector fields $z \in W^2_N(\operatorname{div};\Omega)$ whose distributional divergence belongs to $L^2(\Omega)$ and whose normal component vanishes on $\Gamma_N = \partial\Omega$. Strong duality applies, i.e., we have $I(u) = D(z)$ for solutions u and z, which is a well-posedness property of the variational problem.

An error estimate follows from coercivity properties of I_h and the crucial discrete duality relation $I_h(u_h) \geq D_h(z_h)$. More precisely, with appropriately defined quasi-interpolants $\mathcal{I}_h u$, that is continuous at midpoints of element sides, and $\mathcal{J}_h z$, for a sufficiently regular solution z of the continuous dual problem, we have

$$\frac{\alpha}{2} \|\Pi_h(u_h - \mathcal{I}_h u)\|^2 \leq I_h(\mathcal{I}_h u) - I_h(u_h) \leq I_h(\mathcal{I}_h u) - D_h(\mathcal{J}_h z).$$

The quasi-interpolants are defined in such a way that we have

$$\nabla_h \mathcal{I}_h u = \Pi_h \nabla u, \quad \operatorname{div} \mathcal{J}_h z = \Pi_h \operatorname{div} z.$$

These relations allow us apply Jensen's inequality, which in the present context has the interpretation of a total-variation diminishing interpolant, and thereby leads to the discrete error estimate

$$\frac{\alpha}{2} \|\Pi_h(u_h - \mathcal{I}_h u)\| \le c h^{1/2},$$

provided that $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$, $u \in L^\infty(\Omega)$, and $\gamma r' \ge 2$ or $\gamma \ge 0$ if $r = 1$. The estimate implies the error bound

$$\|u - \Pi_h u_h\| \le c h^{1/2},$$

where $\Pi_h u_h$ can be replaced by u_h provided that the sequence $(u_h)_{h>0}$ is uniformly bounded in $L^\infty(\Omega)$. The convergence rate $\mathcal{O}(h^{1/2})$ coincides with the rate for Crouzeix–Raviart finite element methods, cf. [18, 10], and is quasi-optimal for the approximation of a generic function in $BV(\Omega) \cap L^\infty(\Omega)$. The optimal rate can in general not be obtained with continuous finite element methods [7, 8]. Note that since $\gamma r = 0$ is allowed the approach is not a pure penalty method. On the other hand, it does not suffer from locking effects for large values of γr owing to the use of quadrature in the jump contributions. We refer the reader to [14] for discretizations using finite difference methods.

General error analysis. The analysis summarized for the numerical approximation of the total variation-regularized problem by discontinuous methods can be generalized in several ways. For this, we consider a convex mimimization problem defined via (a suitable extension of) the functional

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

on a Sobolev space $W^{1,p}_D(\Omega)$ of functions with vanishing traces on $\Gamma_D \subset \partial\Omega$. The dual formulation is given by a maximization of the functional

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

in a space $W^q_N(\operatorname{div};\Omega)$ of vector fields in $L^q(\Omega;\mathbb{R}^d)$ whose normal components vanish on $\Gamma_N = \partial\Omega \setminus \Gamma_D$ and whose distributional divergence belongs to $L^q(\Omega)$, where $q = p'$ is the conjugate exponent of p.

A class of discontinuous Galerkin discretizations of the primal problem is given by the functionals

$$I_h(u_h) = \int_\Omega \phi(\nabla_h u_h) + \psi_h(x, \Pi_h u_h) \,\mathrm{d}x \\ + \frac{1}{r} \int_{\mathcal{S}_h \setminus \Gamma_N} \alpha_{\mathcal{S}}^{-r} |[\![u_h]\!]_h|^r \,\mathrm{d}s + \frac{1}{s} \int_{\mathcal{S}_h \setminus \Gamma_D} \beta_{\mathcal{S}}^s |\{u_h\}_h|^s \,\mathrm{d}s$$

with suitable exponents $r, s \geq 1$ and weights $\alpha_\mathcal{S}, \beta_\mathcal{S} : \mathcal{S}_h \to \mathbb{R}_{\geq 0}$ on the element sides. The quantities $\{u_h\}$ are on every side the average of traces of u_h from adjacent elements, its mean $\{u_h\}_h$ coincides for piecewise affine functions with the evaluation at the midpoint of S, i.e.,

$$\{u_h\}_h = \{u_h\}(x_S) = \lim_{\varepsilon\to 0} \frac{1}{2}\big(u_h(x + \varepsilon n_S) + u_h(x_S - \varepsilon n_S)\big).$$

We show that a discrete duality argument leads to the discrete dual functional

$$\begin{aligned} D_h(z_h) = &- \int_\Omega \phi^*(\Pi_{0,h} z_h) + \psi_h^*(x, \operatorname{div}_h z_h) \,\mathrm{d}x \\ &- \frac{1}{r'} \int_{\mathcal{S}_h \setminus \Gamma_N} \alpha_\mathcal{S}^{r'} |\{z_h \cdot n\}|^{r'} \,\mathrm{d}s - \frac{1}{s'} \int_{\mathcal{S}_h \setminus \Gamma_D} \beta_\mathcal{S}^{-s'} |[\![z_h \cdot n]\!]|^{s'} \,\mathrm{d}s. \end{aligned}$$

This functional is defined on a broken Raviart–Thomas finite element space $\mathcal{RT}^{0,dg}(\mathcal{T}_h)$. Here, jumps of the normal component of z_h across element sides are penalized which corresponds to the presence of the averages $\{u_h\}$ on element sides in the discrete primal problem. The duality of jumps and averages on element sides is a result of an elementwise integration by parts and the elementary formula for inner sides

$$[\![u_h z_h \cdot n_S]\!] = [\![u_h]\!]\{z_h \cdot n_S\} + \{u_h\}[\![z_h \cdot n_S]\!],$$

that relates jumps of products to products of jumps and averages. Together with Fenchel's inequality it leads to the important discrete duality relation

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

for solutions u_h and z_h of the discrete primal and dual problems, respectively. As above, this inequality is important for an error analysis. In particular, it provides full control on nonconformity errors which otherwise require the use of a Strang lemma or suitable reconstruction operatos, cf., e.g., [19, 12]. Here, these error contributions are entirely controlled via structure-preserving features of the discretizations.

Nonlinear Dirichet and obstacle problems. For a class of nonlinear Dirichlet problems with linear low order terms given by

$$\psi(x, s) = -f(x)s$$

we obtain with an appropriate coercivity functional σ the general and constant-free discrete error estimate

$$\begin{aligned} \int_\Omega \sigma(\nabla_h u_h, \nabla_h \mathcal{I}_h u) \,\mathrm{d}x \leq &\int_\Omega \big(D\phi^*(z) - D\phi^*(\Pi_h \mathcal{J}_h z)\big) \cdot (z - \Pi_h \mathcal{J}_h z) \,\mathrm{d}x \\ &+ \frac{1}{r'} \|\alpha_\mathcal{S}\{\mathcal{J}_h z \cdot n_\mathcal{S}\}\|_{L^{r'}(\mathcal{S}_h \setminus \Gamma_N)}^{r'} + \frac{1}{s'} \|\beta_\mathcal{S}\{\mathcal{I}_h u\}_h\|_{L^{s'}(\mathcal{S}_h \setminus \Gamma_D)}^{s'}, \end{aligned}$$

with quasi-interpolants $\mathcal{I}_h u$ and $\mathcal{J}_h z$ of sufficiently regular primal and dual solutions u and z, respectively. The concepts also apply to obstacle problems, for which the low order term is given by

$$\psi(x,s) = -f(x)s + I_{\mathbb{R}_{\geq 0}}(s),$$

with the indicator function $I_{\mathbb{R}_{\geq 0}}$ that enforces the solution u of the primal problem to be nonnegative. In the case of a quadratic functional $\phi(v) = |v|^2/2$ and with $r = s = 2$ we obtain for regular solutions $u \in W^{1,2}_D(\Omega) \cap W^{2,2}(\Omega)$ the constant-free discrete error estimate

$$\begin{aligned}\frac{1}{2}\|\nabla_h(u_h - \mathcal{I}_h u)\|^2 \leq \frac{1}{2}&\|z - \Pi_h \mathcal{J}_h z\|^2 \\ &+ \|f + \Delta u\|\big(\|u - \mathcal{I}_h u\| + h\|\nabla_h(u - \mathcal{I}_h u)\|\big) \\ &+ \frac{1}{2}\|\alpha_{\mathcal{S}}\{\mathcal{J}_h z \cdot n_{\mathcal{S}}\}\|^2_{L^2(\mathcal{S}_h \setminus \Gamma_N)} + \frac{1}{2}\|\beta_{\mathcal{S}}\{\mathcal{I}_h u\}_h\|^2_{L^2(\mathcal{S}_h \setminus \Gamma_D)}.\end{aligned}$$

The right-hand side is of quadratic order if

$$\big\|\alpha_{\mathcal{S}} h_{\mathcal{S}}^{-3/2}\big\|_{L^\infty(\mathcal{S}_h)} + \big\|\beta_{\mathcal{S}} h_{\mathcal{S}}^{-3/2}\big\|_{L^\infty(\mathcal{S}_h)} \leq c_{\mathcal{T}},$$

i.e., if $\alpha_{\mathcal{S}} = \beta_{\mathcal{S}} = \mathcal{O}(h_{\mathcal{S}}^{3/2})$. Note that $\alpha_{\mathcal{S}} > 0$ is needed for well-posedness of the discrete problem while if $\beta_{\mathcal{S}} = 0$ then the contribution to D_h involving the jumps $[\![z_h \cdot n_{\mathcal{S}}]\!]$ becomes an indicator functional and the space $\mathcal{RT}^{0,dg}(\mathcal{T}_h)$ has to be replaced by $\mathcal{RT}^0_N(\mathcal{T}_h)$.

Relations to other methods. In contrast to established discontinuous Galerkin methods for elliptic problems as in [2, 3, 15, 21], which are typically derived from strong formulations of partial differential equations, we obtain here more restrictive conditions on the penalty parameters in the case of differentiable elliptic equations to obtain quasi-optimal error estimates. In these cases the simple interior penalty approach realized by the discrete minimization problems I_h is inconsistent with the weak formulations of the corresponding partial differential equations. While the variational approach for penalty based discontinuous Galerkin methods for variational problems considered in [13] shows that convergence is guaranteed under mild conditions, our numerical experiments confirm that they are not sufficient to obtain optimal convergence rates. For the nondifferentiable total-variation problem our discretizations do not require penalizations and our approach is conistent with a natural discretization of the total-variation norm. Generally, the error analysis used here only uses optimality conditions in terms of first order system and subdifferentials. Another advantage of our error analysis based on duality arguments is that it provides explicit estimates that do not require absorbing terms and hence precisely determine the role of the parameters involved in the discontinuous Galerkin discretization. Additionally, it leads to a realistic and optimal regularity condition in terms of the dual solution. Throughout this article we use simplicial partitions which is important for the error analysis. Our duality arguments transfer verbatimly to general classes of polyhedral partitions.

