Variational problems with gradient constraints:
A priori and *a posteriori* error identities[*]

Harbir Antil[†1], Sören Bartels[‡2], Alex Kaltenbach[§3], and Rohit Khandelwal[¶4]

[2]Department of Applied Mathematics, University of Freiburg, Hermann–Herder–Str. 10, 79104 Freiburg, GERMANY

[1,4]Department of Mathematical Sciences and the Center for Mathematics and Artificial Intelligence (CMAI), George Mason University, Fairfax, VA 22030, USA.

[3]Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 135, 10623 Berlin, GERMANY

October 24, 2024

Abstract

In this paper, on the basis of a (Fenchel) duality theory on the continuous level, we derive an *a posteriori* error identity for arbitrary conforming approximations of a primal formulation and a dual formulation of variational problems involving gradient constraints. In addition, on the basis of a (Fenchel) duality theory on the discrete level, we derive an *a priori* error identity that applies to the approximation of the primal formulation using the Crouzeix–Raviart element and to the approximation of the dual formulation using the Raviart–Thomas element, and leads to error decay rates that are optimal with respect to the regularity of a dual solution.

Keywords: Variational problems with gradient constraints, Elasto-plastic torsion; Crouzeix–Raviart element; Raviart–Thomas element, *a priori* error identity; *a posteriori* error identity.

AMS MSC (2020): 49J40; 49M29; 65N30; 65N15; 65N50.

1. Introduction

The present article aims to study the following variational problem with gradient constraints:

$$\min_{v \in K} \left\{ I(v) := \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - (f,v)_\Omega - \langle g,v \rangle_{\Gamma_N} \right\} \tag{1.1}$$

over the convex set

$$K := \left\{ v \in W^{1,\infty}(\Omega) \mid |\nabla v| \le \zeta \text{ a.e. in } \Omega,\ v = u_D \text{ a.e. on } \Gamma_D \right\}.$$

Here, $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, is a bounded simplicial Lipschitz domain with (topological) boundary $\partial\Omega$ that is split into the Dirichlet (*i.e.*, Γ_D) and Neumann (*i.e.*, Γ_N) parts. The functions $f \in L^1(\Omega)$, $g \in W^{-1,1}(\Gamma_N)$, $u_D \in W^{1,\infty}(\Gamma_D)$ represent the load, the Neumann and the Dirichlet boundary data, respectively. Apart from that, the function $\zeta \in L^\infty(\Omega)$ represents the gradient obstacle function. Under generic assumptions on the data (*cf.* Section 3), by standard arguments, one can establish that (1.1) admits a unique solution (*cf.* [12, 21]).

[†]Email: `hantil@gmu.edu`
[‡]Email: `bartels@mathematik.uni-freiburg.de`
[§]Email: `kaltenbach@math.tu-berlin.de`
[¶]Email: `rkhandel@gmu.edu`

This model problem arises in various applications, *e.g.*, (i) *elasto-plastic torsion in structural engineering* (*cf.* [16, Chap. 2]), (ii) *stochastic optimal control* (*cf.* [30]), (iii) *mathematical finance* (*cf.* [29]). Despite it being of significant interest, there exist (a) no provably convergent solvers, which also take discretization into account, for this problem and (b) no *a posteriori* or *a priori* error estimates in full generality. The goal of this paper is to close both these gaps. During this process, several new tools, which are of interest on their own, have been developed.

A key challenge in developing a solver for (1.1) is on how to handle the projection onto the constraint set K. This is highly non-trivial no matter what discretization is used. To overcome this, motivated by [6], we consider a primal-dual formulation. The main contributions are stated next:

1.1 Main contributions

1. This article works under *full generality* of problem (1.1): more precisely, for the data in (1.1), we only assume that $f \in L^1(\Omega)$, $g \in W^{-1,1}(\Gamma_N)$, $\zeta \in L^\infty(\Omega)$, and $u_D \in W^{1,\infty}(\Gamma_D)$ such that there exists a trace lift $\widehat{u}_D \in W^{1,\infty}(\Omega)$ satisfying $\|\frac{\nabla \widehat{u}_D}{\zeta}\|_{\infty,\Omega} < \infty$. In this context, the Sobolev space $W^{1,\infty}(\Omega)$ turns out to provide the correct norm topology on the convex set K, so that well-established convex duality methods (*e.g.*, the celebrated Fenchel duality theorem (*cf.* [13, Rem. 4.2, (4.21), p. 61])) can be applied.

2. A *thorough characterization* of the (Fenchel) dual problem (in the sense of [13, Rem. 4.2, p. 60/61]) as well as convex optimality relations are provided in Theorem 3.1, including a strong duality relation. The derived (Fenchel) dual problem is defined on $(L^\infty(\Omega))^*$ (*i.e.*, the dual space of essentially bounded Lebesgue measurable functions $L^\infty(\Omega)$), which is isometrically isomorphic to $(\mathrm{ba}(\Omega))^d$ (*i.e.*, the space of bounded and finitely additive vector measures).

3. At the continuous level, for arbitrary conforming approximations of the primal and dual problem, an *a posteriori* error estimator (*i.e.*, *primal-dual gap estimator*) is derived in Lemma 4.1. This is followed by the *primal-dual total error* in Lemma 4.3. Both these error quantities are shown to be exactly equal by the *primal-dual gap identity* in Theorem 4.5.

4. The results from the second and third bullets above also hold at the *discrete level for appropriate finite element discretizations*. In particular, Theorem 5.2 establishes a strong discrete duality relation and convex optimality relations. Theorem 5.4 shows the reconstruction of the discrete primal solution from the discrete dual solution via an *inverse generalized Marini formula*.

5. Discrete primal-dual gap estimator and discrete primal-dual total errors are respectively derived in Lemmas 6.1 and 6.3. They are again shown to be equal in Theorem 6.5. This is followed by *a priori error estimates* in Theorem 6.6.

6. The outer loop of the main algorithm drives the primal-dual gap estimator to zero (*e.g.*, via mesh refinement). The inner loop consists of a sub-problem solve (dual problem solve). An example of inner loop via gradient flow is provided in Algorithm 7.1. Other alternatives include Newton or quasi-Newton methods. The convergence of inner loop is established in Proposition 7.2.

1.2 Related contributions

The existing body of literature largely either focuses on the theoretical analysis or on the numerical analysis of simplified problems: The contributions [7, 10, 8] provide a numerical study of (1.1). They operate under an simplified scenario, in which $f \in L^1(\Omega)$ and $\zeta \in L^\infty(\Omega)$ are constants and $\partial\Omega = \Gamma_D$ with $u_D = 0$. Under these assumptions, problem (1.1) simplifies to the classical obstacle problem, where the obstacle function is given via the distance function to boundary $\partial\Omega$ (*cf.* [19, Thm. 1.2]). They derive *a priori* error estimates using classical techniques and under higher regularity assumptions. The contribution [7] equally derives an *a posteriori* error estimator using a convex duality approach. However, it poses the primal problem on the Sobolev space $W_0^{1,2}(\Omega)$. As we notice in this paper, this could lead to an ill-posed dual problem as, then, the celebrated Fenchel duality theorem (*cf.* [13, Rem. 4.2, (4.21), p. 61]) is not applicable. The recent contribution [9] treats non-constant $f \in L^1(\Omega)$ and imposes the gradient constraint via penalization. *A priori* error estimates between the penalized solution and its approximation are derived. However, the complete *a priori* error estimate is stated as an open problem under [9, Rem. 6.1].

2. Preliminaries

Throughout the entire paper, let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded simplicial Lipschitz domain such that $\partial\Omega$ is divided into two disjoint (relatively) open sets: a Dirichlet part $\Gamma_D \subseteq \partial\Omega$ with $|\Gamma_D| > 0$[1] and a Neumann part $\Gamma_N \subseteq \partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$.

For a (Lebesgue) measurable set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, and (Lebesgue) measurable functions or vector fields $v, w \colon \omega \to \mathbb{R}^\ell$, $\ell \in \mathbb{N}$, we employ the inner product $(v, w)_\omega := \int_\omega v \odot w \, \mathrm{d}x$, whenever the right-hand side is well-defined, where $\odot \colon \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}$ either denotes scalar multiplication or the Euclidean inner product. The integral mean over a (Lebesgue) measurable set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, with $|\omega| > 0$ of an integrable function or vector field $v \colon \omega \to \mathbb{R}^\ell$, $\ell \in \mathbb{N}$, is defined by $\langle v \rangle_\omega := \frac{1}{|\omega|}\int_\omega v \, \mathrm{d}x$.

2.1 Classical function spaces

For $m \in \mathbb{N}$, $p \in [1, \infty]$, and an open set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, we define
$$W^{m,p}(\omega) := \left\{ v \in L^p(\omega) \mid \mathrm{D}^{\boldsymbol{\alpha}} v \in L^p(\omega) \text{ for all } \boldsymbol{\alpha} \in (\mathbb{N}_0)^n \text{ with } |\boldsymbol{\alpha}| \leq m \right\},$$
where $\mathrm{D}^{\boldsymbol{\alpha}} := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \cdot \ldots \cdot \partial x_n^{\alpha_n}}$ and $|\boldsymbol{\alpha}| := \sum_{i=1}^n \alpha_i$ for each multi-index $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N}_0)^n$, equipped with the *Sobolev norm* $\|\cdot\|_{m,p,\omega} := \|\cdot\|_{p,\omega} + |\cdot|_{m,p,\omega}$, where $\|\cdot\|_{p,\omega} := (\int_\omega |\cdot|^p \, \mathrm{d}x)^{\frac{1}{p}}$ and
$$|\cdot|_{m,p,\omega} := \left(\sum_{\boldsymbol{\alpha} \in (\mathbb{N}_0)^n \, : \, 0 < |\boldsymbol{\alpha}| \leq m} \|\mathrm{D}^{\boldsymbol{\alpha}}(\cdot)\|_{p,\omega}^p \right)^{\frac{1}{p}}.$$
Then, for $s \in (0, \infty) \setminus \mathbb{N}$, $p \in [1, \infty]$, and an open set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, the *Sobolev–Slobodeckij semi-norm*, for every $v \in W^{m,p}(\omega)$, is defined by
$$|v|_{s,p,\omega} := \left(\sum_{|\boldsymbol{\alpha}|=m} \int_\omega \int_\omega \frac{|(\mathrm{D}^{\boldsymbol{\alpha}} v)(x) - (\mathrm{D}^{\boldsymbol{\alpha}} v)(y)|^p}{|x-y|^{p\theta + d}} \, \mathrm{d}x \, \mathrm{d}y \right)^{\frac{1}{p}},$$
where $m \in \mathbb{N}_0$ and $\theta \in (0, 1)$ are such that $s = m + \theta$. Then, for $s \in (0, \infty) \setminus \mathbb{N}$ and an open set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, the *Sobolev–Slobodeckij space* is defined by
$$W^{s,p}(\omega) := \left\{ v \in W^{m,p}(\omega) \mid |v|_{s,p,\omega} < \infty \right\},$$
where $m \in \mathbb{N}_0$ and $\theta \in (0, 1)$ are such that $s = m + \theta$.

2.1.1 Integration-by-parts formula

For $m \in \mathbb{N}$, $p, p' \in [1, \infty]$, where $\frac{1}{p} + \frac{1}{p'} = 1$, and an open set $\omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, we define
$$W^{p'}(\mathrm{div}; \Omega) := \left\{ y \in (L^{p'}(\Omega))^n \mid \mathrm{div}\, y \in L^{p'}(\Omega) \right\}.$$
Next, denote by $\mathrm{tr}(\cdot) \colon W^{1,p}(\Omega) \to W^{1-\frac{1}{p},p}(\partial\Omega)$ and $\mathrm{tr}((\cdot) \cdot n) \colon W^{p'}(\mathrm{div}; \Omega) \to (W^{1-\frac{1}{p},p}(\partial\Omega))^*$ the trace operator and the normal trace operator, respectively, where $n \colon \partial\Omega \to \mathbb{S}^{d-1}$ denotes the outward unit normal vector field to $\partial\Omega$. Then, for every $v \in W^{1,p}(\Omega)$ and $y \in W^{p'}(\mathrm{div}; \Omega)$, there holds the integration-by-parts formula (*cf.* [15, Sec. 4.3, (4.12)])
$$(\nabla v, y)_\Omega + (v, \mathrm{div}\, y)_\Omega = \langle \mathrm{tr}(y \cdot n), \mathrm{tr}(v) \rangle_{\partial\Omega}, \tag{2.1}$$
where, for every $\widehat{y} \in (W^{1-\frac{1}{p},p}(\partial\Omega))^*$, $\widehat{v} \in W^{1-\frac{1}{p},p}(\gamma)$, and $\gamma \in \{\Gamma_D, \Gamma_N, \partial\Omega\}$, we abbreviate
$$\langle \widehat{y}, \mathrm{tr}(\widehat{v}) \rangle_\gamma := \langle \widehat{y}, \mathrm{tr}(\widehat{v}) \rangle_{W^{1-\frac{1}{p},p}(\gamma)}. \tag{2.2}$$
Eventually, we employ the notation
$$W_D^{1,p}(\Omega) := \left\{ v \in W^{1,p}(\Omega) \mid \mathrm{tr}(v) = 0 \text{ a.e. on } \Gamma_D \right\},$$
$$W_N^{p'}(\mathrm{div}; \Omega) := \left\{ y \in W^{p'}(\mathrm{div}; \Omega) \mid \langle \mathrm{tr}(y \cdot n), \mathrm{tr}(v) \rangle_{\partial\Omega} = 0 \text{ for all } v \in W_D^{1,p}(\Omega) \right\}.$$
In what follows, we omit writing both $\mathrm{tr}(\cdot)$ and $\mathrm{tr}((\cdot) \cdot n)$ in this context.

---

[1] For a (Lebesgue) measurable set $M \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, we denote by $|M|$ its d-dimensional Lebesgue measure. For a $(d-1)$-dimensional submanifold $M \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, we denote by $|M|$ its $(d-1)$-dimensional Hausdorff measure.

2.1.2 Bounded and finitely additive vector measures

Denote by $\mathcal{M}(\mathrm{d}x;\Omega)$, the σ-algebra of Lebesgue measurable sets. Then, the *space of bounded and finitely additive vector measures on* $\mathcal{M}(\mathrm{d}x;\Omega)$ is defined by (*cf.* [31, Def. 4.1])

$$(\mathrm{ba}(\Omega))^d := \left\{ \nu\colon \mathcal{M}(\mathrm{d}x;\Omega)\to\mathbb{R}^d \,\middle|\, \begin{array}{l} \nu(\emptyset)=0\,,\ |\nu(\Omega)|<\infty\,,\ \nu(A\cup B)=\nu(A)+\nu(B) \\ \text{for all } A,B\in\mathcal{M}(\mathrm{d}x;\Omega) \text{ with } A\cap B=\emptyset \end{array} \right\}.$$

The *total variation norm*, for every $\nu\in(\mathrm{ba}(\Omega))^d$, is defined by

$$|\nu|(\Omega) := \sup_{(A_i)_{i\in\mathbb{N}}\in\pi(\Omega)} \sum_{i\in\mathbb{N}} |\nu(A_i)|\,,$$

where $\pi(\Omega)\subseteq 2^{\mathcal{M}(\mathrm{d}x;\Omega)}$ is the set of all countable, disjoint partitions of Ω into Lebesgue measurable sets. By Lebesgue's measure decomposition theorem, for every $\nu\in(\mathrm{ba}(\Omega))^d$, there exist unique $y\in(L^1(\Omega))^d$ and $\nu^s\in(\mathrm{ba}(\Omega))^d$ with $\nu^s\perp y\otimes\mathrm{d}x^2$, such that (*cf.* [31, Thm. 2.28])

$$\nu = y\otimes\mathrm{d}x + \nu^s \quad \text{in } (\mathrm{ba}(\Omega))^d\,. \tag{2.3}$$

The dual space of $(L^\infty(\Omega))^d$ (*i.e.*, $((L^\infty(\Omega))^d)^*$) is isometrically isomorphic to $(\mathrm{ba}(\Omega))^d$; in the sense that for every $y^*\in((L^\infty(\Omega))^d)^*$, there exists a unique $\nu\in(\mathrm{ba}(\Omega))^d$ such that (*cf.* [31, Thm. 3.1])

$$\langle y^*, y\rangle_{(L^\infty(\Omega))^d} = \int_\Omega y\,\mathrm{d}\nu \quad \text{for all } y\in(L^\infty(\Omega))^d\,,$$

$$\|y^*\|_{((L^\infty(\Omega))^d)^*} = |\nu|(\Omega)\,.$$

2.2 Triangulations and standard finite element spaces

In what follows, we denote by $\{\mathcal{T}_h\}_{h>0}$ a family of uniformly shape regular triangulations of $\Omega\subseteq\mathbb{R}^d$, $d\in\mathbb{N}$, (*cf.* [14]), where $h>0$ refers to the *averaged mesh-size, i.e.,* $h:=(|\Omega|/\mathrm{card}(\mathcal{N}_h))^{\frac{1}{d}}$, where $\mathcal{N}_h$ contains the vertices of $\mathcal{T}_h$. We define the following sets of sides:

$$\mathcal{S}_h := \mathcal{S}_h^i\cup\mathcal{S}_h^{\partial\Omega}\,,$$
$$\mathcal{S}_h^i := \{T\cap T' \mid T,T'\in\mathcal{T}_h\,, \dim_{\mathscr{H}}(T\cap T')=d-1\}\,,$$
$$\mathcal{S}_h^{\partial\Omega} := \{T\cap\partial\Omega \mid T\in\mathcal{T}_h\,, \dim_{\mathscr{H}}(T\cap\partial\Omega)=d-1\}\,,$$
$$\mathcal{S}_h^\gamma := \{S\in\mathcal{S}_h^{\partial\Omega} \mid \mathrm{int}(S)\subseteq\gamma\} \text{ for } \gamma\in\{\Gamma_D,\Gamma_N\}\,,$$

where the Hausdorff dimension is defined by $\dim_{\mathscr{H}}(M):=\inf\{d'\geq 0 \mid \mathscr{H}^{d'}(M)=0\}$ for all $M\subseteq\mathbb{R}^d$. It is also assumed that the triangulations $\{\mathcal{T}_h\}_{h>0}$ and boundary parts Γ_D and Γ_N are chosen such that $\mathcal{S}_h^{\partial\Omega}=\mathcal{S}_h^{\Gamma_D}\dot{\cup}\mathcal{S}_h^{\Gamma_N}$, *e.g.*, in the case $d=2$, the boundary parts $\overline{\Gamma}_D$ and $\overline{\Gamma}_N$ touch only in vertices.

