Error analysis for a Crouzeix–Raviart approximation of the obstacle problem

Sören Bartels*[1] and Alex Kaltenbach†[2]

[1]Department of Applied Mathematics, University of Freiburg, Hermann–Herder–Straße 10, 79104 Freiburg

[2]Department of Applied Mathematics, University of Freiburg, Ernst–Zermelo–Straße 1, 79104 Freiburg

February 3, 2023

Abstract

In the present paper, we study a Crouzeix–Raviart approximation of the obstacle problem, which imposes the obstacle constraint in the midpoints (i.e., barycenters) of the elements of a triangulation. We establish a priori error estimates imposing natural regularity assumptions, which are optimal, and the reliability and efficiency of a primal-dual type a posteriori error estimator for general obstacles and involving data oscillation terms stemming only from the right-hand side. The theoretical findings are supported by numerical experiments.

Keywords: Obstacle problem; Crouzeix–Raviart element; a priori error analysis; a posteriori error analysis.

AMS MSC (2020): 35J20; 49J40; 49M29; 65N30; 65N15; 65N50.

1. Introduction

The obstacle problem is a prototypical example of a non-smooth convex minimization problem with an inequality constraint that leads to a variational inequality. It has countless applications, e.g., in the contexts of fluid filtration in porous media, constrained heating, elasto-plasticity, optimal control, and financial mathematics, cf. [14, 29]. It is deeply related to models in free boundary value problems, the study of minimal surfaces, and the capactity of a set in potential theory, cf. [14]. The problem is to find the equilibrium position of an elastic membrane whose boundary is held fixed and which is constrained to lie above a given obstacle.

More precisely, given an external force $f \in L^2(\Omega)$ and an obstacle $\chi \in W^{1,2}(\Omega)$ with $\operatorname{tr}\chi \leq 0$ on Γ_D, where $\Gamma_D \subseteq \partial\Omega$ denotes the Dirichlet part of the topological boundary $\partial\Omega$, the obstacle problem seeks for a minimizer $u \in W^{1,2}_D(\Omega) \coloneqq \{v \in W^{1,2}(\Omega) \mid \operatorname{tr} v = 0 \text{ in } \Gamma_D\}$ of the energy functional $I\colon W^{1,2}_D(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\nabla v\|_\Omega^2 - (f, v)_\Omega + I_K(v)\,, \tag{1.1}$$

where

$$K \coloneqq \big\{v \in W^{1,2}_D(\Omega) \mid v \geq \chi \text{ a.e. in } \Omega\big\}\,,$$

and $I_K\colon W^{1,2}_D(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is defined by $I_K(v) \coloneqq 0$ if $v \in K$ and $I_K(v) \coloneqq +\infty$ else.

*Email: bartels@mathematik.uni-freiburg.de

†Email: alex.kaltenbach@mathematik.uni-freiburg.de

1.1 Related contributions

The finite element approximation of (1.1) has been intensively analyzed by numerous authors: First contributions addressing the a priori and a posteriori error analysis of approximations of (1.1) using the conforming Lagrange finite element can be found in [26, 1, 32, 37, 38, 13, 43, 17, 42, 10, 11, 9, 48, 27], imposing the obstacle constraint in the nodes of a triangulation, and in [34, 28], enforcing the obstacle constraint in the limit via a penalization approach. We refer to [17] for a short review. Contributions addressing the a priori and a posteriori error analysis of an approximation of (1.1) deploying Discontinuous Galerkin (DG) type methods can be found in [30, 18, 5], equally imposing the obstacle constraint in the nodes of a triangulation. The first contribution addressing the a priori error analysis of an approximation of (1.1) in two dimensions deploying the Crouzeix–Raviart element can be found in [47] and imposes the obstacle constraint in the midpoints (i.e., barycenters) of elements of a triangulation. In [6], for homogeneous Dirichlet boundary data and zero obstacle, this result was extended to arbitrary dimensions. In [15], an a priori and a posterior error analysis of an approximation of (1.1) deploying the Crouzeix–Raviart element which imposes the obstacle constraint in the integral mean values of element sides of a triangulation was carried out.

1.2 New contributions

Inspired by [47] as well as recent contributions [5, 6, 7], different from the contribution [15], we treat an approximation of the obstacle problem (1.1) deploying the Crouzeix–Raviart element that imposes the obstacle constraint in the midpoints (i.e., barycenters) of elements of a triangulation. More precisely, given a family of regular triangulations $\mathcal{T}_h$, $h>0$, setting $f_h \coloneqq \Pi_h f \in \mathcal{L}^0(\mathcal{T}_h)$ and for $\chi_h \in \mathcal{L}^0(\mathcal{T}_h)$ approximating $\chi \in W^{1,2}(\Omega)$, our discrete obstacle problem seeks for a minimizer $u_h^{cr} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ of the functional $I_h^{cr}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, for every $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) \coloneqq \tfrac{1}{2}\|\nabla_h v_h\|_\Omega^2 - (f_h, \Pi_h v_h)_\Omega + I_{K_h^{cr}}(v_h)\,, \tag{1.2}$$

where

$$K_h^{cr} \coloneqq \left\{ v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h) \mid \Pi_h v_h \ge \chi_h \text{ a.e. in } \Omega \right\},$$

and $I_{K_h^{cr}}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$ is defined by $I_{K_h^{cr}}(v_h) \coloneqq 0$ if $v_h \in K_h^{cr}$ and $I_{K_h^{cr}}(v_h) \coloneqq +\infty$ else. Here, $\mathcal{L}^0(\mathcal{T}_h)$ denotes the space of element-wise constant functions, $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ the Crouzeix–Raviart finite element space, i.e., the space of element-wise affine functions that are continuous in the midpoints (i.e., barycenters) of inner element sides and that vanish in the midpoints of element sides that belong to Γ_D, $\nabla_h\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h) \to \mathcal{L}^0(\mathcal{T}_h)$ the element-wise gradient and $\Pi_h\colon L^2(\Omega) \to \mathcal{L}^0(\mathcal{T}_h)$ the (local) L^2-projection operator onto element-wise constant functions. Imposing the obstacle constraint in the midpoints of elements follows a systematic approximation procedure for general convex minimization problems deploying the Crouzeix–Raviart element introduced in [6, 7] and has the advantage that the resulting discrete convex minimization problem generates discrete convex duality relations that are analogous to those in the continuous setting –up to non-conforming modifications– and that enable a systematic a priori error analysis and a posteriori error analysis:

- In [6], a systematic procedure for the derivation of a priori error estimates for convex minimization problems deploying the Crouzeix–Raviart element based on (discrete) convex duality relations was proposed. Following this systematic procedure, with comparably little effort, we derive a priori error estimates, which are optimal for natural regularity assumptions and also apply in arbitrary dimensions. More precisely, our a priori error estimates exploit that the discrete primal-dual gap controls the convexity measure of (1.2) and the concavity measure of its dual functional, i.e., that for every $v_h \in K_h^{cr}$ and $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$, it holds

$$\tfrac{1}{2}\|\nabla_h v_h - \nabla_h u_h^{cr}\|_\Omega^2 + (-\overline{\lambda}_h^{cr}, \Pi_h(v_h - u_h^{cr}))_\Omega + \tfrac{1}{2}\|\Pi_h y_h - \Pi_h z_h^{rt}\|_\Omega^2 \le I_h^{cr}(v_h) - D_h^{rt}(y_h)\,, \tag{1.3}$$

  where $\mathcal{RT}^0_N(\mathcal{T}_h)$ denotes the Raviart–Thomas finite element space, i.e., the space of element-wise affine vector fields that have continuous constant normal components on element sides that vanish on Γ_N, $z_h^{rt} \in \mathcal{RT}^0_N(\mathcal{T}_h)$ the unique discrete dual solution, i.e., the maximizer of the discrete dual energy functional $D_h^{cr}\colon \mathcal{RT}_N^0(\mathcal{T}_h)\to\mathbb{R}\cup\{-\infty\}$, and $\overline{\lambda}_h^{cr}\in\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$ the unique discrete Lagrange multiplier satisfying $\overline{\lambda}_h^{cr}\leq 0$ a.e. in Ω and for all $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$

$$(\overline{\lambda}_h^{cr},\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega-(\nabla_h u_h^{cr},\nabla_h v_h)_\Omega\,.$$

If $\chi_h\coloneqq\Pi_h I_{cr}\chi\in\mathcal{L}^0(\mathcal{T}_h)$, where $I_{cr}\colon W^{1,2}(\Omega)\to\mathcal{S}^{1,cr}(\Omega)$ denotes the Crouzeix–Raviart quasi-interpolation operator, then $I_{cr}u\in K_h^{cr}$. Thus, under natural regularity assumptions, i.e., $u,\chi\in W^{2,2}(\Omega)$, the choices $v_h=I_{cr}u\in K_h^{cr}$ and $y_h=I_{rt}z\in\mathcal{RT}_N^0(\mathcal{T}_h)$, where $z=\nabla u\in W^{1,2}(\Omega;\mathbb{R}^d)\cap W_N^2(\mathrm{div};\Omega)$ denotes the dual solution, i.e., the maximizer of the dual energy functional $D\colon L^2(\Omega;\mathbb{R}^d)\to\mathbb{R}\cup\{-\infty\}$, and $I_{rt}\colon W^{1,2}(\Omega;\mathbb{R}^d)\cap W_N^2(\mathrm{div};\Omega)\to\mathcal{RT}_N^0(\mathcal{T}_h)$ the Raviart–Thomas quasi-interpolation operator, are admissible in (1.3) and lead to quasi-optimal a priori error estimates.

- In [7], a systematic procedure for the derivation of reliable, quasi-constant-free a posteriori error estimates for convex minimization problems deploying the Crouzeix–Raviart element based on (discrete) convex duality relations was proposed. Following this systematic procedure, we derive a posteriori error estimates, which, by definition, are reliable and constant-free. Apart from that, we establish the efficiency of these a posteriori error estimates for general obstacles $\chi\in W^{1,2}(\Omega)$. More precisely, our a posteriori error estimates exploit that the primal-dual gap controls the convexity measure of (1.1) and the concavity measure of its dual functional, i.e., that for every $v\in K$ and $y\in L^2(\Omega;\mathbb{R}^d)$, it holds

$$\tfrac{1}{2}\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega+\tfrac{1}{2}\|y-z\|_\Omega^2\leq I(v)-D(y)\,,\tag{1.4}$$

where $\Lambda\in(W_D^{1,2}(\Omega))^*$ is the unique Lagrange multiplier satisfying $\Lambda\leq 0$ in $(W_D^{1,2}(\Omega))^*$ and for all $v\in W_D^{1,2}(\Omega)$

$$\langle\Lambda,v\rangle_\Omega=(f,v)_\Omega-(\nabla u,\nabla v)_\Omega\,.$$

For the a posteriori error estimate (1.4) being practicable it is necessary to have a sufficiently accurate and computationally cheap procedure to obtain an approximation $y\in L^2(\Omega;\mathbb{R}^d)$ of the dual solution $z=\nabla u\in L^2(\Omega;\mathbb{R}^d)$ at hand. In the case $f=f_h\in\mathcal{L}^0(\mathcal{T}_h)$, the discrete dual solution $z_h^{rt}\in\mathcal{RT}_N^0(\mathcal{T}_h)$ is admissible in (1.4) and leads to a constant-free reliable and efficient a posteriori error estimator $\eta_h^2\coloneqq I-D(z_h^{rt})\colon W_D^{1,2}(\Omega)\to\mathbb{R}$, which has similarities to the residual type a posteriori estimator derived in [42] but is simper and avoids jump terms of the obstacle that arise in the efficiency analysis in [42]. In particular, note that the discrete dual solution can cheaply be computed via the generalized Marini formula

$$z_h^{rt}=\nabla_h u_h^{cr}+\frac{\overline{\lambda}_h^{cr}-f_h}{d}(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})\quad\text{ in }\mathcal{RT}_N^0(\mathcal{T}_h)\,.\tag{1.5}$$

A typical choice of $v\in K$ is obtained via nodal averaging $u_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and truncating to enforce the continuous obstacle constraint. Moreover, any conforming approximation $u_h\in K$ can be used such as a continuous Lagrange approximation $u_h^c\in K_h^c\coloneqq K\cap\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, so that our analysis also implies the full reliability and efficiency error analysis for continuous Lagrange approximations, even for general obstacles and oscillation terms only stemming from the right-hand side since lumping is not needed in our analysis.

As a whole, our approach brings together and extends ideas and concepts from [13, 43, 42, 10, 11, 15], and leads to a full error analysis.

1.3 Outline

This article is organized as follows: In Section 2, we introduce the employed notation and the relevant finite element spaces. In Section 3, we give a brief review of the continuous and the discrete obstacle problem. In Section 4, we derive a priori error estimates for the Crouzeix–Raviart approximation (1.2) of (1.1), which are optimal for natural regularity assumptions. In Section 5, we establish the reliability and efficiency of a so-called primal-dual a posteriori error estimator. In Section 6, we confirm our theoretical findings via numerical experiments. In the Appendix A, we derive local efficiency estimates for a Crouzeix–Raviart approximation of (1.1).

2. PRELIMINARIES

Throughout the entire article, if not otherwise specified, we always denote by $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, a bounded polyhedral Lipschitz domain, whose topological boundary $\partial\Omega$ is disjointly divided into a closed Dirichlet part Γ_D, for which we always assume that $|\Gamma_D| > 0$[1], and a Neumann part Γ_N, i.e., $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\emptyset = \Gamma_D \cap \Gamma_N$. We use $c, C > 0$ to denote generic constants, that may change from line to line, but are not depending on the crucial quantities. For (Lebesgue) measurable functions $u, v\colon \Omega \to \mathbb{R}$ and a (Lebesgue) measurable set $M \subseteq \Omega$, we employ the product

$$(u,v)_M \coloneqq \int_M u\, v \,\mathrm{d}x\,,$$

whenever the right-hand side is well-defined. Analogously, for (Lebesgue) measurable vector fields $z, y\colon \Omega \to \mathbb{R}^d$ and a (Lebesgue) measurable set $M \subseteq \Omega$, we write $(z,y)_M \coloneqq \int_M z \cdot y \,\mathrm{d}x$.

2.1 *Standard function spaces*

For $l \in \mathbb{N}$, we employ the standard notations[2]

$$\begin{aligned} W^{1,2}_D(\Omega;\mathbb{R}^l) &\coloneqq \big\{ v \in L^2(\Omega;\mathbb{R}^l) \mid \nabla v \in L^2(\Omega;\mathbb{R}^{l\times d}),\ \mathrm{tr}\, v = 0 \text{ in } L^2(\Gamma_D;\mathbb{R}^l) \big\}\,,\\ W^2_N(\mathrm{div};\Omega) &\coloneqq \big\{ y \in L^2(\Omega;\mathbb{R}^d) \mid \mathrm{div}\, y \in L^2(\Omega),\ \mathrm{tr}\, y \cdot n = 0 \text{ in } W^{-\frac{1}{2},2}(\Gamma_N) \big\}\,, \end{aligned}$$

$W^{1,2}(\Omega;\mathbb{R}^l) \coloneqq W^{1,2}_D(\Omega;\mathbb{R}^l)$ if $\Gamma_D = \emptyset$, and $W^2(\mathrm{div};\Omega) \coloneqq W^2_N(\mathrm{div};\Omega)$ if $\Gamma_N = \emptyset$, where we denote by $\mathrm{tr}\colon W^{1,2}(\Omega;\mathbb{R}^l) \to L^2(\partial\Omega;\mathbb{R}^l)$ and by $\mathrm{tr}(\cdot)\cdot n\colon W^2(\mathrm{div};\Omega) \to W^{-\frac{1}{2},2}(\partial\Omega)$, the trace and normal trace operator, resp. In particular, we predominantly omit $\mathrm{tr}(\cdot)$ in this context. In addition, we employ the abbreviations $L^2(\Omega) \coloneqq L^2(\Omega;\mathbb{R}^1)$, $W^{1,2}(\Omega) \coloneqq W^{1,2}(\Omega;\mathbb{R}^1)$, and $W^{1,2}_D(\Omega) \coloneqq W^{1,2}_D(\Omega;\mathbb{R}^1)$, as well as $\|\cdot\|_\Omega \coloneqq \|\cdot\|_{L^2(\Omega;\mathbb{R}^l)}$, $\|\cdot\|_{*,\Omega} \coloneqq \|\cdot\|_{(W^{1,2}_D(\Omega;\mathbb{R}^l))^*}$, and $\langle\cdot,\cdot\rangle_\Omega \coloneqq \langle\cdot,\cdot\rangle_{W^{1,2}_D(\Omega;\mathbb{R}^l)}$, $l \in \mathbb{N}$.

2.2 *Triangulations and standard finite element spaces*

Throughout the entire paper, we denote by $\mathcal{T}_h$, $h > 0$, a family of regular, i.e., uniformly shape regular and conforming, triangulations of $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, cf. [25]. Here, $h > 0$ refers to the average mesh-size, i.e., if we set $h_T \coloneqq \mathrm{diam}(T)$ for all $T \in \mathcal{T}_h$, then, we have that $h \coloneqq \frac{1}{\mathrm{card}(\mathcal{T}_h)} \sum_{T\in\mathcal{T}_h} h_T$. For every element $T \in \mathcal{T}_h$, we denote by $\rho_T > 0$, the supremum of diameters of inscribed balls. We assume that there exists a constant $\omega_0 > 0$, independent of $h > 0$, such that $\max_{T\in\mathcal{T}_h} h_T \rho_T^{-1} \le \omega_0$. The smallest such constant is called the chunkiness of $(\mathcal{T}_h)_{h>0}$. Also note that, in what follows, all constants may depend on the chunkiness, but are independent of $h > 0$. For every $T \in \mathcal{T}_h$, let ω_T denote the patch of T, i.e., the union of all elements of $\mathcal{T}_h$ touching T. We assume that $\mathrm{int}(\omega_T)$ is connected for all $T \in \mathcal{T}_h$. Under these assumptions, $|T| \sim |\omega_T|$ uniformly in $T \in \mathcal{T}_h$ and $h > 0$, and the number of elements in ω_T and patches to which an element T belongs to are uniformly bounded with respect to $T \in \mathcal{T}_h$ and $h > 0$. We define the sides of $\mathcal{T}_h$ in the following way: an interior side is the closure of the non-empty relative interior of $\partial T \cap \partial T'$, where $T, T' \in \mathcal{T}_h$ are adjacent elements. For an interior side $S \coloneqq \partial T \cap \partial T' \in \mathcal{S}_h$, where $T, T' \in \mathcal{T}_h$, we employ the notation $\omega_S \coloneqq T \cup T'$. A boundary side is the closure of the non-empty relative interior of $\partial T \cap \partial\Omega$, where $T \in \mathcal{T}_h$ denotes a boundary element of $\mathcal{T}_h$. For a boundary side $S \coloneqq \partial T \cap \partial\Omega$, we employ the notation $\omega_S \coloneqq T$. By $\mathcal{S}_h^i$ and $\mathcal{S}_h$, we denote the sets of all interior sides and the set of all sides, respectively. Finally, we define the maximum mesh-size $h_{\max} \coloneqq \max_{T\in\mathcal{T}_h} h_T$ and for every $S \in \mathcal{S}_h$, we define $h_S \coloneqq \mathrm{diam}(S)$. For (Lebesgue) measurable functions $u, v\colon \mathcal{S}_h \to \mathbb{R}$ and $\mathcal{M}_h \subseteq \mathcal{S}_h$, we employ the product

$$(u,v)_{\mathcal{M}_h} \coloneqq \sum_{S\in\mathcal{M}_h} (u,v)_S\,, \qquad \text{where } (u,v)_S \coloneqq \int_S uv \,\mathrm{d}s\,,$$

whenever all integrals are well-defined. Analogously, for (Lebesgue) measurable vector fields $z, y\colon \mathcal{S}_h \to \mathbb{R}^d$ and $\mathcal{M}_h \subseteq \mathcal{S}_h$, we write $(z,y)_{\mathcal{M}_h} \coloneqq \sum_{S\in\mathcal{M}_h} (z,y)_S$, where $(z,y)_S \coloneqq \int_S z\cdot y \,\mathrm{d}s$.

[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] Here, $W^{-\frac{1}{2},2}(\partial\Omega) \coloneqq (W^{\frac{1}{2},2}(\partial\Omega))^*$ and $W^{-\frac{1}{2},2}(\Gamma_N) \coloneqq (W^{\frac{1}{2},2}(\Gamma_N))^*$.

For $k \in \mathbb{N}\cup\{0\}$ and $T \in \mathcal{T}_h$, let $\mathcal{P}_k(T)$ denote the set of polynomials of maximal degree k on T. Then, for $k \in \mathbb{N}\cup\{0\}$ and $l \in \mathbb{N}$, the sets of continuous and element-wise polynomial functions or vector fields, respectively, are defined by

$$\begin{aligned}\mathcal{S}^k(\mathcal{T}_h)^l &\coloneqq \big\{v_h \in C^0(\overline{\Omega};\mathbb{R}^l) \ \mid v_h|_T \in \mathcal{P}_k(T)^l \text{ for all } T \in \mathcal{T}_h\big\}\,,\\ \mathcal{L}^k(\mathcal{T}_h)^l &\coloneqq \big\{v_h \in L^\infty(\Omega;\mathbb{R}^l) \ \mid v_h|_T \in \mathcal{P}_k(T)^l \text{ for all } T \in \mathcal{T}_h\big\}\,.\end{aligned}$$

The element-wise constant mesh-size function $h_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T}_h)$ is defined by $h_{\mathcal{T}}|_T \coloneqq h_T$ for all $T \in \mathcal{T}_h$. The side-wise constant mesh-size function $h_{\mathcal{S}} \in \mathcal{L}^0(\mathcal{S}_h)$ is defined by $h_{\mathcal{S}}|_S \coloneqq h_S$ for all $S \in \mathcal{S}_h$. Denoting by $\mathcal{N}_h$, the set of all vertices (or nodes) of $\mathcal{T}_h$, for every $T \in \mathcal{T}_h$ and $S \in \mathcal{S}_h$, we denote by $x_T \coloneqq \frac{1}{d+1}\sum_{z\in\mathcal{N}_h\cap T} z$ and $x_S \coloneqq \frac{1}{d}\sum_{z\in\mathcal{N}_h\cap S} z$, the midpoints (i.e., barycenters) of T and S, respectively. The (local) L^2-projection operator onto element-wise constant functions or vector fields, respectively, is denoted by

$$\Pi_h\colon L^1(\Omega;\mathbb{R}^l) \to \mathcal{L}^0(\mathcal{T}_h)^l\,.$$

For each $v_h \in \mathcal{L}^1(\mathcal{T}_h)^l$, it holds $\Pi_h v_h|_T = v_h(x_T)$ in T for all $T \in \mathcal{T}_h$. There exists a constant $c_\Pi > 0$, depending only on the chunkiness $\omega_0 > 0$, such that for every $v \in L^2(\Omega;\mathbb{R}^l)$, $l \in \mathbb{N}$, and $T \in \mathcal{T}_h$, cf. [25, Theorem 18.16], it holds

(L0.1) $\|\Pi_h v\|_T \le \|v\|_T$,

(L0.2) $\|v - \Pi_h v\|_T \le c_\Pi\, h_T\, \|\nabla v\|_T$ if $v \in W^{1,2}(T;\mathbb{R}^l)$.