Outline. The outline of this article is as follows. In Section 2 we introduce notation and define appropriate finite element spaces. A discrete duality theory is provided in Section 3. The application to nonlinear Dirichlet problems, total variation minimization, and elliptic obstacle problems is discussed in the subsequent Sections 4-6. Numerical experiments are presented in Section 7.

2. Notation and finite element spaces

For a sequence of regular triangulations $(\mathcal{T}_h)_{h>0}$, where $h > 0$ refers to a maximal mesh-size that tends to zero, the set of elementwise polynomial functions or vector fields of maximal polynomial degree $k \geq 0$ is defined by

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

We let $\Pi_h : L^1(\Omega;\mathbb{R}^\ell) \to \mathcal{L}^0(\mathcal{T}_h)^\ell$ denote the L^2 projection onto elementwise constant functions or vector fields and note that Π_h is self-adjoint, i.e.,

$$\int_\Omega \Pi_h f g \,\mathrm{d}x = \int_\Omega f \Pi_h g \,\mathrm{d}x$$

for all $f, g \in L^1(\Omega)$. We let $\mathcal{S}_h$ denote the set of sides of elements and define the mesh-size function $h_\mathcal{S}|_S = h_S = \operatorname{diam}(S)$ for all sides $S \in \mathcal{S}_h$. We let $n_\mathcal{S} : \mathcal{S}_h \to \mathbb{R}^d$ denote a unit vector field given for every side $S \in \mathcal{S}_h$ by

$$n_\mathcal{S}|_S = n_S$$

for a fixed unit normal n_S on S which is assumed to coincide with the outer unit normal if $S \subset \partial\Omega$. The jump and average on a side S of a function $v_h \in \mathcal{L}^k(\mathcal{T}_h)^\ell$ are for $x \in S$ defined for inner sides via

$$[\![v_h]\!](x) = \lim_{\varepsilon \to 0} \big(v_h(x - \varepsilon n_S) - v_h(x + \varepsilon n_S) \big),$$
$$\{v_h\}(x) = \lim_{\varepsilon \to 0} \frac{1}{2} \big(v_h(x - \varepsilon n_S) + v_h(x + \varepsilon n_S) \big).$$

For $S \subset \partial\Omega$ we set

$$[\![v_h]\!] = \{v_h\} = v_h.$$

The integral means of jumps and averages are denoted by

$$[\![v_h]\!]_h = |S|^{-1} \int_S [\![v_h]\!] \,\mathrm{d}s, \quad \{v_h\}_h = |S|^{-1} \int_S \{v_h\} \,\mathrm{d}s,$$

which in case of elementwise affine functions coincides with the evaluation at the midpoint x_S for every $S \in \mathcal{S}_h$. We define the space of discontinuous, piecewise linear functions via

$$\mathcal{S}^{1,dg}(\mathcal{T}_h) = \mathcal{L}^1(\mathcal{T}_h).$$

A space of discontinuous vector fields is given by

$$\mathcal{R}T^{0,dg}(\mathcal{T}_h) = \mathcal{L}^0(\mathcal{T}_h)^d + (\operatorname{id} - x_\mathcal{T})\mathcal{L}^0(\mathcal{T}_h),$$

where id is the identity and $x_{\mathcal{T}} = \Pi_h \operatorname{id} \in \mathcal{L}^0(\mathcal{T}_h)^d$ the elementwise constant vector field that coincides with the midpoint x_T on every $T \in \mathcal{T}_h$. Differential operators on these spaces are defined elementwise, indicated by a subscript h, i.e., we have

$$\nabla_h v_h|_T = \nabla(v_h|_T), \quad \operatorname{div}_h z_h|_T = \operatorname{div}(z_h|_T)$$

for $v_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$, $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ and all $T \in \mathcal{T}_h$. The operators are also applied to weakly differentiable functions and vector fields in which case they coincide with the weak gradient and the weak divergence. By construction, any vector field $y_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ has a piecewise constant normal component $y_h \cdot n_L$ along straight lines L with normal n_L. Subspaces of elementwise affine functions and vector fields with certain continuity properties on element sides are given by

$$\mathcal{S}^{1,cr}_D(\mathcal{T}_h) = \{v_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h) : [\![v_h]\!]_h|_S = 0 \text{ for all } S \in \mathcal{S}_h \setminus \Gamma_N\},$$

and

$$\mathcal{RT}^0_N(\mathcal{T}_h) = \{y_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h) : [\![y_h \cdot n_S]\!]_h|_S = 0 \text{ for all } S \in \mathcal{S}_h \setminus \Gamma_D\},$$

which coincide with low order Crouzeix–Raviart and Raviart–Thomas finite element spaces introduced in [20, 26]. These spaces provide quasi-interpolation operators

$$\mathcal{I}_h : W^{1,p}_D(\Omega) \to \mathcal{S}^{1,cr}_D(\mathcal{T}_h), \quad \mathcal{J}_h : W^q_N(\operatorname{div};\Omega) \to \mathcal{RT}^0_N(\mathcal{T}_h),$$

with the projection properties

$$\nabla_h \mathcal{I}_h v = \Pi_h \nabla v, \quad \operatorname{div} \mathcal{J}_h y = \Pi_h \operatorname{div} y,$$

and the interpolation estimates

$$\begin{aligned} \|v - \mathcal{I}_h v\|_{L^p(\Omega)} &\le c_{\mathcal{I},1} h \|\nabla v\|_{L^p(\Omega)}, \\ \|v - \mathcal{I}_h v\|_{L^p(\Omega)} + h\|\nabla_h(v - \mathcal{I}_h v)\|_{L^p(\Omega)} &\le c_{\mathcal{I},2} h^2 \|D^2 v\|_{L^p(\Omega)}, \end{aligned}$$

for $v \in W^{1,p}_D(\Omega)$ with $1 \le p \le \infty$, and

$$\|y - \mathcal{J}_h y\|_{L^q(\Omega)} \le c_{\mathcal{J}} h \|\nabla y\|_{L^q(\Omega)}$$

for $y \in W^q_N(\operatorname{div};\Omega)$ with $1 \le q \le \infty$. We always assume that h is sufficiently small so that we have $\|\mathcal{I}_h v\|_{L^p(\Omega)} \le \sqrt{2}\|v\|_{W^{1,p}(\Omega)}$ and $\|\mathcal{J}_h y\|_{L^q(\Omega)} \le \sqrt{2}\|y\|_{W^{1,q}(\Omega)}$. We refer the reader to [11, 12, 5] for details. Elementary calculations lead to the identities

$$[\![v_h y_h \cdot n_S]\!] = \begin{cases} [\![v_h]\!]\{y_h \cdot n_S\} + \{v_h\}[\![y_h \cdot n_S]\!] & \text{if } S \not\subset \partial\Omega, \\ [\![v_h]\!]\{y_h \cdot n_S\} & \text{if } S \subset \Gamma_D, \\ \{v_h\}[\![y_h \cdot n_S]\!] & \text{if } S \subset \Gamma_N. \end{cases}$$

By carrying out an elementwise integration by parts we thus find that for $v_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ we have

$$(1)\qquad \begin{aligned} &\int_\Omega v_h \operatorname{div} y_h \,\mathrm{d}x + \int_\Omega \nabla_h v_h \cdot y_h \,\mathrm{d}x \\ &\qquad = \int_{\mathcal{S}_h\setminus\Gamma_N} [\![v_h]\!]_h \{y_h \cdot n_S\} \,\mathrm{d}s + \int_{\mathcal{S}_h\setminus\Gamma_D} \{v_h\}_h [\![y_h\cdot n_S]\!] \,\mathrm{d}s. \end{aligned}$$

If $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ then the terms on the right-hand side are equal to zero. We furthermore note that if an elementwise constant vector field $y_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ satisfies

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

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ then it belongs to $\mathcal{RT}^0_N(\mathcal{T}_h)$. This fact follows from an elementwise integration by parts with $v_h = \varphi_S$ for the Crouzeix–Raviart basis functions φ_S associated with sides $S \in \mathcal{S}_h \setminus \Gamma_D$. To bound functionals defined by integrals on the skeleton $\mathcal{S}_h$ we use the discrete trace inequality

$$(2)\qquad \|h_\mathcal{S} \psi_h\|^s_{L^s(\mathcal{S}_h)} \le c_\mathcal{T} \|\psi_h\|^s_{L^s(\Omega)},$$

for a piecewise linear function $\psi_h \in \mathcal{L}^1(\mathcal{T}_h)$, $s \ge 1$, and a constant $c_\mathcal{T}$ that depends on the geometry of $\mathcal{T}_h$.

3. Discrete conjugation

We collect the jump and average terms needed for the discontinuous Galerkin discretizations in functionals J_h and K_h. The results of this section apply to general classes of regular polyhedral partitions.

Definition 3.1 (Jumps and averages)**.** *Let $r, s \ge 1$ and let $\alpha_\mathcal{S}, \beta_\mathcal{S} : \mathcal{S}_h \to \mathbb{R}_{\ge 0}$ be piecewise constant. For $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ and $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ define*

$$\begin{aligned} J_h(u_h) &= \frac{1}{r}\|\alpha_\mathcal{S}^{-1} [\![u_h]\!]_h\|^r_{L^r(\mathcal{S}_h\setminus\Gamma_N)} + \frac{1}{s}\|\beta_\mathcal{S}\{u_h\}_h\|^s_{L^s(\mathcal{S}_h\setminus\Gamma_D)}, \\ K_h(z_h) &= \frac{1}{r'}\|\alpha_\mathcal{S}\{z_h\cdot n_\mathcal{S}\}\|^{r'}_{L^{r'}(\mathcal{S}_h\setminus\Gamma_N)} + \frac{1}{s'}\|\beta_\mathcal{S}^{-1}[\![z_h\cdot n_\mathcal{S}]\!]\|^{s'}_{L^{s'}(\mathcal{S}_h\setminus\Gamma_D)}, \end{aligned}$$

where we require $[\![u_h]\!]_h = 0$ if $\alpha_\mathcal{S} = 0$ and $[\![z_h\cdot n_\mathcal{S}]\!] = 0$ if $\beta_\mathcal{S} = 0$. For $r = 1$ or $s = 1$ the functionals $(1/r')\|\cdot\|^{r'}_{L^{r'}}$ or $(1/s')\|\cdot\|^{s'}_{L^{s'}}$ are interpreted as indicator functionals $I_{K_1(0)}$ of the closed unit ball $K_1(0)$.