For $k\in\mathbb{N}\cup\{0\}$ and $T\in\mathcal{T}_h$, let $\mathbb{P}^k(T)$ denote the set of polynomials of maximal degree k on T. Then, for $k\in\mathbb{N}\cup\{0\}$, the set of element-wise polynomial functions is defined by

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

For $\ell\in\mathbb{N}$, the (local) L^2-projection $\Pi_h\colon (L^1(\Omega))^\ell\to(\mathcal{L}^0(\mathcal{T}_h))^\ell$ onto element-wise constant functions or vector fields, respectively, for every $v\in(L^1(\Omega))^\ell$ is defined by $\Pi_h v|_T:=\langle v\rangle_T$ for all $T\in\mathcal{T}_h$.

For $m\in\mathbb{N}\cup\{0\}$ and $S\in\mathcal{S}_h$, let $\mathbb{P}^m(S)$ denote the set of polynomials of maximal degree m on S. Then, for $m\in\mathbb{N}\cup\{0\}$, the set of side-wise polynomial functions is defined by

$$\mathcal{L}^m(\mathcal{S}_h) := \{v_h\in L^\infty(\cup\mathcal{S}_h) \mid v_h|_T\in\mathbb{P}^m(S) \text{ for all } S\in\mathcal{S}_h\}\,.$$

For $\ell\in\mathbb{N}$, the (local) L^2-projection $\pi_h\colon (L^1(\cup\mathcal{S}_h))^\ell\to(\mathcal{L}^0(\mathcal{S}_h))^\ell$ onto side-wise constant functions or vector fields, respectively, for every $v\in(L^1(\cup\mathcal{S}_h))^\ell$ is defined by $\pi_h v|_S:=\langle v\rangle_S$ for all $S\in\mathcal{S}_h$.

For every $v_h\in\mathcal{L}^k(\mathcal{T}_h)$ and $S\in\mathcal{S}_h$, the *jump across* S is defined by

$$[\![v_h]\!]_S := \begin{cases} v_h|_{T_+} - v_h|_{T_-} & \text{if } S\in\mathcal{S}_h^i\,, \text{ where } T_+,T_-\in\mathcal{T}_h \text{ satisfy } \partial T_+\cap\partial T_-=S\,, \\ v_h|_T & \text{if } S\in\mathcal{S}_h^{\partial\Omega}\,, \text{ where } T\in\mathcal{T}_h \text{ satisfies } S\subseteq\partial T\,. \end{cases}$$

---

^{2}For every $y\in(L^1(\Omega))^d$, we employ the notation $\langle y\otimes\mathrm{d}x,\widehat{y}\rangle_{(L^\infty(\Omega))^d}:=\int_\Omega y\cdot\widehat{y}\,\mathrm{d}x$ for all $\widehat{y}\in(L^\infty(\Omega))^d$.

For every $y_h \in (\mathcal{L}^k(\mathcal{T}_h))^d$ and $S \in \mathcal{S}_h$, the *normal jump across S* is defined by

$$[\![y_h \cdot n]\!]_S := \begin{cases} y_h|_{T_+} \cdot n_{T_+} + y_h|_{T_-} \cdot n_{T_-} & \text{if } S \in \mathcal{S}_h^i\,, \text{ where } T_+, T_- \in \mathcal{T}_h \text{ satisfy } \partial T_+ \cap \partial T_- = S\,, \\ y_h|_T \cdot n & \text{if } S \in \mathcal{S}_h^{\partial\Omega}\,, \text{ where } T \in \mathcal{T}_h \text{ satisfies } S \subseteq \partial T\,, \end{cases}$$

where, for every $T \in \mathcal{T}_h$, $n_T \colon \partial T \to \mathbb{S}^{d-1}$ denotes the outward unit normal vector field to T.

2.2.1 Crouzeix–Raviart element

The *Crouzeix–Raviart finite element space* (*cf.* [11]) is defined as

$$\mathcal{S}^{1,cr}(\mathcal{T}_h) := \left\{ v_h \in \mathcal{L}^1(\mathcal{T}_h) \mid \pi_h[\![v_h]\!] = 0 \text{ a.e. on } \cup \mathcal{S}_h^i \right\}.$$

The Crouzeix–Raviart finite element space with homogeneous Dirichlet boundary condition on Γ_D is defined by

$$\mathcal{S}_D^{1,cr}(\mathcal{T}_h) := \left\{ v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h) \mid \pi_h v_h = 0 \text{ a.e. on } \cup \mathcal{S}_h^{\Gamma_D} \right\}.$$

Denote by $\varphi_S \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying $\varphi_S(x_{S'}) = \delta_{S,S'}$ for all $S, S' \in \mathcal{S}_h$, a basis of $\mathcal{S}^{1,cr}(\mathcal{T}_h)$. Then, the (Fortin) quasi-interpolation operator $\Pi_h^{cr} \colon W^{1,1}(\Omega) \to \mathcal{S}^{1,cr}(\mathcal{T}_h)$ (*cf.* [15, Secs. 36.2.1, 36.2.2]), for every $v \in W^{1,1}(\Omega)$ defined by

$$\Pi_h^{cr} v := \sum_{S \in \mathcal{S}_h} \langle v \rangle_S \, \varphi_S\,, \tag{2.4}$$

preserves averages of gradients and moments (on sides), *i.e.*, for every $v \in W^{1,1}(\Omega)$, it holds that

$$\nabla_h \Pi_h^{cr} v = \Pi_h \nabla v \quad \text{a.e. in } \Omega\,, \tag{2.5}$$
$$\pi_h \Pi_h^{cr} v = \pi_h v \quad \text{a.e. on } \cup \mathcal{S}_h\,, \tag{2.6}$$

where $\nabla_h \colon \mathcal{L}^1(\mathcal{T}_h) \to (\mathcal{L}^0(\mathcal{T}_h))^d$ is defined by $(\nabla_h v_h)|_T := \nabla(v_h|_T)$ for all $v_h \in \mathcal{L}^1(\mathcal{T}_h)$ and $T \in \mathcal{T}_h$.

2.2.2 Raviart–Thomas element

The *(lowest order) Raviart–Thomas finite element space* (*cf.* [26]) is defined as

$$\mathcal{RT}^0(\mathcal{T}_h) := \left\{ y_h \in (\mathcal{L}^1(\mathcal{T}_h))^d \mathrel{\Big|} \begin{array}{l} y_h|_T \cdot n_T = \text{const on } \partial T \text{ for all } T \in \mathcal{T}_h\,, \\ [\![y_h \cdot n]\!]_S = 0 \text{ on } S \text{ for all } S \in \mathcal{S}_h^i \end{array} \right\}.$$

The Raviart–Thomas finite element space with homogeneous normal boundary condition on Γ_N is defined by

$$\mathcal{RT}_N^0(\mathcal{T}_h) := \left\{ y_h \in \mathcal{RT}^0(\mathcal{T}_h) \mid y_h \cdot n = 0 \text{ a.e. on } \Gamma_N \right\}.$$

Denote by $\psi_S \in \mathcal{RT}^0(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying $\psi_S|_{S'} \cdot n_{S'} = \delta_{S,S'}$ on S' for all $S' \in \mathcal{S}_h$, a basis of $\mathcal{RT}^0(\mathcal{T}_h)$, where n_S is the unit normal vector on S pointing from T_- to T_+ if $T_+, T_- \in \mathcal{T}_h$ with $S = \partial T_+ \cap \partial T_-$. Then, the (Fortin) quasi-interpolation operator $\Pi_h^{rt} \colon V_{p,q}(\Omega) := \{y \in (L^p(\Omega))^d \mid \operatorname{div} y \in L^q(\Omega)\} \to \mathcal{RT}^0(\mathcal{T}_h)$ (*cf.* [14, Sec. 16.1]), where $p > 2$ and $q > \frac{2d}{d+2}$, for every $y \in V_{p,q}(\Omega)$ defined by

$$\Pi_h^{rt} y := \sum_{S \in \mathcal{S}_h} \langle y \cdot n_S \rangle_S \, \psi_S\,, \tag{2.7}$$

preserves averages of divergences and normal traces (on sides), *i.e.*, for every $y \in V_{p,q}(\Omega)$, it holds that

$$\operatorname{div} \Pi_h^{rt} y = \Pi_h \operatorname{div} y \quad \text{a.e. in } \Omega\,, \tag{2.8}$$
$$\Pi_h^{rt} y \cdot n = \pi_h y \cdot n \quad \text{a.e. on } \cup \mathcal{S}_h\,. \tag{2.9}$$

2.2.3 Discrete integration-by-parts formula

For every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}^0(\mathcal{T}_h)$, there holds the *discrete integration-by-parts formula*

$$(\nabla_h v_h, \Pi_h y_h)_\Omega + (\Pi_h v_h, \operatorname{div} y_h)_\Omega = (\pi_h v_h, y_h \cdot n)_{\partial\Omega}\,. \tag{2.10}$$

3. Gradient constrained variational problem

In this section, we discuss a gradient constrained variational problem.

• *Primal problem.* Given $f \in L^1(\Omega)$, $g \in W^{-1,1}(\Gamma_N)$, $\zeta \in L^\infty(\Omega)$ such that $\operatorname{ess\,inf}_{x \in \Omega} \zeta(x) > 0$, and $u_D \in W^{1,\infty}(\Gamma_D)$ such that there exists a trace lift $\widehat{u}_D \in W^{1,\infty}(\Omega)$ satisfying $\|\frac{\nabla \widehat{u}_D}{\zeta}\|_{\infty,\Omega} < 1$, the *primal problem* is given via the minimization of $I: W^{1,\infty}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,\infty}(\Omega)$ defined by

$$
\begin{aligned}
I(v) &:= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 + I_K(v) - (f,v)_\Omega - \langle g,v\rangle_{\Gamma_N} \\
&:= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 + I_{K_\zeta(0)}^\Omega(\nabla v) - (f,v)_\Omega - \langle g,v\rangle_{\Gamma_N} + I_{\{u_D\}}^{\Gamma_D}(v)\,,
\end{aligned}
\tag{3.1}
$$

where

$$
K := \left\{ v \in W^{1,\infty}(\Omega) \mid |\nabla v| \le \zeta \text{ a.e. in } \Omega\,,\ v = u_D \text{ a.e. on } \Gamma_D \right\},
$$

$I_K := I_{K_\zeta(0)}^\Omega \circ \nabla: W^{1,\infty}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, $I_{K_\zeta(0)}^\Omega: (L^\infty(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{y} \in (L^\infty(\Omega))^d$, is defined by

$$
I_{K_\zeta(0)}^\Omega(\widehat{y}) := \begin{cases} 0 & \text{if } |\widehat{y}| \le \zeta \text{ a.e. in } \Omega\,, \\ +\infty & \text{else}\,, \end{cases}
$$

and $I_{\{u_D\}}^{\Gamma_D}: W^{1,\infty}(\Gamma_D) \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{v} \in W^{1,\infty}(\Gamma_D)$, is defined by

$$
I_{\{u_D\}}^{\Gamma_D}(\widehat{v}) := \begin{cases} 0 & \text{if } \widehat{v} = u_D \text{ a.e. on } \Gamma_D\,, \\ +\infty & \text{else}\,. \end{cases}
$$

Since the functional (3.1) is proper, strictly convex, weakly coercive, and lower semi-continuous, the direct method in the calculus of variations yields the existence of a unique minimizer $u \in K$, called *primal solution*. We reserve the notation $u \in K$ for the primal solution.

• *Dual problem.* A *(Fenchel) dual problem* (in the sense of [13, Rem. 4.2, p. 60/61]) to the minimization of (3.1) is given via the maximization of $D: (\mathrm{ba}(\Omega))^d \to \mathbb{R} \cup \{-\infty\}$, for every $\nu = y \otimes \mathrm{d}x + \nu^s \in (\mathrm{ba}(\Omega))^d$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (*cf.* (2.3)), defined by[3]

$$
\begin{aligned}
D(\nu) := &-\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x - |\tfrac{1}{\zeta}\nu^s|(\Omega) - I_{K^*}(\nu) \\
&+ \langle \nu, \nabla \widehat{u}_D\rangle_{(L^\infty(\Omega))^d} - (f,\widehat{u}_D)_\Omega - \langle g,\widehat{u}_D\rangle_{\Gamma_N}\,,
\end{aligned}
\tag{3.2}
$$

where

$$
K^* := \left\{ \nu \in (\mathrm{ba}(\Omega))^d \mid \langle \nu, \nabla v\rangle_{(L^\infty(\Omega))^d} = (f,v)_\Omega + \langle g,v\rangle_{\Gamma_N} \text{ for all } v \in W_D^{1,\infty}(\Omega) \right\},
$$

$I_{K^*}: (\mathrm{ba}(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{\nu} \in (\mathrm{ba}(\Omega))^d$, is defined by

$$
I_{K^*}(\widehat{\nu}) := \begin{cases} 0 & \text{if } \widehat{\nu} \in K^*\,, \\ +\infty & \text{else}\,, \end{cases}
$$

and $\phi^*: \Omega \times \mathbb{R}^d \to \mathbb{R}$, for a.e. $x \in \Omega$ and every $s \in \mathbb{R}^d$, defined by

$$
\phi^*(x,s) = \begin{cases} \tfrac{1}{2}|s|^2 & \text{if } |s| \le \zeta(x)\,, \\ \zeta(x)|s| - \tfrac{\zeta(x)^2}{2} & \text{if } |s| > \zeta(x)\,, \end{cases}
$$

denotes the Fenchel conjugate to $\phi: \Omega \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ (with respect to the second argument), for a.e. $x \in \Omega$ and every $t \in \mathbb{R}^d$, defined by

$$
\phi(x,t) := \tfrac{1}{2}|t|^2 + \begin{cases} 0 & \text{if } |t| \le \zeta(x)\,, \\ +\infty & \text{if } |t| > \zeta(x)\,. \end{cases}
$$

---

[3]Here, $\tfrac{1}{\zeta}\nu^s \in (\mathrm{ba}(\Omega))^d$ is defined by $(\tfrac{1}{\zeta}\nu^s)(A) := \int_A \tfrac{1}{\zeta}\,\mathrm{d}\nu^s$ for all $A \in \mathcal{M}(\mathrm{d}x;\Omega)$.

Appealing to [27, Thm. 2], for every $y \in W^1(\mathrm{div}; \Omega)$, we have that

$$D(y \otimes \mathrm{d}x) = -\int_\Omega \phi^*(\cdot, y)\,\mathrm{d}x - I^\Omega_{\{-f\}}(\mathrm{div}\,y) - I^{\Gamma_N}_{\{g\}}(y \cdot n) + \langle y, u_D\rangle_{\partial\Omega} - \langle g, u_D\rangle_{\Gamma_N}\,,$$

where $I^\Omega_{\{-f\}} \colon (L^1(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{v} \in L^1(\Omega)$, is defined by

$$I^\Omega_{\{-f\}}(\widehat{v}) := \begin{cases} 0 & \text{if } \widehat{v} = -f \text{ a.e. in } \Omega\,, \\ +\infty & \text{else}\,, \end{cases}$$

and $I^{\Gamma_N}_{\{g\}} \colon W^{-1,1}(\partial\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{v} \in W^{-1,1}(\partial\Omega)$, is defined by

$$I^{\Gamma_N}_{\{g\}}(\widehat{v}) := \begin{cases} 0 & \text{if } \langle \widehat{v}, v\rangle_{\partial\Omega} = \langle g, v\rangle_{\Gamma_N} \text{ for all } v \in W^{1,\infty}_D(\Omega)\,, \\ +\infty & \text{else}\,. \end{cases}$$

The identification of the (Fenchel) dual problem (in the sense of [13, Rem. 4.2, p. 60/61]) to the minimization of (3.1) with the maximization of (3.2) can be found in the proof of the following result that also establishes the validity of a strong duality relation and convex optimality relations.

Theorem 3.1 (strong duality and convex optimality relations). *The following statements apply:*
(i) A (Fenchel) dual problem to the minimization of (3.1) is given via the maximization of (3.2).
(ii) There exists a maximizer $\mu = z \otimes \mathrm{d}x + \mu^s \in (\mathrm{ba}(\Omega))^d$, where $z \in (L^1(\Omega))^d$ and $\mu^s \in (\mathrm{ba}(\Omega))^d$ with $\mu^s \perp z \otimes \mathrm{d}x$ (cf. (2.3)), of (3.2) satisfying the admissibility condition

$$\langle \mu, \nabla v\rangle_{(L^\infty(\Omega))^d} = (f, v)_\Omega + \langle g, v\rangle_{\Gamma_N} \quad \text{for all } v \in W^{1,\infty}_D(\Omega)\,. \tag{3.3}$$

In addition, there holds a strong duality relation, *i.e., we have that*

$$I(u) = D(\mu)\,. \tag{3.4}$$

(iii) There hold convex optimality relations, *i.e., we have that*

$$\nabla u = D_t \phi^*(\cdot, z) = \begin{cases} z & \text{if } |z| \leq \zeta\,, \\ \zeta \frac{z}{|z|} & \text{if } |z| > \zeta \end{cases} \quad \text{a.e. in } \Omega\,, \tag{3.5}$$

$$\left|\tfrac{1}{\zeta}\mu^s\right|(\Omega) = \langle \mu^s, \nabla u\rangle_{(L^\infty(\Omega))^d}\,. \tag{3.6}$$

Remark 3.2. *(i) By the standard equality condition in the Fenchel–Young inequality (cf. [13, Prop. 5.1, p. 21]), the convex optimality relation (3.5) is equivalent to*

$$z \cdot \nabla u = \phi^*(\cdot, z) + \phi(\cdot, \nabla u) \quad \text{a.e. in } \Omega\,. \tag{3.7}$$

In addition, by [13, Cor. 5.2(5.8), p. 22], the convex optimality relation (3.5) is equivalent to

$$z \in \partial_t\phi(\cdot, \nabla u) \quad \text{a.e. in } \Omega\,. \tag{3.8}$$

(ii) If $\mu = z \otimes \mathrm{d}x \in (\mathrm{ba}(\Omega))^d$, i.e., $\mu^s = 0$ in the Lebesgue decomposition (2.3), then from the admissibility condition (3.3), it follows that $z \in W^1(\mathrm{div}; \Omega)$ with

$$\mathrm{div}\,z = -f \quad \text{a.e. in } \Omega\,, \tag{3.9}$$

$$\langle z \cdot n, v\rangle_{\partial\Omega} = \langle g, v\rangle_{\Gamma_N} \quad \text{for all } v \in W^{1,\infty}_D(\Omega)\,. \tag{3.10}$$

Proof (of Theorem 3.1). ad (i). First, we introduce the functionals $G \colon (L^\infty(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$ and $F \colon W^{1,\infty}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $y \in (L^\infty(\Omega))^d$ and $v \in W^{1,\infty}(\Omega)$, respectively, defined by

$$G(y) := \int_\Omega \phi(\cdot, \nabla v)\,\mathrm{d}x$$
$$= \tfrac{1}{2}\|y\|^2_{2,\Omega} + I^\Omega_{K_\zeta(0)}(y)\,,$$
$$F(v) := -(f, v)_\Omega - \langle g, v\rangle_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(v)\,.$$

Then, for every $v \in W^{1,\infty}(\Omega)$, we have that

$$I(v) = G(\nabla v) + F(v)\,.$$

Thus, in accordance with [13, Rem. 4.2, p. 60/61], the (Fenchel) dual problem to the minimization of (3.1) is given via the maximization of $D\colon (\mathrm{ba}(\Omega))^d \to \mathbb{R} \cup \{-\infty\}$, for every $\nu \in (\mathrm{ba}(\Omega))^d$ defined by

$$D(\nu) := -G^*(\nu) - F^*(-\nabla^*\nu)\,, \tag{3.11}$$

where $\nabla^*\colon (\mathrm{ba}(\Omega))^d \to (W^{1,\infty}(\Omega))^*$ denotes the adjoint operator to $\nabla\colon W^{1,\infty}(\Omega) \to (L^\infty(\Omega))^d$.