The element-wise gradient operator $\nabla_h\colon \mathcal{L}^1(\mathcal{T}_h)^l \to \mathcal{L}^0(\mathcal{T}_h)^{l\times d}$, $l \in \mathbb{N}$, for every $v_h \in \mathcal{L}^1(\mathcal{T}_h)^l$, is defined by $\nabla_h v_h|_T \coloneqq \nabla(v_h|_T)$ in T for all $T \in \mathcal{T}_h$.

2.2.1 Crouzeix–Raviart element

The Crouzeix–Raviart finite element space, introduced in [19], consists of element-wise affine functions that are continuous in the midpoints of inner element sides, i.e.,[3]

$$\mathcal{S}^{1,cr}(\mathcal{T}_h) \coloneqq \big\{v_h \in \mathcal{L}^1(\mathcal{T}_h) \mid [\![v_h]\!]_S(x_S) = 0 \text{ for all } S \in \mathcal{S}_h^i\big\}\,.$$

Crouzeix–Raviart finite element functions that vanish in the midpoints of boundary element sides that correspond to the Dirichlet boundary Γ_D are contained in the space

$$\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \coloneqq \big\{v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h) \mid v_h(x_S) = 0 \text{ for all } S \in \mathcal{S}_h \cap \Gamma_D\big\}\,.$$

In particular, we have that $\mathcal{S}^{1,cr}_D(\mathcal{T}_h) = \mathcal{S}^{1,cr}(\mathcal{T}_h)$ if $\Gamma_D = \emptyset$. A basis of $\mathcal{S}^{1,cr}(\mathcal{T}_h)$ is given by functions $\varphi_S \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying the Kronecker property $\varphi_S(x_{S'}) = \delta_{S,S'}$ for all $S, S' \in \mathcal{S}_h$. A basis of $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ is given by $\varphi_S \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, $S \in \mathcal{S}_h \setminus \Gamma_D$. There exists a constant $c_P^{cr} > 0$, depending only on the chunkiness $\omega_0 > 0$, such that for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, it holds

$$\|v_h\|_\Omega \le c_P^{cr}\, \|\nabla_h v_h\|_\Omega\,. \tag{2.1}$$

The quasi-interpolation operator $I_{cr}\colon W^{1,2}_D(\Omega) \to \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I_{cr}v \coloneqq \sum_{S\in\mathcal{S}_h} v_S\,\varphi_S\,, \qquad \text{where } v_S \coloneqq \fint_S v\,\mathrm{d}s\,, \tag{2.2}$$

preserves averages of gradients, i.e., $\nabla_h(I_{cr}v) = \Pi_h(\nabla v)$ in $\mathcal{L}^0(\mathcal{T}_h)^d$ for every $v \in W^{1,2}_D(\Omega)$. There exists a constant $c_{cr} > 0$, depending only on the chunkiness $\omega_0 > 0$, such that for every $v \in W^{1,2}_D(\Omega)$ and $T \in \mathcal{T}_h$, cf. [22, Remark 4.4 & Theorem 4.6], it holds

(CR.1) $\|\nabla_h I_{cr}v\|_T \le \|\nabla v\|_T$,

(CR.2) $\|v - I_{cr}v\|_T \le c_{cr}\, h_T\, \|\nabla v\|_{\omega_T}$,

(CR.3) $\|v - I_{cr}v\|_T + h_T\|\nabla(v - I_{cr}v)\|_T \le c_{cr}\, h_T^2\, \|D^2 v\|_{\omega_T}$ if $v \in W^{2,2}(T)$.

[3] Here, for every $S \in \mathcal{S}_h^i$, $[\![v_h]\!]_S \coloneqq v_h|_{T_+} - v_h|_{T_-}$ on S, where $T_+, T_- \in \mathcal{T}_h$ satisfy $\partial T_+ \cap \partial T_- = S$, and for every $S \in \mathcal{S}_h \cap \partial\Omega$, $[\![v_h]\!]_S \coloneqq v_h|_T$ on S, where $T \in \mathcal{T}_h$ satisfies $S \subseteq \partial T$.

2.2.2 Raviart–Thomas element

The lowest order Raviart–Thomas finite element space, introduced in [41], consists of element-wise affine vector fields that have continuous constant normal components on inner elements sides, i.e.,[4]

$$\begin{aligned}\mathcal{R}T^0(\mathcal{T}_h) \coloneqq \big\{ y_h \in \mathcal{L}^1(\mathcal{T}_h)^d \mid y_h|_T \cdot n_T = \text{const on } \partial T \text{ for all } T \in \mathcal{T}_h\,,\\ [\![y_h \cdot n]\!]_S = 0 \text{ on } S \text{ for all } S \in \mathcal{S}_h^i \big\}\,.\end{aligned}$$

Raviart–Thomas finite element functions that have vanishing normal components on the Neumann boundary Γ_N are contained in the space

$$\mathcal{R}T^0_N(\mathcal{T}_h) \coloneqq \big\{ y_h \in \mathcal{R}T^0(\mathcal{T}_h) \mid y_h \cdot n = 0 \text{ on } \Gamma_N \big\}\,.$$

In particular, we have that $\mathcal{R}T^0_N(\mathcal{T}_h) = \mathcal{R}T^0(\mathcal{T}_h)$ if $\Gamma_N = \emptyset$. A basis of $\mathcal{R}T^0(\mathcal{T}_h)$ is given by vector fields $\psi_S \in \mathcal{R}T^0(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, satisfying the Kronecker property $\psi_S|_{S'} \cdot n_{S'} = \delta_{S,S'}$ on S' for all $S' \in \mathcal{S}_h$, where n_S for all $S \in \mathcal{S}_h$ is the unit normal vector on S pointing from T_- to T_+ if $T_+ \cap T_- = S \in \mathcal{S}_h$. A basis of $\mathcal{R}T^0_N(\mathcal{T}_h)$ is given by $\psi_S \in \mathcal{R}T^0_N(\mathcal{T}_h)$, $S \in \mathcal{S}_h \setminus \Gamma_N$. The quasi-interpolation operator $I_{rt} \colon W^{1,2}(\Omega;\mathbb{R}^d) \cap W^2_N(\mathrm{div};\Omega) \to \mathcal{R}T^0_N(\mathcal{T}_h)$, for every $y \in W^{1,2}(\Omega;\mathbb{R}^d) \cap W^2_N(\mathrm{div};\Omega)$ defined by

$$I_{rt} y \coloneqq \sum_{S \in \mathcal{S}_h} y_S\, \psi_S\,, \qquad \text{where } y_S \coloneqq \fint_S y \cdot n_S \,\mathrm{d}s\,, \tag{2.3}$$

preserves averages of divergences, i.e., $\mathrm{div}(I_{rt} y) = \Pi_h(\mathrm{div}\, y)$ in $\mathcal{L}^0(\mathcal{T}_h)$ for every $y \in W^{1,2}(\Omega;\mathbb{R}^d)$ $\cap W^2_N(\mathrm{div};\Omega)$. There exists a constant $c_{rt} > 0$, depending only on the chunkiness $\omega_0 > 0$, such that for every $y \in W^{1,2}(\Omega;\mathbb{R}^d) \cap W^2_N(\mathrm{div};\Omega)$ and $T \in \mathcal{T}_h$, cf. [25, Theorem 16.4], it holds

(RT.1) $\|I_{rt} y\|_T \le c_{rt}\,(\|I_{rt} y\|_T + h_T\,\|\nabla y\|_T)\,,$
(RT.2) $\|y - I_{rt} y\|_T \le c_{rt}\, h_T\,\|\nabla y\|_T\,,$
(RT.3) $\|\mathrm{div}\,(y - I_{rt} y)\|_T \le c_{rt}\, h_T\,\|\mathrm{div}\, y\|_T\,.$

2.2.3 Discrete integration-by-parts formula

An element-wise integration-by-parts implies that for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{R}T^0(\mathcal{T}_h)$, we have the *discrete integration-by-parts formula*

$$(\nabla_h v_h, \Pi_h y_h)_\Omega + (\Pi_h v_h,\, \mathrm{div}\, y_h)_\Omega = (v_h, y_h \cdot n)_{\partial\Omega}\,. \tag{2.4}$$

Here, we used that $y_h \in \mathcal{R}T^0(\mathcal{T}_h)$ has continuous constant normal components on inner element sides, i.e., $y_h|_T \cdot n_T = \text{const}$ on ∂T for every $T \in \mathcal{T}_h$ and $[\![y_h \cdot n]\!]_S = 0$ on S for every $S \in \mathcal{S}_h^i$, as well as that the jumps of $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ across inner element sides have vanishing integral mean, i.e., $\int_S [\![v_h]\!]_S \,\mathrm{d}s = [\![v_h]\!]_S(x_S) = 0$ for all $S \in \mathcal{S}_h^i$. In particular, for any $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $y_h \in \mathcal{R}T^0_N(\mathcal{T}_h)$, (2.4) reads

$$(\nabla_h v_h, \Pi_h y_h)_\Omega = -(\Pi_h v_h,\, \mathrm{div}\, y_h)_\Omega\,. \tag{2.5}$$

In [16, 5, 6, 7], the discrete integration-by-parts formula (2.5) formed a cornerstone in the derivation of a discrete convex duality theory and, as such, also plays a central role in the hereinafter analysis. For instance, cf. [6, Lemma 2.1], for every $v \in W^{1,2}_D(\Omega)$ and $y \in W^{1,2}(\Omega;\mathbb{R}^d) \cap W^2_N(\mathrm{div};\Omega)$, (2.5) enables to exchange quasi-interpolation operators via

$$(\mathrm{div}\, y, v - \Pi_h I_{cr} v)_\Omega = -(\nabla v, y - \Pi_h I_{rt} y)_\Omega\,. \tag{2.6}$$

In addition, cf. [8, Section 2.4], there holds the orthogonal decomposition

$$\mathcal{L}^0(\mathcal{T}_h)^d = \ker(\mathrm{div}|_{\mathcal{R}T^0_N(\mathcal{T}_h)}) \oplus \nabla_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))\,. \tag{2.7}$$

[4]For every $S \in \mathcal{S}_h^i$, $[\![y_h \cdot n]\!]_S \coloneqq y_h|_{T_+} \cdot n_{T_+} + y_h|_{T_-} \cdot n_{T_-}$ on S, where $T_+, T_- \in \mathcal{T}_h$ satisfy $\partial T_+ \cap \partial T_- = S$, and 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, and for every $S \in \mathcal{S}_h \cap \partial\Omega$, $[\![y_h \cdot n]\!]_S \coloneqq y_h|_T \cdot n$ on S, where $T \in \mathcal{T}_h$ satisfies $S \subseteq \partial T$ and $n \colon \partial\Omega \to \mathbb{S}^{d-1}$ denotes the outward unit normal vector field to Ω.

3. Obstacle problem

In this section, we discuss the continuous and the discrete obstacle problem.

3.1 Continuous obstacle problem

Primal problem. Given a force $f \in L^2(\Omega)$ and an obstacle $\chi \in W^{1,2}(\Omega)$ with $\chi \leq 0$ on Γ_D, the (continuous) obstacle problem is defined via the minimization of $I\colon W^{1,2}_D(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\nabla v\|_\Omega^2 - (f, v)_\Omega + I_K(v)\,, \tag{3.1}$$

where

$$K \coloneqq \big\{v \in W^{1,2}_D(\Omega) \mid v \geq \chi \text{ a.e. in } \Omega\big\}\,,$$

and $I_K : W^{1,2}_D(\Omega) \to \mathbb{R}\cup\{+\infty\}$ is given via $I_K(v) \coloneqq 0$ if $v \in K$ and $I_K(v) \coloneqq +\infty$ else. In what follows, we refer to the minimization of the functional (3.1) as the *primal problem*. Since the functional (3.1) is proper, strictly convex, weakly coercive, and lower semi-continuous, cf. [3, Theorem 5.1], the direct method in the calculus of variations, cf. [20], yields the existence of a unique minimizer $u \in K$, called the *primal solution*. In what follows, we reserve the notation $u \in K$ for the primal solution. Since the functional (3.1) is not Fréchet differentiable, the optimality conditions associated with the primal problem are not given via a variational equality. Instead, they are given via a variational inequality. In fact, cf. [3, Theorem 5.2], $u \in K$ is minimal for (3.1) if and only if for every $v \in K$

$$(\nabla u, \nabla u - \nabla v)_\Omega \leq (f, u - v)_\Omega\,. \tag{3.2}$$

Dual problem. Appealing to [24, Section 2.4, p. 84 ff.], the *dual problem* to the obstacle problem is defined via the maximization of $D\colon L^2(\Omega;\mathbb{R}^d) \to \mathbb{R} \cup \{-\infty\}$, for every $y \in L^2(\Omega;\mathbb{R}^d)$ defined by

$$D(y) \coloneqq -\tfrac{1}{2}\|y\|_\Omega^2 - I_K^*(\operatorname{Div}(y) + F)\,, \tag{3.3}$$

where $I_K^*\colon (W^{1,2}_D(\Omega))^* \to \mathbb{R}\cup\{+\infty\}$ is defined by $I_K^*(v^*) \coloneqq 0$ if $\langle v^*, v\rangle_{W^{1,2}_D(\Omega)} \leq 0$ for all $v \in K$ and $I_K^*(v^*) \coloneqq +\infty$ else, $\operatorname{Div}\colon L^2(\Omega;\mathbb{R}^d) \to (W^{1,2}_D(\Omega))^*$ is defined by $\langle \operatorname{Div} y, v\rangle_\Omega \coloneqq -(y, \nabla v)_\Omega$ for all $y \in L^2(\Omega;\mathbb{R}^d)$ and $v \in W^{1,2}_D(\Omega)$, and $F \in (W^{1,2}_D(\Omega))^*$ is defined by $\langle F, v\rangle_\Omega \coloneqq (f, v)_\Omega$ for all $v \in W^{1,2}_D(\Omega)$. For every $y \in W^2_N(\operatorname{div};\Omega)$, there holds the representation

$$D(y) \coloneqq -\tfrac{1}{2}\|y\|_\Omega^2 - (\operatorname{div} y + f, \chi)_\Omega - I_-(\operatorname{div} y + f)\,, \tag{3.4}$$

where $I_-\colon L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is defined by $I_-(g) \coloneqq 0$ if $g \in L^2(\Omega)$ with $g \leq 0$ a.e. in Ω and $I_-(g) \coloneqq +\infty$ else. Moreover, in [24, Section 2.4, p. 84 ff.], it is shown that there exists a unique maximizer $z \in L^2(\Omega;\mathbb{R}^d)$ of (3.3), called the *dual solution*, and a *strong duality relation*, i.e.,

$$I(u) = D(z)\,, \tag{3.5}$$

applies. In addition, there hold the *convex optimality relations*

$$z = \nabla u \quad \text{ in } L^2(\Omega;\mathbb{R}^d)\,, \tag{3.6}$$

$$\langle \operatorname{Div} z + F, u\rangle_\Omega = I_K^*(\operatorname{Div} z + F)\,. \tag{3.7}$$

Augmented problem. Due to [3, Theorem 5.2], there exists a Lagrange multiplier $\Lambda \in (W^{1,2}_D(\Omega))^*$ with $\Lambda \leq 0$ in $(W^{1,2}_D(\Omega))^*$, i.e., $\langle \Lambda, v\rangle_\Omega \leq 0$ for all $v \in W^{1,2}_D(\Omega)$ with $v \geq 0$ for a.e. Ω, such that for every $v \in W^{1,2}_D(\Omega)$, it holds the *augmented problem*

$$(\nabla u, \nabla v)_\Omega + \langle \Lambda, v\rangle_\Omega = (f, v)_\Omega\,, \tag{3.8}$$

i.e., $\Lambda = \operatorname{Div} z + f$ in $(W^{1,2}_D(\Omega))^*$. Then, cf. [3, Theorem 5.2], there holds the *complementary condition*

$$\langle \Lambda, u\rangle_\Omega = I_K^*(\Lambda)\,. \tag{3.9}$$

If there exists $\lambda \in L^2(\Omega)$ such that $\langle \Lambda, v\rangle_\Omega = (\lambda, v)_\Omega$ for all $v \in W^{1,2}_D(\Omega)$, cf. [36], then (3.9) reads

$$\lambda(u - \chi) = 0 \quad \text{ a.e. in } \Omega\,. \tag{3.10}$$

3.2 Discrete obstacle problem

Discrete primal problem. Given a force $f\in L^2(\Omega)$ and an obstacle $\chi\in W^{1,2}(\Omega)$ such that $\chi\leq 0$ on Γ_D, with $f_h\coloneqq \Pi_h f\in \mathcal{L}^0(\mathcal{T}_h)$ and $\chi_h\in \mathcal{L}^0(\mathcal{T}_h)$ approximating χ, the discrete obstacle problem is defined via the minimization of $I_h^{cr}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$, for every $v_h\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h)\coloneqq \tfrac{1}{2}\|\nabla_h v_h\|_\Omega^2-(f_h,\Pi_h v_h)_\Omega+I_{K_h^{cr}}(v_h)\,, \tag{3.11}$$

where

$$K_h^{cr}\coloneqq \big\{v_h\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\mid \Pi_h v_h\geq \chi_h\text{ a.e. in }\Omega\big\}\,,$$

and $I_{K_h^{cr}}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$ is given via $I_{K_h^{cr}}(v_h)\coloneqq 0$ if $v_h\in K_h^{cr}$ and $I_{K_h^{cr}}(v_h)\coloneqq +\infty$ else. In what follows, we refer to the minimization of the functional (3.11) as the *discrete primal problem*. Since the functional (3.11) is proper, strictly convex, weakly coercive, and lower semi-continuous, the direct method in the calculus of variations, cf. [20], yields the existence of a unique minimizer $u_h^{cr}\in K_h^{cr}$, called the *discrete primal solution*. In what follows, we reserve the notation $u_h^{cr}\in K_h^{cr}$ for the discrete primal solution. In addition, $u_h^{cr}\in K_h^{cr}$ is the unique minimizer of (3.11) if and only if for every $v_h\in K_h^{cr}$, it holds

$$(\nabla_h u_h^{cr},\nabla_h u_h^{cr}-\nabla_h v_h)_\Omega\leq (f_h,\Pi_h u_h^{cr}-\Pi_h v_h)_\Omega\,. \tag{3.12}$$

Discrete dual problem. Appealing to [6, Subsection 4.1], the *discrete dual problem* to the discrete obstacle problem is defined via the maximization of $D_h^{rt}\colon \mathcal{R}T^0_N(\mathcal{T}_h)\to\mathbb{R}\cup\{-\infty\}$, for every $y_h\in \mathcal{R}T^0_N(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h)\coloneqq -\tfrac{1}{2}\|\Pi_h y_h\|_\Omega^2-(\operatorname{div}y_h+f_h,\chi_h)_\Omega-I_-(\operatorname{div}y_h+f_h)\,. \tag{3.13}$$

Discrete augmented problem. The *discrete augmented problem*, similar to the augmented problem (3.8), seeks for a *discrete Lagrange multiplier* $\overline{\lambda}_h^{cr}\in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ such that $\overline{\lambda}_h^{cr}\leq 0$ a.e. in Ω and for every $v_h\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, it holds

$$(\overline{\lambda}_h^{cr},\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega-(\nabla_h u_h^{cr},\nabla_h v_h)_\Omega\,. \tag{3.14}$$

The following proposition establishes the well-posedness of the discrete augmented problem (3.14).

Proposition 3.1. *The following statements apply:*

(i) The discrete augmented problem is well-posed, i.e., there exists a unique discrete Lagrange multiplier $\overline{\lambda}_h^{cr}\in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ that satisfies (3.14).

(ii) The discrete Lagrange multiplier $\overline{\lambda}_h^{cr}\in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ satisfies $\overline{\lambda}_h^{cr}\leq 0$ a.e. in Ω and the discrete complementarity condition

$$\overline{\lambda}_h^{cr}(\Pi_h u_h^{cr}-\chi_h)=0\qquad\text{in }\mathcal{L}^0(\mathcal{T}_h)\,. \tag{3.15}$$

Remark 3.2. *The discrete complementarity condition (3.15) is a discrete analogue of the (continuous) variational complementarity condition (3.9) and the (continuous) point-wise complememtarity condition (3.10), respectively.*

Proof (of Proposition 3.1). ad (i). We relax the obstacle constraint via a penalization scheme, i.e., for every $\varepsilon>0$, we consider the minimization of $I_{h,\varepsilon}^{cr}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}$, for every $v_h\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ defined by

$$I_{h,\varepsilon}^{cr}(v_h)\coloneqq \tfrac{1}{2}\|\nabla_h v_h\|_\Omega^2-(f_h,\Pi_h v_h)_\Omega+\tfrac{\varepsilon^{-2}}{2}\|(\Pi_h v_h-\chi_h)_-\|_\Omega^2\,.$$

Since for any $\varepsilon>0$, $I_{h,\varepsilon}^{cr}\colon \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\to\mathbb{R}$ is continuous, strictly convex, and weakly coercive, the direct method in the calculus of variation yields the existence of a unique minimizer $u_{h,\varepsilon}^{cr}\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, which, for every $v_h\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, abbreviating $\lambda_{h,\varepsilon}^{cr}\coloneqq \varepsilon^{-2}(\Pi_h u_{h,\varepsilon}^{cr}-\chi_h)_-\in \mathcal{L}^0(\mathcal{T}_h)$, satisfies

$$(\nabla_h u_{h,\varepsilon}^{cr},\nabla_h v_h)_\Omega+(\lambda_{h,\varepsilon}^{cr},\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega\,. \tag{3.16}$$

Due to the minimality of $u_{h,\varepsilon}^{cr}\in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, we find that $I_{h,\varepsilon}^{cr}(u_{h,\varepsilon}^{cr})\leq I_{h,\varepsilon}^{cr}(u_h^{cr})$ and, as a consequence, using that $(\Pi_h u_h^{cr}-\chi_h)_-=0$ a.e. in Ω, that

$$\tfrac{1}{2}\|\nabla_h u_{h,\varepsilon}^{cr}\|_\Omega^2+\tfrac{\varepsilon^2}{2}\|\lambda_{h,\varepsilon}^{cr}\|_\Omega^2\le \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_\Omega^2+(f_h,\Pi_h(u_{h,\varepsilon}^{cr}-u_h^{cr}))_\Omega\,. \tag{3.17}$$

Using the κ-Young inequality $ab\le \frac{1}{4\kappa}a^2+\kappa b^2$, valid for all $a,b\ge 0$ and $\kappa>0$, (L0.1), and the discrete Poincaré inequality (2.1), for every $\varepsilon>0$, we find that

$$|(f_h,\Pi_h u_{h,\varepsilon}^{cr})_\Omega|\le \tfrac{1}{4\kappa}\,\|f_h\|_\Omega^2+\kappa\,(c_P^{cr})^2\,\|\nabla_h u_{h,\varepsilon}^{cr}\|_\Omega^2\,. \tag{3.18}$$