To show that the functionals J_h and K_h are in discrete duality we let ϕ^* and ψ^* be the convex conjugates of the convex functions ϕ and ψ, i.e.,

$$\phi^*(y) = \sup_{v\in\mathbb{R}^n} y\cdot v - \phi(v), \quad \psi^*(x,s) = \sup_{t\in\mathbb{R}} s\,t - \psi(x,t).$$

For simple power functionals the conjugate is determined by the conjugate exponent, i.e., for a factor $c \ge 0$ and an exponent $\sigma \ge 1$ we have

$$g(v) = \frac{1}{\sigma} c^\sigma |v|^\sigma \quad \iff \quad g^*(w) = \begin{cases} \frac{1}{\sigma'} c^{-\sigma'} |w|^{\sigma'} & \text{for } \sigma > 1, \\ I_{K_1(0)}(c^{-1} w) & \text{for } \sigma = 1, \end{cases}$$

where $\sigma' = \sigma/(\sigma - 1)$ and $I_{K_1(0)}$ is the indicator functional of the closed unit ball around 0. The definition of g^* leads to Fenchel's inequality

$$v \cdot w \le g(v) + g^*(w), \tag{3}$$

where equality holds if and only if $w = Dg(v)$ or equivalently $v = Dg^*(w)$.

Proposition 3.2 (Discrete conjugation). *For $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ and $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ define*

$$V_h(u_h) = \int_\Omega \phi(\nabla_h u_h) \,\mathrm{d}x + J_h(u_h),$$
$$W_h(z_h) = -\int_\Omega \phi^*(\Pi_{0,h} z_h) \,\mathrm{d}x - K_h(z_h).$$

Given any $\overline{u}_h \in \mathcal{L}^0(\mathcal{T}_h)$ we have that

$$\inf \big\{ V_h(u_h) : u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h),\, \Pi_h u_h = \overline{u}_h \big\} \\ \ge \sup \big\{ W_h(z_h) - (\overline{u}_h, \operatorname{div}_h z_h) : z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h) \big\}.$$

Proof. For arbitrary $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ and $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ with $\Pi_h u_h = \overline{u}_h$ we use the integration-by-parts formula (1) to verify that

$$\begin{aligned} W_h(z_h) - (\overline{u}_h, \operatorname{div}_h z_h) = &-\int_\Omega \phi^*(\Pi_{0,h} z_h) \,\mathrm{d}x + \int_\Omega \nabla_h u_h \cdot \Pi_h z_h \,\mathrm{d}x \\ &- \frac{1}{r'} \int_{\mathcal{S}_h \setminus \Gamma_N} \alpha_{\mathcal{S}}^{r'} |\{z_h \cdot n_{\mathcal{S}}\}|^{r'} \,\mathrm{d}s - \int_{\mathcal{S}_h \setminus \Gamma_N} [\![u_h]\!]_h \{z_h \cdot n_{\mathcal{S}}\} \,\mathrm{d}s \\ &- \frac{1}{s'} \int_{\mathcal{S}_h \setminus \Gamma_D} \beta_{\mathcal{S}}^{-s'} |[\![z_h \cdot n_{\mathcal{S}}]\!]|^{s'} \,\mathrm{d}s - \int_{\mathcal{S}_h \setminus \Gamma_D} \{u_h\}_h [\![z_h \cdot n_{\mathcal{S}}]\!] \,\mathrm{d}s. \end{aligned}$$

With Fenchel's inequality we deduce that

$$-\phi^*(\Pi_{0,h} z_h) + \nabla_h u_h \cdot \Pi_h z_h \le \phi(\nabla_h u_h),$$

and

$$-\frac{1}{r'} \alpha_{\mathcal{S}}^{r'} |\{z_h \cdot n_{\mathcal{S}}\}|^{r'} + (-[\![u_h]\!]_h)\{z_h \cdot n_{\mathcal{S}}\} \le \frac{1}{r} \alpha_{\mathcal{S}}^{-r} |[\![u_h]\!]_h|^r,$$

as well as

$$-\frac{1}{s'} \beta_{\mathcal{S}}^{-s'} |[\![z_h \cdot n_{\mathcal{S}}]\!]|^{s'} + (-\{u_h\}_h)[\![z_h \cdot n_{\mathcal{S}}]\!] \le \frac{1}{s} \beta_{\mathcal{S}}^{s} |\{u_h\}_h|^s.$$

On combining the estimates and noting that u_h and z_h are arbitrary, we deduce the statement. □

The discrete convex conjugates lead to a canonical discrete dual problem.

Theorem 3.3 (Discrete duality)**.** *For $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ let*

$$I_h(u_h) = \int_\Omega \phi(\nabla_h u_h) + \psi_h(x, \Pi_h u_h) \,\mathrm{d}x + J_h(u_h),$$

where $\psi_h : \Omega \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is elementwise constant with respect to the first argument. Then with the discrete dual functional defined for $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$ by

$$D_h(z_h) = -\int_\Omega \phi^*(\Pi_h z_h) + \psi_h^*(x, \operatorname{div}_h z_h) \,\mathrm{d}x - K_h(z_h)$$

we have

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

Proof. With the inequality of Proposition 3.2 we have, using $\overline{u}_h = \Pi_h u_h$ that

$$\begin{aligned} &\int_\Omega \phi(\nabla_h u_h) \,\mathrm{d}x + J_h(u_h) + \int_\Omega \psi_h(x, \overline{u}_h) \,\mathrm{d}x \\ &\ge -\int_\Omega \phi^*(\Pi_{0,h} z_h) \,\mathrm{d}x - K_h(z_h) - (\overline{u}_h, \operatorname{div}_h z_h) + \int_\Omega \psi_h(x, \overline{u}_h) \,\mathrm{d}x. \end{aligned}$$

Fenchel's inequality shows that on every $T \in \mathcal{T}_h$ we have

$$\overline{u}_h \operatorname{div}_h z_h \le \psi_h(x, \overline{u}_h) + \psi_h^*(x, \operatorname{div}_h z_h).$$

This implies the asserted inequality. □

A strong duality relation can be established under additional conditions. We consider a particular but typical definition of the penalty terms. The formula stated in the following proposition provides a discrete dual solution via a simple postprocessing of the discreteprimal solution and generalizes a result from [25].

Proposition 3.4 (Reconstruction and strong duality)**.** *Assume that J_h and K_h are defined with $r = s = 2$ and $\beta_\mathcal{S} = 0$ and that ϕ and ψ are continuously differentiable. If $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ is minimal for I_h then the vector field*

$$\widetilde{z}_h = D\phi(\nabla_h u_h) + d^{-1} D\psi_h(\Pi_h u_h)(\mathrm{id} - x_\mathcal{T})$$

belongs to $\mathcal{RT}^0_N(\mathcal{T}_h)$ and is maximal for D_h with $I_h(u_h) = D_h(\widetilde{z}_h)$.

Proof. We note that u_h solves the discrete Euler–Lagrange equations

$$\begin{aligned} \int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h v_h \,\mathrm{d}x + \int_\Omega D\psi_h(\Pi_h u_h) v_h \,\mathrm{d}x \\ = -\int_{\mathcal{S}_h \setminus \Gamma_N} \alpha_\mathcal{S}^{-2} [\![u_h]\!]_h [\![v_h]\!]_h \,\mathrm{d}s \end{aligned}$$

for all $v_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ and the term on the right-hand side vanishes if $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. To show that $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ we choose $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ with $-\operatorname{div} y_h = D\psi_h(\Pi_h u_h)$. Then, $\widetilde{z}_h - y_h$ is elementwise constant and we have

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

for all $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. In particular, we deduce that $\widetilde{z}_h - y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ and therefore $\widetilde{z}_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$. Given a side $S \in \mathcal{S}_h \setminus \Gamma_D$ with adjacent element $T \in \mathcal{T}_h$ we let $\varphi_{S,T} \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ be the function that is supported on T, vanishes in those midpoints of sides that do not belong to S, and satisfies $\varphi_{S,T}(x_S) = 1$. We have $[\![\varphi_{S,T}]\!]_h = -1$ on S and the discrete Euler–Lagrange equations and an integration by parts yield that

$$\int_S \alpha_S^{-2} [\![u_h]\!]_h \,\mathrm{d}s = \int_S D\phi(\nabla_h u_h) \cdot n_S \,\mathrm{d}s + \frac{1}{d} \int_S D\psi_h(\Pi_h u_h)(x - x_T) \cdot n_S \,\mathrm{d}s$$

where we used $\int_T \varphi_{S,T} \,\mathrm{d}x = |T|/(d+1) = (x - x_T) \cdot n_S |S|/d$ for every $x \in S$. Since $\widetilde{z}_h \cdot n_{\mathcal{S}}$ is constant and continuous on $\mathcal{S}_h$ we deduce that $\alpha_{\mathcal{S}}^{-2}[\![u_h]\!]_h = \{\widetilde{z}_h \cdot n_{\mathcal{S}}\}$. Using the identities $\Pi_h \widetilde{z}_h = D\phi(\nabla_h u_h)$ and $\operatorname{div} \widetilde{z}_h = D\psi_h(\Pi_h u_h)$ and noting that these imply equality in (3) we find that

$$\begin{aligned} \phi(\nabla_h u_h) &= \Pi_h \widetilde{z}_h \cdot \nabla_h u_h - \phi^*(\Pi_h \widetilde{z}_h), \\ \psi_h(\Pi_h u_h) &= \operatorname{div} \widetilde{z}_h \, \Pi_h u_h - \psi_h^*(\operatorname{div} \widetilde{z}_h). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} -\int_\Omega \phi^*(\Pi_h \widetilde{z}_h) + \psi_h^*(\operatorname{div}\widetilde{z}_h) \,\mathrm{d}x &= \int_\Omega \phi(\nabla_h u_h) + \psi_h(\Pi_h u_h) \,\mathrm{d}x \\ &\quad - \int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h u_h + D\psi_h(\Pi_h u_h) \Pi_h u_h \,\mathrm{d}x. \end{aligned}$$

Using the discrete Euler–Lagrange equation with $v_h = u_h$ we find that

$$-\int_\Omega D\phi(\nabla_h u_h) \cdot \nabla_h u_h + D\psi_h(\Pi_h u_h) \Pi_h u_h \,\mathrm{d}x = \int_{\mathcal{S}_h \setminus \Gamma_N} \alpha_{\mathcal{S}}^{-2} [\![u_h]\!]_h^2 \,\mathrm{d}s.$$

By combining the last two identities and incorporating $\alpha_{\mathcal{S}}^{-2}[\![u_h]\!]_h = \{\widetilde{z}_h \cdot n_{\mathcal{S}}\}$ we deduce that $D_h(\widetilde{z}_h) = I_h(u_h)$. □

4. Nonlinear Dirichlet problems

We derive an error estimate for a class of nonlinear Dirichlet problems with linear low order terms. We say that $\phi \in C^1(\mathbb{R}^d)$ is *σ-coercive* if there exists a nonnegative functional $\sigma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ such that for all $a, b \in \mathbb{R}^d$ we have

$$\phi(a) + D\phi(a) \cdot (b - a) + \sigma(a; b) \leq \phi(b).$$

The low order term is assumed to be given by the function

$$\psi(x, s) = -f(x)s$$

for some given $f \in L^{p'}(\Omega)$. We assume below that the corresponding continuous problems, defined with

$$\begin{aligned} I(u) &= \int_\Omega \phi(\nabla u) \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x, \\ D(z) &= -\int_\Omega \phi^*(z) \,\mathrm{d}x - I_{-f}(\operatorname{div} z), \end{aligned}$$

are in strong duality and refer the reader to [4, 27] for sufficient conditions and general statements. We have that the indicator functional $\psi^*(x,t) = I_{\{-f(x)\}}(t)$ enforces the constraint $\operatorname{div} z = -f$ and that the discrete primal and dual problem are given by the functionals

$$\begin{aligned} I_h(u_h) &= \int_\Omega \phi(\nabla_h u_h)\,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x + J_h(u_h), \\ D_h(z_h) &= -\int_\Omega \phi^*(\Pi_h z_h)\,\mathrm{d}x - I_{\{-f_h\}}(\operatorname{div} z_h) - K_h(z_h), \end{aligned}$$

where we assume $f_h = \Pi_h f$. The indicator functional $I_{\{-f_h\}}$ enforces the constraint $\operatorname{div} z_h = -f_h$.