First, resorting to [28, Thm. 1], for every $\nu = y \otimes \mathrm{d}x + \nu^s \in (\mathrm{ba}(\Omega))^d$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (cf. (2.3)), we find that

$$G^*(\nu) = \int_\Omega \phi^*(\cdot, y)\,\mathrm{d}x + (I_C)^*(\nu^s)\,, \tag{3.12}$$

where

$$C := \{y \in (L^\infty(\Omega))^d \mid G(y) < +\infty\} = \{y \in (L^\infty(\Omega))^d \mid |y| \le \zeta \text{ a.e. in } \Omega\}\,,$$

$I_C\colon (L^\infty(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$, for every $\hat{y} \in (L^\infty(\Omega))^d$, is defined by

$$I_C(\hat{y}) := \begin{cases} 0 & \text{if } \hat{y} \in C\,, \\ +\infty & \text{else}\,, \end{cases}$$

and, thus,

$$(I_C)^*(\nu^s) = \sup_{y \in C} \langle \nu^s, y \rangle_{(L^\infty(\Omega))^d} = |\tfrac{1}{\zeta}\nu^s|(\Omega)\,. \tag{3.13}$$

Second, for every $\nu \in (\mathrm{ba}(\Omega))^d$, we have that

$$\begin{aligned} F^*(-\nabla^*\nu) &= \sup_{v \in W^{1,\infty}(\Omega)} \left\{ -\langle \nabla^*\nu, v \rangle_{W^{1,\infty}(\Omega)} + (f,v)_\Omega + \langle g, v \rangle_{\Gamma_N} - I_{\{u_D\}}^{\Gamma_D}(v) \right\} \\ &= \sup_{v \in W_D^{1,\infty}(\Omega)} \left\{ -\langle \nu, \nabla v \rangle_{(L^\infty(\Omega))^d} + (f,v)_\Omega + \langle g, v \rangle_{\Gamma_N} \right\} \\ &\quad - \langle \nu, \nabla \widehat{u}_D \rangle_{(L^\infty(\Omega))^d} + (f, \widehat{u}_D)_\Omega + \langle g, \widehat{u}_D \rangle_{\Gamma_N} \\ &= I_{K^*}(\nu) - \langle \nu, \nabla \widehat{u}_D \rangle_{(L^\infty(\Omega))^d} + (f, \widehat{u}_D)_\Omega + \langle g, \widehat{u}_D \rangle_{\Gamma_N}\,. \end{aligned} \tag{3.14}$$

In summary, using (3.12), (3.13), and (3.14) in (3.11), for every $\nu \in (\mathrm{ba}(\Omega))^d$, we arrive at the claimed representation (3.2).

ad (ii). Since both $G\colon (L^\infty(\Omega))^d \to \mathbb{R} \cup \{+\infty\}$ and $F\colon W^{1,\infty}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ are proper, convex, and lower semi-continuous and since $G\colon (L^\infty(\Omega))^d \to \mathbb{R}$ is continuous at $\nabla \widehat{u}_D \in \mathrm{dom}(G \circ \nabla)$ (recall that $\|\frac{\nabla \widehat{u}_D}{\zeta}\|_{\infty,\Omega} < 1$) with $\widehat{u}_D \in \mathrm{dom}(F)$, *i.e.*, we have that

$$G(y) \to G(\nabla \widehat{u}_D) \quad (y \to \nabla \widehat{u}_D \quad \text{in } (L^\infty(\Omega))^d)\,,$$

by the celebrated Fenchel duality theorem (*cf.* [13, Rem. 4.2, (4.21), p. 61]) and Lebesgue decomposition theorem (*cf.* (2.3)), there exists a maximizer $\mu = z \otimes \mathrm{d}x + \mu^s \in (\mathrm{ba}(\Omega))^d$, where $z \in (L^1(\Omega))^d$ and $\mu^s \in (\mathrm{ba}(\Omega))^d$ with $\mu^s \perp z \otimes \mathrm{d}x$, of (3.11) and a strong duality relation applies, *i.e.*, it holds that

$$I(u) = D(\mu)\,.$$

ad (iii). By the standard (Fenchel) convex duality theory (*cf.* [13, Rem. 4.2, (4.24), (4.25), p. 61]), there hold the convex optimality relations

$$-\nabla^*\mu \in \partial F(u)\,, \tag{3.15}$$
$$\mu \in \partial G(\nabla u)\,. \tag{3.16}$$

While the inclusion (3.15) is equivalent to the admissibility condition (3.3), the inclusion (3.14), by [28, Cor. 1B], is equivalent to

$$z \in \partial_t \phi(\cdot, \nabla u) \quad \text{a.e. in } \Omega\,, \tag{3.17}$$
$$|\tfrac{1}{\zeta}\mu^s|(\Omega) = \langle \mu^s, \nabla u \rangle_{(L^\infty(\Omega))^d}\,. \tag{3.18}$$

Eventually, by [13, Cor. 5.2(5.8), p. 22], (3.17) is equivalent to (3.5). $\qquad \square$

4. *A posteriori* ERROR ANALYSIS

In this section, resorting to convex duality arguments from [5], we derive an *a posteriori* error identity for conforming approximations of the primal problem (3.1) and the dual problem (3.2) at the same time. To this end, we introduce the *primal-dual gap estimator* $\eta^2_{\mathrm{gap}} \colon K \times K^* \to \mathbb{R}$, for every $v \in K$ and $\nu \in K^*$ defined by

$$\eta^2(v, \nu) := I(v) - D(\nu). \tag{4.1}$$

The primal-dual gap estimator (4.1) can be decomposed into two contributions that precisely measure the violation of the convex optimality relations (3.5),(3.6).

Lemma 4.1 (decomposition of the primal-dual gap estimator). *For every $v \in K$ and $\nu = y \otimes \mathrm{d}x$ $+ \nu^s \in K^*$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (cf. (2.3)), we have that*

$$\eta^2_{\mathrm{gap}}(v, \nu) = \eta^2_{\mathrm{gap},A}(v, y) + \eta^2_{\mathrm{gap},B}(v, \nu^s),$$

$$\eta^2_{\mathrm{gap},A}(v, y) := \int_\Omega \{\phi(\cdot, \nabla v) - \nabla v \cdot y + \phi^*(\cdot, y)\}\, \mathrm{d}x,$$

$$\eta^2_{\mathrm{gap},B}(v, \nu^s) := |\tfrac{1}{\zeta}\nu^s|(\Omega) - \langle \nu^s, \nabla v\rangle_{(L^\infty(\Omega))^d}.$$

Remark 4.2 (interpretation of the components of the primal-dual gap estimator).
(i) The estimator $\eta^2_{\mathrm{gap},A}$ measures the violation of the convex optimality relation (3.5);
(ii) The estimator $\eta^2_{\mathrm{gap},B}$ measures the violation of the convex optimality relation (3.6).

Proof (of Lemma 4.1). Using the admissibility condition (3.3), for every $v \in K$ and $\nu = y \otimes \mathrm{d}x + \nu^s \in K^*$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (cf. (2.3)), we find that

$$\begin{aligned}
I(v) - D(\nu) &= \int_\Omega \phi(\cdot, \nabla v)\, \mathrm{d}x - (f, v)_\Omega - \langle g, v\rangle_{\Gamma_N} \\
&\quad + \int_\Omega \phi^*(\cdot, y)\, \mathrm{d}x + |\tfrac{1}{\zeta}\nu^s|(\Omega) \\
&\quad - \langle \nu, \nabla \widehat{u}_D\rangle_{(L^\infty(\Omega))^d} + (f, \widehat{u}_D)_\Omega + \langle g, \widehat{u}_D\rangle_{\Gamma_N} \\
&= \int_\Omega \phi(\cdot, \nabla v)\, \mathrm{d}x - (f, v - \widehat{u}_D)_\Omega - \langle g, v - \widehat{u}_D\rangle_{\Gamma_N} \\
&\quad + \int_\Omega \phi^*(\cdot, y)\, \mathrm{d}x + |\tfrac{1}{\zeta}\nu^s|(\Omega) - \langle \nu, \nabla \widehat{u}_D\rangle_{(L^\infty(\Omega))^d} \\
&= \int_\Omega \phi(\cdot, \nabla v)\, \mathrm{d}x - \langle \nu, \nabla(v - \widehat{u}_D)\rangle_{(L^\infty(\Omega))^d} \\
&\quad + \int_\Omega \phi^*(\cdot, y)\, \mathrm{d}x + |\tfrac{1}{\zeta}\nu^s|(\Omega) - \langle \nu, \nabla \widehat{u}_D\rangle_{(L^\infty(\Omega))^d} \\
&= \int_\Omega \{\phi(\cdot, \nabla v) - \nabla v \cdot y + \phi^*(\cdot, y)\}\, \mathrm{d}x + |\tfrac{1}{\zeta}\nu^s|(\Omega) - \langle \nu^s, \nabla v\rangle_{(L^\infty(\Omega))^d}. \qquad \square
\end{aligned}$$

In the next step, we derive meaningful representations of the *optimal strong convexity measures* for the primal energy functional (3.1) at the primal solution $u \in K$, *i.e.*, for $\rho^2_I \colon K \to [0, +\infty)$, for every $v \in K$ defined by

$$\rho^2_I(v) := I(v) - I(u), \tag{4.2}$$

and for the negative of the dual energy functional (3.2), *i.e.*, for $\rho^2_{-D} \colon K^* \to [0, +\infty)$, for every $\nu \in K^*$ defined by

$$\rho^2_{-D}(\nu) := -D(\nu) + D(\mu), \tag{4.3}$$

which will serve as '*natural*' error quantities in the primal-dual gap identity (*cf.* Theorem 4.5).

Lemma 4.3 (representations of the optimal strong convexity measures). *The following statements apply:*

(i) *For every $v \in K$, we have that*

$$\rho_I^2(v) = \tfrac{1}{2}\|\nabla v - \nabla u\|_{2,\Omega}^2 + \left(\tfrac{|z|}{\zeta} - 1, \zeta^2 - \nabla u \cdot \nabla v\right)_{\{|\nabla u| = \zeta\}}.$$

(ii) *For every $\nu = y \otimes \mathrm{d}x + \nu^s \in K^*$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (cf. (2.3)), we have that*

$$\rho_{-D}^2(\nu) = \tfrac{1}{2}(D_t^2\phi^*\{y, z\}(y - z), y - z)_\Omega + |\tfrac{1}{\zeta}\nu^s|(\Omega) - \langle \nu^s, \nabla u \rangle_{(L^\infty(\Omega))^d},$$

where $D_t^2\phi^*\{\cdot, \cdot\}(\cdot) : \mathbb{R}^d \times \mathbb{R}^d \times \Omega \to \mathbb{R}^{d \times d}$, for a.e. $x \in \Omega$ and every $t, s \in \mathbb{R}^d$, is defined by

$$D_t^2\phi^*\{t, s\}(x) := \int_0^1 D_t^2\phi^*(x, \lambda t + (1 - \lambda)s)\,\mathrm{d}\lambda.$$

Remark 4.4. (i) *Since $\{|z| \geq \zeta\} = \{|\nabla u| = \zeta\}$ (cf. (3.5)) and $\nabla u \cdot \nabla v \leq \zeta^2$ a.e. in $\{|\nabla u| = \zeta\}$ for all $v \in K$, for every $v \in K$, we have that*

$$\left(\tfrac{|z|}{\zeta} - 1\right)(\zeta^2 - \nabla u \cdot \nabla v) \geq 0 \quad \text{a.e. in } \{|\nabla u| = \zeta\}.$$

(ii) *Since for a.e. $x \in \Omega$ and every $t \in \mathbb{R}^d$, we have that*

$$D_t^2\phi^*(x, t) = \begin{cases} \mathrm{I}_{d \times d} & \text{if } |t| \leq \zeta(x), \\ \frac{\zeta(x)}{|t|}(\mathrm{I}_{d \times d} - \frac{t \otimes t}{|t|^2}) & \text{if } |t| > \zeta(x), \end{cases} \tag{4.4}$$

for a.e. $x \in \Omega$ and every $t, s \in \mathbb{R}^d$, by the Cauchy–Schwarz inequality, we observe that

$$D_t^2\phi^*(x, t) : s \otimes s = \begin{cases} |s|^2 & \text{if } |t| \leq \zeta(x), \\ \frac{\zeta(x)}{|t|}(|s|^2 - \frac{(t \cdot s)^2}{|t|^2}) & \text{if } |t| > \zeta(x), \end{cases}$$

$$\geq \begin{cases} |s|^2 & \text{if } |t| \leq \zeta(x), \\ 0 & \text{if } |t| > \zeta(x). \end{cases}$$

Proof (of Lemma 4.3). ad (i). Using the admissibility condition (3.3) and the convex optimality relation (3.6), for every $v \in K$, we find that

$$\begin{aligned} I(v) - I(u) &= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla u\|_{2,\Omega}^2 - (f, v - u)_\Omega - \langle g, v - u \rangle_{\Gamma_N} \\ &= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla u\|_{2,\Omega}^2 - \langle \mu, \nabla v - \nabla u \rangle_{(L^\infty(\Omega))^d} \\ &= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla u\|_{2,\Omega}^2 - (z, \nabla v - \nabla u)_\Omega - \langle \mu^s, \nabla v - \nabla u \rangle_{(L^\infty(\Omega))^d} \\ &= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla u\|_{2,\Omega}^2 - (z, \nabla v - \nabla u)_\Omega + |\tfrac{1}{\zeta}\mu^s|(\Omega) - \langle \mu^s, \nabla v \rangle_{(L^\infty(\Omega))^d}. \end{aligned} \tag{4.5}$$

Next, due to the convex optimality relation (3.5), we have that

$$\begin{aligned} z &= \nabla u & \text{a.e. in } \{|\nabla u| < \zeta\}, \\ z &= \tfrac{|z|}{\zeta}\nabla u & \text{a.e. in } \{|\nabla u| = \zeta\}. \end{aligned} \tag{4.6}$$

Therefore, using (4.6) and the binomial formula in (4.5), we find that

$$\begin{aligned} I(v) - I(u) &= \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla u\|_{2,\Omega}^2 - (\nabla u, \nabla v - \nabla u)_\Omega \\ &\quad + \left(\left(\tfrac{|z|}{\zeta} - 1\right)\nabla u, \nabla u - \nabla v\right)_{\{|\nabla u| = \zeta\}} \\ &\quad + |\tfrac{1}{\zeta}\mu^s|(\Omega) - \langle \mu^s, \nabla v \rangle_{(L^\infty(\Omega))^d} \\ &= \tfrac{1}{2}\|\nabla v - \nabla u\|_{2,\Omega}^2 + \left(\tfrac{|z|}{\zeta} - 1, \zeta^2 - \nabla u \cdot \nabla v\right)_{\{|\nabla u| = \zeta\}} \\ &\quad + |\tfrac{1}{\zeta}\mu^s|(\Omega) - \langle \mu^s, \nabla v \rangle_{(L^\infty(\Omega))^d}, \end{aligned}$$

which is the claimed representation of the optimal strong convexity measure $\rho_I^2 : K \to [0, +\infty)$.

ad (ii). Using the admissibility condition (3.3) and taylor expansion, for every $\nu = y \otimes \mathrm{d}x + \nu^s \in K^*$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (*cf.* (2.3)), we find that

$$
\begin{aligned}
-D(\nu) + D(\mu) &= \int_\Omega \phi^*(\cdot, y)\,\mathrm{d}x - \int_\Omega \phi^*(\cdot, z)\,\mathrm{d}x \\
&\quad + |\tfrac{1}{\zeta}\nu^s|(\Omega) - |\tfrac{1}{\zeta}\mu^s|(\Omega) - \langle \nu - \mu, \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d} \\
&= (D_t\phi^*(\cdot, z), y - z)_\Omega + \tfrac{1}{2}(D_t^2\phi^*\{y, z\}(y - z), y - z)_\Omega \\
&\quad + |\tfrac{1}{\zeta}\nu^s|(\Omega) - |\tfrac{1}{\zeta}\mu^s|(\Omega) - \langle \nu - \mu, \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d}\,.
\end{aligned}
\tag{4.7}
$$

Next, using the convex optimality relations (3.5),(3.6), for every $\nu = y \otimes \mathrm{d}x + \nu^s \in K^*$, where $y \in (L^1(\Omega))^d$ and $\nu^s \in (\mathrm{ba}(\Omega))^d$ with $\nu^s \perp y \otimes \mathrm{d}x$ (*cf.* (2.3)), we observe that

$$
\begin{aligned}
(D_t\phi^*(\cdot, z), y - z)_\Omega - \langle \nu - \mu, \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d} &= (\nabla u, y - z)_\Omega - (y - z, \nabla\widehat{u}_D)_\Omega \\
&\quad - \langle \nu^s - \mu^s, \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d} \\
&= (y - z, \nabla u - \nabla\widehat{u}_D)_\Omega \\
&\quad + \langle \nu^s - \mu^s, \nabla u - \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d} \\
&\quad - \langle \nu^s - \mu^s, \nabla u \rangle_{(L^\infty(\Omega))^d} \\
&= \langle \nu - \mu, \nabla u - \nabla\widehat{u}_D \rangle_{(L^\infty(\Omega))^d} \\
&\quad - \langle \nu^s, \nabla u \rangle_{(L^\infty(\Omega))^d} + |\tfrac{1}{\zeta}\mu^s|(\Omega) \\
&= -\langle \nu^s, \nabla u \rangle_{(L^\infty(\Omega))^d} + |\tfrac{1}{\zeta}\mu^s|(\Omega)\,,
\end{aligned}
\tag{4.8}
$$

where we used in the last step that, due to the admissibility condition (3.3), for every $\widehat{v} \in W_D^{1,\infty}(\Omega)$, we have that

$$
\langle \nu, \nabla\widehat{v} \rangle_{(L^\infty(\Omega))^d} = \langle \mu, \nabla\widehat{v} \rangle_{(L^\infty(\Omega))^d}\,.
$$

Eventually, using (4.8) in (4.7), we conclude that the claimed representation of the optimal strong convexity measure $\rho_{-D}^2 \colon K^* \to [0, +\infty)$ applies. $\qquad\square$

Towards this end, we have everything at our disposal to establish an *a posteriori* error identity that identifies the *primal-dual total error* $\rho_{\mathrm{tot}}^2 \colon K \times K^* \to [0, +\infty)$, for every $v \in K$ and $\nu \in K^*$ defined by

$$
\rho_{\mathrm{tot}}^2(v, \nu) := \rho_I^2(v) + \rho_{-D}^2(\nu)\,,
\tag{4.9}
$$

with the primal-dual gap estimator $\eta_{\mathrm{gap}}^2 \colon K \times K^* \to [0, +\infty)$ (*cf.* (4.1)). This leads to an *a posteriori* error identity, called *primal-dual gap identity*, which can be interpreted as a generalization of the celebrated Prager–Synge identity (*cf.* [25]).