Using (3.18) for $\kappa=\frac{1}{4(c_P^{cr})^2}>0$ in (3.17), for every $\varepsilon>0$, we arrive at

$$\tfrac{1}{4}\|\nabla_h u_{h,\varepsilon}^{cr}\|_\Omega^2+\varepsilon^2\|\lambda_{h,\varepsilon}^{cr}\|_\Omega^2\le \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_\Omega^2-(f_h,\Pi_h u_h^{cr})_\Omega+(c_P^{cr})^2\,\|f_h\|_\Omega^2\,. \tag{3.19}$$

Using the discrete Poincaré inequality (2.1) in (3.19), we find that $(u_{h,\varepsilon}^{cr})_{\varepsilon>0}\subseteq \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ is bounded. Hence, owing to the finite dimensionality of $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we deduce the existence of $\tilde{u}_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ such that, for a not re-labeled subsequence, it holds

$$u_{h,\varepsilon}^{cr}\to\tilde{u}_h^{cr}\quad\text{ in }\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\quad(\varepsilon\to 0)\,. \tag{3.20}$$

Let $E_h^{cr}\colon\mathcal{L}^0(\mathcal{T}_h)\to(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*$ for every $\mu_h\in\mathcal{L}^0(\mathcal{T}_h)$ and $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ be defined by

$$\langle E_h^{cr}\mu_h,v_h\rangle_{\mathcal{S}_D^{1,cr}(\mathcal{T}_h)}\coloneqq(\mu_h,\Pi_h v_h)_\Omega\,. \tag{3.21}$$

Then, from (3.16), also using (L0.1), for every $\varepsilon>0$, it follows that

$$\begin{aligned}\|E_h^{cr}\lambda_{h,\varepsilon}^{cr}\|_{(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*}&=\sup_{v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h);\|v_h\|_\Omega+\|\nabla_h v_h\|_\Omega\le 1}(f_h,\Pi_h v_h)_\Omega-(\nabla_h u_{h,\varepsilon}^{cr},\nabla_h v_h)_\Omega\\&\le\|f_h\|_\Omega+\|\nabla_h u_{h,\varepsilon}^{cr}\|_\Omega\,.\end{aligned} \tag{3.22}$$

Using (3.19) in (3.22), we find that $(E_h^{cr}\lambda_{h,\varepsilon}^{cr})_{\varepsilon>0}\subseteq(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*$ is bounded. Thus, due to the finite dimensionality of $(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*$ and the closedness of the range $R(E_h^{cr})$, there exists $\tilde{\lambda}_h\in\mathcal{L}^0(\mathcal{T}_h)$ with

$$E_h^{cr}\lambda_{h,\varepsilon}^{cr}\to E_h^{cr}\tilde{\lambda}_h\quad\text{ in }(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*\quad(\varepsilon\to 0)\,. \tag{3.23}$$

Next, using (3.19) once more and that, by definition, $\lambda_{h,\varepsilon}^{cr}=\varepsilon^{-2}(\Pi_h u_{h,\varepsilon}^{cr}-\chi_h)_-$, we deduce that

$$\|(\Pi_h u_{h,\varepsilon}^{cr}-\chi_h)_-\|_\Omega^2=\varepsilon^4\,\|\lambda_{h,\varepsilon}^{cr}\|_\Omega^2\to 0\quad(\varepsilon\to 0)\,.$$

Because, on the other hand, due to (3.20), $(\Pi_h u_{h,\varepsilon}^{cr}-\chi_h)_-\to(\Pi_h\tilde{u}_h^{cr}-\chi_h)_-$ in $L^2(\Omega)$ $(\varepsilon\to 0)$, we conclude that $(\Pi_h\tilde{u}_h^{cr}-\chi_h)_-=0$ a.e. in Ω. In other words, we have that

$$\tilde{u}_h^{cr}\in K_h^{cr}\,. \tag{3.24}$$

As a consequence of (3.24), for every $\varepsilon>0$ and $v_h\in K_h^{cr}$, resorting to (3.20) and the minimality of $u_{h,\varepsilon}^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ for $I_{h,\varepsilon}^{cr}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathbb{R}$, we find that

$$I_h^{cr}(\tilde{u}_h^{cr})=\lim_{\varepsilon\to 0}I_h^{cr}(u_{h,\varepsilon}^{cr})\le\lim_{\varepsilon\to 0}I_{h,\varepsilon}^{cr}(u_{h,\varepsilon}^{cr})\le\lim_{\varepsilon\to 0}I_{h,\varepsilon}^{cr}(v_h)=I_h^{cr}(v_h)\,.$$

Hence, due to the uniqueness of $u_h^{cr}\in K_h^{cr}$ as a minimizer of $I_h^{cr}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathbb{R}$, we infer that $\tilde{u}_h^{cr}=u_h^{cr}$ in $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$. By passing for $\varepsilon\to 0$ in (3.16), for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, using (3.20), (3.23), and the definition of $E_h^{cr}\colon\mathcal{L}^0(\mathcal{T}_h)\to(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^*$, cf. (3.21), we conclude that

$$(\nabla_h u_h^{cr},\nabla_h v_h)_\Omega+(\tilde{\lambda}_h,\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega\,. \tag{3.25}$$

Owing to $\mathcal{L}^0(\mathcal{T}_h)=\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))\bigoplus\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^\perp$, there exist unique $\overline{\lambda}_h^{cr}\in\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$ and $\lambda_h\in\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^\perp$ such that $\tilde{\lambda}_h=\overline{\lambda}_h^{cr}+\lambda_h$ in $\mathcal{L}^0(\mathcal{T}_h)$. By the aid of the latter decomposition, for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we conclude from (3.25) that

$$(\nabla_h u_h^{cr},\nabla_h v_h)_\Omega+(\overline{\lambda}_h^{cr},\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega\,.$$

Next, let $\overline{\mu}_h^{cr}\in\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$ be such that for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, it holds

$$(\nabla_h u_h^{cr},\nabla_h v_h)_\Omega+(\overline{\mu}_h^{cr},\Pi_h v_h)_\Omega=(f_h,\Pi_h v_h)_\Omega\,.$$

Then, $\overline{\lambda}_h^{cr}-\overline{\mu}_h^{cr}\in\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))\cap\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))^\perp=\{0\}$ and, thus, $\overline{\lambda}_h^{cr}=\overline{\mu}_h^{cr}$ in $\Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$.

ad (ii). Let $\mathcal{T}_h^{cr} \subseteq \mathcal{T}_h$ be such that $\operatorname{span}(\{\chi_T \mid T \in \mathcal{T}_h^{cr}\}) = \Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$. For each $T \in \mathcal{T}_h^{cr}$, there exists $v_h^T \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ such that $\Pi_h v_h^T = \chi_T$. Next, let $\alpha_T \in \mathbb{R}$ be such that $\Pi_h u_h^{cr} + \alpha_T \Pi_h v_h^T = \Pi_h u_h^{cr} + \alpha_T \chi_T \geq \chi_h$ a.e. in Ω. Then, for $v_h = \alpha_T v_h^T \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ in (3.14), in particular, using (3.12) for $v_h = u_h^{cr} + \alpha_T v_h^T \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we deduce that

$$\alpha_T\,|T|\,\overline{\lambda}_h^{cr} = \alpha_T\,[(f_h, \Pi_h v_h^T)_\Omega - (\nabla_h u_h^{cr}, \nabla_h v_h^T)_\Omega] \leq 0 \quad \text{ a.e. in } T\,. \tag{3.26}$$

As $T \in \mathcal{T}_h^{cr}$ was arbitrary and $\overline{\lambda}_h^{cr} \in \Pi_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))$, we conclude from (3.26) that $\overline{\lambda}_h^{cr} \leq 0$ a.e. in Ω. Eventually, for $T \in \mathcal{T}_h^{cr}$ such that $\Pi_h u_h^{cr} > \chi_h$ a.e. in T, there exists some $\alpha_T < 0$ such that $\Pi_h u_h^{cr} + \alpha_T \Pi_h v_h^T = \Pi_h u_h^{cr} + \alpha_T \chi_T \geq \chi_h$ a.e. in Ω. For this $\alpha_T < 0$ in (3.26), also using that $\overline{\lambda}_h^{cr} \leq 0$ a.e. in Ω, we arrive at

$$0 \leq \alpha_T\,|T|\,\overline{\lambda}_h^{cr} = \alpha_T\,[(f_h, \Pi_h v_h^T)_\Omega - (\nabla_h u_h^{cr}, \nabla_h v_h^T)_\Omega] \leq 0 \quad \text{ a.e. in } T\,,$$

so that $\overline{\lambda}_h^{cr} = 0$ a.e. in T. In other words, it holds (3.15). □

Given a discrete Lagrange multiplier $\overline{\lambda}_h^{cr} \in \mathcal{L}^0(\mathcal{T}_h)$ satisfying (3.14), we define the *discrete flux*

$$z_h^{rt} \coloneqq \nabla_h u_h^{cr} + \frac{\overline{\lambda}_h^{cr} - f_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \in \mathcal{L}^1(\mathcal{T}_h)^d\,, \tag{3.27}$$

which, by definition, satisfies

$$\Pi_h z_h^{rt} = \nabla_h u_h^{cr} \quad \text{ in } \mathcal{L}^0(\mathcal{T}_h)^d\,. \tag{3.28}$$

The following proposition proves that the discrete flux is admissible in the discrete dual problem and even a discrete dual solution.

Proposition 3.3. *The following statements apply:*

(i) The discrete flux $z_h^{rt} \in \mathcal{L}^1(\mathcal{T}_h)^d$ satisfies $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ and

$$\operatorname{div} z_h^{rt} = \overline{\lambda}_h^{cr} - f_h \quad \text{ in } \mathcal{L}^0(\mathcal{T}_h)\,. \tag{3.29}$$

In particular, it holds $\operatorname{div} z_h^{rt} + f_h \leq 0$ a.e. in Ω, i.e., $I_-(\operatorname{div} z_h^{rt} + f_h) = 0$.

(ii) The discrete flux $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ is a maximizer of (3.13) and discrete strong duality, i.e., $I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})$, applies. In addition, there holds the discrete complementary condition

$$(\operatorname{div} z_h^{rt} + f_h)\,(\Pi_h u_h^{cr} - \chi_h) = 0 \quad \text{ in } \mathcal{L}^0(\mathcal{T}_h)\,. \tag{3.30}$$

Proof. ad (i). Since, due to $|\Gamma_D| > 0$, $\operatorname{div}\colon \mathcal{RT}_N^0(\mathcal{T}_h) \to \mathcal{L}^0(\mathcal{T}_h)$ is surjective, there exists some $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ such that $\operatorname{div} y_h = \overline{\lambda}_h^{cr} - f_h$ in $\mathcal{L}^0(\mathcal{T}_h)$. Then, using the discrete integration-by-parts formula (2.5) and (3.14), for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we find that

$$(\Pi_h y_h, \nabla_h v_h)_\Omega = -(\operatorname{div} y_h, \Pi_h v_h)_\Omega = (f_h - \overline{\lambda}_h^{cr}, \Pi_h v_h)_\Omega = (\nabla_h u_h^{cr}, \nabla_h v_h)_\Omega\,. \tag{3.31}$$

Using (3.28) in (3.31), for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we arrive at

$$(y_h - z_h^{rt}, \nabla_h v_h)_\Omega = (\Pi_h(y_h - z_h^{rt}), \nabla_h v_h)_\Omega = 0\,. \tag{3.32}$$

On the other hand, we have that $\operatorname{div}(y_h - z_h^{rt}) = 0$ a.e. in T for all $T \in \mathcal{T}_h$, so that $y_h - z_h^{rt} \in \mathcal{L}^0(\mathcal{T}_h)^d$. Hence, by (3.32) and the orthogonal decomposition (2.7), we conclude that $y_h - z_h^{rt} \in \nabla_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h)) \subseteq \mathcal{RT}_N^0(\mathcal{T}_h)$ and, thus, $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$, since already $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$.

ad (ii). The discrete optimality relation (3.30) follows from (3.29) and (3.15). In consequence, it remains to establish the strong duality relation. Using (3.30), the discrete integration-by-parts formula (2.5), (3.28), and $I_-(\operatorname{div} z_h^{rt} + f_h) = 0$, we observe that

$$\begin{aligned}
I_h^{cr}(u_h^{cr}) &= \tfrac{1}{2}\|\nabla_h u_h^{cr}\|_\Omega^2 - (f_h, \Pi_h u_h^{cr})_\Omega \\
&= \tfrac{1}{2}\|\Pi_h z_h^{rt}\|_\Omega^2 + (\operatorname{div} z_h^{rt}, \Pi_h u_h^{cr})_\Omega - (\operatorname{div} z_h^{rt} + f_h, \chi_h)_\Omega \\
&= \tfrac{1}{2}\|\Pi_h z_h^{rt}\|_\Omega^2 - (\Pi_h z_h^{rt}, \nabla_h u_h^{cr})_\Omega - (\operatorname{div} z_h^{rt} + f_h, \chi_h)_\Omega \\
&= -\tfrac{1}{2}\|\Pi_h z_h^{rt}\|_\Omega^2 - (\operatorname{div} z_h^{rt} + f_h, \chi_h)_\Omega = D_h^{rt}(z_h^{rt})\,.
\end{aligned}$$
□

4. A priori error analysis

In this section, we establish a priori error estimates for the discrete primal problem.

Theorem 4.1. *If $u,\chi\in W^{2,2}(\Omega)$, i.e., $z\in W^{1,2}(\Omega;\mathbb{R}^d)$ and $\lambda:=f+\operatorname{div} z\in L^2(\Omega)$, and $\chi_h:=\Pi_h I_{cr}\chi \in \mathcal{L}^0(\mathcal{T}_h)$, then there exists a constant $c>0$, depending only on the chunkiness $\omega_0>0$, such that*

$$\begin{aligned}\|\nabla_h I_{cr}u-\nabla_h u_h^{cr}\|_\Omega^2+(-\overline{\lambda}_h^{cr},\Pi_h(I_{cr}u-u_h^{cr}))_\Omega+\|\Pi_h I_{rt}z-\Pi_h z_h^{rt}\|_\Omega^2\\ \le c\,h_{\max}^2\,(\|D^2u\|_\Omega^2+\|D^2\chi\|_\Omega^2+\|\lambda\|_\Omega^2)\,.\end{aligned}$$

Proof. Using that, owing to the discrete augmented problem (3.14), $\frac{1}{2}a^2-\frac{1}{2}b^2=\frac{1}{2}(a-b)^2+b(a-b)$ for all $a,b\in\mathbb{R}$, and the strong concavity of (3.13), for every $v_h\in K_h^{cr}$ and $y_h\in\mathcal{R}T^0(\mathcal{T}_h)$, it holds

$$\begin{aligned}\tfrac{1}{2}\|\nabla_h v_h-\nabla_h u_h^{cr}\|_\Omega^2+(-\overline{\lambda}_h^{cr},\Pi_h(v_h-u_h^{cr}))_\Omega&=I_h^{cr}(v_h)-I_h^{cr}(u_h^{cr})\,,\\ \tfrac{1}{2}\|\Pi_h y_h-\Pi_h z_h^{rt}\|_\Omega^2&\le D_h^{rt}(z_h^{rt})-D_h^{rt}(y_h)\,,\end{aligned}$$

that $I_{cr}u\in K_h^{cr}$, as $⨍_S(u-\chi)\,\mathrm{d}s\ge 0$ and $\varphi_S(x_T)=\frac{1}{d+1}$ for all $T\in\mathcal{T}_h$ and $S\in\mathcal{S}_h$ with $S\subseteq\partial T$, that $I_{rt}z\in\mathcal{R}T^0_N(\mathcal{T}_h)$ with $\operatorname{div} I_{rt}z+f_h=\Pi_h(\operatorname{div} z+f)=\Pi_h\lambda\le 0$ a.e. in Ω, the discrete strong duality relation $I_h^{cr}(u_h^{cr})=D_h^{rt}(z_h^{rt})$, cf. Proposition 3.3 (ii), $\nabla_h I_{cr}u=\Pi_h\nabla u$ in $\mathcal{L}^0(\mathcal{T}_h)^d$, (L0.1), and the strong duality relation $I(u)=D(z)$, cf. (3.5), we find that

$$\begin{aligned}\tfrac{1}{2}\|\nabla_h I_{cr}u-\nabla_h u_h^{cr}\|_\Omega^2&+(-\overline{\lambda}_h^{cr},\Pi_h(I_{cr}u-u_h^{cr}))_\Omega+\tfrac{1}{2}\|\Pi_h I_{rt}z-\Pi_h z_h^{rt}\|_\Omega^2\\ &\le I_h^{cr}(I_{cr}u)-D_h^{rt}(I_{rt}z)\\ &\le I(u)+(f,u-\Pi_h I_{cr}u)_\Omega-D_h^{rt}(I_{rt}z)\\ &=-\tfrac{1}{2}\|z\|_\Omega^2+(f,u-\Pi_h I_{cr}u)_\Omega+\tfrac{1}{2}\|\Pi_h I_{rt}z\|_\Omega^2\\ &\quad-(\operatorname{div} z+f,\chi)_\Omega+(\operatorname{div} I_{rt}z+f_h,\Pi_h I_{cr}\chi)_\Omega\,.\end{aligned}\tag{4.1}$$

Next, using in (4.1) the exchange of quasi-interpolation operators (2.6) and $z=\nabla u$, cf. (3.6), i.e.,

$$(\operatorname{div} z,u-\Pi_h I_{cr}u)_\Omega=-(z,z-\Pi_h I_{rt}z)_\Omega=-\|z\|_\Omega^2+(z,\Pi_h I_{rt}z)_\Omega\,,$$

$\operatorname{div} I_{rt}z+f_h=\Pi_h\lambda$ in $\mathcal{L}^0(\mathcal{T}_h)$, and $\operatorname{div} z+f=\lambda$ in $L^2(\Omega)$, abbreviating $\tilde{u}:=u-\chi\in W^{2,2}(\Omega)$, we get

$$\begin{aligned}\tfrac{1}{2}\|\nabla_h I_{cr}u-\nabla_h u_h^{cr}\|_\Omega^2&+(-\overline{\lambda}_h^{cr},\Pi_h(I_{cr}u-u_h^{cr}))_\Omega+\tfrac{1}{2}\|\Pi_h I_{rt}z-\Pi_h z_h^{rt}\|_\Omega^2\\ &\le(\lambda,\tilde{u}-\Pi_h I_{cr}\tilde{u})_\Omega+\tfrac{1}{2}\|z\|_\Omega^2-(z,\Pi_h I_{rt}z)_\Omega+\tfrac{1}{2}\|\Pi_h I_{rt}z\|_\Omega^2\\ &=(\lambda,\tilde{u}-I_{cr}\tilde{u})_\Omega+(\lambda,I_{cr}\tilde{u}-\Pi_h I_{cr}\tilde{u})_\Omega+\tfrac{1}{2}\|z-\Pi_h I_{rt}z\|_\Omega^2\\ &=:I_h^1+I_h^2+I_h^3\,.\end{aligned}\tag{4.2}$$

As a consequence, it remains to estimate the terms I_h^1, I_h^2 and I_h^3:

ad I_h^1. Using (CR.3), we obtain

$$I_h^1\le\|\lambda\|_\Omega\|\tilde{u}-I_{cr}\tilde{u}\|_\Omega\le c_{cr}\,h_{\max}^2\,\|\lambda\|_\Omega\,\|D^2\tilde{u}\|_\Omega\,.\tag{4.3}$$

ad I_h^2. Using that $I_{cr}\tilde{u}-\Pi_h I_{cr}\tilde{u}=\nabla_h I_{cr}\tilde{u}\cdot(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})$ in $\mathcal{L}^1(\mathcal{T}_h)^d$, $\lambda=0$ in $\{\tilde{u}>0\}$, $\nabla\tilde{u}=0$ in $\{\tilde{u}=0\}$, and (CR.3), we obtain

$$I_h^2\le(\lambda,(\nabla_h I_{cr}\tilde{u}-\nabla\tilde{u})\cdot(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d}))_\Omega\le c_{cr}\,h_{\max}^2\,\|\lambda\|_\Omega\,\|D^2\tilde{u}\|_\Omega\,.\tag{4.4}$$

ad I_h^3. Using (L0.1), (L0.2), and (RT.2), we obtain

$$I_h^3\le\|z-\Pi_h z\|_\Omega^2+\|\Pi_h(z-I_{rt}z)\|_\Omega^2\le(c_\Pi^2+c_{rt}^2)\,h_{\max}^2\,\|\nabla z\|_\Omega^2\,.\tag{4.5}$$

Combining (4.2)–(4.5), we arrive at the claimed a priori error estimate. □

Corollary 4.2. *If $u,\chi\in W^{2,2}(\Omega)$, then there exists a constant $c>0$, depending only on the chunkiness $\omega_0>0$, such that*

$$\|\nabla_h u_h^{cr}-\nabla u\|_\Omega^2\le c\,h_{\max}^2\,(\|D^2u\|_\Omega^2+\|D^2\chi\|_\Omega^2+\|\lambda\|_\Omega^2)\,.$$

Proof. Resorting to (CR.2), the assertion follows from Theorem 4.1, exploiting that for all $v_h\in K_h^{cr}$

$$(-\overline{\lambda}_h^{cr},\Pi_h(v_h-u_h^{cr}))_\Omega=(\nabla_h u_h^{cr},\nabla_h v_h-\nabla_h u_h^{cr})_\Omega-(f_h,\Pi_h(v_h-\nabla_h u_h^{cr}))_\Omega\ge 0\,.\qquad\square$$

5. A posteriori error analysis

In this section, we examine the *primal-dual a posteriori error estimator* $\eta_h^2\colon W_D^{1,2}(\Omega)\to\mathbb{R}$, for every $v\in W_D^{1,2}(\Omega)$ defined by

$$\begin{aligned}\eta_h^2(v)&\coloneqq \eta_{A,h}^2(v)+\eta_{B,h}^2(v)+\eta_{C,h}^2\,,\\ \eta_{A,h}^2(v)&\coloneqq \|\nabla v-\nabla_h u_h^{cr}\|_\Omega^2\,,\\ \eta_{B,h}^2(v)&\coloneqq (-\overline{\lambda}_h^{cr},\Pi_h(v-\chi))_\Omega\,,\\ \eta_{C,h}^2&\coloneqq \tfrac{1}{d^2}\|h_{\mathcal{T}}(f_h-\overline{\lambda}_h^{cr})\|_\Omega^2\,.\end{aligned}\tag{5.1}$$

for reliability and efficiency.

Remark 5.1. *(i) The estimator η_h^2 appeared in a similar form in [15] in a Crouzeix–Raviart approximation of the obstacle problems, imposing the obstacle constraint at the barycenters of element sides. However, imposing the obstacle constraint at the barycenters of elements leads to a simplified form compared to the estimator in [15].*

(ii) The estimator $\eta_{A,h}^2$ provides control over the flux relation (3.6).

(iii) The estimator $\eta_{B,h}^2$ measures the discrepancy in the complementary condition (3.9), cf. [11].

(iv) The estimator $\eta_{C,h}^2$ measures the irregularity of the dual solution, i.e., $\operatorname{div} z\notin L^2(\Omega)$.