Proposition 4.1 (Error estimate)**.** *Assume that $\phi \in C^1(\mathbb{R}^d)$ is strictly convex and σ-coercive and assume that strong duality holds for the continuous problem. For the minimizer $u \in W^{1,p}_D(\Omega)$ of I and the discrete minimizer $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ for I_h we have with a solution $z \in W^q_N(\operatorname{div};\Omega)$ of the dual problem satisfying the regularity condition $z \in W^{1,1}(\Omega;\mathbb{R}^d)$ that*

$$\begin{aligned} \int_\Omega \sigma(\nabla_h u_h; \nabla_h \mathcal{I}_h u)\,\mathrm{d}x \le \int_\Omega \big(D\phi^*(z) - D\phi^*(\Pi_h \mathcal{J}_h z)\big) \cdot (z - \Pi_h \mathcal{J}_h z)\,\mathrm{d}x \\ + J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z). \end{aligned}$$

Proof. The minimality of u_h implies that

$$\delta_h^2 = \int_\Omega \sigma(\nabla_h u_h; \nabla_h \mathcal{I}_h u)\,\mathrm{d}x \le I_h(\mathcal{I}_h u) - I_h(u_h).$$

With the duality relation $I_h(u_h) \ge D_h(z_h) \ge D_h(\mathcal{J}_h z)$ we infer that

$$\begin{aligned} \delta_h^2 &\le I_h(\mathcal{I}_h u) - D_h(\mathcal{J}_h z) \\ &= \int_\Omega \phi(\nabla_h \mathcal{I}_h u)\,\mathrm{d}x - \int_\Omega f_h \Pi_h \mathcal{I}_h u\,\mathrm{d}x + \int_\Omega \phi^*(\Pi_h \mathcal{J}_h z)\,\mathrm{d}x \\ &\quad + J_h(\mathcal{I}_h u_h) + K_h(\mathcal{J}_h z), \end{aligned}$$

where we used that $\operatorname{div} \mathcal{J}_h z = -f_h$. Jensen's inequality in combination with $\nabla_h \mathcal{I}_h u = \Pi_h \nabla u$ and the strong duality relation $I(u) = D(z)$ lead to

$$\int_\Omega \phi(\nabla_h \mathcal{I}_h u)\,\mathrm{d}x \le \int_\Omega \phi(\nabla u)\,\mathrm{d}x = -\int_\Omega \phi^*(z)\,\mathrm{d}x + \int_\Omega f u\,\mathrm{d}x.$$

This implies that we have

$$\begin{aligned} \delta_h^2 \le -\int_\Omega \phi^*(z)\,\mathrm{d}x + \int_\Omega f u - f_h \Pi_h \mathcal{I}_h u\,\mathrm{d}x + \int_\Omega \phi^*(\mathcal{J}_h z)\,\mathrm{d}x \\ + J_h(\mathcal{I}_h u_h) + K_h(\mathcal{J}_h z). \end{aligned}$$

Since $\operatorname{div} z = -f$ and $\operatorname{div} \mathcal{J}_h z = -f_h$ it follows from the integration-by-parts formula (1) and the identity $\nabla_h \mathcal{I}_h u = \Pi_h \nabla u$ that

$$\int_\Omega fu - f_h \Pi_h \mathcal{I}_h u \,\mathrm{d}x = \int_\Omega z \cdot \nabla u - \mathcal{J}_h z \cdot \nabla_h \mathcal{I}_h u \,\mathrm{d}x$$
$$= \int_\Omega (z - \Pi_h \mathcal{J}_h z) \cdot \nabla u \,\mathrm{d}x.$$

We use that $z = D\phi(\nabla u)$ and hence $\nabla u = D\phi^*(z)$, i.e.,

$$\int_\Omega fu - f_h \Pi_h \mathcal{I}_h u \,\mathrm{d}x = \int_\Omega D\phi^*(z) \cdot (z - \Pi_h \mathcal{J}_h z) \,\mathrm{d}x.$$

The convexity of ϕ^* provides the relation

$$\phi^*(\Pi_h \mathcal{J}_h z) + D\phi^*(\Pi_h \mathcal{J}_h z) \cdot (z - \Pi_h \mathcal{J}_h z) \le \phi^*(z). \tag{4}$$

On combining the inequalities we find that

$$\delta_h^2 \le \int_\Omega \big(D\phi^*(z) - D\phi^*(\Pi_h \mathcal{J}_h z)\big) \cdot (z - \Pi_h \mathcal{J}_h z) \,\mathrm{d}x$$
$$+ J_h(\mathcal{I}_h u_h) + K_h(\mathcal{J}_h z),$$

which implies the asserted estimate. □

Remark 4.2. *The estimate of the proposition can be improved by incorporating a coercivity property of ϕ^* in* (4).

Under additional conditions a convergence rate can be deduced. To illustrate this we assume for simplicity the Lipschitz property

$$\|D\phi^*(v) - D\phi^*(w)\|_{L^q(\Omega)} \le c_\phi \|v - w\|_{L^q(\Omega)}$$

which can be replaced, e.g., by a local Lipschitz estimate.

Corollary 4.3 (Lipschitz differentiability)**.** *In addition to the assumptions of Proposition 4.1 assume that $D\phi^*$ is Lipschitz continuous. Then we have*

$$\int_\Omega \sigma(\nabla_h u_h; \nabla_h \mathcal{I}_h u) \,\mathrm{d}x \le c_\phi \|z - \Pi_h \mathcal{J}_h z\|_{L^q(\Omega)}^2$$
$$+ c_{\mathcal{T}} \|h_{\mathcal{S}}^{-1} \alpha_{\mathcal{S}}^{r'}\|_{L^\infty(\mathcal{S})} \|\mathcal{J}_h z\|_{L^{r'}(\Omega)}^{r'} + c_{\mathcal{T}} \|h_{\mathcal{S}}^{-1} \beta_{\mathcal{S}}^{s}\|_{L^\infty(\mathcal{S})} \|\mathcal{I}_h u\|_{L^s(\Omega)}^{s}.$$

In particular, if $z \in W^{1,q}(\Omega;\mathbb{R}^d)$, $u \in W^{1,p}(\Omega)$, and $\alpha_{\mathcal{S}} = c_\alpha h_{\mathcal{S}}^{\gamma}$, $\beta_{\mathcal{S}} = c_\beta h_{\mathcal{S}}^{\sigma}$ with $\gamma r', \sigma s \ge 3$ and $r' \le q$, $s \le p$ then the right-hand side is of quadratic order.

Proof. The estimate is an immediate consequence of Proposition 4.1 noting that $[\![\mathcal{I}_h u]\!]_h = 0$ for all $S \in \mathcal{S}_h \setminus \Gamma_N$ and $[\![\mathcal{J}_h z \cdot n_{\mathcal{S}}]\!]_h = 0$ for all $S \in \mathcal{S}_h \setminus \Gamma_D$ and the inequalities (2). □

5. Total-variation minimization

Setting $\Gamma_N = \partial\Omega$ and $\Gamma_D = \emptyset$ we consider the minimization of the functional

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

in the set of all $u \in BV(\Omega) \cap L^2(\Omega)$. We refer the reader to [1, 6] for definitions and properties of the variational problem. The dual formulation consists in determining $z \in W^2_N(\mathrm{div};\Omega)$ which is maximal for

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

In particular, we have that $z \in L^\infty(\Omega;\mathbb{R}^d)$ and strong duality applies, i.e., for solutions u and z we have

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

cf., e.g., [23]. The discrete primal problem seeks $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ which is minimial for

$$I_h(u_h) = \int_\Omega |\nabla_h u_h| \,\mathrm{d}x + J_h(u_h) + \frac{\alpha}{2}\|\Pi_h(u_h-g)\|^2.$$

The functionals I_h approximate I under moderate conditions on the discretization parameters r and p.

Proposition 5.1 (Γ-convergence)**.** *Assume $r \ge 1$, $s \le 2$, and that*

$$\|\alpha_{\mathcal{S}}^{r'} h_{\mathcal{S}}^{-1}\|_{L^\infty(\mathcal{S}_h)} + \|\beta_{\mathcal{S}}^{s} h_{\mathcal{S}}^{-1}\|_{L^\infty(\mathcal{S}_h)} \to 0$$

as $h \to 0$ where the first term is replaced by $\|\alpha_{\mathcal{S}}\|_{L^\infty(\mathcal{S}_h)}$ if $r = 1$. We then have $I_h \to I$ in the sense of Γ convergence with respect to strong convergence in $L^1(\Omega)$.