Theorem 4.5 (primal-dual gap identity)**.** *For every $v \in K$ and $\nu \in K^*$, we have that*

$$
\rho_{\mathrm{tot}}^2(v, \nu) = \eta_{\mathrm{gap}}^2(v, \nu)\,.
$$

Proof. Combining the definitions (4.1), (4.2), (4.3), (4.9), and the strong duality relation (3.4), for every $v \in K$ and $\nu \in K^*$, we find that

$$
\begin{aligned}
\rho_{\mathrm{tot}}^2(v, \nu) &= \rho_I^2(v) + \rho_{-D}^2(\nu) \\
&= I(v) - I(u) + D(\mu) - D(\nu) \\
&= I(v) - D(\nu) \\
&= \eta_{\mathrm{gap}}^2(v, \nu)\,. \qquad\square
\end{aligned}
$$

The primal-dual gap identity (*cf.* Theorem 4.5) requires to approximate the primal and the dual problem at the same time in a conforming way –a potentially computationally expensive task. On the basis of orthogonality relations between the Crouzeix–Raviart and the Raviart–Thomas element, in the subsequent section, we transfer all convex duality relations from Section 3 to a discrete level to arrive at a discrete reconstruction formula that allows us to approximate the primal and the dual problem at the same time using only the Crouzeix–Raviart element.

5. Discrete gradient constrained variational problem

In this section, we discuss a discrete gradient constrained variational problem.

- *Discrete primal problem.* Let $f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $g_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, $\zeta_h \in \mathcal{L}^0(\mathcal{T}_h)$ with ess $\inf_{x\in\Omega}\zeta_h(x) > 0$, and $u_D^h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$ such that there exists a trace lift $\widehat{u}_D \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ satisfying $\|\frac{\nabla \widehat{u}_D^h}{\zeta_h}\|_{\infty,\Omega} < 1$ be approximations of $f \in L^1(\Omega)$, $g \in W^{-1,1}(\Gamma_N)$, $\zeta \in L^\infty(\Omega)$, and $u_D \in W^{1,\infty}(\Gamma_D)$, respectively. Then, the *discrete primal problem* is given via the minimization of $I_h^{cr}\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R} \cup \{+\infty\}$, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defined by

$$
\begin{aligned}
I_h^{cr}(v_h) &:= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - I_{K_h^{cr}}(v_h) - (f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N} \\
&:= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - I_{K_{\zeta_h}(0)}^\Omega(\nabla_h v_h) - (f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N} + I_{\{u_D^h\}}^{\Gamma_D}(v_h)\,,
\end{aligned}
\tag{5.1}
$$

where

$$
K_h^{cr} := \left\{ v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h) \mid |\nabla_h v_h| \leq \zeta_h \text{ a.e. in } \Omega,\ \pi_h v_h = u_D^h \text{ a.e. on } \Gamma_D \right\},
$$

$I_{K_h^{cr}} := I_{K_{\zeta_h}(0)}^\Omega \circ \nabla_h\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, $I_{K_{\zeta_h}(0)}^\Omega\colon (\mathcal{L}^0(\mathcal{T}_h))^d \to \mathbb{R}\cup\{+\infty\}$, for every $\widehat{y}_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$, is defined by

$$
I_{K_{\zeta_h}(0)}^\Omega(\widehat{y}_h) := \begin{cases} 0 & \text{if } |\widehat{y}_h| \leq \zeta_h \text{ a.e. in } \Omega, \\ +\infty & \text{else}, \end{cases}
$$

and $I_{\{u_D^h\}}^{\Gamma_D}\colon \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D}) \to \mathbb{R}\cup\{+\infty\}$, for every $\widehat{v}_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$, is defined by

$$
I_{\{u_D^h\}}^{\Gamma_D}(\widehat{v}_h) := \begin{cases} 0 & \text{if } \widehat{v}_h = u_D^h \text{ a.e. on } \Gamma_D, \\ +\infty & \text{else}. \end{cases}
$$

Since the functional (5.1) is proper, strictly convex, weakly coercive, and lower semi-continuous, the direct method in the calculus of variations yields the existence of a unique minimizer $u_h^{cr} \in K_h^{cr}$, called *discrete primal solution*. We reserve the notation $u_h^{cr} \in K_h^{cr}$ for the discrete primal solution.

- *Discrete dual problem.* A *(Fenchel) dual problem* (in the sense of [13, Rem. 4.2, p. 60/61]) to the minimization of (5.1) is given via the maximization of $D_h^{rt}\colon \mathcal{R}T^0(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, for every $y_h \in \mathcal{R}T^0(\mathcal{T}_h)$ defined by

$$
\begin{aligned}
D_h^{rt}(y_h) &:= -\int_\Omega \phi_h^*(\cdot, \Pi_h y_h)\,\mathrm{d}x + (y_h \cdot n, u_D^h)_{\Gamma_D} - I_{K_h^{rt,*}}(y_h) \\
&:= -\int_\Omega \phi_h^*(\cdot, \Pi_h y_h)\,\mathrm{d}x + (y_h \cdot n, u_D^h)_{\Gamma_D} - I_{\{-f_h\}}^\Omega(\operatorname{div} y_h) - I_{\{g_h\}}^{\Gamma_N}(y_h \cdot n),
\end{aligned}
\tag{5.2}
$$

where

$$
K_h^{rt,*} := \left\{ y_h \in \mathcal{R}T^0(\mathcal{T}_h) \mid \operatorname{div} y_h = -f_h \text{ a.e. in } \Omega,\ y_h \cdot n = g_h \text{ a.e. on } \Gamma_N \right\},
$$

$I_{K_h^{rt,*}} := I_{\{-f_h\}}^\Omega \circ \operatorname{div} + I_{\{g_h\}}^{\Gamma_N}((\cdot) \cdot n)\colon \mathcal{R}T^0(\mathcal{T}_h) \to \mathbb{R} \cup \{+\infty\}$, $I_{\{g_h\}}^{\Gamma_N}\colon \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}) \to \mathbb{R} \cup \{+\infty\}$, for every $\widehat{v}_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, is defined by

$$
I_{\{g_h\}}^{\Gamma_N}(\widehat{v}_h) := \begin{cases} 0 & \text{if } \widehat{v}_h = g_h \text{ a.e. on } \Gamma_N, \\ +\infty & \text{else}, \end{cases}
$$

and $\phi_h^*\colon \Omega \times \mathbb{R}^d \to \mathbb{R}$, for a.e. $x \in \Omega$ and every $s \in \mathbb{R}^d$, defined by

$$
\phi_h^*(x, s) = \begin{cases} \tfrac{1}{2}|s|^2 & \text{if } |s| \leq \zeta_h(x), \\ \zeta_h(x)|s| - \tfrac{\zeta_h(x)^2}{2} & \text{if } |s| > \zeta_h(x), \end{cases}
$$

denotes the Fenchel conjugate to $\phi_h\colon \Omega \times \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ (with respect to the second argument),

for a.e. $x \in \Omega$ and every $t \in \mathbb{R}^d$, defined by

$$\phi_h(x,t) := \tfrac{1}{2}|t|^2 + \begin{cases} 0 & \text{if } |t| \leq \zeta_h(x)\,, \\ +\infty & \text{if } |t| > \zeta_h(x)\,. \end{cases}$$

Remark 5.1. *Note that the discrete dual problem is not only defined on finitely additive vector measures but on a subset of finitely additive vector measures that are absolutely continuous with respect to the d-dimensional Lebesgue measure* $\mathrm{d}x$. *Hence, if* $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $g = g_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, $u_D = u_D^h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$, *and* $\zeta = \zeta_h \in \mathcal{L}^0(\mathcal{T}_h)$, *vector fields that are admissible in the discrete dual problem, i.e.,* $y_h \in K_h^{rt,*}$, *are also are admissible the continuous dual problem, i.e.,* $y_h \in K^*$.

The identification of the (Fenchel) dual problem (in the sense of [13, Rem. 4.2, p. 60/61]) to the minimization of (5.1) with the maximization of (5.2) can be found in the proof of the following result that also establishes the validity of a strong duality relation and convex optimality relations.

Theorem 5.2 (strong duality and convex duality relations). *The following statements apply:*

(i) The (Fenchel) dual problem to the minimization of (5.1) *is given via the maximization of* (5.2).

(ii) There exists a maximizer $z_h^{rt} \in \mathcal{R}T^0(\mathcal{T}_h)$ *of* (5.2) *satisfying the discrete admissibility conditions*

$$\operatorname{div} z_h^{rt} = -f_h \qquad a.e.\ in\ \Omega\,, \tag{5.3}$$

$$z_h^{rt} \cdot n = g_h \qquad a.e.\ on\ \Gamma_N\,. \tag{5.4}$$

In addition, there holds a discrete strong duality relation, i.e., we have that

$$I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})\,. \tag{5.5}$$

(iii) If a discrete strong duality relation, then there hold the discrete convex optimality relation

$$\nabla_h u_h^{cr} = D_t \phi_h^*(\cdot, \Pi_h z_h^{rt}) = \begin{cases} \Pi_h z_h^{rt} & \text{if } |\Pi_h z_h^{rt}| \leq \zeta_h\,, \\ \zeta_h \dfrac{\Pi_h z_h^{rt}}{|\Pi_h z_h^{rt}|} & \text{if } |\Pi_h z_h^{rt}| > \zeta_h \end{cases} \qquad a.e.\ in\ \Omega\,. \tag{5.6}$$

Remark 5.3 (equivalent discrete convex optimality relations). *By the standard equality condition in the Fenchel–Young inequality (cf. [13, Prop. 5.1, p. 21]), the discrete convex optimality relation* (5.6) *is equivalent to*

$$\Pi_h z_h^{rt} \cdot \nabla_h u_h^{cr} = \phi_h^*(\cdot, \Pi_h z_h^{rt}) + \phi_h(\cdot, \nabla_h u_h^{cr}) \qquad a.e.\ in\ \Omega\,. \tag{5.7}$$

In addition, according to [13, Cor. 5.2(5.8), p. 22], the discrete convex optimality relation (5.6) *is equivalent to*

$$\Pi_h z_h^{rt} \in \partial_t \phi_h(\cdot, \nabla_h u_h^{cr}) \qquad a.e.\ in\ \Omega\,. \tag{5.8}$$

Proof (of Theorem 5.2). ad (i). First, we introduce the functionals $G_h \colon (\mathcal{L}^0(\mathcal{T}_h))^d \to \mathbb{R} \cup \{+\infty\}$ and $F_h \colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R} \cup \{+\infty\}$, for every $\overline{y}_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$ and $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, respectively, defined by

$$\begin{aligned} G_h(\overline{y}_h) &:= \int_\Omega \phi_h(\cdot, \overline{y}_h)\,\mathrm{d}x \\ &= \tfrac{1}{2}\|\overline{y}_h\|_{2,\Omega}^2 - I_{K_{\zeta_h}(0)}^\Omega(\overline{y}_h)\,, \\ F_h(v_h) &:= -(f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N} + I_{\{u_D^h\}}^{\Gamma_D}(v_h)\,. \end{aligned}$$

Then, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, we have that

$$I_h^{cr}(v_h) = G_h(\nabla_h v_h) + F_h(v_h)\,.$$

Thus, in accordance with [13, Rem. 4.2, p. 60], the (Fenchel) dual problem to the minimization of (5.1) is given via the maximization of $D_h^0 \colon (\mathcal{L}^0(\mathcal{T}_h))^d \to \mathbb{R} \cup \{-\infty\}$, for every $\overline{y}_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$ defined by

$$D_h^0(\overline{y}_h) := -G_h^*(\overline{y}_h) - F^*(-\nabla_h^* \overline{y}_h)\,, \tag{5.9}$$

where $\nabla_h^*\colon (\mathcal{L}^0(\mathcal{T}_h))^d \to (\mathcal{S}^{1,cr}(\mathcal{T}_h))^*$ denotes the adjoint operator to $\nabla_h\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to (\mathcal{L}^0(\mathcal{T}_h))^d$.

First, using element-wise the definition of the Fenchel conjugate of $\phi_h\colon \Omega \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ (with respect to the second argument), for every $\overline{y}_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$, we have that

$$G_h^*(\overline{y}_h) = \int_\Omega \phi_h^*(\cdot, \overline{y}_h)\, \mathrm{d}x. \tag{5.10}$$

Second, for every $\overline{y}_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$, using the lifting lemma [5, Lemma A.1 in the case $\Gamma_C = \emptyset$] and the discrete integration-by-parts formula (2.10), we have that

$$
\begin{aligned}
F_h^*(-\nabla_h^*\overline{y}_h) &\tag{5.11}\\
&= \sup_{v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)} \left\{ -(\overline{y}_h, \nabla_h v_h)_\Omega + (f_h, \Pi_h v_h)_\Omega + (g_h, \pi_h v_h)_{\Gamma_N} - I_{\{u_D^h\}}^{\Gamma_D}(v_h) \right\}\\
&= \sup_{v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)} \left\{ -(\overline{y}_h, \nabla_h v_h)_\Omega + (f_h, \Pi_h v_h)_\Omega + (g_h, \pi_h v_h)_{\Gamma_N} \right\}\\
&\quad - (\overline{y}_h, \nabla_h \widehat{u}_D^h)_\Omega + (f_h, \Pi_h \widehat{u}_D^h)_\Omega + (g_h, \pi_h \widehat{u}_D^h)_{\Gamma_N}\\
&= \begin{cases} \left. \begin{array}{l} I_{\{-f_h\}}^\Omega(\operatorname{div} y_h) + I_{\{g_h\}}^{\Gamma_N}(y_h \cdot n) \\ -(y_h, \nabla_h \widehat{u}_D^h)_\Omega + (f_h, \Pi_h \widehat{u}_D^h)_\Omega + (g_h, \pi_h \widehat{u}_D^h)_{\Gamma_N} \end{array} \right\} & \begin{array}{l}\text{if there exists } y_h \in \mathcal{R}T^0(\mathcal{T}_h) \\ \text{such that } \overline{y}_h = \Pi_h y_h \text{ a.e. in } \Omega, \end{array} \\ +\infty & \text{else}, \end{cases}\\
&= \begin{cases} \left. \begin{array}{l} I_{\{-f_h\}}^\Omega(\operatorname{div} y_h) + I_{\{g_h\}}^{\Gamma_N}(y_h \cdot n) \\ -(y_h \cdot n, u_D^h)_{\Gamma_D} \end{array} \right\} & \begin{array}{l}\text{if there exists } y_h \in \mathcal{R}T^0(\mathcal{T}_h) \\ \text{such that } \overline{y}_h = \Pi_h y_h \text{ a.e. in } \Omega, \end{array} \\ +\infty & \text{else}. \end{cases}
\end{aligned}
$$

Hence, using (5.10) and (5.11) in (5.9), for every $\overline{y}_h = \Pi_h y_h \in \Pi_h(\mathcal{R}T^0(\mathcal{T}_h))$, where $y_h \in \mathcal{R}T^0(\mathcal{T}_h)$, we arrive at the claimed representation (5.2). Since $D_h^0 = -\infty$ in $(\mathcal{L}^0(\mathcal{T}_h))^d \setminus \Pi_h(\mathcal{R}T^0(\mathcal{T}_h))$, we can restrict to $\Pi_h(\mathcal{R}T^0(\mathcal{T}_h))$. Finally, we simply define the functional $D_h^{rt}\colon \mathcal{R}T^0(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, for every $y_h \in \mathcal{R}T^0(\mathcal{T}_h)$, by

$$D_h^{rt}(y_h) := D_h^0(\Pi_h y_h).$$

ad (ii). Since $G_h\colon (\mathcal{L}^0(\mathcal{T}_h))^d \to \mathbb{R} \cup \{+\infty\}$ and $F_h\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$ are proper, convex, and lower semi-continuous and $G_h\colon (\mathcal{L}^0(\mathcal{T}_h))^d \to \mathbb{R} \cup \{+\infty\}$ is continuous at $\nabla_h \widehat{u}_D^h \in \operatorname{dom}(G_h \circ \nabla_h)$ (recall that $\|\frac{\nabla_h \widehat{u}_D^h}{\zeta_h}\|_{\infty,\Omega} < 1$) with $\widehat{u}_D^h \in \operatorname{dom}(F_h)$, *i.e.*, we have that

$$G_h(\overline{y}_h) \to G_h(\nabla_h \widehat{u}_D^h) \quad \left(\overline{y}_h \to \nabla_h \widehat{u}_D^h \quad \text{in } (\mathcal{L}^0(\mathcal{T}_h))^d\right),$$

by the celebrated Fenchel duality theorem (*cf.* [13, Rem. 4.2, (4.21), p. 61]), there exists a maximizer $\overline{z}_h^0 \in (\mathcal{L}^0(\mathcal{T}_h))^d$ of (5.9) and a strong duality relation applies, *i.e.*, it holds that

$$I_h^{cr}(u_h^{cr}) = D_h^0(\overline{z}_h^0).$$

Since $I_h^{cr}(u_h^{cr}) < +\infty$ and $\operatorname{dom}(-D_h^0) \subseteq \Pi_h(\mathcal{R}T^0(\mathcal{T}_h))$, there exists vector field $z_h^{rt} \in \mathcal{R}T^0(\mathcal{T}_h)$ such that $\overline{z}_h^0 = \Pi_h z_h^{rt}$ a.e. in Ω. In particular, we have that $D_h^0(\overline{z}_h^0) = D_h^{rt}(z_h^{rt})$, so that $z_h^{rt} \in \mathcal{R}T^0(\mathcal{T}_h)$ is a maximizer of (5.2) and a discrete strong duality relation applies, *i.e.*, it holds that

$$I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt}).$$

ad (iii). By the standard (Fenchel) convex duality theory (*cf.* [13, Rem. 4.2, (4.24), (4.25), p. 61]), there hold the convex optimality relations

$$-\nabla_h^* \Pi_h z_h^{rt} \in \partial F_h(u_h^{cr}), \tag{5.12}$$

$$\Pi_h z_h^{rt} \in \partial G_h(\nabla_h u_h^{cr}). \tag{5.13}$$

The inclusion (5.13) is equivalent to the discrete convex optimality relation (5.6). The inclusion (5.12) is equivalent to the discrete admissibility condition (5.3). $\qquad\square$

Given a discrete dual solution (*i.e.*, a maximizer of (5.2)) and a Lagrange multiplier associated with the discrete admissibility condition (5.3), which we obtain as a by-product in the computation of the discrete dual solution and which is actually a placeholder for the (local) L^2-projection of the discrete primal solution, the discrete primal solution is immediately available via an *inverse generalized Marini formula* (*cf.* [23, 2, 1, 6]).