5.1 Reliability

In this subsection, we identify error quantities that are controlled by the a posteriori error estimator $\eta_h^2\colon W_D^{1,2}(\Omega)\to\mathbb{R}$, cf. (5.1). In doing so, we combine two different but related approaches: first, we resort to first-order relations based on (discrete) convex duality, leading to constant-free estimates; second, we resort to second-order relations based on the (discrete) augmented problems, leading to estimates for further error quantities that are not covered be the first approach.

5.1.1 Reliability based on (discrete) convex duality

In this subsection, we follow the procedure for the derivation of, by definition, reliable and constant-free a posteriori error estimates based on (discrete) convex duality outlined in the introduction.

Lemma 5.2. *For every $v\in K$, we have that*

$$\begin{aligned}\tfrac12\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega+\tfrac12\|z_h^{rt}-z\|_\Omega^2&\le \tfrac12\|\nabla v-z_h^{rt}\|_\Omega^2+\eta_{B,h}^2(v)+(f_h-f,v-\chi)_\Omega\\ &\le \eta_{A,h}^2(v)+\eta_{B,h}^2(v)+\eta_{C,h}^2+(f_h-f,v-\chi)_\Omega\,.\end{aligned}$$

Proof. Using that, owing to the augmented problem (3.8), $\frac12a^2-\frac12b^2=\frac12(a-b)^2+b(a-b)$ for all $a,b\in\mathbb{R}$, and the strong concavity of (3.3), for every $v\in K$ and $y\in L^2(\Omega;\mathbb{R}^d)$, it holds

$$\tfrac12\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega=I(v)-I(u)\,,\tag{5.2}$$

$$\tfrac12\|y-z\|_\Omega^2\le D(z)-D(y)\,,\tag{5.3}$$

the strong duality relation $I(u)=D(z)$, cf. (3.5), that $z_h^{rt}\in W_N^2(\operatorname{div};\Omega)$ with $\operatorname{div} z_h^{rt}+f_h=\overline{\lambda}_h^{cr}\le 0$ a.e. in $\mathcal{L}^0(\mathcal{T}_h)$, cf. (3.29), integration-by-parts, and $\Pi_h z_h^{rt}=\nabla_h u_h^{cr}$ in $\mathcal{L}^0(\mathcal{T}_h)^d$, cf. (3.28), we get

$$\begin{aligned}\tfrac12\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega+\tfrac12\|z_h^{rt}-z\|_\Omega^2&\le I(v)-D(z_h^{rt})\\ &=\tfrac12\|\nabla v\|_\Omega^2+(\operatorname{div} z_h^{rt}-\overline{\lambda}_h^{cr},v)_\Omega+(f_h-f,v)_\Omega\\ &\quad+\tfrac12\|z_h^{rt}\|_\Omega^2+(\operatorname{div} z_h^{rt}+f_h,\chi)_\Omega-(f_h-f,\chi)\\ &=\tfrac12\|\nabla v\|_\Omega^2-(z_h^{rt},\nabla v)_\Omega+\tfrac12\|z_h^{rt}\|_\Omega^2+\eta_{B,h}^2(v)+(f_h-f,v-\chi)_\Omega\\ &=\tfrac12\|\nabla v-z_h^{rt}\|_\Omega^2+\eta_{B,h}^2(v)+(f_h-f,v-\chi)_\Omega\\ &\le \eta_{A,h}^2(v)+\eta_{B,h}^2(v)+\|z_h^{rt}-\Pi_h z_h^{rt}\|_\Omega^2+(f_h-f,v-\chi)_\Omega\,.\end{aligned}$$

Due to $z_h^{rt}-\Pi_h z_h^{rt}=\frac{\overline{\lambda}_h^{cr}-f_h}{d}(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})$ in $\mathcal{L}^1(\mathcal{T}_h)^d$, cf. (3.27), we conclude the assertion. □

Remark 5.3. *(i) For every $v \in K$, due to (3.2), we have that*

$$\langle -\Lambda, v-u\rangle_\Omega = (\nabla u, \nabla v - \nabla u)_\Omega - (f, v-u)_\Omega \geq 0\,.$$

(ii) Since $\overline{\lambda}_h^{cr} \leq 0$ a.e. in Ω (cf. Proposition 3.1 (ii)) and $\Pi_h(v-\chi)\geq 0$ a.e. in Ω for all $v \in K$, for every $v\in K$, we have that $\eta^2_{B,h}(v)\geq 0$ and, thus, $\eta_h^2(v)\geq 0$, since, then,

$$(-\overline{\lambda}_h^{cr})\Pi_h(v-\chi)\geq 0 \qquad \text{a.e. in } \Omega\,.$$

Remark 5.4. *The reliability estimate in Lemma 5.2 is entirely constant-free.*

Remark 5.5 (Improved reliability)**.** *If we have that $v = v_h \in \mathcal{S}^1_D(\mathcal{T}_h)\cap K$ in Lemma 5.2, then, given $z_h^{rt} - \Pi_h z_h^{rt} \perp \nabla v_h - \nabla_h u_h^{cr}$ in $L^2(\Omega;\mathbb{R}^d)$, we arrive at the improved reliability estimate*

$$\tfrac{1}{2}\|\nabla v_h - \nabla u\|_\Omega^2 + \langle -\Lambda, v_h - u\rangle_\Omega + \tfrac{1}{2}\|z_h^{rt} - z\|^2_\Omega \leq \tfrac{1}{2}\eta^2_{A,h}(v_h) + \eta^2_{B,h}(v_h) + \tfrac{1}{2}\eta^2_{C,h} + (f_h - f, v-\chi)_\Omega\,.$$

Corollary 5.6. *If $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, then the following statements apply:*

(i) For every $v \in K$, it holds

$$\tfrac{1}{2}\|\nabla v - \nabla u\|_\Omega^2 + \langle -\Lambda, v - u\rangle_\Omega + \tfrac{1}{2}\|z_h^{rt} - z\|^2_\Omega \leq \tfrac{1}{2}\|\nabla v - z_h^{rt}\|^2_\Omega + \eta^2_{B,h}(v) \leq \eta_h^2(v)\,.$$

(ii) For every $v_h \in \mathcal{S}^1_D(\mathcal{T}_h)\cap K$, it holds

$$\tfrac{1}{2}\|\nabla v_h - \nabla u\|_\Omega^2 + \langle -\Lambda, v_h - u\rangle_\Omega + \tfrac{1}{2}\|z_h^{rt} - z\|^2_\Omega \leq \tfrac{1}{2}\eta^2_{A,h}(v_h) + \eta^2_{B,h}(v_h) + \tfrac{1}{2}\eta^2_{C,h}\,.$$

Proof. For the claim (i), we refer to Lemma 5.2. For the claim (ii), we refer to Remark 5.5. □

The a posteriori error estimator $\eta_h^2 \colon W^{1,2}_D(\Omega)\to\mathbb{R}$, cf. (5.1), furthermore, controls the error between the continuous Lagrange multiplier $\Lambda \in (W^{1,2}_D(\Omega))^*$, defined by (3.8), and the discrete Lagrange multiplier $\overline{\lambda}_h^{cr} \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$, defined by (3.14), measured in the Sobolev dual norm. To this end, we introduce the $(W^{1,2}_D(\Omega))^*$-representation $\overline{\Lambda}_h^{cr}\in (W^{1,2}_D(\Omega))^*$ of $\overline{\lambda}_h^{cr} \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$\langle \overline{\Lambda}_h^{cr}, v\rangle_\Omega \coloneqq (\overline{\lambda}_h^{cr}, \Pi_h v)_\Omega\,.$$

Lemma 5.7. *The following statements apply:*

(i) If we set $\mathrm{osc}_h(f) \coloneqq \|h_{\mathcal{T}}(f_h - f)\|^2_\Omega$, cf. Theorem A.1, then

$$\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,\Omega} \leq \|\nabla_h u_h^{cr} - \nabla u\|_\Omega + \eta_{C,h} + c_\Pi\,(\mathrm{osc}_h(f))^{\frac{1}{2}}\,. \tag{5.4}$$

(ii) If $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, then for every $v\in K$, it holds

$$\|\overline{\Lambda}_h^{cr} - \Lambda\|^2_{*,\Omega} \leq 9\,\eta^2_{A,h}(v) + 6\,\eta^2_{B,h}(v) + 9\,\eta^2_{C,h} \leq 9\,\eta^2_h(v)\,. \tag{5.5}$$

Proof. ad (i). For every $v\in W^{1,2}_D(\Omega)$ satisfying $\|v\|_\Omega + \|\nabla v\|_\Omega \leq 1$, using (3.8), (3.29), integration-by-parts, (3.27), $f - f_h \perp \Pi_h v$ in $L^2(\Omega)$, and (L0.2), it holds

$$\begin{aligned}\langle \overline{\Lambda}_h^{cr} - \Lambda, v\rangle_\Omega &= (f_h + \operatorname{div} z_h^{rt}, v)_\Omega - (f,v)_\Omega + (\nabla u, \nabla v)_\Omega\\ &= (\nabla u - z_h^{rt}, \nabla v)_\Omega + (f_h - f, v)_\Omega\\ &= (\nabla u - \nabla_h u_h^{cr}, \nabla v)_\Omega + (\tfrac{1}{d}(f_h - \overline{\lambda}_h^{cr})(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}), \nabla v)_\Omega\\ &\quad + (f_h - f, v - \Pi_h v)_\Omega\\ &\leq \|\nabla_h u_h^{cr} - \nabla u\|_\Omega + \eta_{C,h} + c_\Pi\,(\mathrm{osc}_h(f))^{\frac{1}{2}}\,.\end{aligned} \tag{5.6}$$

Taking the supremum with respect to every $v \in W^{1,2}_D(\Omega)$ satisfying $\|v\|_\Omega + \|\nabla v\|_\Omega \leq 1$ in (5.6), we conclude the assertion.

ad (ii). Due to (i) and $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, for every $v \in W^{1,2}_D(\Omega)$, it holds

$$\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,\Omega} \leq \|\nabla v - \nabla u\|_\Omega + \eta^2_{A,h}(v) + \eta^2_{C,h}\,. \tag{5.7}$$

Then, resorting in (5.7) to Lemma 5.2, for every $v\in K$, we find that

$$\begin{aligned}\|\overline{\Lambda}_h^{cr} - \Lambda\|^2_{*,\Omega} &\leq 3\,[\|\nabla v - \nabla u\|^2_\Omega + \eta^2_{A,h}(v) + \eta^2_{C,h}]\\ &\leq 3\,[3\,\eta^2_{A,h}(v) + 2\,\eta^2_{B,h}(v) + 3\,\eta^2_{C,h}]\,.\end{aligned}$$

□

Remark 5.8. *If $v = v_h \in \mathcal{S}^1_D(\mathcal{T}_h) \cap K$ in Lemma 5.2, then, given Remark 5.6 (i), we arrive at the improved reliability estimate*

$$\|\overline{\Lambda}_h^{cr} - \Lambda\|^2_{*,\Omega} \leq 6\,\eta_h^2(v)\,.$$

5.1.2 *Reliability based on variational equations*

Following an approach which resorts to the (discrete) augmented problems, i.e., (3.8) and (3.14), it is possible to establish the following reliability result, which identifies additional quantities that are controlled by the a posteriori error estimator $\eta_h^2\colon W^{1,2}_D(\Omega) \to \mathbb{R}$, cf. (5.1).

Lemma 5.9. *For every $v \in W^{1,2}_D(\Omega)$ and $\varepsilon, \tilde{\varepsilon} > 0$, we have that*

$$\begin{aligned}(\tfrac{1}{2} - \varepsilon\, c_{cr}^2 - \tilde{\varepsilon}\, c_\Pi^2)\|\nabla v - \nabla u\|^2_\Omega + \langle -\Lambda, v-u\rangle_\Omega + \tfrac{1}{2}\|\nabla u - \nabla_h u_h^{cr}\|^2_\Omega + (-\overline{\lambda}_h^{cr}, \Pi_h(u-\chi))_\Omega \\ \leq \tfrac{1}{2}\eta^2_{A,h}(v) + \eta^2_{B,h}(v) + \tfrac{d^2}{4\varepsilon}\,\eta^2_{C,h} + \tfrac{1}{4\tilde{\varepsilon}}\,\mathrm{osc}_h(f)\,.\end{aligned}$$

From Lemma 5.9 we can immediately deduce the following reliability results.

Corollary 5.10. *The following statements apply:*

(i) For every $v \in W^{1,2}_D(\Omega)$, it holds

$$\begin{aligned}\tfrac{1}{4}\|\nabla v - \nabla u\|^2_\Omega + \langle -\Lambda, v-u\rangle_\Omega + \tfrac{1}{2}\|\nabla u - \nabla_h u_h^{cr}\|^2_\Omega + (-\overline{\lambda}_h^{cr}, \Pi_h(u-\chi))_\Omega \\ \leq \tfrac{1}{2}\eta^2_{A,h}(v) + \eta^2_{B,h}(v) + 2\,c_{cr}^2\,d^2\,\eta^2_{C,h} + 2\,c_\Pi^2\,\mathrm{osc}_h(f)\,.\end{aligned}$$

(ii) For every $v \in W^{1,2}_D(\Omega)$, it holds

$$\begin{aligned}\langle -\Lambda, v-u\rangle_\Omega + \tfrac{1}{2}\|\nabla u - \nabla_h u_h^{cr}\|^2_\Omega + (-\overline{\lambda}_h^{cr}, \Pi_h(u-\chi))_\Omega \\ \leq \tfrac{1}{2}\eta^2_{A,h}(v) + \eta^2_{B,h}(v) + c_{cr}^2\,d^2\,\eta^2_{C,h} + c_\Pi^2\,\mathrm{osc}_h(f)\,.\end{aligned}$$

Proof. The claim (i) follows from Lemma 5.9 for $\varepsilon = \frac{1}{8c_{cr}^2} > 0$ and $\tilde{\varepsilon} = \frac{1}{8c_\Pi^2} > 0$. The claim (ii) follows from Lemma 5.9 for $\varepsilon = \frac{1}{4c_{cr}^2} > 0$ and $\tilde{\varepsilon} = \frac{1}{4c_\Pi^2} > 0$. ☐

Having Corollary 5.10 (ii) at hand, by analogy with Lemma 5.7 (ii), we arrive at the following reliability result for the error between the continuous and the discrete Lagrange multiplier measured in the dual norm.

Corollary 5.11. *For every $v \in K$, we have that*

$$\begin{aligned}\|\overline{\Lambda}_h^{cr} - \Lambda\|^2_{*,\Omega} &\leq 3\,\eta^2_{A,h}(v) + 6\,\eta^2_{B,h}(v) + 6(1 + c_{cr}^2\,d^2)\,\eta^2_{C,h} + 12\,c_\Pi^2\,\mathrm{osc}_h(f) \\ &\leq 6(1 + c_{cr}^2\,d^2)\,\eta_h^2(v) + 12\,c_\Pi^2\,\mathrm{osc}_h(f)\,.\end{aligned}$$

Proof. Appealing to Lemma 5.7 (5.4),

$$\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,\Omega} \leq \|\nabla u - \nabla_h u_h^{cr}\|_\Omega + \eta_{C,h} + c_\Pi\,(\mathrm{osc}_h(f))^{\frac{1}{2}}\,. \tag{5.8}$$

Then, resorting in (5.8) to Corollary 5.10 (ii), for every $v \in K$, we find that

$$\begin{aligned}\|\overline{\Lambda}_h^{cr} - \Lambda\|^2_{*,\Omega} &\leq 3\,[\|\nabla u - \nabla_h u_h^{cr}\|^2_\Omega + \eta^2_{C,h} + c_\Pi\,\mathrm{osc}_h(f)] \\ &\leq 3\,[\eta^2_{A,h}(v) + 2\,\eta^2_{B,h}(v) + 2\,(1 + c_{cr}^2\,d^2)\,\eta^2_{C,h} + 4\,c_\Pi^2\,\mathrm{osc}_h(f)]\,,\end{aligned}$$

which is the claimed reliability estimate. ☐

Proof (of Lemma 5.9). Resorting to (3.8), for every $v \in W^{1,2}_D(\Omega)$, since $f - f_h \perp \Pi_h(u-v)$ in $L^2(\Omega)$, we find that

$$\begin{aligned}\langle \Lambda, u-v\rangle_\Omega + (\nabla u, \nabla u - \nabla v)_\Omega &= (f, u-v)_\Omega \\ &= (f - f_h, u-v)_\Omega + (f_h, u-v)_\Omega \\ &= (f - f_h, u - v - \Pi_h(u-v))_\Omega + (f_h, u-v)_\Omega\,.\end{aligned} \tag{5.9}$$

Resorting to (3.14), for every $v\in W^{1,2}_D(\Omega)$, we find that

$$\begin{aligned}(\overline{\lambda}_h^{cr},\Pi_h I_{cr}(u-v))_\Omega+(\nabla_h u_h^{cr},\nabla I_{cr}(u-v))_\Omega&=(f_h,\Pi_h I_{cr}(u-v))_\Omega\\&=-(f_h-\overline{\lambda}_h^{cr},u-v-I_{cr}(u-v))_\Omega\\&\quad+(\overline{\lambda}_h^{cr},\Pi_h I_{cr}(u-v)-(u-v))_\Omega\\&\quad+(f_h,u-v)_\Omega\,,\end{aligned}\tag{5.10}$$

i.e., owing to $\nabla_h I_{cr}(u-v)=\Pi_h\nabla(u-v)$ in $\mathcal{L}^0(\mathcal{T}_h)^d$ and $\nabla_h u_h^{cr}\in\mathcal{L}^0(\mathcal{T}_h)^d$, for every $v\in W^{1,2}_D(\Omega)$

$$\begin{aligned}(\overline{\lambda}_h^{cr},u-v)_\Omega+(\nabla u_h^{cr},\nabla u-\nabla v)_\Omega&=-(f_h-\overline{\lambda}_h^{cr},u-v-I_{cr}(u-v))_\Omega\\&\quad+(f_h,u-v)_\Omega\,.\end{aligned}\tag{5.11}$$

If we subtract (5.11) from (5.9), we arrive at

$$\begin{aligned}\langle\Lambda-\overline{\Lambda}_h^{cr},u-v\rangle_\Omega+(\nabla u-\nabla_h u_h^{cr},\nabla u-\nabla v)_\Omega&=(f-f_h,u-v-\Pi_h(u-v))_\Omega\\&\quad+(f_h-\overline{\lambda}_h^{cr},u-v-I_{cr}(u-v))_\Omega\,.\end{aligned}\tag{5.12}$$

The binomial theorem shows that

$$(\nabla u-\nabla_h u_h^{cr},\nabla u-\nabla v)_\Omega+\tfrac{1}{2}\|\nabla v-\nabla_h u_h^{cr}\|_\Omega^2=\tfrac{1}{2}\|\nabla u-\nabla_h u_h^{cr}\|_\Omega^2+\tfrac{1}{2}\|\nabla v-\nabla u\|_\Omega^2\,.\tag{5.13}$$

Resorting to (CR.2), (L0.2) and the ε-Young inequality $ab\le\frac{1}{4\varepsilon}a^2+\varepsilon b^2$, valid for all $a,b\ge 0$ and $\varepsilon>0$, for every $\varepsilon,\tilde{\varepsilon}>0$, we find that

$$|(f-f_h,u-v-\Pi_h(u-v))_\Omega|\le\tfrac{1}{4\tilde{\varepsilon}}\operatorname{osc}_h(f)+\tilde{\varepsilon}\,c_\Pi^2\,\|\nabla v-\nabla u\|_\Omega^2\,,\tag{5.14}$$

$$|(f_h-\overline{\lambda}_h^{cr},u-v-I_{cr}(u-v))_\Omega|\le\tfrac{1}{4\varepsilon}\,\|h_\mathcal{T}(f_h-\overline{\lambda}_h^{cr})\|_\Omega^2+\varepsilon\,c_{cr}^2\,\|\nabla v-\nabla u\|_\Omega^2\,.\tag{5.15}$$

Therefore, combining (5.12)–(5.14), we conclude the claimed inequality. □

Given the findings of Lemma 5.2, Lemma 5.9, and Lemma 5.7, we introduce the error measure $\rho_h^2\colon W^{1,2}_D(\Omega)\to\mathbb{R}$, for every $v\in W^{1,2}_D(\Omega)$ defined by

$$\begin{aligned}\rho_h^2(v)&:=\tfrac{1}{2}\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega\\&\quad+\|\nabla u-\nabla_h u_h^{cr}\|_\Omega^2+(-\overline{\lambda}_h^{cr},\Pi_h(u-\chi))_\Omega+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2\,.\end{aligned}\tag{5.16}$$

Theorem 5.12 (Reliability)**.** *There exist constants $c_{rel},c_{osc}>0$, depending only on the chunkiness $\omega_0>0$, such that for every $v\in K$, we have that*

$$\rho_h^2(v)\le c_{rel}\,\eta_h^2(v)+c_{osc}\operatorname{osc}_h(f)\,.$$

Proof. Immediate consequence of Lemma 5.2, Lemma 5.9, and Lemma 5.7. □

Remark 5.13 (Comments on the reliability constant $c_{rel}>0$)**.**

(i) Appealing to Corollary 5.11, for every $v\in K$, we have that

$$\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2\le 6\,(1+c_{cr}^2\,d^2)\,\eta_h^2(v)+12\,c_\Pi^2\operatorname{osc}_h(f)\,.$$

If $f=f_h\in\mathcal{L}^0(\mathcal{T}_h)$, then Lemma 5.7 yields that for every $v\in K$, we have that

$$\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2\le\min\{9,6\,(1+c_{cr}^2\,d^2)\}\,\eta_h^2(v)\,.$$

(ii) Appealing to Corollary 5.10, for every $v\in K$, we have that

$$\begin{aligned}&\tfrac{1}{2}\|\nabla v-\nabla u\|_\Omega^2+2\langle-\Lambda,v-u\rangle_\Omega+\|\nabla u-\nabla_h u_h^{cr}\|_\Omega^2+2(-\overline{\lambda}_h^{cr},\Pi_h(u-\chi))_\Omega\\&\quad\le\max\{2,4\,c_{cr}^2\,d^2\}\,\eta_h^2(v)+2\,c_\Pi^2\operatorname{osc}_h(f)\,.\end{aligned}$$

If $f=f_h\in\mathcal{L}^0(\mathcal{T}_h)$, then Lemma 5.2 yields that for every $v\in K$, we have that

$$\tfrac{1}{2}\|\nabla v-\nabla u\|_\Omega^2+\langle-\Lambda,v-u\rangle_\Omega+\tfrac{1}{2}\|z_h^{rt}-z\|_\Omega^2\le\eta_h^2(v)\,.$$

(iii) Combining (i) and (ii), we find that

$$c_{rel}\le\max\{2,4\,c_{cr}^2\,d^2\}+6(1+c_{cr}^2\,d^2)\,,\quad c_{osc}\le 14\,c_\Pi^2\,.$$

5.2 Efficiency

In this subsection, we show the efficiency of the a posteriori error estimator $\eta_h^2\colon W_D^{1,2}(\Omega)\to\mathbb{R}$, cf. (5.1), with respect to the error measure $\rho_h^2\colon W_D^{1,2}(\Omega)\to\mathbb{R}$, cf. (5.16).