Proof. (i) To show that $I(u) \le \liminf I_h(u_h)$ for a sequence $(u_h)_{h>0}$ with $I_h(u_h) \le c$ we first note that

$$\begin{aligned}|Du_h|(\Omega) &= \|\nabla u_h\|_{L^1(\Omega)} + \|[\![u_h]\!]\|_{L^1(\mathcal{S}_h\setminus\Gamma_N)}\\ &\le \|\nabla u_h\|_{L^1(\Omega)} + \|[\![u_h]\!]_h\|_{L^1(\mathcal{S}_h\setminus\Gamma_N)} + \|h_{\mathcal{S}}[\![\nabla_h u_h]\!]\|_{L^1(\mathcal{S}_h)}\\ &\le \|\nabla u_h\|_{L^1(\Omega)} + \|\alpha_{\mathcal{S}}\|_{L^{r'}(\mathcal{S}_h\setminus\Gamma_N)}\|\alpha_{\mathcal{S}}^{-1}[\![u_h]\!]_h\|_{L^r(\mathcal{S}_h\setminus\Gamma_N)} + c_{\mathcal{T}}\|\nabla_h u_h\|_{L^1(\Omega)}.\end{aligned}$$

Since $\|\alpha_{\mathcal{S}}\|^{r'}_{L^{r'}(\mathcal{S}_h\setminus\Gamma_N)} \le \|\alpha_{\mathcal{S}}^{r'} h_{\mathcal{S}}^{-1}\|_{L^\infty(\mathcal{S}_h\setminus\Gamma_N)} c_{\mathcal{T}}|\Omega|$ and since

$$\|u_h\|_{L^1(\Omega)} = \|\Pi_h u_h\|_{L^1(\Omega)} + h\|\nabla_h u_h\|_{L^1(\Omega)}$$

we find that $(u_h)_{h>0}$ is bounded in $BV(\Omega)$. We let $u \in BV(\Omega)$ be an appropriate accumulation point so that $u_h \to u$ in $L^1(\Omega)$ and $\Pi_h u_h \rightharpoonup u$ in $L^2(\Omega)$. Using that for $\psi \in C_0^\infty(\Omega;\mathbb{R}^d)$ we have

$$\begin{aligned}\int_\Omega \Pi_h u_h \operatorname{div}\psi \,\mathrm{d}x &= -\int_\Omega \nabla_h u_h \cdot \psi \,\mathrm{d}x + \int_{\mathcal{S}_h\setminus\Gamma_N} [\![u_h]\!]_h \mathcal{J}_h\psi \cdot n_{\mathcal{S}} \,\mathrm{d}s\\ &\quad + \int_\Omega \nabla_h u_h \cdot (\psi - \mathcal{J}_h\psi)\,\mathrm{d}x,\end{aligned}$$

where

$$\int_{\mathcal{S}_h} [\![u_h]\!]_h \mathcal{J}_h \psi \,\mathrm{d}s \le \|\alpha_{\mathcal{S}}\|_{L^{r'}(\mathcal{S}_h)} \|\alpha_{\mathcal{S}}^{-1} [\![u_h]\!]_h\|_{L^r(\mathcal{S}_h \setminus \Gamma_N)} \|\mathcal{J}_h \psi\|_{L^\infty(\mathcal{S}_h)}$$

tends to zero owing to the conditions on $\alpha_{\mathcal{S}}$. If $\|\psi\|_{L^\infty(\Omega)} \le 1$ then this leads to

$$\int_\Omega u \operatorname{div} \psi \,\mathrm{d}x \le \liminf_{h\to 0} \|\nabla_h u_h\|_{L^1(\Omega)}$$

and in particular to the bound

$$|\mathrm{D}u|(\Omega) \le \liminf_{h\to 0} \|\nabla_h u_h\|_{L^1(\Omega)}.$$

Since $\Pi_h(u_h - g_h) \rightharpoonup (u-g)$ in $L^2(\Omega)$ and $J_h(u_h) \ge 0$ we deduce that $I(u) \le \liminf_{h\to 0} I_h(u_h)$.
(ii) To prove that for every $u \in BV(\Omega) \cap L^2(\Omega)$ there exists a sequence $(u_h)_{h>0}$ with $u_h \rightharpoonup u$ in $L^2(\Omega)$ and $I(u) = \lim_{h\to 0} I_h(u_h)$ we use the intermediate density of continuous finite element functions in $BV(\Omega)\cap L^2(\Omega)$ to obtain a sequence $(u_h)_{h>0}$ with $\|\alpha_{\mathcal{S}} [\![u_h]\!]\|_{L^r(\mathcal{S}_h \setminus \Gamma_N)} = 0$ and which converges intermediately in $BV(\Omega)$, weakly in $L^2(\Omega)$, and strongly in $L^1(\Omega)$ to u. The condition on $\beta_{\mathcal{S}}$ yields that $\|\beta_{\mathcal{S}}\{u_h\}\|_{L^s(\mathcal{S}_h)} \to 0$ as $h \to 0$. Altogether, this implies the attainment result $I(u) = \lim_{h\to 0} I_h(u_h)$. □

Remark 5.2. *For $r = 1$ the condition $\|\alpha_{\mathcal{S}}\|_{L^\infty(\mathcal{S}_h)} \to 0$ corresponds to the use of quadrature in the definition of the penalty terms. If instead of the mean of the jump $[\![u_h]\!]_h$ the full jump $[\![u_h]\!]$ is used in the definition of J_h, then if $r = 1$ it suffices to require that $\alpha_{\mathcal{S}} = 1$ since the functional I_h then involves the exact term $|Du_h|(\Omega)$. Our error estimate below shows that the condition on $\alpha_{\mathcal{S}}$ can be weakened if a regularity condition is satisfied.*

For an error estimate the discrete dual functional is required. It consists in maximizing the functional

$$D_h(z_h) = -I_{K_1(0)}(\Pi_{0,h} z_h) - \frac{1}{2\alpha}\|\operatorname{div} z_h + \alpha g_h\|^2 + \frac{\alpha}{2}\|g_h\|^2 - K_h(z_h)$$

in the set of vector fields $z_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$.

Proposition 5.3 (Error estimate)**.** *Assume that $g \in L^\infty(\Omega)$ and that there exists a Lipschitz continuous solution $z \in W^2_N(\mathrm{div};\Omega) \cap W^{1,\infty}(\Omega)$ for the dual problem. Moreover, suppose that*

$$\|h_{\mathcal{S}}^{-1} \alpha_{\mathcal{S}}^{r'}\|_{L^\infty(\mathcal{S}_h)} + \|h_{\mathcal{S}}^{-1} \beta_{\mathcal{S}}^{s}\|_{L^\infty(\mathcal{S}_h)} \le ch,$$

where the first term can be omitted if $r = 1$ and $0 < \alpha_{\mathcal{S}} \le 1$. Then, for the solutions $u \in BV(\Omega)\cap L^2(\Omega)$ and $u_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ of the primal and discrete primal problem we have

$$\|u - \Pi_h u_h\| \le c h^{1/2} M_{u,z,g},$$

with a factor $M_{u,z,g}$ that depends on $\alpha > 0$, $\|u\|_{L^\infty(\Omega)} \le \|g\|_{L^\infty(\Omega)}$, $\|g\|_{L^2(\Omega)}$, and $\|\nabla z\|_{L^\infty(\Omega)}$.

Proof. (i) By the coercivity of the discrete functional I_h and the discrete duality relation $\inf I_h \geq \sup D_h$ we have

$$\frac{\alpha}{2}\|\Pi_h(v_h - u_h)\|^2 \leq I_h(v_h) - I_h(u_h) \leq I_h(v_h) - D_h(y_h)$$

for every $v_h \in \mathcal{S}^{1,dg}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}^{0,dg}(\mathcal{T}_h)$.
(ii) By noting that $u \in L^\infty(\Omega)$ and choosing regularizations $(u_\varepsilon)_\varepsilon > 0$ of u we construct a quasi-interpolant $\widetilde{u}_h = \lim_{\varepsilon\to 0} \mathcal{I}_h u_\varepsilon \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ satisfying

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

In particular, we have that $[\![\widetilde{u}_h]\!]_h = 0$ on inner element sides $S \in \mathcal{S}_h \setminus \partial\Omega$ and hence

$$\begin{aligned}I_h(\widetilde{u}_h) &= \|\nabla_h \widetilde{u}_h\|_{L^1(\Omega)} + \frac{\alpha}{2}\|\Pi_h(\widetilde{u}_h - g)\|^2 + \frac{1}{s}\|\beta_{\mathcal{S}}\{\widetilde{u}_h\}_h\|^s_{L^s(\mathcal{S}_h)}\\ &\leq I(u) + \frac{\alpha}{2}\big(\|\Pi_h(\widetilde{u}_h - g)\|^2 - \|u-g\|^2\big) + \frac{1}{s}\|h_{\mathcal{S}}^{-1}\beta_{\mathcal{S}}^s\|_{L^\infty(\mathcal{S}_h)} c_{\mathcal{T}}\|\widetilde{u}_h\|^s_{L^s(\Omega)}.\end{aligned}$$

Abbreviating $\overline{u}_h = \Pi_h \widetilde{u}_h$ and $g_h = \Pi_h g$ we have

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

Incorporating the bound $\|\mathcal{I}_h u\|_{L^s(\Omega)} \leq c_s\|u\|_{L^\infty(\Omega)}$, these identities imply that

$$\begin{aligned}I_h(\widetilde{u}_h) \leq I(u) &+ \frac{\alpha}{2}\|\overline{u}_h - u\|_{L^1(\Omega)}\|\overline{u}_h + u - 2g\|_{L^\infty(\Omega)}\\ &- \frac{\alpha}{2}\|g - g_h\|^2 + c_{\mathcal{T}} c_s^s \|h_{\mathcal{S}}^{-1}\beta_{\mathcal{S}}^s\|_{L^\infty(\mathcal{S}_h)}\|u\|^s_{L^\infty(\Omega)}.\end{aligned}$$

(iii) With $L = \|\nabla z\|_{L^\infty(\Omega)}$ we have that $\|\mathcal{J}_h z\|_{L^\infty(\Omega)} \leq \varrho_h = 1 + chL$. Hence, for $\widetilde{z}_h = \varrho_h^{-1}\mathcal{J}_h z \in \mathcal{RT}^0_N(\mathcal{T}_h)$ we have $|\Pi_h \widetilde{z}_h(x_T)| \leq 1$ as well as $\operatorname{div} \widetilde{z}_h = \varrho_h^{-1}\Pi_h \operatorname{div} z$. With these relations we deduce that

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

where we used Jensen's inequality, $\varrho_h \geq 1$, and $\|g_h\|^2 - \|g\|^2 = \|g - g_h\|^2$. In the case $r = 1$ we note that $|z| \leq 1$ implies that $|\widetilde{z}_h \cdot n_{\mathcal{S}}| \leq 1$ and hence since $\alpha_{\mathcal{S}}^{-1} \ge 1$ that $K_h(\widetilde{z}_h) = 0$. If $r > 1$ we have

$$K_h(\widetilde{z}_h) = \frac{1}{r'} \|\alpha_S \{\widetilde{z}_h \cdot n_{\mathcal{S}}\}\|_{L^{r'}(\mathcal{S}_h)}^{r'} \le c_{\mathcal{T}} \|h_{\mathcal{S}}^{-1} \alpha_{\mathcal{S}}^{r'}\|_{L^\infty(\mathcal{S}_h)} \|z\|_{L^{r'}(\Omega)}^{r'}.$$