Theorem 5.4 (inverse generalized Marini formula)**.** *Given a discrete dual solution $z_h^{rt} \in \mathcal{R}T^0(\mathcal{T}_h)$ and a Lagrange multiplier $\overline{\lambda}_h \in \mathcal{L}^0(\mathcal{T}_h)$ such that for every $(y_h, \overline{\eta}_h)^\top \in \mathcal{R}T_N^0(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$, it holds that*

$$-(D_t\phi_h^*(\cdot, \Pi_h z_h^{rt}), \Pi_h y_h)_\Omega - (\overline{\lambda}_h, \operatorname{div} y_h)_\Omega = -(y_h \cdot n, u_D^h)_{\Gamma_D}, \tag{5.14}$$

$$(\operatorname{div} z_h^{rt}, \overline{\eta}_h)_\Omega = -(f_h, \overline{\eta}_h)_\Omega, \tag{5.15}$$

the discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ is given via the inverse generalized Marini formula

$$u_h^{cr} = \overline{\lambda}_h + D_t\phi_h^*(\cdot, \Pi_h z_h^{rt}) \cdot (\operatorname{id}_{\mathbb{R}^d} - \Pi_h \operatorname{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega. \tag{5.16}$$

In particular, from (5.16), it follows that

$$\Pi_h u_h^{cr} = \overline{\lambda}_h \quad \text{a.e. in } \Omega. \tag{5.17}$$

Remark 5.5 (well-posedness of the non-linear saddle point formulation (5.14),(5.15))**.**

(i) *While equation (5.15) is equivalent to the discrete admissibility condition (5.3), setting $\overline{\lambda}_h := \Pi_h u_h^{cr} \in \mathcal{L}^0(\mathcal{T}_h)$ and using the discrete convex optimality relation (5.6), the equation (5.14) is an immediate consequence of the discrete integration-by-parts formula (2.10) together with $\pi_h u_h^{cr} = u_D^h$ a.e. on Γ_D.*

(ii) *The inverse generalized Marini formula (cf. (5.16)) together with the regularity of the discrete dual energy functional (5.2) motivates to, first, compute the discrete dual solution (e.g., using a semi-implicit discretized L^2-gradient flow) and, then, to compute the discrete primal solution for free via the inverse generalized Marini formula (cf. (5.16)).*

Proof (of Theorem 5.4). To begin with, we introduce the abbreviation

$$\widehat{u}_h^{cr} := \overline{\lambda}_h + D_t\phi_h^*(\cdot, \Pi_h z_h^{rt}) \cdot (\operatorname{id}_{\mathbb{R}^d} - \Pi_h \operatorname{id}_{\mathbb{R}^d}) \in \mathcal{L}^1(\mathcal{T}_h). \tag{5.18}$$

Then, by definition (5.18), we have that

$$\nabla_h \widehat{u}_h^{cr} = D_t\phi_h^*(\cdot, \Pi_h z_h^{rt}) \quad \text{a.e. in } \Omega, \tag{5.19}$$

$$\Pi_h \widehat{u}_h^{cr} = \overline{\lambda}_h \quad \text{a.e. in } \Omega. \tag{5.20}$$

From the equations (5.14),(5.20) and $\nabla_h u_h^{cr} = D_t\phi_h^*(\cdot, \Pi_h z_h^{rt})$ (*cf.* (5.6)), for every $y_h \in \mathcal{R}T_N^0(\mathcal{T}_h)$, it follows that

$$(\nabla_h u_h^{cr}, \Pi_h y_h)_\Omega + (\Pi_h \widehat{u}_h^{cr}, \operatorname{div} y_h)_\Omega - (u_D^h, y_h \cdot n)_{\Gamma_D} = 0. \tag{5.21}$$

Resorting to the discrete integration-by-parts formula (2.10) in (5.21), for every $y_h \in \mathcal{R}T_N^0(\mathcal{T}_h)$, we find that

$$\begin{aligned}
(\widehat{u}_h^{cr} - u_h^{cr}, \operatorname{div} y_h)_\Omega &= (\Pi_h \widehat{u}_h^{cr} - \Pi_h u_h^{cr}, \operatorname{div} y_h)_\Omega \\
&= (\Pi_h \widehat{u}_h^{cr}, \operatorname{div} y_h)_\Omega + (\nabla_h u_h^{cr}, \Pi_h y_h)_\Omega - (y_h \cdot n, \pi_h u_h^{cr})_{\Gamma_D} \\
&= 0.
\end{aligned} \tag{5.22}$$

Since, on the other hand, we have that $\nabla_h(\widehat{u}_h^{cr} - u_h^{cr}) = 0$ a.e. in Ω, *i.e.*, $\widehat{u}_h^{cr} - u_h^{cr} \in \mathcal{L}^0(\mathcal{T}_h)$, (5.22) implies that

$$\widehat{u}_h^{cr} - u_h^{cr} \in (\operatorname{div}(\mathcal{R}T_N^0(\mathcal{T}_h)))^\perp, \tag{5.23}$$

where the orthogonality needs to be understood in $\mathcal{L}^0(\mathcal{T}_h)$ equipped with $(\cdot, \cdot)_\Omega$. Due to $|\Gamma_D| > 0$, and, thus, $\Gamma_N \neq \partial\Omega$, we have that

$$\operatorname{div}(\mathcal{R}T_N^0(\mathcal{T}_h)) = \mathcal{L}^0(\mathcal{T}_h),$$

so that from (5.23), we conclude that $\widehat{u}_h^{cr} = u_h^{cr}$ in $\mathcal{S}^{1,cr}(\mathcal{T}_h)$. $\qquad\square$

6. *A priori* error analysis

In this section, resorting to the discrete convex duality relations established in Section 5, following [5], we derive an *a priori* error identity for the discrete primal problem and the discrete dual problem at the same time. From this *a priori* error identity, we derive an *a priori* error estimate with an explicit error decay rate that is quasi-optimal. To this end, we proceed analogously to the continuous setting and start with examining the *discrete primal-dual gap estimator* $\eta^2_{\text{gap},h} \colon K^{cr}_h \times K^{rt,*}_h \to [0,+\infty)$, for every $v_h \in K^{cr}_h$ and $y_h \in K^{rt,*}_h$ defined by

$$\eta^2_{\text{gap},h}(v_h, y_h) := I^{cr}_h(v_h) - D^{rt}_h(y_h). \tag{6.1}$$

The discrete primal-dual gap estimator has the following integral representation.

Lemma 6.1. *For every $v_h \in K^{cr}_h$ and $y_h \in K^{rt,*}_h$, we have that*

$$\eta^2_{\text{gap},h}(v_h, y_h) := \int_\Omega \left\{ \phi^*_h(\cdot, \Pi_h y_h) - \Pi_h y_h \cdot \nabla_h v_h + \phi_h(\cdot, \nabla_h v_h) \right\} \mathrm{d}x.$$

Remark 6.2 (Interpretation of the discrete primal-dual gap estimator). *The estimator $\eta^2_{\text{gap},h}$ measures the violation of the discrete convex optimality relations* (5.6).

Proof (of Lemma 6.1). Using the discrete admissibility condition (5.3), the discrete integration-by-parts formula (2.10), and the binomial formula, for every $v_h \in K^{cr}_h$ and $y_h \in K^{rt,*}_h$, we find that

$$
\begin{aligned}
I^{cr}_h(v_h) - D^{rt}_h(y_h) &= \tfrac{1}{2}\|\nabla_h v_h\|^2_{2,\Omega} - (f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N} \\
&\quad + \int_\Omega \phi^*_h(\cdot, \Pi_h y_h)\,\mathrm{d}x - (y_h \cdot n, u^h_D)_{\Gamma_D} \\
&= \int_\Omega \phi_h(\cdot, \nabla_h v_h)\,\mathrm{d}x + (\operatorname{div} y_h, \Pi_h v_h)_\Omega - (y_h \cdot n, \pi_h v_h)_{\Gamma_N} \\
&\quad + \int_\Omega \phi^*_h(\cdot, \Pi_h y_h)\,\mathrm{d}x - (y_h \cdot n, \pi_h v_h)_{\Gamma_D} \\
&= \int_\Omega \left\{ \phi^*_h(\cdot, \Pi_h y_h) - \Pi_h y_h \cdot \nabla_h v_h + \phi_h(\cdot, \nabla_h v_h) \right\} \mathrm{d}x,
\end{aligned}
$$

which is the claimed integral representation of $\eta^2_{\text{gap},h} \colon K^{cr}_h \times K^{rt,*}_h \to \mathbb{R}$. ∎

In the next step, we derive meaningful representations of the *discrete optimal strong convexity measures* for the discrete primal energy functional (5.1) at the discrete primal solution $u^{cr}_h \in K^{cr}_h$, i.e., for $\rho^2_{I^{cr}_h} \colon K^{cr}_h \to [0,+\infty)$, for every $v_h \in K^{cr}_h$ defined by

$$\rho^2_{I^{cr}_h}(v_h) := I^{cr}_h(v_h) - I^{cr}_h(u^{cr}_h), \tag{6.2}$$

and for the negative of the discrete dual energy functional (5.2), i.e., for $\rho^2_{-D^{rt}_h} \colon K^{rt,*}_h \to [0,+\infty)$, for every $y_h \in K^{rt,*}_h$ defined by

$$\rho^2_{-D^{rt}_h}(y_h) := -D^{rt}_h(y_h) + D^{rt}_h(z^{rt}_h), \tag{6.3}$$

which will serve as 'natural' error quantities in the discrete primal-dual gap identity (*cf.* Theorem 6.5).

Lemma 6.3 (representations of the discrete optimal strong convexity measures). *The following statements apply:*

(i) For every $v_h \in K^{cr}_h$, we have that

$$\rho^2_{I^{cr}_h}(v_h, u^{cr}_h) = \tfrac{1}{2}\|\nabla_h v_h - \nabla_h u^{cr}_h\|^2_{2,\Omega} + \left(\tfrac{|\Pi_h z^{rt}_h|}{\zeta_h} - 1, \zeta^2_h - \nabla_h u^{cr}_h \cdot \nabla_h v_h\right)_{\{|\nabla_h u^{cr}_h| = \zeta_h\}}.$$

(ii) For every $y_h \in K^{rt,*}_h$, we have that

$$\rho^2_{-D^{rt}_h}(y_h, z^{rt}_h) = \tfrac{1}{2}(D^2_t \phi^*_h\{\Pi_h y_h, \Pi_h z^{rt}_h\}(\Pi_h y_h - \Pi_h z^{rt}_h), \Pi_h y_h - \Pi_h z^{rt}_h)_\Omega.$$

Remark 6.4. *(i) Since $\{|\Pi_h z_h^{rt}| \geq \zeta_h\} = \{|\nabla_h u_h^{cr}| = \zeta_h\}$ (cf. (5.6)) and $\nabla_h u_h^{cr} \cdot \nabla_h v_h \leq \zeta_h^2$ a.e. in $\{|\nabla_h u_h^{cr}| = \zeta_h\}$ for all $v_h \in K_h^{cr}$, for every $v_h \in K_h^{cr}$, we have that*

$$\left(\frac{|\Pi_h z_h^{rt}|}{\zeta_h} - 1 \right)(\zeta_h^2 - \nabla_h u_h^{cr} \cdot \nabla_h v_h) \geq 0 \quad \text{a.e. in } \{|\nabla_h u_h^{cr}| = \zeta_h\}.$$

(ii) Analogously to (4.4), since for a.e. $x \in \Omega$ and every $t \in \mathbb{R}^d$, we have that

$$D_t^2 \phi_h^*(x, t) = \begin{cases} \mathrm{I}_{d \times d} & \text{if } |t| \leq \zeta_h(x), \\ \frac{\zeta_h(x)}{|t|}(\mathrm{I}_{d \times d} - \frac{t \otimes t}{|t|^2}) & \text{if } |t| > \zeta_h(x), \end{cases} \tag{6.4}$$

for a.e. $x \in \Omega$ and every $t, s \in \mathbb{R}^d$, by the Cauchy–Schwarz inequality, we observe that

$$D_t^2 \phi_h^*(x, t) : s \otimes s = \begin{cases} |s|^2 & \text{if } |t| \leq \zeta_h(x), \\ \frac{\zeta_h(x)}{|t|}(|s|^2 - \frac{(t \cdot s)^2}{|t|^2}) & \text{if } |t| > \zeta_h(x) \end{cases}$$

$$\geq \begin{cases} |s|^2 & \text{if } |t| \leq \zeta_h(x), \\ 0 & \text{if } |t| > \zeta_h(x). \end{cases}$$

Proof (of Lemma 6.3). **ad (i).** Using the discrete integration-by-parts formula (2.10), for every $v_h \in K_h^{cr}$, due to $\pi_h v_h = \pi_h u_h^{cr}$ a.e. on Γ_D, we find that

$$\begin{aligned}
I_h^{cr}(v_h) - I_h^{cr}(u_h^{cr}) &= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_{2,\Omega}^2 \\
&\quad - (f_h, \Pi_h v_h - \Pi_h u_h^{cr})_\Omega - (g_h, \pi_h v_h - \pi_h u_h^{cr})_{\Gamma_N} \\
&= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_{2,\Omega}^2 \\
&\quad + (\operatorname{div} y_h, \Pi_h v_h - \Pi_h u_h^{cr})_\Omega - (y_h \cdot n, \pi_h v_h - \pi_h u_h^{cr})_{\partial\Omega} \\
&= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_{2,\Omega}^2 - (\Pi_h z_h^{rt}, \nabla_h v_h - \nabla_h u_h^{cr})_\Omega.
\end{aligned}$$

Next, using that, due to the discrete convex optimality relation (5.6), we have that

$$\begin{aligned}
\Pi_h z_h^{rt} &= \nabla_h u_h^{cr} & \text{a.e. in } \{|\nabla_h u_h^{cr}| < \zeta_h\}, \\
\Pi_h z_h^{rt} &= \frac{|\Pi_h z_h^{rt}|}{\zeta_h} \nabla_h u_h^{cr} & \text{a.e. in } \{|\nabla_h u_h^{cr}| = \zeta_h\}.
\end{aligned}$$

Therefore, using the binomial formula, we find that

$$\begin{aligned}
I_h^{cr}(v_h) - I_h^{cr}(u_h^{cr}) &= \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 - \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_{2,\Omega}^2 - (\nabla_h u_h^{cr}, \nabla_h v_h - \nabla_h u_h^{cr})_\Omega \\
&\quad + \left(\left(\frac{|\Pi_h z_h^{rt}|}{\zeta_h} - 1 \right) \nabla_h u_h^{cr}, \nabla_h u_h^{cr} - \nabla_h v_h \right)_{\{|\nabla_h u_h^{cr}| = \zeta_h\}} \\
&= \tfrac{1}{2}\|\nabla_h v_h - \nabla_h u_h^{cr}\|_{2,\Omega}^2 + \left(\left(\frac{|\Pi_h z_h^{rt}|}{\zeta_h} - 1 \right), \zeta_h^2 - \nabla_h u_h^{cr} \cdot \nabla_h v_h \right)_{\{|\nabla_h u_h^{cr}| = \zeta_h\}},
\end{aligned}$$

which is the claimed representation of $\rho_{I_h^{cr}}^2 \colon K_h^{cr} \to [0, +\infty)$.

ad (ii). Using the discrete admissibility condition (5.3) and Taylor expansion, for every $y_h \in K_h^{rt,*}$, due to $y_h \cdot n = z_h^{rt} \cdot n$ a.e. on Γ_N, we find that

$$\begin{aligned}
-D_h^{rt}(y_h) + D_h^{rt}(z_h^{rt}) &= \int_\Omega \phi_h^*(\cdot, \Pi_h y_h) - \int_\Omega \phi_h^*(\cdot, \Pi_h z_h^{rt}) - (y_h \cdot n - z_h^{rt} \cdot n, u_D^h)_{\Gamma_D} \\
&= (D_t \phi_h^*(\cdot, \Pi_h z_h^{rt}), \Pi_h y_h - \Pi_h z_h^{rt})_\Omega - (y_h \cdot n - z_h^{rt} \cdot n, u_D^h)_{\partial\Omega} \\
&\quad + \tfrac{1}{2}(D_t^2 \phi_h^*\{\Pi_h y_h, \Pi_h z_h^{rt}\}(\Pi_h y_h - \Pi_h z_h^{rt}), \Pi_h y_h - \Pi_h z_h^{rt})_\Omega \\
&= \tfrac{1}{2}(D_t^2 \phi_h^*\{\Pi_h y_h, \Pi_h z_h^{rt}\}(\Pi_h y_h - \Pi_h z_h^{rt}), \Pi_h y_h - \Pi_h z_h^{rt})_\Omega,
\end{aligned}$$

where we used that, by the discrete convex optimality condition (5.6), the discrete integration-by-parts formula (2.10), and the discrete admissibility condition (5.3), for every $y_h \in K_h^{rt,*}$, it holds that

$$\begin{aligned}
(D_t \phi_h^*(\cdot, \Pi_h z_h^{rt}), \Pi_h y_h - \Pi_h z_h^{rt})_\Omega &= (\nabla_h u_h^{cr}, \Pi_h y_h - \Pi_h z_h^{rt})_\Omega \\
&= (\operatorname{div} z_h^{rt} - \operatorname{div} y_h, \Pi_h u_h^{cr})_\Omega + (y_h \cdot n - z_h^{rt} \cdot n, \pi_h u_h^{cr})_{\partial\Omega} \\
&= (y_h \cdot n - z_h^{rt} \cdot n, u_D^h)_{\partial\Omega},
\end{aligned}$$

which is the claimed representation of $\rho_{-D_h^{rt}}^2 \colon K_h^{rt,*} \to [0, +\infty)$. $\qquad\square$

Eventually, we have everything at our disposal to establish a discrete *a posteriori* error identity that identifies the *discrete primal-dual total error* $\rho_{\mathrm{tot},h}^2 \colon K_h^{cr} \times K_h^{rt,*} \to [0,+\infty)$, for every $v_h \in K_h^{cr}$ and $y_h \in K_h^{rt,*}$ defined by

$$\rho_{\mathrm{tot},h}^2(v_h, y_h) := \rho_{I_h^{cr}}^2(v_h) + \rho_{-D_h^{rt}}^2(y_h)\,, \tag{6.5}$$

with the discrete primal-dual gap estimator (6.1).