Theorem 5.14 (Efficiency)**.** *There exist constants $c_{\mathit{eff}}, c_{osc}>0$, depending on the chunkiness $\omega_0>0$, such that for every $v\in W_D^{1,2}(\Omega)$, we have that*

$$\eta_h^2(v)\le c_{\mathit{eff}}\,\rho_h^2(v)+c_{osc}\,\mathrm{osc}_h(f)\,.$$

Proof. Apparently, for every $v\in W_D^{1,2}(\Omega)$, we have that

$$\eta_{A,h}^2(v)\le 2\,[\|\nabla v-\nabla u\|_\Omega^2+\|\nabla u-\nabla_h u_h^{cr}\|_\Omega^2]\,,$$

In addition, appealing to Lemma A.3 (A.7), there exists a constant $c>0$, depending only on the chunkiness $\omega_0>0$, such that

$$\eta_{C,h}^2\le c\,[\|\nabla_h u_h^{cr}-\nabla u\|_\Omega^2+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+\mathrm{osc}_h(f)]\,.$$

Eventually, for every $v\in W_D^{1,2}(\Omega)$, using Young's and Poincaré's inequality, we find that

$$\begin{aligned}\eta_{B,h}^2(v)&=(-\overline{\lambda}_h^{cr},\Pi_h(u-\chi))_\Omega+\langle-\Lambda,v-u\rangle_\Omega+\langle\Lambda-\overline{\Lambda}_h^{cr},v-u\rangle_\Omega\\&\le(-\overline{\lambda}_h^{cr},\Pi_h(u-\chi))_\Omega+\langle-\Lambda,v-u\rangle_\Omega\\&\quad+\tfrac{1}{2}\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+\tfrac{1+c_P^2}{2}\|\nabla v-\nabla u\|_\Omega^2\,,\end{aligned}$$

where $c_P>0$ denotes the Poincaré constant. ☐

Remark 5.15. *Since the discrete primal solution $u_h^{cr}\in K_h^{cr}$ is neither an admissible approximation of the primal solution $u\in K$ in Theorem 5.12 nor in Theorem 5.14, since, in general,*

$$u_h^{cr}\notin W_D^{1,2}(\Omega)\quad\text{ and }\quad u_h^{cr}\ngeq\chi\text{ a.e. in }\Omega\,,$$

it is necessary to post-process $u_h^{cr}\in K_h^{cr}$. In the numerical experiments, cf. Section 6, we employ the post-processed function $v_h=\max\{I_h^{av}u_h^{cr},\chi\}$, where $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$ is a node-averaging quasi-interpolation operator, cf. Subsection A.1, which, by the Sobolev chain rule, satisfies

$$v_h\in W_D^{1,2}(\Omega)\quad\text{ and }\quad v_h\ge\chi\text{ a.e. in }\Omega\,,$$

i.e., $v_h\in K$. Note that $\mathrm{tr}\,v_h=0$ in Γ_D due to $I_h^{av}u_h^{cr}=0$ on Γ_D and $\chi\le 0$ in Γ_D. In addition, using the best-approximation property of $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$, cf. Proposition A.4, we have that

$$\begin{aligned}\|\nabla v_h-\nabla_h u_h^{cr}\|_\Omega&\le\|\nabla I_h^{av}u_h^{cr}-\nabla u_h^{cr}\|_\Omega+\|\nabla I_h^{av}u_h^{cr}-\nabla\chi\|_{\{I_h^{av}u_h^{cr}<\chi\}}\\&\le 2\,\|\nabla I_h^{av}u_h^{cr}-\nabla u_h^{cr}\|_\Omega+\|\nabla u_h^{cr}-\nabla\chi\|_{\{I_h^{av}u_h^{cr}<\chi\}}\\&\le c\,\|\nabla u_h^{cr}-\nabla u\|_\Omega+\|\nabla u-\nabla\chi\|_{\{I_h^{av}u_h^{cr}<\chi\}}\,,\end{aligned}$$

and, thus,

$$\|\nabla v_h-\nabla u\|_\Omega\le c\,\|\nabla u_h^{cr}-\nabla u\|_\Omega+\|\nabla u-\nabla\chi\|_{\{I_h^{av}u_h^{cr}<\chi\}}\,.$$

In other words, the error between $v_h\in K$ and $u_h^{cr}\in K_h^{cr}$ (and $u\in K$, respectively) is controlled by the error between $u_h^{cr}\in K_h^{cr}$ and $u\in K$ plus a contribution capturing the violation of the continuous obstacle constraint by $u_h^{cr}\in K_h^{cr}$.

Remark 5.16. *Appealing to Theorem A.1 and Lemma A.3 (A.7), there exists a constant $c>0$, depending only on the chunkiness $\omega_0>0$, such that for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we have that*

$$\|\nabla_h v_h-\nabla_h u_h^{cr}\|_\Omega^2+\|h_{\mathcal{T}}(f_h-\overline{\lambda}_h^{cr})\|_\Omega^2\le c\,[\|\nabla_h v_h-\nabla u\|_\Omega^2+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+\mathrm{osc}_h(f)]\,.$$

Thus, it is possible to establish efficiency estimates for parts of the primal-dual a posteriori error estimator, which also apply for non-conforming functions, cf. Appendix A.

6. Numerical experiments

In this section, we confirm the theoretical findings of Section 4 and Section 5 via numerical experiments. All experiments were conducted using the finite element software package FEniCS (version 2019.1.0), cf. [39]. All graphics are generated using the Matplotlib (version 3.5.1) library, cf. [33].

6.1 Implementation details

We approximate the discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and the associated discrete Lagrange multiplier $\overline{\lambda}_h^{cr} \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ jointly satisfying the discrete augmented problem (3.14) via the primal-dual active set strategy interpreted as a super-linear converging semi-smooth Newton method, cf. [3, Subsection 5.3.1] or [31]. For sake of completeness, we will briefly outline important implementation details related with this strategy.

We fix an ordering of the element sides $(S_i)_{i=1,\dots,N_h^{cr}}$ and an ordering of the elements $(T_i)_{i=1,\dots,N_h^0}$, where $N_h^{cr} \coloneqq \mathrm{card}(\mathcal{S}_h)$ and $N_h^0 \coloneqq \mathrm{card}(\mathcal{T}_h)$, such that[5]

$$\mathrm{span}(\{\chi_{T_i} \mid i = 1, \dots, N_h^{cr,0}\}) = \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))\,,$$

where $N_h^{cr,0} = \dim(\Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))) \in \{N_h^0, N_h^0 - 1\}$ because of $\mathrm{codim}_{\mathcal{L}^0(\mathcal{T}_h)}(\Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))) = 1$, cf. [8, Corollary 3.2]. Then, if we define the matrices

$$\begin{aligned}
\mathrm{S}_h^{cr} &\coloneqq ((\nabla_h\varphi_{S_i}, \nabla_h\varphi_{S_j})_\Omega)_{i,j=1,\dots,N_h^{cr}} \in \mathbb{R}^{N_h^{cr}\times N_h^{cr}}\,,\\
\mathrm{P}_h^{cr,0} &\coloneqq ((\Pi_h\varphi_{S_i}, \chi_{T_j})_\Omega)_{i=1,\dots,N_h^{cr},j=1,\dots,N_h^{cr,0}} \in \mathbb{R}^{N_h^{cr}\times N_h^{cr,0}}\,,
\end{aligned}$$

and, assuming for the entire section that $\chi_h \coloneqq \Pi_h I_{cr}\chi \in \mathcal{L}^0(\mathcal{T}_h)$, the vectors

$$\begin{aligned}
\mathrm{X}_h^{cr} &\coloneqq ((I_{cr}\chi, \varphi_{S_i})_\Omega)_{i=1,\dots,N_h^{cr}} \in \mathbb{R}^{N_h^{cr}}\,,\\
\mathrm{F}_h^{0} &\coloneqq ((f_h, \chi_{T_i})_\Omega)_{i=1,\dots,N_h^{cr,0}} \in \mathbb{R}^{N_h^{cr,0}}\,,
\end{aligned}$$

the same argumentation as in [3, Lemma 5.3] shows that the discrete augmented problem (3.14) is equivalent to finding vectors $(\mathrm{U}_h^{cr}, \mathrm{L}_h^{cr})^\top \in \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}$ such that

$$\begin{aligned}
\mathrm{S}_h^{cr}\mathrm{U}_h^{cr} + \mathrm{P}_h^{cr,0}\mathrm{L}_h^{cr} &= \mathrm{P}_h^{cr,0}\mathrm{F}_h^0 &&\text{in } \mathbb{R}^{N_h^{cr}}\,,\\
\mathscr{C}_h(\mathrm{U}_h^{cr}, \mathrm{L}_h^{cr}) &= 0_{\mathbb{R}^{N_h^{cr,0}}} &&\text{in } \mathbb{R}^{N_h^{cr,0}}\,,
\end{aligned} \tag{6.1}$$

where for given $\alpha > 0$, the mapping $\mathscr{C}_h \colon \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}} \to \mathbb{R}^{N_h^{cr,0}}$ for every $(\mathrm{U}_h, \mathrm{L}_h)^\top \in \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}$ is defined by

$$\mathscr{C}_h(\mathrm{U}_h, \mathrm{L}_h) \coloneqq \mathrm{L}_h - \min\left\{0, \mathrm{L}_h + \alpha(\mathrm{P}_h^{cr,0})^\top(\mathrm{U}_h - \mathrm{X}_h^{cr})\right\} \quad \text{in } \mathbb{R}^{N_h^{cr,0}}\,.$$

More precisely, the discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and the associated discrete Lagrange multiplier $\overline{\lambda}_h^{cr} \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ jointly satisfying the discrete augmented problem (3.14) as well as $(\mathrm{U}_h^{cr}, \mathrm{L}_h^{cr})^\top \in \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}$, respectively, are related by[6]

$$\begin{aligned}
u_h^{cr} &= \sum_{i=1}^{N_h^{cr}} (\mathrm{U}_h^{cr}\cdot \mathrm{e}_i)\varphi_{S_i} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)\,,\\
\overline{\lambda}_h^{cr} &= \sum_{i=1}^{N_h^{cr,0}} (\mathrm{L}_h^{cr}\cdot \mathrm{e}_i)\chi_{T_i} \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))\,.
\end{aligned}$$

Next, define the mapping $\mathscr{F}_h \colon \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}} \to \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}$ for every $(\mathrm{U}_h, \mathrm{L}_h)^\top \in \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}$ by

$$\mathscr{F}_h(\mathrm{U}_h, \mathrm{L}_h) \coloneqq \begin{bmatrix} \mathrm{S}_h^{cr}\mathrm{U}_h + \mathrm{P}_h^{cr,0}(\mathrm{L}_h - \mathrm{F}_h^0) \\ \mathscr{C}_h(\mathrm{U}_h, \mathrm{L}_h) \end{bmatrix} \quad \text{in } \mathbb{R}^{N_h^{cr}} \times \mathbb{R}^{N_h^{cr,0}}\,.$$

---

[5] In practice, the element $\widehat{T} \in \mathcal{T}_h$ for which $\mathbb{R}\chi_{\widehat{T}} \perp \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$ is found via searching and erasing a zero column (if existent) in the matrix $((\Pi_h\varphi_{S_i}, \chi_T)_\Omega)_{i=1,\dots,N_h^{cr}, T\in\mathcal{T}_h} \in \mathbb{R}^{N_h^{cr}\times N_h^0}$ leading to $\mathrm{P}_h^{cr,0} \in \mathbb{R}^{N_h^{cr}\times N_h^{cr,0}}$.

[6] Here, for every $i = 1, \dots, N$, $N \in \{N_h^{cr}, N_h^{cr,0}\}$, we denote by $\mathrm{e}_i = (\delta_{ij})_{j=1,\dots,N} \in \mathbb{R}^N$, the i-th unit vector.

Then, the non-linear system (6.1) is equivalent to finding $(\mathrm{U}_h^{cr},\mathrm{L}_h^{cr})^\top \in \mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ such that

$$\mathscr{F}_h(\mathrm{U}_h^{cr},\mathrm{L}_h^{cr}) = 0_{\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}} \quad \text{ in } \mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}\,.$$

By analogy with [4, Theorem 5.11], one finds that the mapping $\mathscr{F}_h\colon \mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}\to\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ is Newton-differentiable at every $(\mathrm{U}_h,\mathrm{L}_h)^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ and, with the *(active) set*

$$\mathscr{A}_h \coloneqq \mathscr{A}_h(\mathrm{U}_h,\mathrm{L}_h)\coloneqq \big\{i\in\{1,\dots,N_h^{cr,0}\}\mid (\mathrm{L}_h+\alpha(\mathrm{P}_h^{cr,0})^\top(\mathrm{U}_h-\mathrm{X}_h^{cr}))\cdot \mathrm{e}_i<0\big\}\,,$$

we have that

$$D\mathscr{F}_h(\mathrm{U}_h,\mathrm{L}_h)\coloneqq \begin{bmatrix} \mathrm{S}_h^{cr} & \mathrm{P}_h^{cr,0}\\ (\mathrm{P}_h^{cr,0})^\top I_{\mathscr{A}_h} & I_{\mathscr{A}_h^c}\end{bmatrix} \quad \text{ in } \mathbb{R}^{N_h^{cr}+N_h^{cr,0}}\times\mathbb{R}^{N_h^{cr}+N_h^{cr,0}}\,,$$

where $I_{\mathscr{A}_h}, I_{\mathscr{A}_h^c}\coloneqq I_{N_h^{cr,0}\times N_h^{cr,0}}-I_{\mathscr{A}_h}\in\mathbb{R}^{N_h^{cr,0}}\times\mathbb{R}^{N_h^{cr,0}}$ for every $i,j\in\{1,\dots,N_h^{cr,0}\}$ are defined by $(I_{\mathscr{A}_h})_{ij}=1$ if $i=j\in\mathscr{A}_h$ and $(I_{\mathscr{A}_h})_{ij}=0$ else.

For a given iterate $(\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1})^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$, one step of the semi-smooth Newton method, cf. [3, Subsection 5.3.1] or [31], determines a direction $(\delta\mathrm{U}_h^{k-1},\delta\mathrm{L}_h^{k-1})^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ such that

$$D\mathscr{F}_h(\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1})(\delta\mathrm{U}_h^{k-1},\delta\mathrm{L}_h^{k-1})^\top = -\mathscr{F}_h(\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1}) \quad \text{ in } \mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}\,. \tag{6.2}$$

Setting $(\mathrm{U}_h^{k},\mathrm{L}_h^{k})^\top\coloneqq(\mathrm{U}_h^{k-1}+\delta\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1}+\delta\mathrm{L}_h^{k-1})^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ and $\mathscr{A}_h^{k-1}\coloneqq\mathscr{A}_h(\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1})$, the linear system (6.2) can equivalently be re-written as

$$\begin{aligned} \mathrm{S}_h^{cr}\mathrm{U}_h^k+\mathrm{P}_h^{cr,0}\mathrm{L}_h^k &= \mathrm{P}_h^{cr,0}\mathrm{F}_h^0 && \text{ in } \mathbb{R}^{N_h^{cr}}\,,\\ I_{(\mathscr{A}_h^{k-1})^c}\mathrm{L}_h^k &= 0_{\mathbb{R}^{N_h^{cr,0}}} && \text{ in } \mathbb{R}^{N_h^{cr,0}}\,,\\ I_{\mathscr{A}_h^{k-1}}\mathrm{U}_h^k &= I_{\mathscr{A}_h^{k-1}}\mathrm{X}_h^{cr} && \text{ in } \mathbb{R}^{N_h^{cr,0}}\,. \end{aligned} \tag{6.3}$$

The semi-smooth Newton method can, thus, equivalently be formulated in the following form, which is a version of a primal-dual active set strategy.

Algorithm 6.1 (Primal-dual active set strategy)**.** *Choose parameters $\alpha>0$ and $\varepsilon_{\mathrm{STOP}}>0$. Moreover, let $(\mathrm{U}_h^0,\mathrm{L}_h^0)^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ and set $k=1$. Then, for every $k\in\mathbb{N}$:*

(i) Define the most recent active set

$$\mathscr{A}_h^{k-1}\coloneqq\mathscr{A}_h(\mathrm{U}_h^{k-1},\mathrm{L}_h^{k-1})\coloneqq\big\{i\in\{1,\dots,N_h^{cr,0}\}\mid(\mathrm{L}_h^{k-1}+\alpha(\mathrm{P}_h^{cr,0})^\top(\mathrm{U}_h^{k-1}-\mathrm{X}_h^{cr}))\cdot\mathrm{e}_i<0\big\}\,.$$

(ii) Compute the iterate $(\mathrm{U}_h^k,\mathrm{L}_h^k)^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ such that

$$\begin{bmatrix} \mathrm{S}_h^{cr} & \mathrm{P}_h^{cr,0}\\ (\mathrm{P}_h^{cr,0})^\top I_{\mathscr{A}_h^{k-1}} & I_{(\mathscr{A}_h^{k-1})^c}\end{bmatrix}\begin{bmatrix}\mathrm{U}_h^k\\ \mathrm{L}_h^k\end{bmatrix}=\begin{bmatrix}\mathrm{P}_h^{cr,0}\mathrm{F}_h^0\\ I_{\mathscr{A}_h^{k-1}}\mathrm{X}_h^{cr}\end{bmatrix}.$$

(iii) Stop if $|\mathrm{U}_h^k-\mathrm{U}_h^{k-1}|\le\varepsilon_{\mathrm{STOP}}$; otherwise, increase $k\to k+1$ and continue with step (i).

Remark 6.2 (Important implementation details)**.** *(i) Algorithm 6.1 converges super-linearly if $(U_h^0,L_h^0)^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$ is sufficiently close to the solution $(U_h^{cr},L_h^{cr})^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$, cf. [31, Theorem 3.1]. Since the Newton-differentiability only holds in finite-dimensional situations and deteriorates as the dimension increases, the condition on the initial guess becomes more critical for increasing dimensions.*

(ii) The degrees of freedom related to the entries $L_h^k|_{(\mathscr{A}_h^{k-1})^c}$ can be eliminated from the linear system of equations in Algorithm 6.1, step (ii) (see also (6.3)).

(iii) Since only a finite number of active sets are possible, the algorithm terminates within a finite number of iterations at the exact solution $(\mathrm{U}_h^{cr},\mathrm{L}_h^{cr})^\top\in\mathbb{R}^{N_h^{cr}}\times\mathbb{R}^{N_h^{cr,0}}$. For this reason, in practice, the stopping criterion in step (iii) is reached with $|\mathrm{U}_h^{k^*}-\mathrm{U}_h^{k^*-1}|=0$ for some $k^*\in\mathbb{N}$, in which case, one has that $\mathrm{U}_h^{k^*}=\mathrm{U}_h^{cr}$, provided $\varepsilon_{\mathrm{STOP}}>0$ is sufficiently small.

(iv) The linear system emerging in each semi-smooth Newton step (cf. Algorithm 6.1, step (ii)) is solved using a sparse direct solver from SciPy *(version 1.8.1), cf. [46].*

(v) Global convergence of the algorithm and monotonicity, i.e., $\mathrm{U}_h^k\ge\mathrm{U}_h^{k-1}\ge\mathrm{X}_h^{cr}$ for $k\ge 3$ can be proved if $\mathrm{S}_h^{cr}\in\mathbb{R}^{N_h^{cr}\times N_h^{cr}}$ is an M-Matrix, cf. [31, Theorem 3.2].

(vi) Classical active set strategies define $\mathscr{A}_h^{k-1}\coloneqq\{i\in\{1,\dots,N_h^{cr,0}\}\mid\mathrm{L}_h^{k-1}\cdot\mathrm{e}_i<0\}$, which corresponds to the formal limit $\alpha\to\infty$.

6.2 Numerical experiments for a priori error analysis

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

For our numerical experiments, we choose $\Omega := (-\frac{3}{2}, \frac{3}{2})^2$, $\Gamma_D := \partial\Omega$, $f := -2 \in L^2(\Omega)$, $\chi := 0 \in W^{1,2}(\Omega)$, and as a manufactured solution, the function $u \in W^{1,2}(\Omega)$, for every $x \in \Omega$ defined by

$$u(x) := \begin{cases} \frac{|x|^2}{2} - \ln(|x|) - \frac{1}{2} & \text{if } x \in \Omega \setminus B_1^d(0) \\ 0 & \text{else} \end{cases} ,$$

As a result, appealing to Theorem 4.1, we can expect the convergence rate 1.

We construct an initial triangulation $\mathcal{T}_{h_0}$, where $h_0 = \frac{1}{\sqrt{2}}$, by subdividing a rectangular Cartesian grid into regular triangles with different orientations. Finer triangulations $\mathcal{T}_{h_k}$, $k = 1, \dots, 7$, where $h_{k+1} = \frac{h_k}{2}$ for all $k = 1, \dots, 7$, are obtained by regular subdivision of the previous grid: each triangle is subdivided into four equal triangles by connecting the midpoints of the edges, i.e., applying the red-refinement rule, cf. [45].

For the resulting series of triangulations $\mathcal{T}_k := \mathcal{T}_{h_k}$, $k = 1, \dots, 7$, we apply the primal-dual active set strategy (cf. Algorithm 6.1) to compute the discrete primal solution $u_k^{cr} := u_{h_k}^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_k)$, $k = 1, \dots, 7$, the discrete Lagrange multiplier $\overline{\lambda}_k^{cr} := \overline{\lambda}_{h_k}^{cr} \in \Pi_{h_k}(\mathcal{S}_D^{1,cr}(\mathcal{T}_k))$, and, subsequently, resorting to (3.27), the discrete dual solution $z_k^{rt} := z_{h_k}^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_k)$, $k = 1, \dots, 7$. Afterwards, we compute the error quantities

$$\left.\begin{aligned} e_{u,k} &:= \|\nabla_{h_k} u_k^{cr} - \nabla u\|_\Omega \,, \\ e_{I_{cr}u,k} &:= \|\nabla_{h_k} u_k^{cr} - \nabla_{h_k} I_{cr} u\|_\Omega \,, \\ e_{z,k} &:= \|z_k^{rt} - z\|_\Omega \,, \\ e_{I_{rt}z,k} &:= \|\Pi_{h_k} z_k^{rt} - \Pi_{h_k} I_{rt} z\|_\Omega \,, \\ e_{\lambda,u,k} &:= (-\overline{\lambda}_k^{cr}, \Pi_{h_k}(u - u_k^{cr}))_\Omega \,, \\ e_{\lambda,I_{cr}u,k} &:= (-\overline{\lambda}_k^{cr}, \Pi_{h_k}(I_{cr}u - u_k^{cr}))_\Omega \,, \end{aligned}\right\} \quad k = 1, \dots, 7 \,. \tag{6.4}$$

As estimation of the convergence rates, the experimental order of convergence (EOC)

$$\texttt{EOC}_k(e_k) := \frac{\log(e_k/e_{k-1})}{\log(h_k/h_{k-1})} \,, \quad k = 2, \dots, 7 \,,$$

where for every $k = 1, \dots, 7$, we denote by e_k, either e_u^k, $e_{I_{cr}u}^k$, e_z^k, $e_{I_{rt}z}^k$, $e_{\lambda,u}^k$, $e_{\lambda,I_{cr}u}^k$, $e_{\lambda,u}^{tot,k} := e_{\lambda,u}^k + e_u^k$, or $e_{\lambda,I_{cr}u}^{tot,k} := e_{\lambda,I_{cr}u}^k + e_{I_{cr}u}^k$, respectively, is recorded.