(iv) We are now in position to combine the previous estimates. The choices $v_h = \widetilde{u}_h$ and $y_h = \widetilde{z}_h$ lead to

$$\begin{aligned}
\frac{\alpha}{2} \|\Pi_h(\widetilde{u}_h - u_h)\|^2 &\le I_h(\widetilde{u}_h) - D_h(\widetilde{z}_h) \\
&\le I(u) + \frac{\alpha}{2} \|\overline{u}_h - u\|_{L^1(\Omega)} \|\overline{u}_h + u - 2g\|_{L^\infty(\Omega)} \\
&\quad - D(z) + |\varrho_h^{-1} - 1| \|\operatorname{div} z\| \|g\| \\
&\quad + c_{\mathcal{T}} \delta_r \|h_{\mathcal{S}}^{-1} \alpha_{\mathcal{S}}^{r'}\|_{L^\infty(\mathcal{S}_h)} \|z\|_{L^{r'}(\Omega)}^{r'} + c_{\mathcal{T}} \|h_{\mathcal{S}}^{-1} \beta_{\mathcal{S}}^{s}\|_{L^\infty(\mathcal{S}_h)} \|u\|_{L^s(\Omega)}^{s},
\end{aligned}$$

where $\delta_r = 0$ if $r = 1$ and $\delta_r = 1$ otherwise. Using $I(u) = D(z)$, the estimate $1 - \varrho_h^{-1} \le chL$, the approximation properties of $\widetilde{u}_h$, and the conditions of the proposition show that

$$\frac{\alpha}{2} \|\Pi_h(\widetilde{u}_h - u_h)\|^2 \le ch \widetilde{M}_{u,z,g}^2.$$

(v) With the estimate

$$\begin{aligned}
\|u - \Pi_h \widetilde{u}_h\|^2 &\le \|u - \Pi_h \widetilde{u}_h\|_{L^\infty(\Omega)} \|u - \Pi_h \widetilde{u}_h\|_{L^1(\Omega)} \\
&\le (1 + c_d) \|u\|_{L^\infty(\Omega)} \big(\|u - \widetilde{u}_h\|_{L^1(\Omega)} + \|\widetilde{u}_h - \Pi_h \widetilde{u}_h\|_{L^1(\Omega)} \big) \\
&\le (1 + c_d) \|u\|_{L^\infty(\Omega)} h \big(c|Du|(\Omega) + \|\nabla_h \widetilde{u}_h\|_{L^1(\Omega)} \big),
\end{aligned}$$

we deduce the asserted error bound. □

Remarks 5.4. *(i) If $(u_h)_{h>0}$ is uniformly bounded in $L^\infty(\Omega)$ then using that*

$$\|u_h - \Pi_h u_h\| \le 2h^{1/2} \|\nabla_h u_h\|_{L^1(\Omega)} \|u_h\|_{L^\infty(\Omega)}$$

we may replace $\Pi_h u_h$ by u_h in the error estimate of Proposition 5.3.
(ii) A reduced convergence rate is expected if z fails to be Lipschitz continuous. If only $u \in L^\infty(\Omega) \cap BV(\Omega$ is assumed then a convergence rate $\mathcal{O}(h^{1/4})$ can be established, cf. [9, 18, 10].

6. Obstacle problem

A model obstacle problem is defined by the functional

$$I(u) = \frac{1}{2} \int_\Omega |\nabla u|^2 \,\mathrm{d}x - \int_\Omega f u \,\mathrm{d}x + I_{\mathbb{R}_{\ge 0}}(u)$$

for $u \in W_D^{1,p}(\Omega)$. The dual functional is given by

$$D(z) = -\frac{1}{2} \int_\Omega |z|^2 \,\mathrm{d}x - I_{\mathbb{R}_{\le 0}}(f + \operatorname{div} z)$$

for vector fields $z \in W^2_N(\mathrm{div};\Omega)$. We have the strong duality relation $I(u) = D(z)$ for solutions u and z and the pointwise complementarity principle that if $u > 0$ then $f + \mathrm{div}\, z = 0$. The discrete functionals are given by

$$I_h(u_h) = \frac{1}{2}\int_\Omega |\nabla_h u_h|^2 \,\mathrm{d}x - \int_\Omega f_h \Pi_h u_h \,\mathrm{d}x + I_{\mathbb{R}_{\ge 0}}(\Pi_h u_h) + J_h(u_h),$$

and

$$D_h(z_h) = -\frac{1}{2}\int_\Omega |\Pi_h z_h|^2 \,\mathrm{d}x - I_{\mathbb{R}_{\le 0}}(f_h + \mathrm{div}\, z_h) - K_h(z_h).$$

Owing to Theorem 3.3 we have that $I_h(u_h) \ge D_h(z_h)$. We assume that the functionals J_h and K_h are defined with the parameters and quantities

$$r = s = 2, \quad \alpha_\mathcal{S} = c_\alpha h_\mathcal{S}^\gamma, \quad \beta_\mathcal{S} = c_\beta h_\mathcal{S}^\gamma$$

for parameters $\gamma \ge 3/2$, $\alpha > 0$, and $\beta \ge 0$.

Proposition 6.1 (Error estimate)**.** *Assume that $u \in W^{1,2}_D(\Omega) \cap W^{2,2}(\Omega)$. Then we have that*

$$\|\nabla_h(u_h - u)\|^2 \le c_\mathcal{J}^2 h^2 \|Dz\|^2 + 2c_\mathcal{I}^2 h^2 \|f + \mathrm{div}\, z\| \|D^2 u\| + c_\mathcal{T} c_{\alpha,\beta}^2 h^{2\gamma-1}\big(\|u\|^2_{W^{1,2}(\Omega)} + \|z\|^2_{W^{1,2}(\Omega)}\big).$$

Proof. We first note that the quasi-interpolants $\mathcal{I}_h u$ and $\mathcal{J}_h z$ are well defined and admissible in the discrete primal and dual problems, respectively, i.e., we have

$$\mathcal{I}_h u(x_T) = \frac{1}{d+1}\sum_{S \subset \partial\mathbb{T}} \int_S u(s)\,\mathrm{d}s \ge 0,$$

for every $T \in \mathcal{T}_h$ and $f_h + \mathrm{div}\,\mathcal{J}_h z = \Pi_h(f + \mathrm{div}\, z) \le 0$. The coercivity of I_h and the discrete duality relation $I_h(u_h) \ge D_h(\mathcal{J}_h z)$ lead to

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

By Jensen's inequality and $\nabla_h \mathcal{I}_h u = \Pi_h \nabla u$ we have $\|\nabla_h \mathcal{I}_h u\| \le \|\nabla u\|$ and with the strong duality relation $I(u) = D(z)$ we infer that

$$\begin{aligned}\delta_h^2 &\le \frac{1}{2}\|\nabla u\|^2 - (f_h, \mathcal{I}_h u) + J_h(\mathcal{I}_h u) + \frac{1}{2}\|\Pi_h \mathcal{J}_h z\|^2 + K_h(\mathcal{J}_h z)\\ &= -\frac{1}{2}\|z\|^2 + (f,u) - (f_h, \Pi_h \mathcal{I}_h u_h) + J_h(\mathcal{I}_h u) + \frac{1}{2}\|\Pi_h \mathcal{J}_h z\|^2 + K_h(\mathcal{J}_h z).\end{aligned}$$

The binomial formula $a^2 - b^2 = 2b(a-b) + (a-b)^2$ and the identities $f_h = \Pi_h f$ and $z = \nabla u$ lead to the estimate

$$\begin{aligned}\delta_h^2 &\le (z, \Pi_h \mathcal{J}_h z - z) + \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 + (f, u - \Pi_h \mathcal{I}_h u_h) + J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z)\\ &= (\nabla u, \Pi_h \mathcal{J}_h z - z) + \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 + (f, u - \Pi_h \mathcal{I}_h u) + J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z).\end{aligned}$$

With the relation $\Pi_h \nabla u = \nabla_h \mathcal{I}_h u$, an integration by parts, and $\mathrm{div}\,\mathcal{J}_h z = \Pi_h \,\mathrm{div}\, z$ we obtain the identities

$$(\nabla u, \Pi_h \mathcal{J}_h z - z) = (\nabla_h \mathcal{I}_h u, \mathcal{J}_h z) - (\nabla u, z) = (\mathrm{div}\, z, u - \Pi_h \mathcal{I}_h u).$$

Using this and the abbreviation $\lambda = f + \operatorname{div} z$ show that we have

$$\begin{aligned}\delta_h^2 &\le \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 + (f + \operatorname{div} z, u - \Pi_h \mathcal{I}_h u) + J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z)\\ &= \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 + (\lambda, u - \mathcal{I}_h u) + (\lambda, \mathcal{I}_h u - \Pi_h \mathcal{I}_h u) + J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z)\end{aligned}$$

We note that $\mathcal{I}_h u|_T - \Pi_h \mathcal{I}_h u(x_T) = \nabla_h \mathcal{I}_h u|_T \cdot (x - x_T)$ and that on the element contact set

$$\mathcal{C}_T = \{x \in T : u(x) = 0\}$$

we have $\nabla u|_{\mathcal{C}_T} = 0$ and $\lambda|_{T\setminus\mathcal{C}_T} = 0$. Hence, it follows that

$$\int_T \lambda(\mathcal{I}_h u - \Pi_h \mathcal{I}_h u)\,\mathrm{d}x = \int_{\mathcal{C}_T} \lambda\,(x - x_T)\cdot\nabla(\mathcal{I}_h u - u)\,\mathrm{d}x$$

for every $T \in \mathcal{T}_h$. We thus obtain the estimate

$$\begin{aligned}\delta_h^2 \le \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 &+ \|\lambda\|\big(\|u - \mathcal{I}_h u\| + h\|\nabla_h(u - \mathcal{I}_h u)\|\big)\\ &+ J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z).\end{aligned}$$

For the side functionals J_h and K_h we have, owing to the continuity properties of $\mathcal{I}_h u$ and $\mathcal{J}_h z$ that

$$\begin{aligned}J_h(\mathcal{I}_h u) + K_h(\mathcal{J}_h z) &= \frac{c_\alpha^2}{2}\|h_\mathcal{S}^\gamma\{\mathcal{I}_h u\}_h\|^2_{L^2(\mathcal{S}_h\setminus\Gamma_D)} + \frac{c_\beta^2}{2}\|h_\mathcal{S}^\gamma\{\mathcal{J}_h z\cdot n_\mathcal{S}\}\|^2_{L^2(\mathcal{S}_h\setminus\Gamma_N)}\\ &\le \frac{1}{2}c_\mathcal{T} c_{\alpha,\beta}^2 h^{2\gamma-1}\big(\|\mathcal{I}_h u\|^2 + \|\mathcal{J}_h z\|^2\big).\end{aligned}$$

By combining the previous estimates we arrive at

$$\begin{aligned}\delta_h^2 \le \frac{1}{2}\|\Pi_h \mathcal{J}_h z - z\|^2 &+ \|\lambda\|\big(\|u - \mathcal{I}_h u\| + h\|\nabla_h(u - \mathcal{I}_h u)\|\big)\\ &+ \frac{1}{2}c_\mathcal{T} h^{2\gamma-1} c_{\alpha,\beta}^2\big(\|\mathcal{I}_h u\|^2 + \|\mathcal{J}_h z\|^2\big).\end{aligned}$$

With basic stability properties of the quasi-interpolation operators as operators from $W^{1,2}(\Omega;\mathbb{R}^\ell) \to L^2(\Omega;\mathbb{R}^\ell)$ we deduce the asserted error bound. □

Remark 6.2. *By defining discontinuous Galerkin methods with certain consistency properties it is possible to derive optimal convergence rates with a penalty term that only involves the factor h^{-1}, cf. [28]. The approach followed here applies to a large class of variational problems and allows for a simple error analysis.*

7. Numerical experiments

We verify in this section the theoretical results and discuss the role of the parameters involved in the discontinuous Galerkin discretizations.