Theorem 6.5 (discrete primal-dual gap identity). *For every $v_h \in K_h^{cr}$ and $y_h \in K_h^{rt,*}$, we have that*

$$\rho_{\mathrm{tot},h}^2(v_h, y_h) = \eta_{\mathrm{gap},h}^2(v_h, y_h)\,.$$

Proof. Using the definitions (6.1), (6.2), (6.3), (6.3), and the discrete strong duality relation (5.5), for every $v_h \in K_h^{cr}$ and $y_h \in K_h^{rt,*}$, we find that

$$
\begin{aligned}
\rho_{\mathrm{tot},h}^2(v_h, y_h) &= \rho_{I_h^{cr}}^2(v_h, u_h^{cr}) + \rho_{-D_h^{rt}}^2(y_h, z_h^{rt}) \\
&= I_h^{cr}(v_h) - I_h^{cr}(u_h^{cr}) + D_h^{rt}(z_h^{rt}) - D_h^{rt}(y_h) \\
&= I_h^{cr}(v_h) - D_h^{rt}(y_h) \\
&= \eta_{\mathrm{gap},h}^2(v_h, y_h)\,. \qquad\qquad\qquad\qquad\qquad \square
\end{aligned}
$$

If we insert the (Fortin) quasi-interpolations (2.4) and (2.7) of the primal and the dual solution (assuming that the singular part vanishes, *i.e.*, $\mu^s = 0$), respectively, in the discrete primal-dual identity (*cf.* Theorem 6.5), we arrive at an *a priori* error identity, which allows us to derive error decay rates in dependence of the regularity of the dual solution.

Theorem 6.6 (*a priori* error identity and error decay rates). *If $f_h := \Pi_h f \in \mathcal{L}^0(\mathcal{T}_h)$, $g_h := \pi_h g \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, $u_D^h := \pi_h u_D \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$, and $\zeta_h := \Pi_h \zeta \in \mathcal{L}^0(\mathcal{T}_h)$, the following statements apply:*

(i) If $\mu^s = 0$ and $z \in V_{p,q}(\Omega)$, where $p > 2$ and $q > \frac{2d}{d+2}$, then $\Pi_h^{cr} u \in K_h^{cr}$, $\Pi_h^{rt} z \in K_h^{rt,*}$, and

$$
\rho_{\mathrm{tot},h}^2(\Pi_h^{cr} u, \Pi_h^{rt} z) = \int_\Omega \left\{ \phi_h^*(\cdot, \Pi_h \Pi_h^{rt} z) - \Pi_h \Pi_h^{rt} z \cdot \Pi_h \nabla u + \phi_h(\cdot, \Pi_h \nabla u) \right\} \mathrm{d}x
$$

$$
\leq \tfrac{1}{2}(\sqrt{d}+1)\, \|z - \Pi_h \Pi_h^{rt} z\|_{2,\Omega}^2 + \tfrac{3}{2} \|\zeta_h - \zeta\|_{2,\Omega}^2\,.
$$

(ii) If, in addition, $z \in (W^{\nu,2}(\Omega))^d$ and $\zeta \in W^{\nu,2}(\Omega)$, where $\nu \in (0,1]$, then

$$
\rho_{\mathrm{tot},h}^2(\Pi_h^{cr} u, \Pi_h^{rt} z) \leq c\, h^{2\nu} \left(|z|_{\nu,2,\Omega}^2 + |\zeta|_{\nu,2,\Omega}^2 \right)\,.
$$

Proof. ad (i). First, using (2.5) and (2.6), we observe that

$$
\begin{aligned}
|\nabla_h \Pi_h^{cr} u| = |\Pi_h \nabla u| \leq \Pi_h |\nabla u| \leq \zeta_h &\qquad \text{a.e. in } \Omega\,, \\
\pi_h \Pi_h^{cr} u = \pi_h u = \pi_h u_D = u_D^h &\qquad \text{a.e. on } \Gamma_D\,,
\end{aligned}
$$

i.e., it holds that $\Pi_h^{cr} u \in K_h^{cr}$. Second, using (2.8) and (2.9), we observe that

$$
\begin{aligned}
\mathrm{div}\, \Pi_h^{rt} z = \Pi_h \mathrm{div}\, z = \Pi_h(-f) = -f_h &\qquad \text{a.e. in } \Omega\,, \\
\Pi_h^{rt} z \cdot n = \pi_h(z \cdot n) = \pi_h g = g_h &\qquad \text{a.e. on } \Gamma_N\,,
\end{aligned}
$$

i.e., it holds that $\Pi_h^{rt} z \in K_h^{rt,*}$. Then, using Theorem 6.5 together with Lemma 6.3 as well as (2.5), abbreviating $\widetilde{z}_h := \Pi_h \Pi_h^{rt} z \in (\mathcal{L}^0(\mathcal{T}_h))^d$, we find that

$$
\rho_{\mathrm{tot},h}^2(\Pi_h^{cr} u, \Pi_h^{rt} z) = \int_\Omega \left\{ \phi_h^*(\cdot, \widetilde{z}_h) - \widetilde{z}_h \cdot \Pi_h \nabla u + \phi_h(\cdot, \Pi_h \nabla u) \right\} \mathrm{d}x\,. \tag{6.6}
$$

Due to $\phi_h(\cdot, \Pi_h \nabla u) = \frac{1}{2} |\Pi_h \nabla u|^2$ a.e. in Ω and $\phi(\cdot, \nabla u) = \frac{1}{2} |\nabla u|^2$ a.e. in Ω, by Jensen's inequality, it holds that $\int_\Omega \phi_h(\cdot, \Pi_h \nabla u)\, \mathrm{d}x \leq \int_\Omega \phi(\cdot, \nabla u)\, \mathrm{d}x$, so that from (6.6), we infer that

$$
\rho_{\mathrm{tot},h}^2(\Pi_h^{cr} u, \Pi_h^{rt} z) \leq \int_\Omega \left\{ \phi_h^*(\cdot, \widetilde{z}_h) - \widetilde{z}_h \cdot \nabla u + \phi(\cdot, \nabla u) \right\} \mathrm{d}x\,. \tag{6.7}
$$

Then, using in (6.7) the convex optimality relations (3.5),(3.7), that by the convexity of $\phi^* \in C^1(\mathbb{R}^d)$, it holds that

$$-\phi^*(\cdot, z) \leq -\phi^*(\cdot, \widetilde{z}_h) - D_t\phi^*(\cdot, \widetilde{z}_h) \cdot (z - \widetilde{z}_h) \quad \text{a.e. in } \Omega\,,$$

we find that

$$\rho_{\text{tot},h}^2(\Pi_h^{cr}u, \Pi_h^{rt}z) \leq \int_\Omega \left\{\phi_h^*(\cdot, \widetilde{z}_h) + (z - \widetilde{z}_h) \cdot D_t\phi^*(\cdot, z) - \phi^*(\cdot, z)\right\} \mathrm{d}x$$
$$= (D_t\phi^*(\cdot, z) - D_t\phi^*(\cdot, \widetilde{z}_h), z - \widetilde{z}_h)_\Omega + \int_\Omega \left\{\phi_h^*(\cdot, \widetilde{z}_h) - \phi^*(\cdot, \widetilde{z}_h)\right\} \mathrm{d}x \qquad (6.8)$$
$$=: I_h^1 + I_h^2\,.$$

So, it is only left to estimate I_h^1 and I_h^2:

ad I_h^1. Using the Taylor expansion

$$D_t\phi^*(\cdot, z) = D_t\phi^*(\cdot, \widetilde{z}_h) + \tfrac{1}{2}D_t^2\phi^*\{z, \widetilde{z}_h\}(z - \widetilde{z}_h) \quad \text{a.e. in } \Omega\,,$$

and that, by (4.4), it holds that $|D_t^2\phi^*(x,t)| \leq \sqrt{d} + 1$ for a.e. $x \in \Omega$ and all $t \in \mathbb{R}^d$, we obtain

$$I_h^1 = \tfrac{1}{2}(D_t^2\phi^*\{z, \widetilde{z}_h\}(z - \widetilde{z}_h), z - \widetilde{z}_h)_\Omega$$
$$\leq \tfrac{1}{2}(\sqrt{d} + 1)\|z - \widetilde{z}_h\|_{2,\Omega}^2\,. \qquad (6.9)$$

ad I_h^2. Due to $\phi_h^*(\cdot, \widetilde{z}_h) = \zeta_h|\widetilde{z}_h| - \tfrac{1}{2}\zeta_h^2$ a.e. in $\{|\widetilde{z}_h| > \zeta_h\}$, $\phi^*(\cdot, \widetilde{z}_h) = \zeta|\widetilde{z}_h| - \tfrac{1}{2}\zeta^2$ a.e. in $\{|\widetilde{z}_h| > \zeta\}$, $\phi_h^*(\cdot, \widetilde{z}_h) = \tfrac{1}{2}|\widetilde{z}_h|^2$ a.e. in $\{|\widetilde{z}_h| \leq \zeta_h\}$, and $\phi^*(\cdot, \widetilde{z}_h) = \tfrac{1}{2}|\widetilde{z}_h|^2$ a.e. in $\{|\widetilde{z}_h| \leq \zeta\}$, it holds that

$$I_h^2 = \tfrac{1}{2}(2|\widetilde{z}_h| - \zeta - \zeta_h, \zeta_h - \zeta)_{\{|\widetilde{z}_h|>\max\{\zeta_h,\zeta\}\}}$$
$$- \tfrac{1}{2}\||\widetilde{z}_h| - \zeta_h\|_{2,\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}}^2 + \tfrac{1}{2}\||\widetilde{z}_h| - \zeta\|_{2,\{\zeta<|\widetilde{z}_h|\leq\zeta_h\}}^2\,. \qquad (6.10)$$

Using that $|\widetilde{z}_h|, \zeta_h \perp \zeta_h - \zeta$ in $L^2(T)$ for all $T \in \mathcal{T}_h$ and that $\{|\widetilde{z}_h| > \zeta_h\}$ can be decomposed into elements from $\mathcal{T}_h$ as $|\widetilde{z}_h|, \zeta_h \in \mathcal{L}^0(\mathcal{T}_h)$, we find that

$$\tfrac{1}{2}(2|\widetilde{z}_h| - \zeta - \zeta_h, \zeta_h - \zeta)_{\{|\widetilde{z}_h|>\max\{\zeta_h,\zeta\}\}} = \tfrac{1}{2}(2|\widetilde{z}_h| - \zeta - \zeta_h, \zeta_h - \zeta)_{\{|\widetilde{z}_h|>\zeta_h\}}$$
$$- \tfrac{1}{2}(2|\widetilde{z}_h| - \zeta - \zeta_h, \zeta_h - \zeta)_{\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}}$$
$$\leq \tfrac{1}{2}\|\zeta_h - \zeta\|_{\{|\widetilde{z}_h|>\zeta_h\}}^2$$
$$+ \tfrac{1}{2}\||\widetilde{z}_h| - \zeta\|_{2,\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}}\|\zeta_h - \zeta\|_{2,\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}} \qquad (6.11)$$
$$+ \tfrac{1}{2}\||\widetilde{z}_h| - \zeta_h\|_{2,\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}}\|\zeta_h - \zeta\|_{2,\{\zeta_h<|\widetilde{z}_h|\leq\zeta\}}$$
$$\leq \tfrac{3}{2}\|\zeta_h - \zeta\|_{2,\{|\widetilde{z}_h|>\zeta_h\}}^2\,,$$

Then, from (6.11) in (6.10), we deduce that

$$I_h^2 \leq \tfrac{3}{2}\|\zeta_h - \zeta\|_{2,\Omega}^2\,. \qquad (6.12)$$

Eventually, combining (6.12) and (6.9) in (6.8), we conclude that the claimed *a priori* error estimate applies.

ad (ii). From (i), using Jensen's inequality and the fractional approximation properties of Π_h (*cf.* [14, Thm. 18.16]) and Π_h^{rt} (*cf.* [14, Thms. 16.4, 16.6]), we conclude that

$$\rho_{\text{tot},h}^2(\Pi_h^{cr}u, \Pi_h^{rt}z) \leq c\|z - \Pi_h z\|_{2,\Omega}^2 + c\|\Pi_h(z - \Pi_h^{rt}z)\|_{2,\Omega}^2 + c\|\zeta - \Pi_h\zeta\|_{2,\Omega}^2$$
$$\leq c\|z - \Pi_h z\|_{2,\Omega}^2 + c\|z - \Pi_h^{rt}z\|_{2,\Omega}^2 + c\|\zeta - \Pi_h\zeta\|_{2,\Omega}^2$$
$$\leq ch^{2\nu}\left(|z|_{\nu,2,\Omega}^2 + |\zeta|_{\nu,2,\Omega}^2\right),$$

which is the claimed *a priori* error estimate. $\qquad \square$

7. Numerical experiments

In this section, we review the theoretical findings of Section 4 and Section 6 via numerical experiments. All experiments were conducted using the finite element software package FEniCS (version 2019.1.0, *cf.* [22]). All graphics were generated using the Matplotlib library (version 3.5.1, *cf.* [18]).

7.1 Implementation details regarding the optimization procedure

The iterative maximization of the discrete dual energy functional (5.2) is realized using a semi-implicit discretized L^2-gradient flow modified with a residual stopping criterion guaranteeing the necessary accuracy in the optimization procedure. For the semi-implicit treatment, we introduce the auxiliary function $\varphi_h\colon \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}$, for a.e. $x \in \Omega$ and every $r \geq 0$ defined by

$$\varphi_h(x, r) := \begin{cases} \frac{1}{2}r^2 & \text{if } r \leq \zeta_h(x), \\ \zeta_h(x)r - \frac{\zeta_h(x)^2}{2} & \text{if } r > \zeta_h(x). \end{cases}$$

Algorithm 7.1 (Semi-implicit discretized L^2-gradient flow). *Let $z_h^0 \in \mathcal{RT}_N^0(\mathcal{T}_h)$ with $\operatorname{div} z_h^0 = -f_h$ a.e. in Ω and $z_h^0 \cdot n = g_h$ a.e. on Γ_N and let $\tau, \varepsilon_{stop}^h > 0$. Then, for every $k \in \mathbb{N}$:*

(i) *Compute the iterates $(z_h^k, \overline{\lambda}_h^k)^\top \in \mathcal{RT}^0(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$ with $z_h^k \cdot n = g_h$ a.e. on Γ_N such that for every $(y_h, \overline{\eta}_h)^\top \in \mathcal{RT}_N^0(\mathcal{T}_h) \times \mathcal{L}^0(\mathcal{T}_h)$, it holds that*

$$\left(\Pi_h \mathrm{d}_\tau z_h^k - \frac{D_t \varphi_h^*(\cdot, |\Pi_h z_h^{k-1}|)}{|\Pi_h z_h^{k-1}|} \Pi_h z_h^k, \Pi_h y_h\right)_\Omega - (\overline{\lambda}_h^k, \operatorname{div} y_h)_\Omega = -(y_h \cdot n, u_D^h)_{\Gamma_D},$$
$$(\operatorname{div} z_h^k, \overline{\eta}_h)_\Omega = -(f_h, \overline{\eta}_h)_\Omega. \tag{7.1}$$

where $\mathrm{d}_\tau z_h^k := \frac{1}{\tau}(z_h^k - z_h^{k-1}) \in \mathcal{RT}_N^0(\mathcal{T}_h)$ denotes the backward difference quotient.

(ii) *Compute the residual $r_h^k \in \mathcal{RT}_N^0(\mathcal{T}_h)$ such that for every $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$, it holds that*

$$(r_h^k, y_h)_\Omega = -\left(\frac{D_t \varphi_h^*(\cdot, |\Pi_h z_h^k|)}{|\Pi_h z_h^k|} \Pi_h z_h^k, \Pi_h y_h\right)_\Omega - (\overline{\lambda}_h^k, \operatorname{div} y_h)_\Omega + (z_h^k \cdot n, u_D^h)_{\Gamma_D}. \tag{7.2}$$

Stop if $\|r_h^k\|_{2,\Omega} \leq \varepsilon_{stop}^h$; otherwise, increase $k \to k+1$ and continue with step (i).

The following proposition establishes the well-posedness (*i.e.*, existence of iterates), stability (*i.e.*, *a priori* bounds), and convergence (*i.e.*, the convergence of the iterates to a discrete dual solution) of Algorithm 7.1.

Proposition 7.2 (well-posedness, stability, and convergence of Algorithm 7.1). *Let the assumptions of Algorithm 7.1 be satisfied. Then, the following statements apply:*

(i) *Algorithm 7.1 is well-posed, i.e., for every $k \in \mathbb{N}$, given the most-recent iterate $z_h^{k-1} \in \mathcal{RT}^0(\mathcal{T}_h)$, there exists a unique iterate $z_h^k \in \mathcal{RT}^0(\mathcal{T}_h)$ solving (7.1).*

(ii) *Algorithm 7.1 is unconditionally strongly stable, i.e., for every $L \in \mathbb{N}$, it holds that*

$$-D_h^{rt}(z_h^L) + \tau \sum_{k=1}^L \|\Pi_h \mathrm{d}_\tau z_h^k\|_{2,\Omega}^2 \leq -D_h^{rt}(z_h^0).$$

(iii) *Algorithm 7.1 terminates after a finite number of steps, i.e., there exists $k^* \in \mathbb{N}$ such that $\|r_h^{k^*}\|_{2,\Omega} \leq \varepsilon_{stop}^h$.*

The proof of Proposition 7.2(ii) is essentially based on the following inequality.