For a series of triangulations $\mathcal{T}_k$, $k = 1, \dots, 7$, obtained by uniform mesh refinement as described above, the EOC is computed and presented in Table 1 and Table 2. In each case, except for $e_k \in \{e_{\lambda,u,k}, e_{\lambda,I_{cr}u,k}\}$, we report a convergence ratio of about $\texttt{EOC}_k(e_k) \approx 1$, $k = 2, \dots, 7$, confirming the optimality of the a priori error estimates established in Theorem 4.1 and Corollary 4.2. For $e_k \in \{e_{\lambda,u,k}, e_{\lambda,I_{cr}u,k}\}$, we report a convergence ratio of about $\texttt{EOC}_k(e_k) \approx 1.5$, $k = 2, \dots, 7$.

k	e_u^k	$\texttt{EOC}_k$	$e_{I_{cr}u}^k$	$\texttt{EOC}_k$	e_z^k	$\texttt{EOC}_k$	$e_{I_{rt}z}^k$	$\texttt{EOC}_k$
1	1.359	—	0.732	—	1.094	—	0.656	—
2	0.787	0.788	0.664	0.141	0.533	1.038	0.453	0.535
3	0.380	1.048	0.324	1.034	0.260	1.033	0.212	1.097
4	0.197	0.948	0.166	0.968	0.131	0.993	0.116	0.872
5	0.099	0.996	0.082	1.008	0.067	0.974	0.059	0.967
6	0.050	0.989	0.042	0.986	0.033	1.010	0.030	0.968
7	0.025	0.998	0.021	1.001	0.017	0.980	0.015	0.993
expected	—	1.00	—	1.00	—	1.00	—	1.00

Table 1: For $e_k \in \{e_u^k, e_{I_{cr}u}^k, e_z^k, e_{I_{rt}z}^k\}$, $k = 2, \dots, 7$: error e_k and experimental order of convergence $\texttt{EOC}_k(e_k)$.

k	$e^k_{\lambda,u}$	$\texttt{EOC}_k$	$e^k_{\lambda,I_{cr}u}$	$\texttt{EOC}_k$	$e^{tot,k}_{\lambda,u}$	$\texttt{EOC}_k$	$e^{tot,k}_{\lambda,I_{cr}u}$	$\texttt{EOC}_k$
1	0.262	—	0.490	—	1.849	—	1.223	—
2	0.144	0.866	0.199	1.300	0.986	0.907	0.863	0.502
3	0.044	1.706	0.072	1.461	0.453	1.123	0.397	1.122
4	0.020	1.133	0.029	1.308	0.226	0.999	0.195	1.024
5	0.006	1.732	0.009	1.636	0.108	1.064	0.092	1.086
6	0.002	1.363	0.003	1.447	0.053	1.024	0.045	1.027
7	0.001	1.618	0.001	1.535	0.026	1.027	0.022	1.036
expected	—	1.00	—	1.00	—	1.00	—	1.00

Table 2: For $e_k \in \{e^k_{\lambda,u}, e^k_{\lambda,I_{cr}u}, e^{tot,k}_{\lambda,u}, e^{tot,k}_{\lambda,I_{cr}u}\}$, $k = 2,\dots,7$: error e_k and experimental order of convergence $\texttt{EOC}_k(e_k)$.

6.3 Numerical experiments for a posteriori error analysis

In this subsection, we confirm the theoretical findings of Section 5. More precisely, we apply the $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$-approximation (3.11) of the obstacle problem (3.1) in an adaptive mesh refinement algorithm based on local refinement indicators $(\eta^2_{h,T})_{T\in\mathcal{T}_h}$ associated with the a posteriori error estimator η^2_h, cf. (5.1). More precisely, for every $v \in W^{1,2}_D(\Omega)$ and $T \in \mathcal{T}_h$, we define

$$\begin{aligned} \eta^2_{A,h,T}(v) &\coloneqq \|\nabla v - \nabla_h u^{cr}_h\|^2_T\,,\\ \eta^2_{B,h,T}(v) &\coloneqq (-\overline{\lambda}^{cr}_h, \Pi_h(v-\chi))_T\,,\\ \eta^2_{C,h,T} &\coloneqq \tfrac{1}{d^2}\|h_{\mathcal{T}}(f_h - \overline{\lambda}^{cr}_h)\|^2_T\,,\\ \eta^2_{h,T}(v) &\coloneqq \eta^2_{A,h,T}(v) + \eta^2_{B,h,T}(v) + \eta^2_{C,h,T}\,. \end{aligned}$$

Before we present numerical experiments, we briefly outline the details of the implementations. In general, we follow the adaptive algorithm, cf. [21]:

Algorithm 6.3 (AFEM)**.** *Let $\varepsilon_{\rm STOP} > 0$, $\theta \in (0,1)$ and $\mathcal{T}_0$ a conforming initial triangulation of Ω. Then, for every $k \in \mathbb{N}\cup\{0\}$:*

('Solve') *Compute the discrete primal solution $u^{cr}_k \coloneqq u^{cr}_{h_k} \in K^{cr}_k \coloneqq K^{cr}_{h_k}$ and the discrete Lagrange multiplier $\overline{\lambda}^{cr}_k \coloneqq \overline{\lambda}^{cr}_{h_k} \in \Pi_{h_k}(\mathcal{S}^{1,cr}_D(\mathcal{T}_k))$ jointly solving the discrete augmented problem (3.14). Post-process $u^{cr}_k \in \mathcal{S}^{1,cr}_D(\mathcal{T}_k)$ to obtain a conforming approximation $v_k \in K$ of the primal solution $u \in K$ and a discrete dual solution $z^{rt}_k \coloneqq z^{rt}_{h_k} \in \mathcal{RT}^0_N(\mathcal{T}_k)$.*

('Estimate') *Compute the local refinement indicators $(\eta^2_{k,T}(v_k))_{T\in\mathcal{T}_k} \coloneqq (\eta^2_{h_k,T}(v_k))_{T\in\mathcal{T}_k}$. If $\eta^2_k(v_k) \coloneqq \eta^2_{h_k}(v_k) \le \varepsilon_{\rm STOP}$, then* STOP*; otherwise, continue with step ('Mark').*

('Mark') *Choose a minimal (in terms of cardinality) subset $\mathcal{M}_k \subseteq \mathcal{T}_k$ such that*

$$\sum_{T\in\mathcal{M}_k} \eta^2_{k,T}(v_k) \ge \theta^2 \sum_{T\in\mathcal{T}_k} \eta^2_{k,T}(v_k)\,.$$

('Refine') *Perform a conforming refinement of $\mathcal{T}_k$ to obtain $\mathcal{T}_{k+1}$ such that each $T \in \mathcal{M}_k$ is refined in $\mathcal{T}_{k+1}$. Increase $k \mapsto k+1$ and continue with step ('Solve').*

Remark 6.4. *(i) The discrete primal solution $u^{cr}_k \in K^{cr}_k$ and the discrete Lagrange multiplier $\overline{\lambda}^{cr}_k \in \Pi_{h_k}(\mathcal{S}^{1,cr}_D(\mathcal{T}_k))$ in step ('Solve') are computed using the primal-dual active set strategy (cf. Algorithm 6.1) for the parameter $\alpha = 1$.*

(ii) The reconstruction of the discrete dual solution $z^{rt}_k \in \mathcal{RT}^0_N(\mathcal{T}_k)$ in step ('Solve') is based on the generalized Marini formula (3.27) and does not entail further computational costs.

(iii) In accordance with Remark 5.15, as a conforming approximation $v_k \in K$ of the primal solution $u \in K$ in step ('Solve'), we employ $v_k = \max\{I^{av}_{h_k} u^{cr}_k, \chi\} \in K$.

(iv) If not otherwise specified, we employ the parameter $\theta = \frac{1}{2}$ in ('Mark').

(v) To find the set $\mathcal{M}_k \subseteq \mathcal{T}_k$ in step ('Mark'), we deploy the Dörfler marking strategy, cf. [23].

(vi) The (minimal) conforming refinement of $\mathcal{T}_k$ with respect to $\mathcal{M}_k$ in step ('Refine') is obtained by deploying the red-green-blue-refinement algorithm, cf. [45].

6.3.1 *Example with corner singularity*

We examine an example from [9]. In this example, we let $\Omega := (-2,2)^2 \setminus ([0,2]\times[-2,0])$, $\Gamma_D := \partial\Omega$, $\Gamma_N := \emptyset$, $f \in L^2(\Omega)$, in polar coordinates, for every $(r,\varphi)^\top \in \mathbb{R}_{>0}\times(0,2\pi)$ defined by

$$f(r,\varphi) := -r^{\frac{2}{3}}\sin(\tfrac{2\varphi}{3})(\tfrac{\gamma_1'(r)}{r} + \gamma_1''(r)) - \tfrac{4}{3}r^{-\frac{1}{3}}\gamma_1'(r)\sin(\tfrac{2\varphi}{3}) - \gamma_2(r)\,,$$

where $\gamma_1,\gamma_2\colon \mathbb{R}_{>0}\to\mathbb{R}$ for every $r\in\mathbb{R}_{>0}$, abbreviating $\overline{r} := 2(r-\frac{1}{4})$, are defined by

$$\gamma_1(r) := \begin{cases} 1 & \text{if } \overline{r} < 0 \\ -6\overline{r}^5 + 15\overline{r}^4 - 10\overline{r}^3 + 1 & \text{if } 0 \le \overline{r} < 1 \\ 0 & \text{if } \overline{r} \ge 1 \end{cases}\,, \qquad \gamma_2(r) := \begin{cases} 0 & \text{if } \overline{r} \le \frac{5}{4} \\ 1 & \text{if } \overline{r} > \frac{5}{4} \end{cases}\,,$$

and $\chi := 0 \in W^{1,2}_D(\Omega)$. Then, the primal solution $u \in K$, in polar coordinates, for every $(r,\varphi)^\top \in \mathbb{R}_{>0}\times(0,2\pi)$ defined by $u(r,\varphi) := r^{\frac{2}{3}}\gamma_1(r)\sin(\frac{2\varphi}{3})$, has a singularity at the origin and, therefore, satisfies $u \notin W^{2,2}(\Omega)$, so that we cannot expect uniform mesh refinement to yield the quasi-optimal linear convergence rate.

The coarsest triangulation $\mathcal{T}_0$ of Figure 2 (and starting triagulation of Algorithm 6.3) consists of 48 halved squares. More precisely, Figure 2 displays the triangulations $\mathcal{T}_k$, $k\in\{0,4,8,12,16,20\}$, generated by the adaptive Algorithm 6.3. The approximate contact zones $\mathcal{C}_k^{cr} := \{\Pi_{h_k}u_k^{cr} = 0\} = \{\overline{\lambda}_k^{cr} < 0\}$, $k\in\{0,4,8,12,16,20\}$, are plotted in white Figure 2 while its complement is shaded[7]. Algorithm 6.3 refines in the complement of the contact zone $\mathcal{C} := \Omega\cap\{|\cdot| > \frac{3}{4}\}$. A refinement towards the origin, where the solution has a singularity in the gradient, and in $\{\frac{1}{4} \le |\cdot| \le \frac{3}{4}\}$, where the solution has large gradients, is reported. This behavior can also be seen in Figure 3, where the discrete primal solution $u_{10}^{cr} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_{10})$, the node-averaged discrete primal solution $I_{h_{10}}^{av}u_{10}^{cr} \in \mathcal{S}^1_D(\mathcal{T}_{10})$, the discrete Lagrange multiplier $\overline{\lambda}_{10}^{cr} \in \Pi_{h_{10}}(\mathcal{S}^{1,cr}_D(\mathcal{T}_{10}))$, and the discrete dual solution $z_{10}^{rt} \in \mathcal{RT}^0_N(\mathcal{T}_{10})$ are plotted on the triangulation $\mathcal{T}_{10}$, which has 1858 degrees of freedom. Figure 1 demonstrates that the adaptive Algorithm 6.3 improves the experimental convergence rate of about $\frac{3}{4}$ for uniform mesh-refinement to the optimal value 1. For uniform mesh-refinement, we expect an asymptotic convergence rate $\frac{3}{4}$ due to the corner singularity. Since not all quantities in the error measure $\rho^2_{h_k}(v_k)$ are computable, in Figure 1, we employ the reduced error measure

$$\tilde{\rho}_k^2(v_k) := \tfrac{1}{2}\|\nabla v_k - \nabla u\|_\Omega^2 + \langle -\Lambda, v_k - u\rangle_\Omega + \|\nabla u - \nabla_{h_k}u_k^{cr}\|_\Omega^2 + (-\overline{\lambda}_k^{cr}, \Pi_{h_k}(u-\chi))_\Omega\,,$$

where we exploit for the computation of the first two terms the identity (5.2).

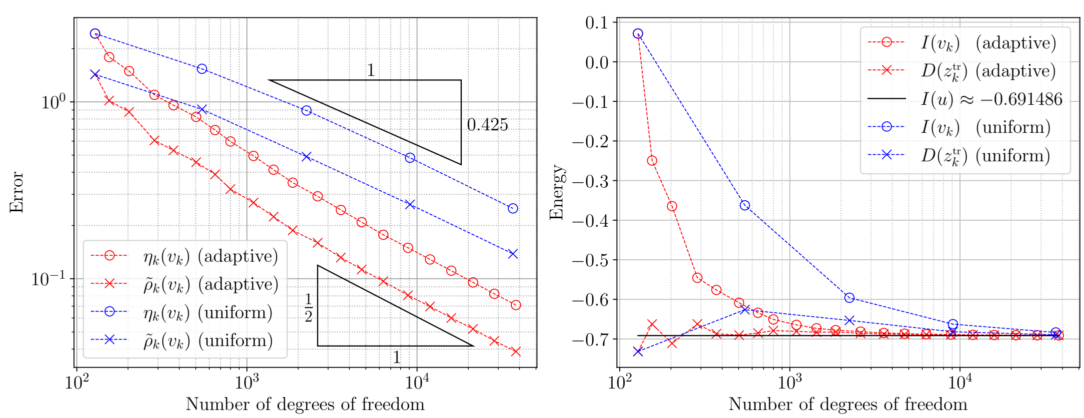

Figure 1: LEFT: Plots of $\eta_k^2(v_k)$ and $\tilde{\rho}_k^2(v_k)$ for $v_k := \max\{I_{h_k}^{av}u_k^{cr},\chi\} \in K$ using adaptive mesh refinement for $k = 0,\ldots,20$ and using uniform mesh refinement for $k = 0,\ldots,4$. RIGHT: Plots of $I(v_k)$, cf. (3.1), for $v_k := \max\{I_{h_k}^{av}u_k^{cr},\chi\} \in K$ and $D(z_k^{rt})$, cf. (3.3), using adaptive mesh refinement for $k = 0,\ldots,20$ and using uniform mesh refinement for $k = 0,\ldots,4$.

[7] we chose this color as in most of the examples the complement of the contact zone is refined and appears darker.

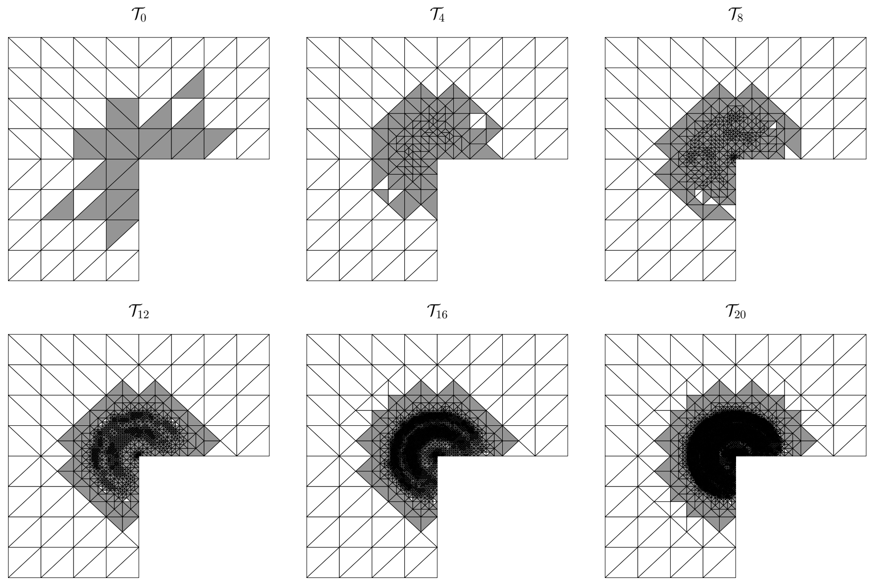

Figure 2: Adaptively refined meshes $\mathcal{T}_k$, $k \in \{0,4,8,12,16,20\}$, with approximate contact zones $\mathcal{C}_k^{cr} \coloneqq \{\Pi_{h_k} u_k^{cr} = 0\} = \{\overline{\lambda}_k^{cr} < 0\}$, $k \in \{0,4,8,12,16,20\}$, shown in white.

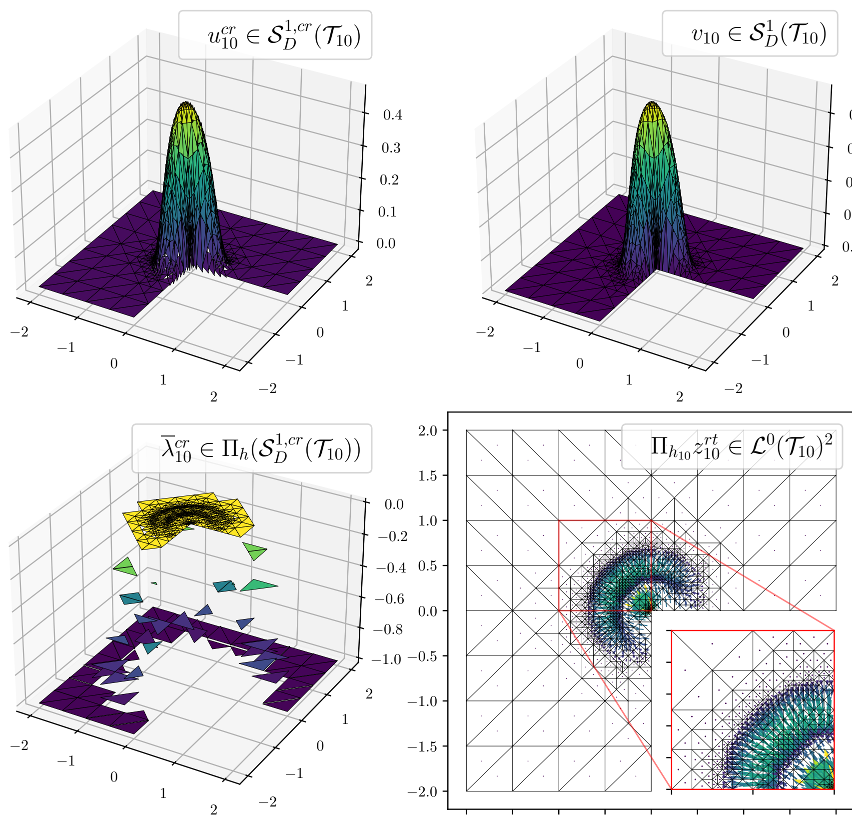

Figure 3: Discrete primal solution $u_{10}^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_{10})$ (upper left), node-averaged discrete primal solution $I_{h_{10}}^{av} u_{10}^{cr} \in \mathcal{S}_D^{1}(\mathcal{T}_{10})$ (upper right), discrete Lagrange multiplier $\overline{\lambda}_{10}^{cr} \in \Pi_{h_{10}}(\mathcal{S}_D^{1,cr}(\mathcal{T}_{10}))$ (lower left), and discrete dual solution $z_{10}^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_{10})$ (lower right) on $\mathcal{T}_{10}$.

6.3.2 *Example with unknown exact solution*

We examine an example from [9]. In this example, we let $\Omega := (-1,1)^2$, $\Gamma_D := \partial\Omega$, $\Gamma_N := \emptyset$, $f = 1 \in L^2(\Omega)$, and $\chi = \operatorname{dist}(\cdot,\partial\Omega) \in W^{1,2}_D(\Omega)$. The primal solution $u \in K$ is not known and cannot be expected to satisfy $u \in W^{2,2}(\Omega)$ inasmuch as $\chi \notin W^{2,2}(\Omega)$, so that uniform mesh refinement is expected to yield a reduced error decay rate compared to the optimal linear error decay rate.

The coarsest triangulation $\mathcal{T}_0$ in Figure 5 (and starting triangulation in Algorithm 6.3) consists of 64 elements. More precisely, Figure 5 displays the triangulations $\mathcal{T}_k$, $k \in \{0,5,10,15,20,25\}$, generated by the adaptive Algorithm 6.3. The approximate contact zones $\mathcal{C}_k^{cr} := \{\Pi_{h_k} u_k^{cr} = \chi_k\} = \{\overline{\lambda}_k^{cr} < 0\}$, $k \in \{0,5,10,15,20,25\}$, where $\chi_k := \Pi_{h_k}\chi \in \mathcal{L}^0(\mathcal{T}_k)$ for every $k \in \{0,5,10,15,20,25\}$, are plotted in white in Figure 5 while their complements are shaded. Note that for every $k \in \mathbb{N}$, we have that $\chi = I_{cr}\chi \in \mathcal{S}^1_D(\mathcal{T}_k)$ and $f = f_{h_k} \in \mathcal{L}^0(\mathcal{T}_k)$.

This example is different from the previous examples; in the sense that the solution and the obstacle are non-smooth along the lines $\mathcal{C} := \{(x,y)^\top \in \Omega \mid x = y \text{ or } x = 1-y\}$. Algorithm 6.3 refines the mesh towards these lines as can be seen in Figure 5. In addition, the approximate contact zones $\mathcal{C}_k^{cr}$, $k \in \{0,\dots,25\}$, reduces to $\mathcal{C}$. This behavior can also be observed in Figure 3, where the discrete primal solution $u_{15}^{cr} \in \mathcal{S}^{1,cr}_D(\mathcal{T}_{15})$, the node-averaged discrete primal solution $I_h^{av} u_{15}^{cr} \in \mathcal{S}^1_D(\mathcal{T}_{15})$, the discrete Lagrange multiplier $\overline{\lambda}_{15}^{cr} \in \Pi_{h_{15}}(\mathcal{S}^{1,cr}_D(\mathcal{T}_{15}))$, and the discrete dual solution $z_{15}^{rt} \in \mathcal{RT}^0_N(\mathcal{T}_{15})$ are plotted on the triangulation $\mathcal{T}_{15}$, which has 3769 degrees of freedom. Algorithm 6.3 improves the experimental convergence rate of about $\frac{1}{2}$ for uniform mesh-refinement to the quasi-optimal value 1. Since not all quantities in the error measure $\rho^2_{h_k}(v_k)$ are computable, in Figure 4, we employ the reduced error measure

$$\tilde{\rho}_k^2(v_k) := \tfrac{1}{2}\|\nabla v_k - \nabla u\|_\Omega^2 + \langle -\Lambda, v_k - u\rangle_\Omega\,,$$

where we exploit for the computation of $\tilde{\rho}_k^2(v_k)$ the identity (5.2) and approximate the value $I(u)$ via Aitken's δ^2-process, cf. [2]. More precisely, we always employ the approximation $I(u) \approx \epsilon_{27}$, where the sequence $(\epsilon_k)_{k\in\mathbb{N};k\geq 2}$, for every $k \in \mathbb{N}$ with $k \geq 2$, is defined by

$$\epsilon_k := \frac{I(v_k)I(v_{k-2}) - I(v_{k-1})^2}{I(v_k) - 2\,I(v_{k-1}) + I(v_{k-2})} \in \mathbb{R}\,.$$

However, it remains unclear whether this is a sufficiently accurate approximation of the exact errors $\rho^2_{h_k}(v_k)$, $k = 0,\dots,25$.