7.1. **Poisson problem.** To verify the optimality of the conditions on the weight function $\alpha_{\mathcal{S}}$ in the error estimates we consider a Poisson problem. The discretized functional reads

$$I_h(u_h) = \frac{1}{2}\int_\Omega |\nabla_h u_h|^2 \,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x + \frac{c_\alpha^{-2}}{2}\int_{\mathcal{S}} h_{\mathcal{S}}^{-2\gamma} |[\![u_h]\!]|^2 \,\mathrm{d}s,$$

subject to homogeneous Dirichlet boundary conditions for u_h on $\Gamma_D = \partial\Omega$. Our parameters correspond to the settings

$$\alpha_{\mathcal{S}} = c_\alpha h_{\mathcal{S}}^\gamma, \quad \beta_{\mathcal{S}} = 0, \quad r = 2, \quad s = 2,$$

where we consider combinations of the parameters

$$\gamma \in \{0.5, 1.0, 1.5, 2.0\}, \quad c_\alpha^{-1} \in \{1.0, 4.0\}.$$

Example 7.1. *Let $d = 2$, $\Omega = (-1,1)^2$, $\Gamma_D = \partial\Omega$, and for $x = (x_1, x_2) \in \Omega$ set*

$$f(x_1, x_2) = 2\pi^2 \sin(\pi x_1)\sin(\pi x_2).$$

Then, the exact solution is given by

$$u(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$$

and satisfies $u \in W^{1,2}_D(\Omega) \cap W^{2,2}(\Omega)$.

The plots in Figure 1 show the experimental errors

$$\|\nabla_h e_h\| = \|\nabla_h(u - u_h)\|$$

versus the number of elements $N = \#\mathcal{T}_h \sim h^{-2}$ for different combinations of parameters γ and c_α. We obsere that the choices $\gamma = 1/2$ and $\gamma = 1$ do in not lead to an experimental optimal convergence rate. The choice $\gamma = 3/2$ leads to linear convergence independently of the choice of the constant factor c_α which is in agreement with the theoretical error estimates.

7.2. **Total-variation minimization.** For a given triangulation we consider the discrete minimization problem defined via the functional

$$I_{h,\varepsilon}(u_h) = \int_\Omega |\nabla_h u_h|_\varepsilon \,\mathrm{d}x + \frac{\alpha}{2}\|\Pi_h u_h - g_h\|^2 + \frac{c_\alpha^{-r}}{r}\int_{\mathcal{S}} h_{\mathcal{S}}^{-\gamma r} |[\![u_h]\!]|_\varepsilon^r \,\mathrm{d}s$$

with the regularized modulus or length $|a|_\varepsilon = (|a|^2 + \varepsilon^2)^{1/2}$ for $a \in \mathbb{R}^\ell$ and $\varepsilon > 0$. Since $0 \le |a|_\varepsilon - |a| \le \varepsilon$ the error estimate of Proposition 5.3 remains valid provided that $\varepsilon \le ch$, we therefore choose $\varepsilon = h$. The definition correponds to the settings

$$\alpha_{\mathcal{S}} = c_\alpha h_{\mathcal{S}}^\gamma, \quad \beta_{\mathcal{S}} = 0.$$

In the following example we consider Dirichlet boundary conditions on $\Gamma_D = \partial\Omega$. While a general existence theory is lacking our error analysis remains valid provided a solution exists, which is the case for the considered setting.

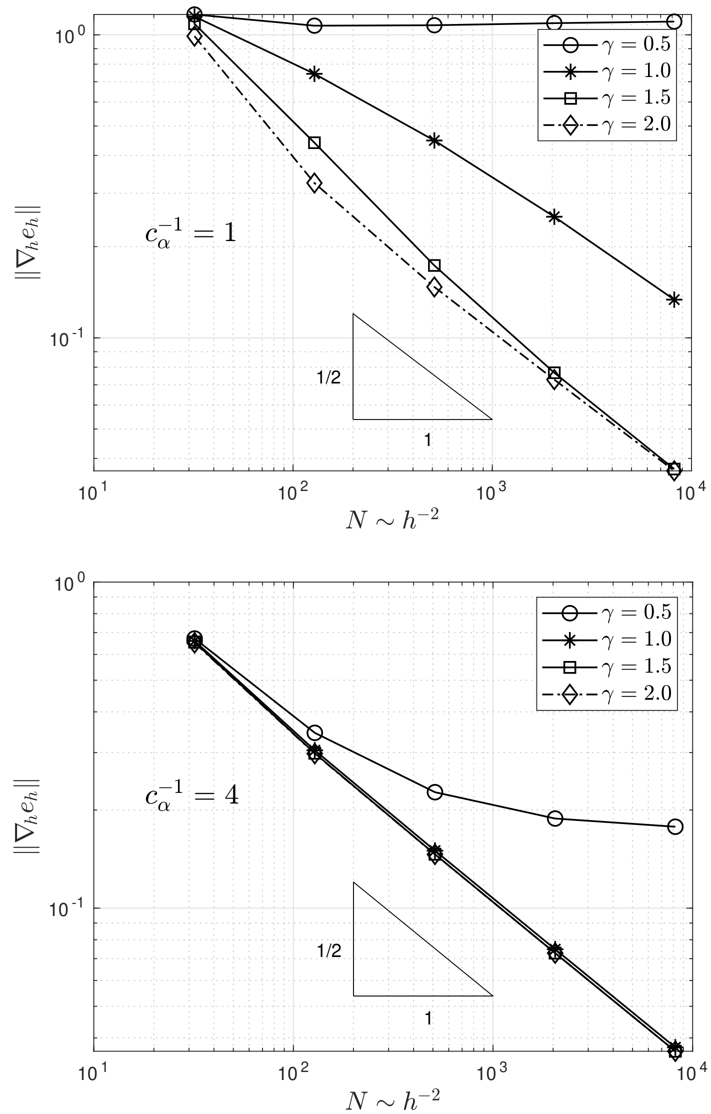

FIGURE 1. Experimental convergence rates in the approximation of the Poisson problem defined in Example 7.1 for different penalty functionals.

Example 7.2. *For $\Omega \subset \mathbb{R}^d$, $\alpha > 0$, and $R > 0$ such that $\overline{B_R(0)} \subset \Omega$, let*

$$g(x) = \chi_{B_R(0)}(x).$$

Then $u(x) = \max\{1 - 2/(\alpha R), 0\}\chi_{B_R(x)}$ is the unique solution of the total variation minimization problem subject to homogeneous Dirichlet conditions *on* $\Gamma_D = \partial\Omega$. *The solution* $z \in W^2(\mathrm{div};\Omega)$ *of the dual problem is given by*

$$z(x) = \begin{cases} R^{-1}x & \textit{for } |x| \le R, \\ -Rx/|x|^2 & \textit{for } |x| \ge R, \end{cases}$$

and satisfies $z \in W^{1,\infty}(\Omega;\mathbb{R}^d)$. *We set* $\alpha = 10$, $R = 1/2$, *and* $\Omega = (-1,1)^2$.

Our numerical approximations are obtained with a semi-implicit discretization of an L^2 gradient flow for $I_{h,\varepsilon}$ with step-size $\tau = 1$ and L^2 stopping criterion $\varepsilon_{\rm stop} = h/100$. We refer the reader to [17, 6] for discussions of iterative methods. The top and bottom plots in Figure 2 show numerical solutions for the parameters

$$\text{(t)} \quad r = 1,\ \gamma = 1,\ c_\alpha = 1, \qquad \text{(b)} \quad r = 2,\ \gamma = 1,\ c_\alpha = 1,$$

on the triangulations $\mathcal{T}_\ell$ with $\ell = 4$ consisting of 2^ℓ halved squares. We observe that the choice $r = 2$ leads to an artificially rounded region, according to the error analysis of Proposition 5.3 they are of comparable accuracy. The analysis showed that the error bound is independent of the γ and c_α if $r = 1$. This is confirmed by the experimental convergence rates shown in Figure 3, where the error quantity

$$\|e_h\|^2 = \|\Pi_h(u - u_h)\|^2$$

is plotted against the number of elements in $\mathcal{T}_\ell$ for combinations of the parameters $r \in \{1,2\}$ and $\gamma \in \{0,1,2\}$ and $c_\alpha^{-1} = 10$. We observe the expected rate $h^{1/2}$ for all combinations except when $\gamma = 0$. In the case $\gamma = 0$ we only observe an error decay if $r = 1$ which confirms the theoretical results but does not lead to the expected optimal convergence rate. Further experiments indicated that this is related to the use of regularization and the approximate iterative solution of the nonlinear systems.

7.3. **Obstacle problem.** We consider an obstacle problem that includes inhomogeneous Dirichlet boundary conditions via a decomposition of the solution and thus leads to the discrete functional

$$\begin{aligned} I_h(u_h) = \frac{1}{2}\int_\Omega |\nabla_h u_h|^2 \,\mathrm{d}x - \int_\Omega f_h u_h \,\mathrm{d}x + I_{\widetilde{\chi}_h}(\Pi_h u_h) \\ + \int_\Omega \nabla_h \mathcal{I}_h \widetilde{u}_D \cdot \nabla_h u_h \,\mathrm{d}x + \frac{c_\alpha^{-2}}{2}\int_{\mathcal{S}} h_{\mathcal{S}}^{-2\gamma} |[\![u_h]\!]|^2 \,\mathrm{d}s, \end{aligned}$$

with the transformed obstacle $\widetilde{\chi}_h = \Pi_h(\chi - \mathcal{I}_h\widetilde{u}_D)$ and subject to homogeneous Dirichlet boundary conditions for u_h on $\Gamma_D = \partial\Omega$. The approximate solution is thus $u_h + u_{D,h}$. Our parameters correspond to the settings

$$\alpha_{\mathcal{S}} = c_\alpha h^{-\gamma}, \quad \beta_{\mathcal{S}} = 0, \quad r = 2, \quad s = 2.$$

We specify the data in following example.