Lemma 7.3. *For a.e. $x \in \Omega$ and every $a, b \in \mathbb{R}^d$, it holds that*

$$\frac{D_t \varphi_h^*(x, |a|)}{|a|} b \cdot (b - a) \geq \varphi_h^*(x, |b|) - \varphi_h^*(x, |a|) + \frac{1}{2} \frac{D_t \varphi_h^*(\cdot, |a|)}{|a|} |b - a|^2.$$

Proof. Follows from [4, Appx. A.2], since $\varphi_h^*(x, \cdot) \in C^1(\mathbb{R}_{\geq 0})$ and $(t \mapsto D_t \varphi_h^*(x, t)/t) \in C^0(\mathbb{R}_{\geq 0})$ is positive and non-decreasing for a.e. $x \in \Omega$. $\square$

Proof (of Proposition 7.2). ad (i). Since $D_t \varphi_h^*(x, t)/t \geq 0$ for a.e. $x \in \Omega$ and all $t \geq 0$, the well-posedness of Algorithm 7.1 is a direct consequence of the Brezzi splitting theorem (*cf.* [3, Thm. 6.4]).

ad (ii). Let $L \in \mathbb{N}$ be arbitrary. Then, for every $k \in \{1, \ldots, L\}$, choosing $y_h = \mathrm{d}_\tau z_h^k \in \mathcal{R}T_N^0(\mathcal{T}_h)$ in (7.1), due to $\operatorname{div}(\mathrm{d}_\tau z_h^k) = 0$ a.e. in Ω, we find that

$$\|\Pi_h \mathrm{d}_\tau z_h^k\|_{2,\Omega}^2 + \left(\tfrac{D_t \varphi_h^*(\cdot, |\Pi_h z_h^{k-1}|)}{|\Pi_h z_h^{k-1}|} \Pi_h z_h^k, \Pi_h \mathrm{d}_\tau z_h^k \right)_\Omega + (\mathrm{d}_\tau z_h^k \cdot n, u_D^h)_{\Gamma_D} = 0 \,. \tag{7.3}$$

According to Lemma 7.3 with $a = \Pi_h z_h^{k-1}|_T \in \mathbb{R}^d$ and $b = \Pi_h z_h^k|_T \in \mathbb{R}^d$ applied for all $T \in \mathcal{T}_h$, for every $k \in \{1, \ldots, L\}$, we have that

$$\tfrac{D_t \varphi_h^*(\cdot, |\Pi_h z_h^{k-1}|)}{|\Pi_h z_h^{k-1}|} \Pi_h z_h^k \cdot \Pi_h \mathrm{d}_\tau z_h^k \geq \mathrm{d}_\tau [\phi_h^*(\cdot, \Pi_h z_h^k)] \quad \text{a.e. in } \Omega \,. \tag{7.4}$$

Using (7.4) in (7.3), for every $k \in \{1, \ldots, L\}$, we arrive at

$$\|\Pi_h \mathrm{d}_\tau z_h^k\|_{2,\Omega}^2 - \mathrm{d}_\tau [D_h^{rt}(z_h^k)] \leq 0 \,. \tag{7.5}$$

Summation of (7.5) with respect to $k \in \{1, \ldots, L\}$, using $\sum_{k=1}^L \mathrm{d}_\tau [D_h^{rt}(z_h^k)] = D_h^{rt}(z_h^L) - D_h^{rt}(z_h^0)$, yields the claimed stability estimate.

ad (iii). Due to (ii), we have that $\|\Pi_h \mathrm{d}_\tau z_h^k\|_{2,\Omega}^2 \to 0$ $(k \to \infty)$. Thus, by $\operatorname{div}(\mathrm{d}_\tau z_h^k) = 0$ a.e. in Ω, the finite-dimensionality of $\mathcal{R}T_N^0(\mathcal{T}_h)$, and the equivalence of norms, it holds that

$$z_h^k - z_h^{k-1} \to 0 \quad \text{in } \mathcal{R}T_N^0(\mathcal{T}_h) \quad (k \to \infty) \,. \tag{7.6}$$

In addition, due to (ii), we have that $D_h^{rt}(z_h^k) \geq D_h^{rt}(z_h^0)$ and, thus, $(z_h^k)_{k \in \mathbb{N}} \subseteq \mathcal{R}T^0(\mathcal{T}_h)$ is bounded. In addition, due to [3, Prop. 7.4] together with $|\Gamma_D| > 0$, for every $k \in \mathbb{N}$, we have that

$$\begin{aligned}
\|\bar{\lambda}_h^k\|_{2,\Omega} &\leq \sup_{y_h \in \mathcal{R}T_N^0(\mathcal{T}_h) : \|y_h\|_{H(\operatorname{div};\Omega)} \leq 1} \left\{ (\bar{\lambda}_h^k, \operatorname{div} y_h)_\Omega \right\} \\
&= \sup_{y_h \in \mathcal{R}T_N^0(\mathcal{T}_h) : \|y_h\|_{H(\operatorname{div};\Omega)} \leq 1} \left\{ \left(\Pi_h \mathrm{d}_\tau z_h^k - \tfrac{D_t \varphi_h^*(\cdot, |\Pi_h z_h^{k-1}|)}{|\Pi_h z_h^{k-1}|} \Pi_h z_h^k, \Pi_h y_h \right)_\Omega + (y_h \cdot n, u_D^h)_{\Gamma_D} \right\} \\
&\leq \|\Pi_h \mathrm{d}_\tau z_h^k\|_{2,\Omega} + \|D_t \varphi_h^*(\cdot, |\Pi_h z_h^k|)\|_{2,\Omega} + c_h \|u_D^h\|_{2,\Gamma_D} \,,
\end{aligned}$$

implying that $(\bar{\lambda}_h^k)_{k \in \mathbb{N}} \subseteq \mathcal{L}^0(\mathcal{T}_h)$ is bounded. Due to the finite-dimensionality of $\mathcal{R}T^0(\mathcal{T}_h)$ and $\mathcal{L}^0(\mathcal{T}_h)$, the Bolzano–Weierstraß theorem yields a subsequences $(z_h^{k_\ell})_{\ell \in \mathbb{N}} \subseteq \mathcal{R}T^0(\mathcal{T}_h)$, $(\bar{\lambda}_h^{k_\ell})_{\ell \in \mathbb{N}} \subseteq \mathcal{L}^0(\mathcal{T}_h)$ as well as limits $\tilde{z}_h \in \mathcal{R}T^0(\mathcal{T}_h)$, $\tilde{\lambda}_h \in \mathcal{L}^0(\mathcal{T}_h)$ such that

$$z_h^{k_\ell} \to \tilde{z}_h \quad \text{in } \mathcal{R}T^0(\mathcal{T}_h) \quad (\ell \to \infty) \,, \tag{7.7}$$

$$\bar{\lambda}_h^{k_\ell} \to \tilde{\lambda}_h \quad \text{in } \mathcal{L}^0(\mathcal{T}_h) \quad (\ell \to \infty) \,. \tag{7.8}$$

Due to $\operatorname{div} z_h^k = -f_h$ a.e. in Ω and $z_h^k \cdot n = g_h$ a.e. on Γ_N for all $k \in \mathbb{N}$, (7.7) implies that

$$\operatorname{div} \tilde{z}_h = -f_h \quad \text{a.e. in } \Omega \,, \tag{7.9}$$

$$\tilde{z}_h \cdot n = g_h \quad \text{a.e. on } \Gamma_N \,. \tag{7.10}$$

In addition, due to (7.6), (7.7) implies that

$$z_h^{k_\ell - 1} \to \tilde{z}_h \quad \text{in } \mathcal{R}T_N^0(\mathcal{T}_h) \quad (\ell \to \infty) \,. \tag{7.11}$$

Thus, using (7.7)–(7.11), by passing for $\ell \to \infty$ in (7.1), for every $y_h \in \mathcal{R}T_N^0(\mathcal{T}_h)$, we obtain

$$\left(\tfrac{D_t \varphi_h^*(\cdot, |\Pi_h \tilde{z}_h|)}{|\Pi_h \tilde{z}_h|} \Pi_h \tilde{z}_h, \Pi_h y_h \right)_\Omega + (\tilde{\lambda}_h, \operatorname{div} y_h)_\Omega + (\tilde{z}_h \cdot n, u_D^h)_{\Gamma_D} = 0 \,, \tag{7.12}$$

which together with (7.9) and (7.10) proves that $\tilde{z}_h \in \mathcal{R}T^0(\mathcal{T}_h)$ is a discrete dual solution and $\tilde{\lambda}_h = \Pi_h u_h^{cr}$ a.e. in Ω. Hence, using (7.6) and (7.12), for every $y_h \in \mathcal{R}T_N^0(\mathcal{T}_h)$, we obtain

$$\begin{aligned}
(r_h^{k_\ell}, y_h)_\Omega &= \left(\tfrac{D_t \varphi_h(\cdot, |\Pi_h z_h^{k_\ell}|)}{|\Pi_h z_h^{k_\ell}|} \Pi_h z_h^{k_\ell}, \Pi_h y_h \right)_\Omega + (\bar{\lambda}_h^{k_\ell}, \operatorname{div} y_h)_\Omega + (y_h \cdot n, u_D^h)_{\Gamma_D} \\
&\to \left(\tfrac{D_t \varphi_h(\cdot, |\Pi_h \tilde{z}_h|)}{|\Pi_h \tilde{z}_h|} \Pi_h \tilde{z}_h, \Pi_h y_h \right)_\Omega + (\tilde{\lambda}_h, \operatorname{div} y_h)_\Omega + (y_h \cdot n, u_D^h)_{\Gamma_D} = 0 \quad (\ell \to \infty) \,,
\end{aligned}$$

i.e., $r_h^{k_\ell} \rightharpoonup 0$ in $\mathcal{R}T_N^0(\mathcal{T}_h)$ $(\ell \to \infty)$, and, thus, by the finite-dimensionality of $\mathcal{R}T_N^0(\mathcal{T}_h)$, $r_h^{k_\ell} \to 0$ in $\mathcal{R}T_N^0(\mathcal{T}_h)$ $(\ell \to \infty)$, implying that $r_h^{k_\ell} \to 0$ in $(L^2(\Omega))^d$ $(\ell \to \infty)$. As this argumentation remains valid for each subsequence of $(r_h^k)_{k \in \mathbb{N}} \subseteq \mathcal{R}T_N^0(\mathcal{T}_h)$, the standard convergence principle yields that $r_h^k \to 0$ in $(L^2(\Omega))^d$ $(k \to \infty)$. In particular, there exists $k^* \in \mathbb{N}$ such that $\|r_h^{k^*}\|_{2,\Omega} \leq \varepsilon_{stop}^h$. $\square$

7.2   Experimental setup

For our numerical experiments, we consider the manufactured solution in [17, Expl. 2.3.2, p. 122]. More precisely, let $\Omega := B_r^2(0) := \{x \in \mathbb{R}^2 \mid |x| < r\}$, $r > 0$, $\Gamma_D := \partial\Omega$ (*i.e.*, $\Gamma_N := \emptyset$), $f \equiv C \in L^1(\Omega)$, $u_D \equiv 0 \in W^{1,\infty}(\Gamma_D)$, and $\zeta \equiv 1 \in L^\infty(\Omega)$. Then, the primal solution $u \in K$ and the dual solution $z \in K^*$, for every $x \in \Omega$, are defined by

$$u(x) := \begin{cases} \frac{C}{4}(r^2 - |x|^2) & \text{if } C \leq \frac{2}{r}, \\ \begin{cases} r - |x| & \text{if } \frac{2}{C} \leq |x| \leq r, \\ -\frac{C}{4}|x|^2 + r - \frac{1}{C} & \text{if } 0 \leq |x| \leq \frac{2}{C}, \end{cases} & \text{if } C \geq \frac{2}{r}, \end{cases},$$

$$z(x) := -\frac{C}{2}x.$$

In what follows, we set $r = 1$ (*i.e.*, $\Omega = B_1^2(0)$). Then, using the `MatLab` (version R2024a, *cf.* [20]) library `DistMesh` (version 1.1, *cf.* [24]), we generate approximative triangulations $\mathcal{T}_{h_i}$, $i = 0, \ldots, 6$, where $h_i \approx 0.36 \times \frac{1}{2^i}$ for all $i = 0, \ldots, 6$, such that $\Omega_{h_i} \subseteq \Omega$, where $\Omega_{h_i} := \text{int}(\cup \mathcal{T}_{h_i})$ for all $i = 0, \ldots, 6$. As approximations of the Dirichlet boundary data serve $u_D^{h_i} := \pi_{h_i} u \in \mathcal{L}^0(\mathcal{S}_{h_i}^{\partial\Omega_{h_i}})$, $i = 0, \ldots, 6$.

For this series of triangulations $\mathcal{T}_{h_i}$, $i = 0, \ldots, 6$, we apply the semi-implicit discretized L^2-gradient flow (*cf.* Algorithm 7.1) with step-size $\tau = 1.0$, stopping parameter $\varepsilon_{stop}^h = 1.0 \times 10^{-4}$, and initial iterate $(z_{h_i}^0, \overline{\lambda}_{h_i}^0)^\top \in \mathcal{RT}_N^0(\mathcal{T}_{h_i}) \times \mathcal{L}^0(\mathcal{T}_{h_i})$ such that for every $(y_{h_i}, \overline{\eta}_{h_i})^\top \in \mathcal{RT}_N^0(\mathcal{T}_{h_i}) \times \mathcal{L}^0(\mathcal{T}_{h_i})$, it holds that

$$(\Pi_{h_i} z_{h_i}^0, \Pi_{h_i} y_{h_i})_{\Omega_{h_i}} - (\overline{\lambda}_{h_i}^0, \text{div } y_{h_i})_{\Omega_{h_i}} = (y_{h_i} \cdot n, u_D^{h_i})_{\partial\Omega_{h_i}},$$

$$(\text{div } z_{h_i}^0, \overline{\eta}_{h_i})_{\Omega_{h_i}} = -(f_{h_i}, \overline{\eta}_{h_i})_{\Omega_{h_i}},$$

to approximate the discrete dual solution $z_{h_i}^{rt} \in K_{h_i}^{rt,*}$, $i = 0, \ldots, 6$, the discrete Lagrange multiplier $\overline{\lambda}_{h_i}^0 \in \mathcal{L}^0(\mathcal{T}_{h_i})$, $i = 0, \ldots, 6$, and, subsequently, using inverse generalized Marini formula (*cf.* (5.16)), the discrete primal solution $u_{h_i}^{cr} \in K_{h_i}^{cr}$, $i = 0, \ldots, 6$.

For determining the convergence rates, the experimental order of convergence (EOC), *i.e.*,

$$\text{EOC}_i(e_i) := \frac{\log(e_i) - \log(e_{i-1})}{\log(h_i) - \log(h_{i-1})}, \quad i = 1, \ldots, 6,$$

where, for every $i = 0, \ldots, 6$, we denote by e_i a generic error quantity.

7.3   Numerical experiments concerning the a priori error analysis

In this subsection, we review the theoretical findings of Section 6.

More precisely, given the experimental setup of Subsection 7.2, we compute the error quantities

$$\left. \begin{aligned} e_i^{\text{tot},h_i} &:= \rho_{\text{tot},h_i}^2(\Pi_{h_i}^{cr} u, \Pi_{h_i}^{rt} z), \\ e_i^{\text{gap},h_i} &:= \eta_{\text{gap},h_i}^2(\Pi_{h_i}^{cr} u, \Pi_{h_i}^{rt} z), \end{aligned} \right\} \quad i = 0, \ldots, 6. \tag{7.13}$$

Inasmuch as $z \in (W^{1,2}(\Omega))^2$ and $\zeta \in W^{1,2}(\Omega)$, Theorem 6.6(ii) predicts the error decay rate $\mathcal{O}(h_i^2) = \mathcal{O}(N_i)$, where $N_i := \dim(\mathcal{RT}^0(\mathcal{T}_{h_i})) + \dim(\mathcal{L}^0(\mathcal{T}_{h_i}))$, $i \in \mathbb{N}$, for the discrete primal-dual total errors (*cf.* (6.5)), which are equal to the discrete primal-dual gap estimators (*cf.* (6.1)), *i.e.*, we expect (*cf.* Theorem 6.6(i))