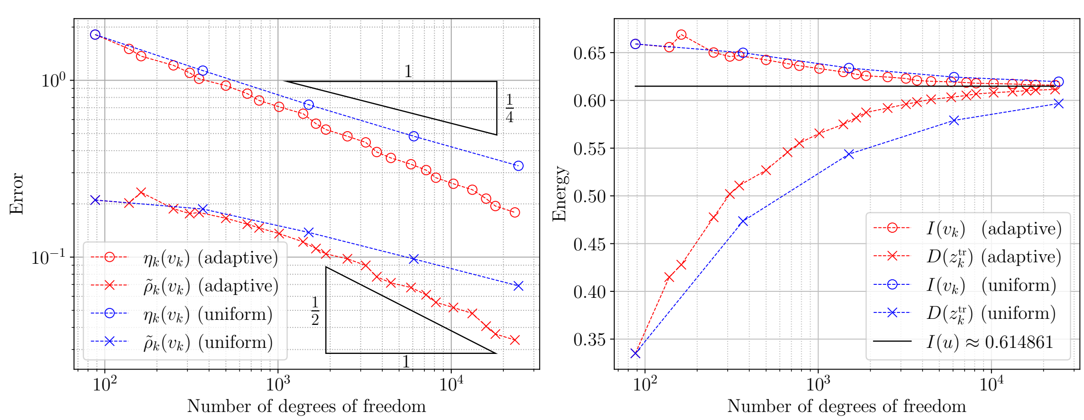

Figure 4: LEFT: Plots of $\eta_k^2(v_k)$ and $\tilde{\rho}_k^2(v_k)$ for $v_k := \max\{I_{h_k}^{av} u_k^{cr}, \chi\} \in K$ using adaptive mesh refinement for $k = 0,\dots,25$ and using uniform mesh refinement for $k = 0,\dots,4$. RIGHT: Plots of $I(v_k)$, cf. (3.1), for $v_k := \max\{I_{h_k}^{av} u_k^{cr}, \chi\} \in K$ and $D(z_k^{rt})$, cf. (3.3), using adaptive mesh refinement for $k = 0,\dots,25$ and using uniform mesh refinement for $k = 0,\dots,4$.

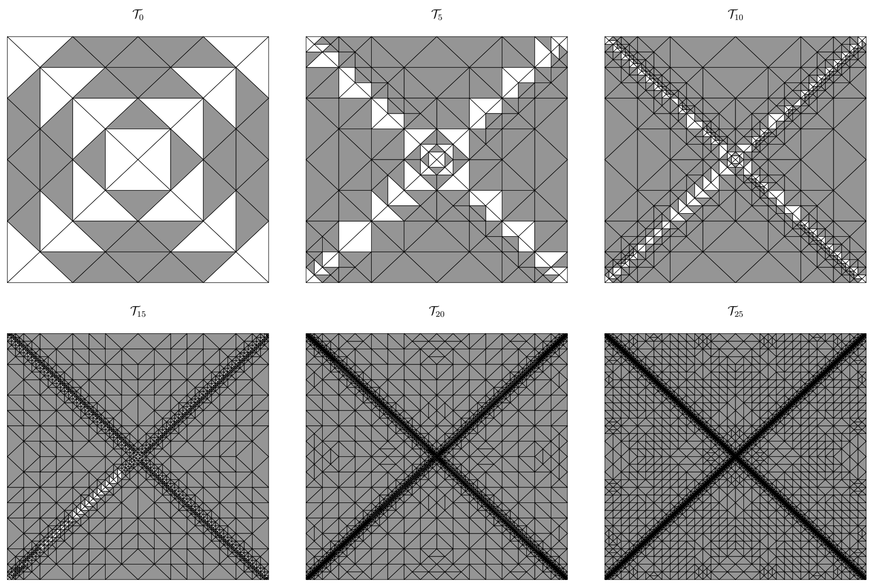

Figure 5: Adaptively refined meshes $\mathcal{T}_k$, $k \in \{0, 5, 10, 15, 20, 25\}$, with approximate contact zones $\mathcal{C}_k^{cr} \coloneqq \{\Pi_{h_k} u_k^{cr} = \Pi_{h_k} \chi_{h_k}\} = \{\overline{\lambda}_k^{cr} < 0\}$, $k \in \{0, 5, 10, 15, 20, 25\}$, shown in white.

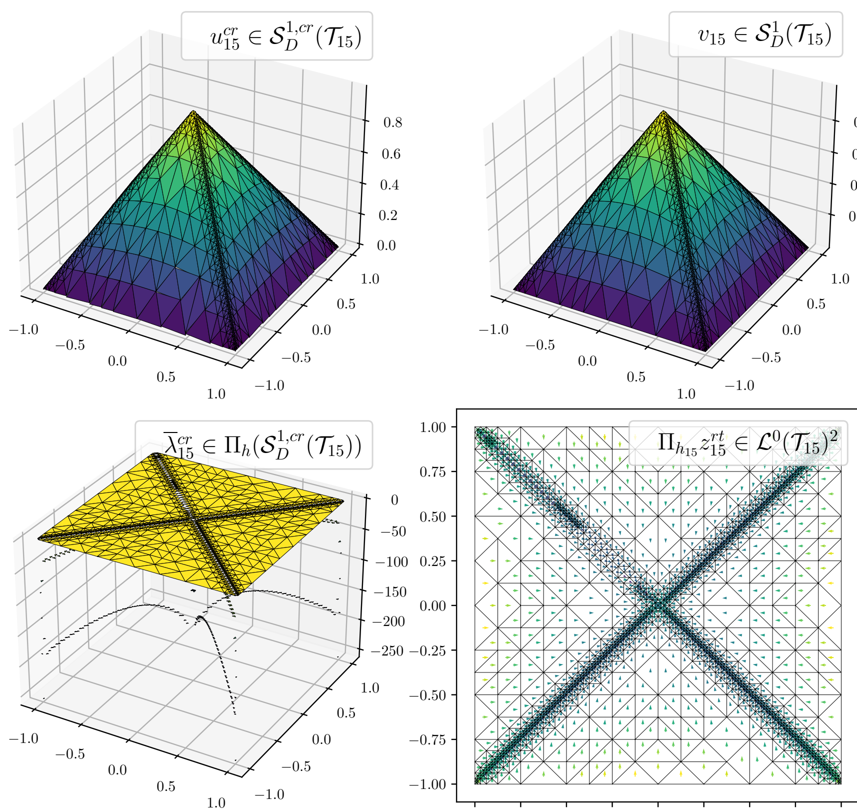

Figure 6: Discrete primal solution $u_{15}^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_{15})$ (upper left), node-averaged discrete primal solution $I_{h_{15}}^{av} u_{15}^{cr} \in \mathcal{S}_D^1(\mathcal{T}_{15})$ (upper right), discrete Lagrange multiplier $\overline{\lambda}_{15}^{cr} \in \Pi_{h_{15}}(\mathcal{S}_D^{1,cr}(\mathcal{T}_{15}))$ (lower left), and discrete dual solution $z_{15}^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_{15})$ (lower right) on $\mathcal{T}_{15}$.

A. Appendix

In this appendix, we derive local efficiency estimates for the Crouzeix–Raviart approximation (3.11) of (3.1), which, in turn, imply the following non-conforming efficiency result.

Theorem A.1. *There exists a constant $c > 0$, depending on the chunkiness $\omega_0 > 0$, such that*

$$\|\nabla_h v_h - \nabla_h u_h^{cr}\|_\Omega^2 \leq c\,[\|\nabla_h v_h - \nabla u\|_\Omega^2 + \|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,\Omega}^2 + \mathrm{osc}_h(f)]\,,$$

where for every $\mathcal{M}_h \subseteq \mathcal{T}_h$, we define $\mathrm{osc}_h(f,\mathcal{M}_h) \coloneqq \sum_{T\in\mathcal{M}_h} \mathrm{osc}_h(f,T)$, where $\mathrm{osc}_h(f,T) \coloneqq \|h_T(f-f_h)\|_T^2$ for every $T \in \mathcal{T}_h$, and $\mathrm{osc}_h(f) \coloneqq \mathrm{osc}_h(f,\mathcal{T}_h)$.

Before we prove Theorem A.1, we will first introduce some technical tools.

A.1 Node-averaging quasi-interpolation operator

The first tool is the *node-averaging quasi-interpolation operator* and its uniform approximation and stability properties, cf. [40, 12, 25].

The node-averaging quasi-interpolation operator $I_h^{av}\colon \mathcal{L}^1(\mathcal{T}_h) \to \mathcal{S}_D^1(\mathcal{T}_h)$, denoting for $z\in\mathcal{N}_h$, by $\mathcal{T}_h(z) \coloneqq \{T \in \mathcal{T}_h \mid z \in T\}$, the set of elements sharing z, for every $v_h \in \mathcal{L}^1(\mathcal{T}_h)$, is defined by

$$I_h^{av} v_h \coloneqq \sum_{z\in\mathcal{N}_h} \langle v_h\rangle_z \varphi_z\,, \qquad \langle v_h\rangle_z \coloneqq \begin{cases} \frac{1}{\mathrm{card}(\mathcal{T}_h(z))}\sum_{T\in\mathcal{T}_h(z)} (v_h|_T)(z) & \text{if } z \in \Omega\cup\Gamma_N \\ 0 & \text{if } z\in\Gamma_D \end{cases}\,,$$

where we denote by $(\varphi_z)_{z\in\mathcal{N}_h}$, the nodal basis of $\mathcal{S}^1(\mathcal{T}_h)$. If $p \in [1,\infty)$, then, there exists a constant $c > 0$, depending on $p \in [1,\infty)$ and the chunkiness $\omega_0 > 0$, such that for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, $T \in \mathcal{T}_h$, and $m \in \{0, 1\}$, cf. [7, Appx. A.2], we have that

(AV.1) $\|v_h - I_h^{av} v_h\|_T \leq c_{av}\, h_T\, \|\nabla_h v_h\|_{\omega_T}$,
(AV.2) $\|\nabla I_h^{av} v_h\|_T \leq c_{av}\, \|\nabla_h v_h\|_{\omega_T}$.

A.2 Local efficiency estimates

The second tool involves the following local efficiency estimates based on standard bubble function techniques, cf. [44].

Lemma A.2. *There exists a constant $c > 0$, depending only the chunkiness $\omega_0 > 0$, such that for every $v_h \in \mathcal{L}^1(\mathcal{T}_h)$, $T \in \mathcal{T}_h$, and $S \in \mathcal{S}_h^i$, respectively, it holds*

$$\|h_T(f_h - \overline{\lambda}_h^{cr})\|_T^2 \leq c\,\|\nabla v_h - \nabla u\|_T^2 + c\,\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,T}^2 + c\,\mathrm{osc}_h(f,T)\,, \tag{A.1}$$

$$h_S\,\|[\![\nabla_h v_h\cdot n]\!]_S\|_S^2 \leq c\,\|\nabla_h v_h - \nabla u\|_{\omega_S}^2 + c\,\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,\omega_S}^2 + c\,\mathrm{osc}_h(f,\omega_S)\,, \tag{A.2}$$

where $\|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,M} \coloneqq \|\overline{\Lambda}_h^{cr} - \Lambda\|_{(W_D^{1,2}(M))^*}$ for every open set $M \subseteq \Omega$.

Proof. ad (A.1). Let $T\in\mathcal{T}_h$ be fixed, but arbitrary. Then, there exists a bubble function $b_T \in W_0^{1,2}(T)$ such that $0 \leq b_T \leq c_b$ in T, $|\nabla b_T| \leq c_b\, h_T^{-1}$ in T, and $\fint_T b_T\,\mathrm{d}x = 1$, where the constant $c>0$ depends only on the chunkiness $\omega_0 > 0$. Using (3.8) and integration-by-parts, taking into account that $\nabla_h v_h \in \mathcal{L}^0(\mathcal{T}_h)^d$ and $b_T \in W_0^{1,2}(T)$ in doing so, for every $\mu \in \mathbb{R}$, we find that

$$(\nabla u - \nabla v_h, \nabla(\mu b_T))_T + \langle \overline{\Lambda}_h^{cr} - \Lambda, \mu b_T\rangle_T = (f - \overline{\lambda}_h^{cr}, \mu b_T)_T\,. \tag{A.3}$$

For the particular choice $\mu = \mu_T \coloneqq h_T(f_h - \overline{\lambda}_h^{cr}) \in \mathbb{R}$ in (A.3) and applying the ε-Young inequality $ab \leq \frac{1}{4\varepsilon}a^2 + \varepsilon b^2$, valid for all $a, b \geq 0$ and $\varepsilon > 0$, also using that $|b_T| \leq c_b$ in T and $h_T|\nabla b_T| \leq c_b$ in T, we observe that

$$\begin{aligned} \|h_T(f_h - \overline{\lambda}_h^{cr})\|_T^2 &= (f - \overline{\lambda}_h^{cr}, h_T\mu_T b_T)_T + (f_h - f, h_T\mu_T b_T)_T \\ &= (\nabla u - \nabla v_h, \nabla(h_T\mu_T b_T))_T + \langle \overline{\Lambda}_h^{cr} - \Lambda, h_T\mu_T b_T\rangle_T + (f_h - f, h_T\mu_T b_T)_T \\ &\leq \tfrac{1}{4\varepsilon}\,[\|\nabla v_h - \nabla u\|_T^2 + \|\overline{\Lambda}_h^{cr} - \Lambda\|_{*,T}^2 + \mathrm{osc}_h(f,T)] + 3\,\varepsilon\, c_b^2\, \|h_T(f_h - \overline{\lambda}_h^{cr})\|_T^2\,. \end{aligned} \tag{A.4}$$

For the particular choice $\varepsilon = \frac{1}{6c_b^2} > 0$ in (A.4), we conclude that (A.1) applies.

ad (A.2). Let $S\in\mathcal{S}_h^i$ be fixed, but arbitrary. Then, there exists a bubble function $b_S\in W_0^{1,p}(\omega_S)$ such that $0\leq b_S\leq c_b$ in ω_S, $|\nabla b_S|\leq c_b\, h_S^{-1}$ in ω_S, and $\fint_S b_S\,\mathrm{d}s=1$, where the constant $c>0$ depends only on the chunkiness $\omega_0>0$. Using (3.8) and integration-by-parts, taking into account that $\nabla_h v_h\in\mathcal{L}^0(\mathcal{T}_h)^d$ and $b_S\in W_0^{1,p}(\omega_S)$ in doing so, for every $\mu\in\mathbb{R}$, we find that

$$(\nabla u-\nabla_h v_h,\nabla(\mu b_S))_{\omega_S}+\langle\overline{\Lambda}_h^{cr}-\Lambda,\mu b_S\rangle_{\omega_S}=(f-\overline{\lambda}_h^{cr},\mu b_S)_{\omega_S}-|S|[\![\nabla v_h\cdot n]\!]_S\mu\,. \tag{A.5}$$

Let $T\in\mathcal{T}_h$ be with $T\subseteq\omega_S$. Then, for the particular choice $\mu=\mu_S:=\frac{|\omega_S|}{|S|}[\![\nabla_h v_h\cdot n]\!]_S\in\mathbb{R}$ in (A.5), $|\omega_S|\leq c_{\omega_0}|S|h_S$ for all $S\in\mathcal{S}_h^i$, where $c_{\omega_0}>0$ depends only on the chunkiness $\omega_0>0$, and applying the ε-Young inequality $ab\leq\frac{1}{4\varepsilon}a^2+\varepsilon b^2$ valid for all $a,b\geq 0$ and $\varepsilon>0$, also using that $|b_S|\leq c_b$ in ω_S and $h_S|\nabla b_S|\leq c_b$ in ω_S, we observe that

$$\begin{aligned}h_S\|[\![\nabla_h v_h\cdot n]\!]_S\|_S^2&=-\tfrac{|\omega_S|}{|S|}[(\nabla u-\nabla_h v_h,\nabla(\mu_S b_S))_{\omega_S}+\langle\overline{\Lambda}_h^{cr}-\Lambda,\mu_S b_S\rangle_{\omega_S}-(f-\overline{\lambda}_h^{cr},\mu_S b_S)_{\omega_S}]\\&\leq\tfrac{c_{\omega_0}}{4\varepsilon}\,[\|\nabla_h v_h-\nabla u\|_{\omega_S}^2+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\omega_S}^2+\|h_{\mathcal{T}}(f_h-\lambda_h^{cr})\|_{\omega_S}^2+\mathrm{osc}_h(f,\omega_S)]\\&\quad+\varepsilon\,c_{\omega_0}\,c_b^2\,h_S\|[\![\nabla_h v_h\cdot n]\!]_S\|_S^2\,.\end{aligned}\tag{A.6}$$

Using (A.6) and (A.1) in (A.5), for sufficiently small $\varepsilon>0$, conclude that (A.2) applies. □

From Lemma A.2 we can derive the following global efficiency result.

Lemma A.3. *There exists a constant $c>0$, depending only the chunkiness $\omega_0>0$, such that for every $v_h\in\mathcal{L}^1(\mathcal{T}_h)$, it holds*

$$\|h_{\mathcal{T}}(f_h-\overline{\lambda}_h^{cr})\|_\Omega^2\leq c\,\|\nabla_h v_h-\nabla u\|_\Omega^2+c\,\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+c\,\mathrm{osc}_h(f)\,,\tag{A.7}$$

$$\|h_{\mathcal{S}}^{1/2}[\![\nabla_h v_h\cdot n]\!]\|_{\mathcal{S}_h^i}^2\leq c\,\|\nabla_h v_h-\nabla u\|_\Omega^2+c\,\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+c\,\mathrm{osc}_h(f)\,.\tag{A.8}$$

Proof. ad (A.7). For $\mu_{\mathcal{T}}b_{\mathcal{T}}:=\sum_{T\in\mathcal{T}_h}\mu_T b_T\in W_D^{1,2}(\Omega)$ in the proof of (A.1), from (A.4) we derive

$$\|h_{\mathcal{T}}(f_h-\overline{\lambda}_h^{cr})\|_\Omega^2=(\nabla u-\nabla v_h,\nabla(h_{\mathcal{T}}\mu_{\mathcal{T}}b_{\mathcal{T}}))_\Omega+\langle\overline{\Lambda}_h^{cr}-\Lambda,h_{\mathcal{T}}\mu_{\mathcal{T}}b_{\mathcal{T}}\rangle_\Omega+(f_h-f,h_{\mathcal{T}}\mu_{\mathcal{T}}b_{\mathcal{T}})_\Omega\,,$$

which together with the ε-Young inequality and $|b_{\mathcal{T}}|+h_{\mathcal{T}}|\nabla b_{\mathcal{T}}|\leq c_b$ in Ω implies (A.7).

ad (A.8). For $\mu_{\mathcal{S}}b_{\mathcal{S}}:=\sum_{S\in\mathcal{S}_h^i}\mu_S b_S\in W_D^{1,2}(\Omega)$ in the proof of (A.2), from (A.6) we derive

$$\|h_{\mathcal{S}}^{1/2}[\![\nabla_h v_h\cdot n]\!]\|_{\mathcal{S}_h^i}^2\leq|(\nabla u-\nabla_h v_h,\nabla(\mu_{\mathcal{S}}b_{\mathcal{S}}))_\Omega+\langle\overline{\Lambda}_h^{cr}-\Lambda,\mu_{\mathcal{S}}b_{\mathcal{S}}\rangle_\Omega-(f-\overline{\lambda}_h^{cr},\mu_{\mathcal{S}}b_{\mathcal{S}})_\Omega|\,,$$

which together with the ε-Young inequality and $|b_{\mathcal{S}}|+h_{\mathcal{S}}|\nabla b_{\mathcal{S}}|\leq c_b$ in Ω implies (A.8). □

A.3 Proof of Theorem A.1

Eventually, we have everything at our disposal to prove Theorem A.1.