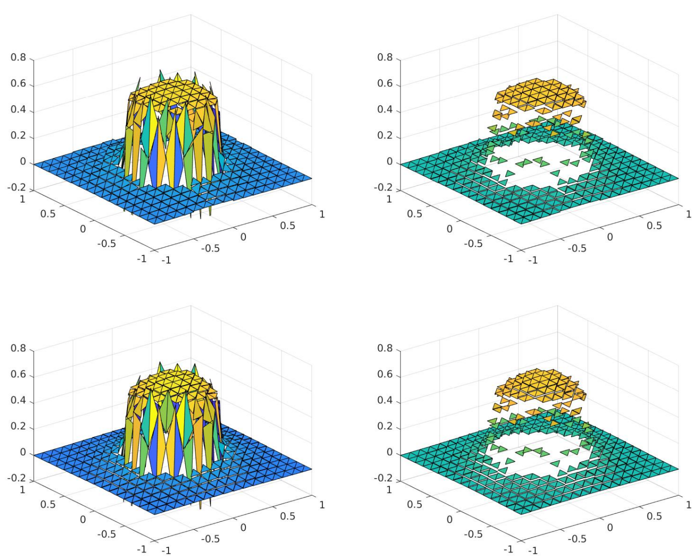

FIGURE 2. Approximations u_h and projections $\Pi_h u_h$ for the total variation minimization problem defined in Example 7.2 for linear (top) and quadratic (bottom) penalty terms.

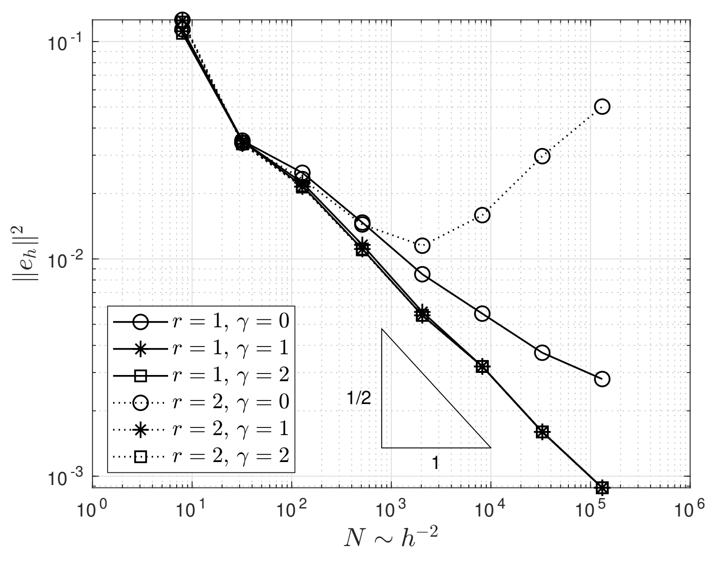

FIGURE 3. Experimental convergence rates in the approximation of the total variation minimization problem defined in Example 7.2 for different penalty terms.

Example 7.3 ([24]). *Let* $\Omega = (-3/2, 3/2)^2$, $f = -2$, $\chi = 0$, *and* $u_D(x) = |x|^2/2 - \log(|x|) - 1/2$ *for* $x \in \Gamma_D = \partial\Omega$. *Then, the exact solution is given* *by*

$$u(x) = \begin{cases} |x|^2/2 - \log(|x|) - 1/2 & \textit{for } |x| \geq 1, \\ 0 & \textit{for } |x| \leq 1. \end{cases}$$

and satisfies $u \in W^{1,2}_D(\Omega) \cap W^{2,2}(\Omega)$.

We solved the discrete minimization problem with a semismooth Newton iteration as in [22] that converged superlinearly towards the stopping criterion that required a correction in the discrete H^1 norm less than $\varepsilon_{\text{stop}} = h$. The left and right plots of Figure 4 show the discontinuous Galerkin approximations for the penalty functionals defined via

$$(\ell) \quad r = 2, \gamma = 1,\ c_\alpha = 1, \qquad (\mathrm{r}) \quad r = 2, \gamma = 3/2,\ c_\alpha = 1.$$

We observe that the jumps along inner edges are smaller for the larger exponent γ. The factor c_α strongly influences the preasymptotic range of the convergence rate which can be observed from Figure 5 where we plotted the approximation errors

$$\|\nabla_h e_h\| = \|\nabla_h (u_h - \mathcal{I}_h u)\|$$

versus the number of elements with a logarithmic scaling on both axes. We obtain the expected linear rate of convergence for $\gamma \geq 3/2$. The decay of the error for $\gamma = 3/2$ is different when $c_\alpha^{-1} = 1$ instead of $c_\alpha^{-1} = 4$.

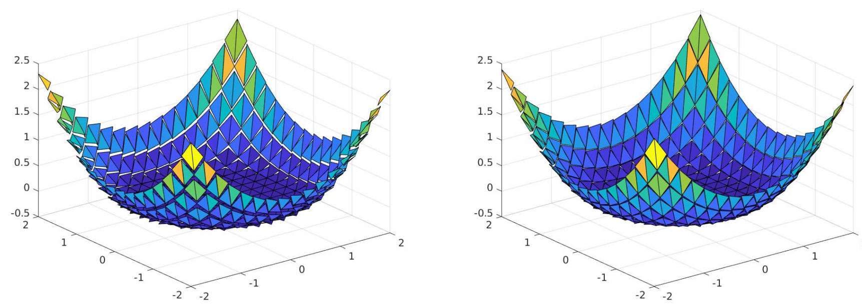

FIGURE 4. Discontinuous Galerkin solution for the obstacle problem defined in Example 7.3 with different penalty functionals.

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, pages xviii+434.

[2] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982. DOI: 10.1137/0719052.

[3] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02. DOI: 10.1137/S0036142901384162.

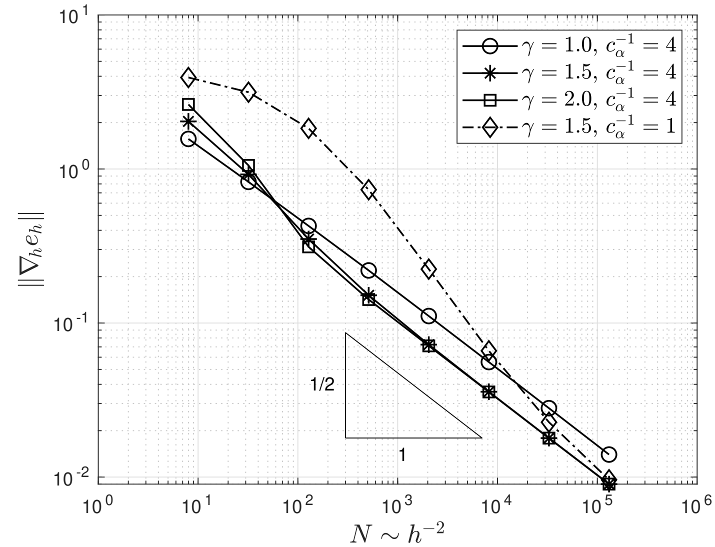

FIGURE 5. Experimental convergence rates in the approximation of the obstacle problem defined in Example 7.3 for different penalty functionals.

[4] H. Attouch, G. Buttazzo, and G. Michaille. Variational analysis in Sobolev and BV spaces, 2006. Applications to PDEs and optimization.

[5] S. Bartels. *Numerical approximation of partial differential equations*, volume 64 of *Texts in Applied Mathematics*. Springer, [Cham], 2016, pages xv+535. DOI: 10.1007/978-3-319-32354-1.

[6] S. Bartels. *Numerical methods for nonlinear partial differential equations*, volume 47 of *Springer Series in Computational Mathematics*. Springer, Cham, 2015, pages x+393. DOI: 10.1007/978-3-319-13797-1.

[7] S. Bartels. Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.*, 50(3):1162–1180, 2012. DOI: 10.1137/11083277X.

[8] S. Bartels, R. H. Nochetto, and A. J. Salgado. A total variation diminishing interpolation operator and applications. *Math. Comp.*, 84(296):2569–2587, 2015. DOI: 10.1090/mcom/2942.

[9] S. Bartels, R. H. Nochetto, and A. J. Salgado. Discrete total variation flows without regularization. *SIAM J. Numer. Anal.*, 52(1):363–385, 2014. DOI: 10.1137/120901544.

[10] S. Bartels. Nonconforming discretizations of convex minimization problems and precise relations to mixed methods, 2020. arXiv: 2002.02359 [math.NA].

[11] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013, pages xiv+685. DOI: 10.1007/978-3-642-36519-5.

[12] 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, pages xviii+397. DOI: 10.1007/978-0-387-75934-0.

[13] A. Buffa and C. Ortner. Compact embeddings of broken Sobolev spaces and applications. *IMA J. Numer. Anal.*, 29(4):827–855, 2009. DOI: 10.1093/imanum/drn038.

[14] C. Caillaud and A. Chambolle. Error estimates for finite differences approximations of the total variation. HAL preprint nr. 02539136, Apr. 2020.

[15] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 38(5):1676–1706, 2000. DOI: 10.1137/S0036142900371003.

[16] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock. An introduction to total variation for image analysis. In *Theoretical foundations and numerical methods for sparse recovery*. Volume 9, Radon Ser. Comput. Appl. Math. Pages 263–340. Walter de Gruyter, Berlin, 2010. DOI: 10.1515/9783110226157.263.

[17] 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. DOI: 10.1007/s10851-010-0251-1.

[18] A. Chambolle and T. Pock. Crouzeix-Raviart approximation of the total variation on simplicial meshes. HAL preprint nr. 01787012, June 2019.

[19] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978, pages xix+530. Studies in Mathematics and its Applications, Vol. 4.

[20] 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.

[21] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012, pages xviii+384. DOI: 10.1007/978-3-642-22980-0.

[22] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888 (2003), 2002. DOI: 10.1137/S1052623401383558.

[23] M. Hintermüller and K. Kunisch. Total bounded variation regularization as a bilaterally constrained optimization problem. *SIAM J. Appl. Math.*, 64(4):1311–1333, 2004. DOI: 10.1137/S0036139903422784.

[24] R. Li, W. Liu, T. Tang, and P. Zhang. Moving mesh finite element methods based on harmonic maps. In *Scientific computing and applications (Kananaskis, AB, 2000)*. Volume 7, Adv. Comput. Theory Pract. Pages 143–156. Nova Sci. Publ., Huntington, NY, 2001.

[25] 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. DOI: 10.1137/0722029.

[26] 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)*, 292–315. Lecture Notes in Math., Vol. 606, 1977.

[27] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970, pages xviii+451.

[28] F. Wang, W. Han, and X.-L. Cheng. Discontinuous Galerkin methods for solving elliptic variational inequalities. *SIAM J. Numer. Anal.*, 48(2):708–733, 2010. DOI: 10.1137/09075891X.

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