$$\text{EOC}_i(e_i^{\text{gap}}) = \text{EOC}_i(e_i^{\text{tot}}) = 2.$$

In Figure 1, for $C \in \{2.5, 5.0, 7.5, 10.0\}$, we report the expected optimal convergence rate of about $\text{EOC}_i(e_i^{\text{tot}}) \approx \text{EOC}_i(e_i^{\text{gap}}) \approx 2$, $i = 1, \ldots, 6$, *i.e.*, an error decay of order $\mathcal{O}(h_i^2) = \mathcal{O}(N_i)$, $i = 1, \ldots, 6$, which confirms the optimality of the error decay rate derived in Theorem 6.6(ii). Moreover, we observe that the *a priori* error identity in Theorem 6.6(i) is approximatively satisfied. Note that, in the case $C \leq 2$, the dual solution is an affine polynomial, so that the latter (up to machine precision) coincide with the discrete dual solution. In addition, in the case $C \leq 2$, the primal solution is a quadratic polynomial, so that, by (2.5), (up to machine precision) we have that $\nabla_h \Pi_h^{cr} u = \nabla_h u_h^{cr}$ a.e. in Ω. For this reason, we did not add error plots for the case $C \leq 2$.
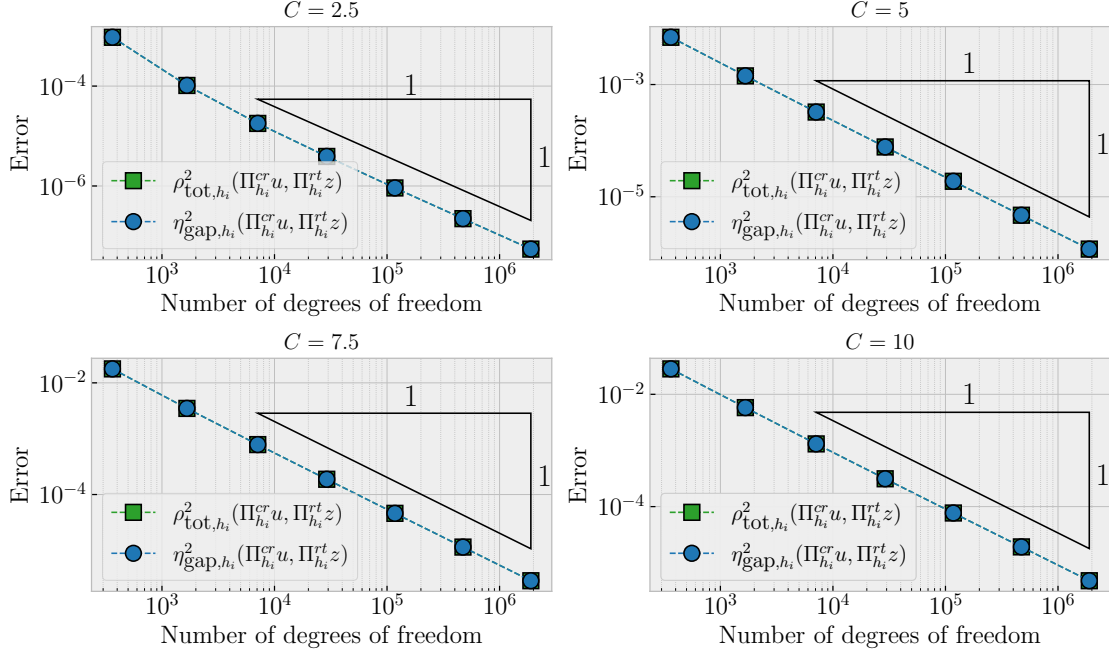
Figure 1: Logarithmic plots of the experimental convergence rates of the error quantities (7.14). For forcing terms $C \in \{2.5, 5.0, 7.5, 10.0\}$, we report the expected quadratic error decay rate, *i.e.*, $\text{EOC}_i(e_i^{\text{tot}}) \approx \text{EOC}_i(e_i^{\text{gap}}) \approx 2$, $i = 1, \ldots, 6$. In addition, we report that the discrete primal-dual error (*cf.* (6.5)) approximatively coincides with the discrete primal-dual gap estimator (*cf.* (6.1)).
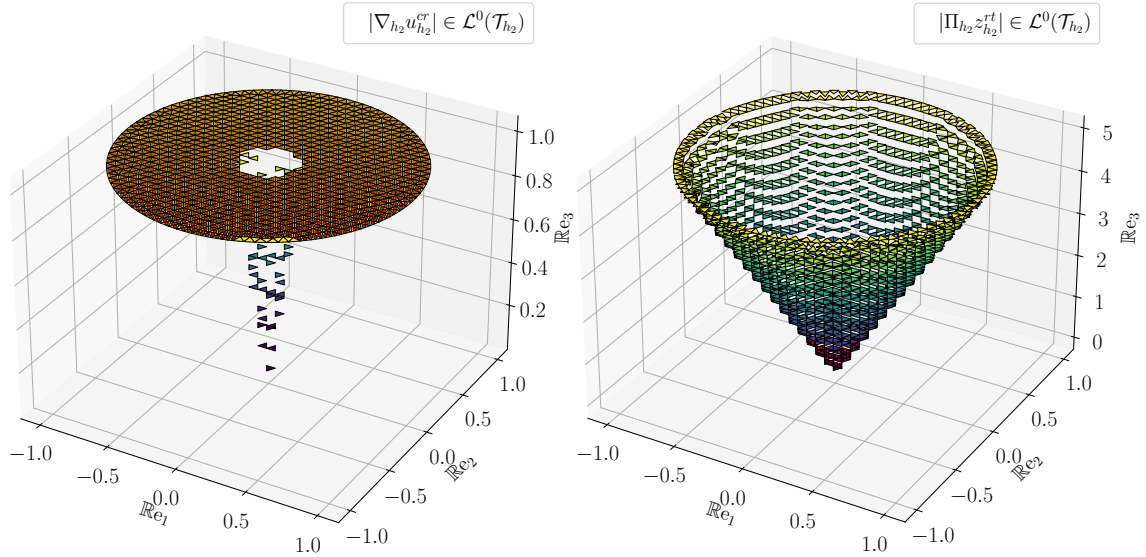

Figure 2: *left*: plot of $|\nabla_{h_2} u_{h_2}^{cr}| \in \mathcal{L}^0(\mathcal{T}_{h_2})$, where red dots mark $T \in \mathcal{T}_{h_2}$ with $|\nabla_{h_2} u_{h_2}^{cr}| = 1$ in T; *right*: plot of $|\Pi_{h_2} z_{h_2}^{rt}| \in \mathcal{L}^0(\mathcal{T}_{h_2})$, where red dots mark $T \in \mathcal{T}_{h_2}$ with $|\Pi_{h_2} z_{h_2}^{rt}| < 1$ in T. In summary, We report that $\{|\nabla_{h_2} u_{h_2}^{cr}| = 1\} = \{|\Pi_{h_2} z_{h_2}^{rt}| \geq 1\}$ and $\{|\nabla_{h_2} u_{h_2}^{cr}| < 1\} = \{|\Pi_{h_2} z_{h_2}^{rt}| < 1\}$ as predicted by the discrete convex optimality relation (5.6).

7.4   Numerical experiments concerning the a posteriori error analysis

In this subsection, we review the theoretical findings of Section 4.

More precisely, given the experimental setup of Subsection 7.2, for every $i = 1, \ldots, 6$, we define the admissible approximation

$$\overline{u}_{h_i}^{cr} := \frac{\Pi_{h_i}^{w1} u_{h_i}^{cr}}{\max\{1, \|\nabla \Pi_{h_i}^{w1} u_{h_i}^{cr}\|_{\infty,\Omega}\}} \in K,$$

where $\Pi_{h_i}^{w1} \colon \mathcal{S}_D^{1,cr}(\mathcal{T}_{h_i}) \to W_D^{1,\infty}(\Omega)$ is a suitable quasi-interpolation operator, *e.g.*,

$$\Pi_{h_i}^{w1} \in \left\{ \Pi_{h_i}^{av,1}, \Pi_{h_i}^{av,2}, \Pi_{h_i}^{sz,1}, \Pi_{h_i}^{sz,2}, \Pi_{h_i}^{L^2,1}, \Pi_{h_i}^{L^2,2} \right\},$$

where, setting $\mathcal{S}_D^k(\mathcal{T}_{h_i}) := \mathcal{L}^k(\mathcal{T}_{h_i}) \cap W_D^{1,\infty}(\Omega)$, for $k \in \{1, 2\}$, we denote by

(i) $\Pi_{h_i}^{av,k} \colon \mathcal{S}_D^{1,cr}(\mathcal{T}_{h_i}) \to \mathcal{S}_D^k(\mathcal{T}_{h_i})$ the *averaging quasi-interpolation operator* from [14, Subsec. 22.4.1];

(ii) $\Pi_{h_i}^{sz,k} \colon \mathcal{S}_D^{1,cr}(\mathcal{T}_{h_i}) \to \mathcal{S}_D^k(\mathcal{T}_{h_i})$ the *Scott–Zhang quasi-interpolation operator* from [14, Appx. A.3];

(iii) $\Pi_{h_i}^{L^2,k} \colon \mathcal{S}_D^{1,cr}(\mathcal{T}_{h_i}) \to \mathcal{S}_D^k(\mathcal{T}_{h_i})$ the *(global) L^2-projection operator* from [14, Subsec. 22.5].

Then, we compute the error quantities

$$\left.\begin{array}{l} e_i^{\mathrm{tot}} := \rho_{\mathrm{tot}}^2(\overline{u}_{h_i}^{cr}, z_{h_i}^{rt}), \\ e_i^{\mathrm{gap}} := \eta_{\mathrm{gap}}^2(\overline{u}_{h_i}^{cr}, z_{h_i}^{rt}), \end{array}\right\} \quad i = 0, \ldots, 6. \tag{7.14}$$

In Figure 3, for $C = 10.0$, we report a reduced (compared to Subsection 7.3) convergence rate of about $\mathrm{EOC}_i(e_i^{\mathrm{tot}}) \approx \mathrm{EOC}_i(e_i^{\mathrm{gap}}) \approx 1$, $i = 1, \ldots, 6$, *i.e.*, an error decay of order $\mathcal{O}(h_i) = \mathcal{O}(N_i^{\frac{1}{2}})$, $i = 1, \ldots, 6$. Moreover, we observe that the *a posteriori* error identity in Theorem 4.5(i) is approximatively satisfied.
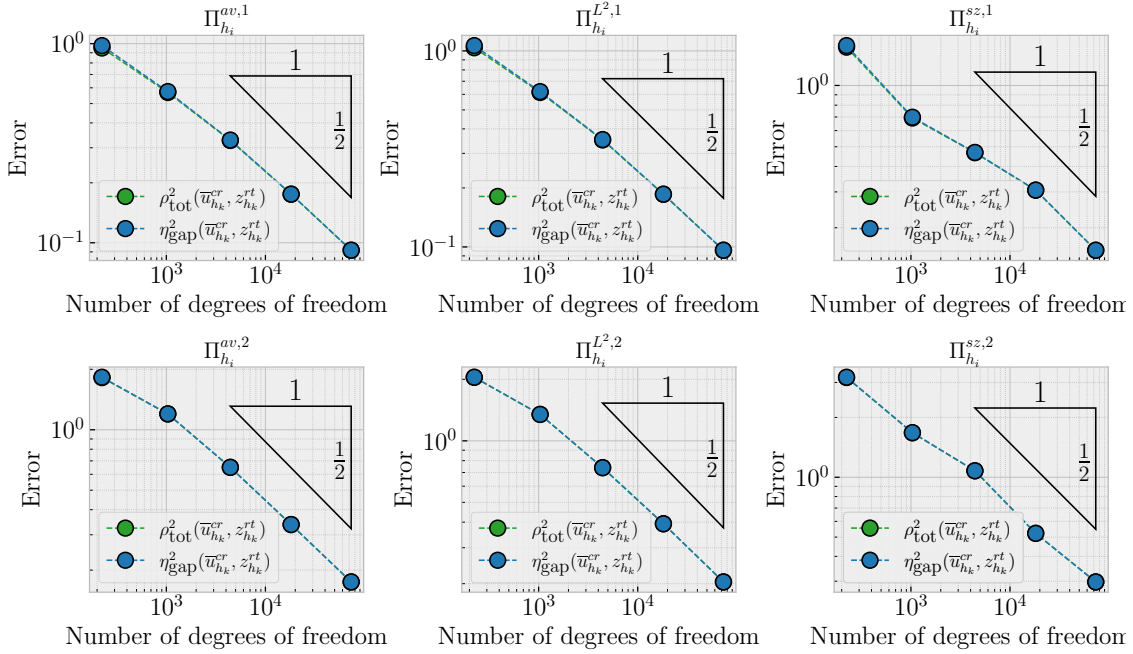

Figure 3: Logarithmic plots of the experimental convergence rates of the error quantities (7.14). For $C = 10.0$ and $\Pi_{h_i}^{w1} \in \{\Pi_{h_i}^{av,1}, \Pi_{h_i}^{av,2}, \Pi_{h_i}^{sz,1}, \Pi_{h_i}^{sz,2}, \Pi_{h_i}^{L^2,1}, \Pi_{h_i}^{L^2,2}\}$, we report a reduced linear error decay rate, *i.e.*, $\mathrm{EOC}_i(e_i^{\mathrm{tot}}) \approx \mathrm{EOC}_i(e_i^{\mathrm{gap}}) \approx 1$, $i = 1, \ldots, 6$. In addition, we report that the primal-dual error (*cf.* (4.9)) approximatively coincides with the primal-dual gap estimator (*cf.* (4.1)).

8. Conclusion and Outlook

A (Fenchel) duality theory for variational problems with gradient constraints was developed. On the basis of this (Fenchel) duality theory, an *a posteriori* error identity was derived, which can be seen as a generalization of the Prager–Sygne identity (*cf.* [25]) for the Poisson problem. To make the *a posteriori* error identity numerically practicable, it is necessary to approximate in a computationally inexpensive way the primal problem and the dual problem at the same time. To this end, exploiting orthogonality relations between the Crouzeix–Raviart element and the Raviart–Thomas element, all (Fenchel) duality relations from the continuous level are transferred to a discrete level and reconstruction of a primal solution from a given dual solution is derived, using the so-called inverse generalized Marini formula. The inverse generalized Marini formula enabled us to approximate the primal problem and the dual problem at the same time using only the Raviart–Thomas element. In addition, the discrete (Fenchel) duality theory allowed us to derive an *a posteriori* error identity also on the discrete level, which, eventually, turned out the be an *a priori* error identity and enabled us to derive explicit error decay rates depending on the regularity of the dual solution and the obstacle function. Numerical experiments were carried out that confirm the optimality of the derived error decay rates and the validity of both the *a priori* error identity and the *a posteriori* error identity. Still open is the question of the accurate post-processing of the discrete primal solution to obtain an admissible approximation of the primal solution. Simple node-averaging and, subsequently, global scaling with the gradient length does not approximate the gradient in $(L^\infty(\Omega))^d$ with the desired accuracy. Instead a quasi-interpolation operator, which either preserves the gradient length or at least approximates the later with the needed accuracy, should be employed. In the end, this would allow us to use the local refinement indicators induced by the primal-dual gap estimator for adaptive mesh-refinement. The latter will be content of future research.

References

[1] T. Arbogast and Z. Chen, On the implementation of mixed methods as nonconforming methods for second-order elliptic problems, *Math. Comput.* **64** no. 211 (1995), 943–972 (English). doi:10.2307/2153478.

[2] D. N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates, *RAIRO, Modélisation Math. Anal. Numér.* **19** (1985), 7–32 (English). doi:10.1051/m2an/1985190100071.

[3] S. Bartels, *Numerical approximation of partial differential equations*, *Texts Appl. Math.* **64**, Cham: Springer, 2016 (English). doi:10.1007/978-3-319-32354-1.

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

[5] S. Bartels, T. Gudi, and A. Kaltenbach, *A priori* and *a posteriori* error identities for the scalar Signorini problem, 2024. doi:10.48550/arXiv.2407.10912

[6] S. Bartels and A. Kaltenbach, Chapter seven - exact a posteriori error control for variational problems via convex duality and explicit flux reconstruction, in *Error Control, Adaptive Discretizations, and Applications, Part 1* (F. Chouly, S. P. Bordas, R. Becker, and P. Omnes, eds.), *Advances in Applied Mechanics* **58**, Elsevier, 2024, pp. 281–361. doi:10.1016/bs.aams.2024.04.001.

[7] M. Bildhauer, M. Fuchs, and S. Repin, The elastic–plastic torsion problem: A posteriori error estimates for approximate solutions, *Numerical Functional Analysis and Optimization* **30** no. 7-8 (2009), 653–664.

[8] F. Chouly, T. Gustafsson, and P. Hild, A Nitsche method for the elastoplastic torsion problem, *ESAIM Math. Model. Numer. Anal.* **57** no. 3 (2023), 1731–1746. doi:10.1051/m2an/2023034.

[9] F. Chouly, T. Gustafsson, and P. Hild, Finite element approximation of penalized elastoplastic torsion problem with nonconstant source term, 08 2024.

[10] F. Chouly and P. Hild, On a finite element approximation for the elastoplastic torsion problem, *Appl. Math. Lett.* **132** (2022), Paper No. 108191, 6. doi:10.1016/j.aml.2022.108191.

[11] M. Crouzeix and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I, *Rev. Franc. Automat. Inform. Rech. Operat., R* **7** no. 3 (1974), 33–76 (English). doi:10.1051/m2an/197307R300331.

[12] G. Duvaut and J.-L. Lions, *Les inéquations en mécanique et en physique*, *Travaux et Recherches Mathématiques* **No. 21**, Dunod, Paris, 1972.

[13] I. Ekeland and R. Témam, *Convex analysis and variational problems.*, unabridged, corrected republication of the 1976 English original ed., *Class. Appl. Math.* **28**, Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999 (English).

[14] A. Ern and J.-L. Guermond, *Finite elements I. Approximation and interpolation*, *Texts Appl. Math.* **72**, Cham: Springer, 2020 (English). doi:10.1007/978-3-030-56341-7.

[15] A. Ern and J.-L. Guermond, *Finite elements II. Galerkin approximation, elliptic and mixed PDEs*, *Texts Appl. Math.* **73**, Cham: Springer, 2021 (English). doi:10.1007/978-3-030-56923-5.

[16] R. Glowinski, *Lectures on numerical methods for nonlinear variational problems*, *Tata Institute of Fundamental Research Lectures on Mathematics and Physics* **65**, Tata Institute of Fundamental Research, Bombay; Springer-Verlag, Berlin-New York, 1980, Notes by M. G. Vijayasundaram and M. Adimurthi.

[17] R. Glowinski, J.-L. Lions, and R. Tremolieres, *Numerical analysis of variational inequalities. Transl. and rev. ed*, *Stud. Math. Appl.* **8**, Elsevier, Amsterdam, 1981 (English).

[18] J. D. Hunter, Matplotlib: A 2d graphics environment, *Computing in Science & Engineering* **9** no. 3 (2007), 90–95. doi:10.1109/MCSE.2007.55.

[19] G. Idone, A. Maugeri, and C. Vitanza, Variational inequalities and the elastic-plastic torsion problem, *J. Optim. Theory Appl.* **117** no. 3 (2003), 489–501. doi:10.1023/A:1023941520452.

[20] T. M. Inc., Matlab version: R2024a, 2024. Available at https://www.mathworks.com.

[21] D. Kinderlehrer and G. Stampacchia, *An introduction to variational inequalities and their applications*, *Pure and Applied Mathematics* **88**, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

[22] A. Logg and G. N. Wells, DOLFIN: automated finite element computing, *ACM Trans. Math. Software* **37** no. 2 (2010), Art. 20, 28. doi:10.1145/1731022.1731030.

[23] 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** no. 3 (1985), 493–496. doi:10.1137/0722029.

[24] P.-O. Persson and G. Strang, A Simple Mesh Generator in MATLAB, *SIAM Review* **46** no. 2 (2004), 329–345. doi:10.1137/S0036144503429121.

[25] W. Prager and J. L. Synge, Approximations in elasticity based on the concept of function space, *Quart. Appl. Math.* **5** (1947), 241–269. doi:10.1090/qam/25902.

[26] P. A. Raviart and J. M. Thomas, A mixed finite element method for 2nd order elliptic problems, Math. Aspects Finite Elem. Meth., Proc. Conf. Rome 1975, Lect. Notes Math. 606, 292-315 (1977)., 1977.

[27] R. T. Rockafellar, Integrals which are convex functionals, *Pac. J. Math.* **24** (1968), 525–539 (English). doi:10.2140/pjm.1968.24.525.

[28] R. T. Rockafellar, Integrals which are convex functionals. II, *Pac. J. Math.* **39** (1971), 439–469 (English). doi:10.2140/pjm.1971.39.439.

[29] S. E. Shreve and H. M. Soner, Optimal investment and consumption with transaction costs, *Ann. Appl. Probab.* **4** no. 3 (1994), 609–692.

[30] H. M. Soner, S. E. Shreve, and N. El Karoui, A free boundary problem related to singular stochastic control: the parabolic case, *Communications in partial differential equations* **16** no. 2-3 (1991), 373–424.

[31] J. Toland, *The dual of $L_\infty(X, \mathcal{L}, \lambda)$, finitely additive measures and weak convergence. A primer*, *SpringerBriefs Math.*, Cham: Springer, 2020 (English). doi:10.1007/978-3-030-34732-1.