Proof (of Theorem A.1). Let $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ be arbitrary and introduce $e_h:=v_h-u_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$. Then, resorting to (3.8), (3.14) and $f-f_h\perp\Pi_h e_h$ in $L^2(\Omega)$, we arrive at

$$\begin{aligned}\|\nabla_h v_h-\nabla_h u_h^{cr}\|_\Omega^2&=(\nabla_h v_h-\nabla_h u_h^{cr},\nabla_h e_h)_\Omega\\&=(\nabla_h v_h,\nabla_h(e_h-I_h^{av}e_h))_\Omega\\&\quad-(f_h-\overline{\lambda}_h^{cr},e_h)_\Omega\\&\quad+(\nabla_h v_h-\nabla u,\nabla I_h^{av}e_h)_\Omega\\&\quad+(f,I_h^{av}e_h)_\Omega-\langle\Lambda,I_h^{av}e_h\rangle_\Omega\\&=(\nabla_h v_h,\nabla_h(e_h-I_h^{av}e_h))_\Omega\\&\quad-(f-\overline{\lambda}_h^{cr},e_h-I_h^{av}e_h)_\Omega\\&\quad+(\nabla_h v_h-\nabla u,\nabla I_h^{av}e_h)_\Omega\\&\quad+(f-f_h,e_h-\Pi_h e_h)_\Omega\\&\quad-\langle\overline{\Lambda}_h^{cr}-\Lambda,I_h^{av}e_h\rangle_\Omega\\&=:I_h^1+I_h^2+I_h^3+I_h^4+I_h^5\,.\end{aligned}\tag{A.9}$$

ad I_h^1. Using that $[\![\nabla_h v_h\cdot n(e_h-I_h^{av}e_h)]\!]_S=[\![\nabla_h v_h\cdot n]\!]_S\{e_h-I_h^{av}e_h\}_S+\{\nabla_h v_h\cdot n\}_S[\![e_h-I_h^{av}e_h]\!]_S$ on S, $\int_S[\![e_h-I_h^{av}e_h]\!]_S\,\mathrm{d}s=0$, and $\{\nabla_h v_h\cdot n\}_S=\text{const}$ on S for all $S\in\mathcal{S}_h^i$, an element-wise integration-by-parts, the discrete trace inequality [25, Lemma 12.8], (AV.1), the ε-Young inequality, and (A.2), for every $\varepsilon>0$, we find that

$$\begin{aligned}I_h^1&=(h_{\mathcal{S}}^{1/2}[\![\nabla_h v_h\cdot n]\!],h_{\mathcal{S}}^{-1/2}\{e_h-I_h^{av}e_h\})_{\mathcal{S}_h^i}\\&\le c_{tr}\,\|h_{\mathcal{S}}^{1/2}[\![\nabla_h v_h\cdot n]\!]\|_{\mathcal{S}_h^i}\|h_{\mathcal{T}}^{-1}(e_h-I_h^{av}e_h)\|_\Omega\\&\le c_{\mathit{eff}}\,\tfrac{c_{tr}^2}{4\varepsilon}\,[\|\nabla_h v_h-\nabla u\|_\Omega^2+\mathrm{osc}_h(f)+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2]+\varepsilon\,c_{av}^2\,\|\nabla_h e_h\|_\Omega^2\,.\end{aligned}\tag{A.10}$$

ad I_h^2. Using the ε-Young inequality, the approximation property of $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$, cf. (AV.1), and (A.1), for every $\varepsilon>0$, we obtain

$$\begin{aligned}I_h^2&\le\tfrac{1}{4\varepsilon}\,\|h_{\mathcal{T}}(f-\overline{\lambda}_h^{cr})\|_\Omega^2+\varepsilon\,\|h_{\mathcal{T}}^{-1}(e_h-I_h^{av}e_h)\|_\Omega^2\\&\le\tfrac{c_{\mathit{eff}}}{4\varepsilon}\,[\|\nabla_h v_h-\nabla u\|_\Omega^2+\mathrm{osc}_h(f)+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2]+\varepsilon\,c_{av}^2\,\|\nabla_h e_h\|_\Omega^2\,.\end{aligned}\tag{A.11}$$

ad I_h^3. Using the ε-Young inequality, the $W^{1,2}$-stability of $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$, cf. (AV.2), and (A.1), for every $\varepsilon>0$, we obtain

$$\begin{aligned}I_h^3&\le\tfrac{1}{4\varepsilon}\,\|\nabla_h v_h-\nabla u\|_\Omega^2+\varepsilon\,\|\nabla I_h^{av}e_h\|_\Omega^2\\&\le\tfrac{1}{4\varepsilon}\,\|\nabla_h v_h-\nabla u\|_\Omega^2+\varepsilon\,c_{av}^2\,\|\nabla_h e_h\|_\Omega^2\,.\end{aligned}\tag{A.12}$$

ad I_h^4. Using the ε-Young inequality and the approximation property of $\Pi_h\colon\mathcal{L}^1(\mathcal{T}_h)\to\mathcal{L}^0(\mathcal{T}_h)$, cf. (L0.1), for every $\varepsilon>0$, we obtain

$$\begin{aligned}I_h^4&\le\tfrac{1}{4\varepsilon}\,\mathrm{osc}_h(f)+\varepsilon\,\|h_{\mathcal{T}}^{-1}(e_h-\Pi_h e_h)\|_\Omega^2\\&\le\tfrac{1}{4\varepsilon}\,\mathrm{osc}_h(f)+\varepsilon\,c_\Pi^2\,\|\nabla_h e_h\|_\Omega^2\,.\end{aligned}\tag{A.13}$$

ad I_h^5. Using the ε-Young inequality, the $W^{1,2}$-stability of $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$, cf. (AV.1) & (AV.2), and the discrete Poincare inequality (2.1), for every $\varepsilon>0$, we obtain

$$\begin{aligned}I_h^5&\le\tfrac{1}{4\varepsilon}\,\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+\varepsilon\,(\|I_h^{av}e_h\|_\Omega^2+\|\nabla I_h^{av}e_h\|_\Omega^2)\\&\le\tfrac{1}{4\varepsilon}\,\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2+\varepsilon\,(1+(c_P^{cr})^2)\,\|\nabla_h e_h\|_\Omega^2\,.\end{aligned}\tag{A.14}$$

Combining (A.10), (A.11), (A.12), (A.13), and (A.14) in (A.9), for every $\varepsilon>0$, we conclude that

$$\begin{aligned}\|\nabla_h v_h-\nabla_h u_h^{cr}\|_\Omega^2\le{}&\tfrac{3+c_{\mathit{eff}}(1+c_{tr}^2)}{4\varepsilon}\,[\|\nabla_h v_h-\nabla u\|_\Omega^2+\mathrm{osc}_h(f)+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2]\\&+\varepsilon\,(3\,c_{av}+c_\Pi^2+1+(c_P^{cr})^2)\,\|\nabla_h e_h\|_\Omega^2\,,\end{aligned}\tag{A.15}$$

For $\varepsilon\coloneqq\frac{1}{2(3\,c_{av}+c_\Pi^2+1+(c_P^{cr})^2)}>0$ and $c\coloneqq(3\,c_{av}+c_\Pi^2+1+(c_P^{cr})^2)(3+c_{\mathit{eff}}(1+c_{tr}^2))>0$ in (A.15), for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, we arrive at

$$\|\nabla_h v_h-\nabla_h u_h^{cr}\|_\Omega^2\le c\,[\|\nabla_h v_h-\nabla u\|_\Omega^2+\mathrm{osc}_h(f)+\|\overline{\Lambda}_h^{cr}-\Lambda\|_{*,\Omega}^2]\,,\tag{A.16}$$

which is the claimed non-conforming efficiency estimate. □

Eventually, the node-averaging quasi-interpolation operator $I_h^{av}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathcal{S}_D^1(\mathcal{T}_h)$ satisfies the following best-approximation result with respect to Sobolev functions.

Proposition A.4. *There exists a constant $c>0$, depending only on the chunkiness $\omega_0>0$, such that for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $T\in\mathcal{T}_h$, it holds*

$$\|\nabla_h v_h-\nabla I_h^{av}v_h\|_T^2\le c\inf_{v\in W_D^{1,2}(\Omega)}\|\nabla_h v_h-\nabla v\|_{\omega_T}^2\,.$$

In particular, for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, $v\in W_D^{1,2}(\Omega)$, and $T\in\mathcal{T}_h$, it holds

$$\|\nabla I_h^{av}v_h-\nabla v\|_T^2\le c\,\|\nabla_h v_h-\nabla v\|_{\omega_T}^2\,.$$

Essential tool in the verification of Proposition A.4 is the following lemma.

Lemma A.5. *There exists a constant $c > 0$, depending only on the chunkiness $\omega_0 > 0$, such that for every $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ and $T \in \mathcal{T}_h$, it holds*

$$\|\nabla_h v_h - \nabla I_h^{av} v_h\|_T^2 \leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} h_T \,\|h_T^{-1}[\![v_h]\!]_S\|_S^2\,,$$

where $\mathcal{S}_h(T) \coloneqq \{S \in \mathcal{S}_h \mid S \cap T \neq \emptyset\}$.

Proof. Appealing to the inverse inequality [25, Lemma 12.1] and the node-based norm equivalence [25, Proposition 12.5], there exists a constant $c > 0$, depending only on the chunkiness $\omega_0 > 0$, such that

$$\begin{aligned}\|\nabla_h v_h - \nabla I_h^{av} v_h\|_T^2 &\leq c\, h_T^{-1}\,\|v_h - I_h^{av} v_h\|_T^2\\ &\leq c\, h_T^{d-2} \sum_{z \in \mathcal{N}_h \cap T} |(v_h|_T)(z) - (I_h^{av} v_h)(z)|^2\,.\end{aligned}\tag{A.17}$$

Next, for every $z \in \mathcal{N}_h \cap T$, we need to distinguish the cases $z \notin \Gamma_D$ and $z \in \Gamma_D$:

Case $z \notin \Gamma_D$. If $z \notin \Gamma_D$, then since each $T' \in \mathcal{T}_h(z)$ can be reached from T via passing through a finite number[8] of interior sides in $\mathcal{S}_h^i(T) \coloneqq \mathcal{S}_h(T) \cap \mathcal{S}_h^i$, using [25, (22.6)], we find that

$$\begin{aligned}|(v_h|_T)(z) - (I_h^{av} v_h)(z)|^2 &\leq c\,\frac{1}{\operatorname{card}(\mathcal{T}_h(z))} \sum_{T' \in \mathcal{T}_h(z)} |(v_h|_T)(z) - (v_h|_{T'})(z)|^2\\ &\leq c \sum_{S \in \mathcal{S}_h^i(T)} |[\![v_h]\!]_S(z)|^2\\ &\leq c \sum_{S \in \mathcal{S}_h^i(T)} h_T^{1-d} \|[\![v_h]\!]_S\|_S^2\,.\end{aligned}\tag{A.18}$$

Case $z \in \Gamma_D$. If $z \in \Gamma_D$, then we need to distinguish the case that $z \in \operatorname{int}\Gamma_D$, i.e., z lies in the relative interior of Γ_D, and $z \in \partial\Gamma_D$, i.e., z lies in the relative boundary of Γ_D:

Subcase $z \in \operatorname{int}\Gamma_D$. If $z \in \operatorname{int}\Gamma_D$, then there exists a boundary side $S \in \mathcal{S}_h(T) \setminus \Gamma_N$ with $z \in S$ and $S \subseteq \partial T$. Thus, resorting to [25, (22.6)], we find that

$$\begin{aligned}|(v_h|_T)(z) - (I_h^{av} v_h)(z)|^2 &= |(v_h|_T)(z)|^2\\ &= |[\![v_h]\!]_S(z)|^2\\ &\leq c\, h_T^{1-d}\, \|[\![v_h]\!]_S\|_S^2\,.\end{aligned}\tag{A.19}$$

Subcase $z \in \partial\Gamma_D$. If $z \in \partial\Gamma_D$, then there exists a boundary side $S \in \mathcal{S}_h(T)$ with $z \in S$, $S \subseteq \partial T$, and either $S \subseteq \Gamma_D$ or $S \subseteq \Gamma_N$. If $S \subseteq \Gamma_D$, then we argue as in (A.19). If $S \subseteq \Gamma_N$, then there exists boundary side $S' \in \mathcal{S}_h(T)\setminus\Gamma_N$ with $z \in S'$ and an element $T' \in \mathcal{T}_h$ with $z \in T'$and $S' \subseteq T'$. If $T' = T$, then we argue as in (A.19). If $T' \neq T$, then since T' can be reached from T via passing through a finite number of interior sides in $\mathcal{S}_h^i(T)$, resorting to [25, (22.6)], we find that

$$\begin{aligned}|(v_h|_T)(z) - (I_h^{av} v_h)(z)|^2 &= |(v_h|_T)(z)|^2\\ &\leq |(v_h|_{T'})(z)|^2 + c \sum_{S \in \mathcal{S}_h^i(T)} |[\![v_h]\!]_S(z)|^2\\ &\leq c\,|[\![v_h]\!]_{S'}(z)|^2 + c \sum_{S \in \mathcal{S}_h^i(T)} |[\![v_h]\!]_S(z)|^2\\ &\leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} h_T^{1-d} \|[\![v_h]\!]_S\|_S^2\,.\end{aligned}\tag{A.20}$$

Eventually, combining (A.18)–(A.20) in (A.17), we conclude the claimed estimate. □

[8] uniformly bounded by a constant depending only on the chunkiness $\omega_0 > 0$.

Proof (of Proposition A.4). Using that $\|[\![v_h]\!]_S\|_{L^\infty(S)} \leq c \fint_S |[\![v_h]\!]_S| \,\mathrm{d}x$ (cf. [25, Lemma 12.1]) as well as $|T| \sim h_T|S|$ for all $T \in \mathcal{T}_h$ and $S \in \mathcal{S}_h(T)$, where the constant $c > 0$ depends only on the chunkiness $\omega_0 > 0$, for every $T \in \mathcal{T}_h$, we infer from Lemma A.5 that

$$\begin{aligned}\|\nabla I_h^{av} v_h - \nabla_h v_h\|_T^2 &\leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} h_T \,\|h_T^{-1}[\![v_h]\!]_S\|_S^2 \\ &\leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} |T| \,(|T|^{-1}\|[\![v_h]\!]_S\|_{L^1(S)})^2 \,.\end{aligned} \tag{A.21}$$

For every $S \in \mathcal{S}_h$, we denote by $\pi_h^S \colon L^1(S) \to \mathbb{R}$, the side-wise (local) L^2-projection operator onto constant functions, for every $w \in L^1(S)$ defined by $\pi_h^S w \coloneqq \fint_S w \,\mathrm{d}s$. Since for every $w \in W^{1,1}(T)$, where $T \in \mathcal{T}_h$ with $T \subseteq \omega_S$, due to the $L^1(S)$-stability of $\pi_h^S \colon L^1(S) \to \mathbb{R}$ and [35, Corollary A.19], it holds

$$\begin{aligned}\|w - \pi_h^S w\|_{L^1(S)} &= \|w - \Pi_h w - \pi_h^S(w - \Pi_h w)\|_{L^1(S)} \\ &\leq 2\,\|w - \Pi_h w\|_{L^1(S)} \\ &\leq c\,\|\nabla w\|_{L^1(T;\mathbb{R}^d)}\,,\end{aligned} \tag{A.22}$$

where $c > 0$ depends only on the chunkiness $\omega_0 > 0$. Next, let $v \in W_D^{1,2}(\Omega)$ be fixed, but arbitrary. Using that $\pi_h^S[\![v_h]\!]_S = [\![v]\!]_S = 0$ in $L^1(S)$ for all $S \in \mathcal{S}_h(T) \setminus \Gamma_N$ and $T \in \mathcal{T}_h$ and (A.22), we find that

$$\begin{aligned}\|[\![v_h]\!]_S\|_{L^1(S)} &= \|[\![v_h - v]\!]_S - \pi_h^S[\![v_h - v]\!]_S\|_{L^1(S)} \\ &\leq \|\nabla v_h - \nabla v\|_{L^1(\omega_S;\mathbb{R}^d)}\,.\end{aligned} \tag{A.23}$$

Then, using in (A.21), (A.23), $|T| \sim |\omega_T| \sim |\omega_S|$ for all $T \in \mathcal{T}_h$ and $S \in \mathcal{S}_h(T)$, where $c > 0$ depends only on the chunkiness $\omega_0 > 0$, and Jensen's inequality, for every $T \in \mathcal{T}_h$, we deduce that

$$\begin{aligned}\|\nabla I_h^{av} v_h - \nabla_h v_h\|_T^2 &\leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} |\omega_S| \,(|\omega_S|^{-1}\|\nabla v_h - \nabla v\|_{L^1(\omega_S;\mathbb{R}^d)})^2 \\ &\leq c \sum_{S \in \mathcal{S}_h(T)\setminus\Gamma_N} \|\nabla v_h - \nabla v\|_{\omega_S}^2 \\ &\leq c\,\|\nabla v_h - \nabla v\|_{\omega_T}^2\,.\end{aligned} \tag{A.24}$$

Eventually, taking in (A.24) the infimum with respect to $v \in W_D^{1,2}(\Omega)$, we conclude the assertion. □

REFERENCES

[1] M. AINSWORTH, J. T. ODEN, and C. Y. LEE, Local a posteriori error estimators for variational inequalities, *Numerical Methods for Partial Differential Equations* **9** no. 1 (1993), 23–33. doi:https://doi.org/10.1002/num.1690090104.

[2] A. C. AITKEN, On bernoulli's numerical solution of algebraic equations, *Proceedings of the Royal Society of Edinburgh* no. 46 (1926), 280–305. doi:10.1017/S0370164600022070.

[3] S. BARTELS, *Numerical methods for nonlinear partial differential equations*, *Springer Series in Computational Mathematics* **47**, Springer, Cham, 2015. doi:10.1007/978-3-319-13797-1.

[4] S. BARTELS, *Numerical approximation of partial differential equations*, *Texts in Applied Mathematics* **64**, Springer, [Cham], 2016. doi:10.1007/978-3-319-32354-1.

[5] S. BARTELS, Error estimates for a class of discontinuous Galerkin methods for nonsmooth problems via convex duality relations, *Math. Comp.* **90** no. 332 (2021), 2579–2602. doi:10.1090/mcom/3656.

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

[7] S. BARTELS and A. KALTENBACH, *Explicit and efficient error estimation for convex minimization problems*, accepted, 2023.

[8] S. Bartels and Z. Wang, Orthogonality relations of Crouzeix-Raviart and Raviart-Thomas finite element spaces, *Numer. Math.* **148** no. 1 (2021), 127–139. doi:10.1007/s00211-021-01199-3.

[9] S. Bartels and C. Carstensen, A convergent adaptive finite element method for an optimal design problem, *Numer. Math.* **108** no. 3 (2008), 359–385. doi:10.1007/s00211-007-0122-x.

[10] D. Braess, A posteriori error estimators for obstacle problems—another look, *Numer. Math.* **101** no. 3 (2005), 415–421. doi:10.1007/s00211-005-0634-1.

[11] D. Braess, R. Hoppe, and J. Schöberl, A posteriori estimators for obstacle problems by the hypercircle method, *Comput. Vis. Sci.* **11** no. 4-6 (2008), 351–362. doi:10.1007/s00791-008-0104-2.

[12] S. C. Brenner, Two-level additive schwarz preconditioners for nonconforming finite element methods, *Mathematics of Computation* **65** no. 215 (1996), 897–921.

[13] H. M. Buss and S. I. Repin, A posteriori error estimates for boundary value problems with obstacles, in *Numerical mathematics and advanced applications (Jyväskylä, 1999)*, World Sci. Publ., River Edge, NJ, 2000, pp. 162–170.

[14] L. A. Caffarelli, The obstacle problem revisited, *J. Fourier Anal. Appl.* **4** no. 4-5 (1998), 383–402. doi:10.1007/BF02498216.

[15] C. Carstensen and K. Köhler, Nonconforming FEM for the obstacle problem, *IMA J. Numer. Anal.* **37** no. 1 (2017), 64–93. doi:10.1093/imanum/drw005.

[16] A. Chambolle and T. Pock, Crouzeix-Raviart approximation of the total variation on simplicial meshes, *J. Math. Imaging Vision* **62** no. 6-7 (2020), 872–899. doi:10.1007/s10851-019-00939-3.

[17] Z. Chen and R. H. Nochetto, Residual type a posteriori error estimates for elliptic obstacle problems, *Numerische Mathematik* **84** (2000), 527–548.

[18] M. Cicuttin, A. Ern, and T. Gudi, Hybrid high-order methods for the elliptic obstacle problem, *J. Sci. Comput.* **83** no. 1 (2020), Paper No. 8, 18. doi:10.1007/s10915-020-01195-z.

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

[20] B. Dacorogna, *Direct methods in the calculus of variations*, second ed., *Applied Mathematical Sciences* **78**, Springer, New York, 2008.

[21] L. Diening and C. Kreuzer, Linear convergence of an adaptive finite element method for the p-Laplacian equation, *SIAM J. Numer. Anal.* **46** no. 2 (2008), 614–638. doi:10.1137/070681508.

[22] L. Diening and M. Růžička, Interpolation operators in Orlicz-Sobolev spaces, *Numer. Math.* **107** no. 1 (2007), 107–129. doi:10.1007/s00211-007-0079-9.

[23] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, *SIAM J. Numer. Anal.* **33** no. 3 (1996), 1106–1124. doi:10.1137/0733054.

[24] I. Ekeland and R. Témam, *Convex analysis and variational problems*, english ed., *Classics in Applied Mathematics* **28**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999, Translated from the French. doi:10.1137/1.9781611971088.

[25] A. Ern and J. L. Guermond, *Finite Elements I: Approximation and Interpolation*, *Texts in Applied Mathematics* no. 1, Springer International Publishing, 2021. doi:10.1007/978-3-030-56341-7.

[26] R. S. Falk, Error estimates for the approximation of a class of variational inequalities, *Math. Comput.* **28** (1974), 963–971.

[27] M. Feischl, M. Page, and D. Praetorius, Convergence of adaptive FEM for some elliptic obstacle problem with inhomogeneous Dirichlet data, *Int. J. Numer. Anal. Model.* **11** no. 1 (2014), 229–253.

[28] D. A. French, S. Larsson, and R. H. Nochetto, Pointwise a posteriori error analysis for an adaptive penalty finite element method for the obstacle problem, *Comput. Methods Appl. Math.* **1** no. 1 (2001), 18–38. doi:10.2478/cmam-2001-0002.

[29] A. Friedman, *Variational principles and free-boundary problems*, second ed., Robert E. Krieger Publishing Co., Inc., Malabar, FL, 1988.

[30] T. Gudi and K. Porwal, A posteriori error control of discontinuous Galerkin methods for elliptic obstacle problems, *Math. Comp.* **83** no. 286 (2014), 579–602. doi:10.1090/S0025-5718-2013-02728-7.

[31] M. Hintermüller, K. Ito, and K. Kunisch, The primal-dual active set strategy as a semismooth newton method, *SIAM Journal on Optimization* **13** no. 3 (2002), 865–888. doi:10.1137/S1052623401383558.

[32] R. H. Hoppe and R. Kornhuber, Adaptive multilevel methods for obstacle problems, *SIAM*

Journal on Numerical Analysis **31** no. 2 (1994), 301–323. doi:10.1137/0731016.

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

[34] C. Johnson, Adaptive finite element methods for the obstacle problem, *Mathematical Models and Methods in Applied Sciences* **02** no. 04 (1992), 483–487. doi:10.1142/S0218202592000284.

[35] A. Kaltenbach and M. Růžička, *Convergence analysis of a Local Discontinuous Galerkin approximation for nonlinear systems with Orlicz-structure*, submitted, 2022.

[36] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Society for Industrial and Applied Mathematics, 2000. doi:10.1137/1.9780898719451.

[37] R. Kornhuber, A posteriori error estimates for elliptic variational inequalities, *Computers & Mathematics with Applications* **31** no. 8 (1996), 49–60. doi:https://doi.org/10.1016/0898-1221(96)00030-2.

[38] R. Kornhuber, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*, *Advances in Numerical Mathematics*, Vieweg+Teubner Verlag, 1997.

[39] A. Logg and G. N. Wells, Dolfin: Automated finite element computing, *ACM Transactions on Mathematical Software* **37** no. 2 (2010). doi:10.1145/1731022.1731030.

[40] P. Oswald, On the robustness of the bpx-preconditioner with respect to jumps in the coefficients, *Mathematics of Computation* **68** no. 226 (1999), 633–650.

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

[42] A. Veeser, Efficient and reliable a posteriori error estimators for elliptic obstacle problems, *SIAM J. Numer. Anal.* **39** no. 1 (2001), 146–167. doi:10.1137/S0036142900370812.

[43] A. Veeser, On a posteriori error estimation for constant obstacle problems, in *Numerical methods for viscosity solutions and applications (Heraklion, 1999)*, *Ser. Adv. Math. Appl. Sci.* **59**, World Sci. Publ., River Edge, NJ, 2001, pp. 221–234. doi:10.1142/9789812799807_0012.

[44] R. Verfürth, A posteriori error estimation and adaptive mesh-refinement techniques, *Journal of Computational and Applied Mathematics* **50** no. 1 (1994), 67–83. doi:https://doi.org/10.1016/0377-0427(94)90290-9.

[45] R. Verfürth, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Oxford University Press, 04 2013. doi:10.1093/acprof:oso/9780199679423.001.0001.

[46] P. Virtanen et al. , SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nature Methods* **17** (2020), 261–272. doi:10.1038/s41592-019-0686-2.

[47] L.-h. Wang, On the error estimate of nonconforming finite element approximation to the obstacle problem, *J. Comput. Math.* **21** no. 4 (2003), 481–490.

[48] A. Weiss and B. I. Wohlmuth, A posteriori error estimator for obstacle problems, *SIAM J. Sci. Comput.* **32** no. 5 (2010), 2627–2658. doi:10.1137/090773921.