Exact a posteriori error control for variational problems via convex duality and explicit flux reconstruction

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

[1]Department of Applied Mathematics, University of Freiburg, Hermann–Herder–Str. 10, 79104 Freiburg
[2]Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 135, 10623 Berlin

February 9, 2024

Abstract

A posteriori error estimates are an important tool to bound discretization errors in terms of computable quantities avoiding regularity conditions that are often difficult to establish. For non-linear and non-differentiable problems, problems involving jumping coefficients, and finite element methods using anisotropic triangulations, such estimates often involve large factors, leading to sub-optimal error estimates. By making use of convex duality arguments, exact and explicit error representations are derived that avoid such effects.

Keywords: A posteriori error estimates; adaptivity; non-conforming methods; convex duality.
AMS MSC (2020): 35Q68; 49M25; 49M29; 65K15; 65N15; 65N50.

1. Introduction

1.1 Sharp error estimation

The derivation of sharp a posteriori error estimates has been an active area of research over several decades. Typical concepts involve the precise characterization of generic constants occurring in residual estimates (*cf.* [10, 1, 83, 50, 88, 40, 52]), the approximation of local problems by higher-order methods (*cf.* [64, 39]), the usage of convex duality relations (*cf.* [66, 84]), and the development of post-processing procedures to obtain equilibrated fluxes (*cf.* [61, 32, 31, 51, 26, 81]). For general discussions of various aspects of a posteriori error estimation, we refer the reader, *e.g.*, to [11, 30, 73, 86, 42]. Recently, fully computable error representations for various convex variational problems have been identified by deriving explicit representation formulas for solutions of dual problems in terms of non-conforming primal approximations. The concept avoids the occurrence of typical constants, applies to a large class of non-quadratic, non-differentiable, constrained, and degenerate problems, for which classical approaches lead to sub-optimal error control. Closely related concepts have been used in the derivation of a posteriori error estimates for mixed and non-conforming methods (*cf.* [8, 7, 4, 45, 2]).

1.2 Prager–Synge identity

A well-known error representation in the context of the *homogeneous Poisson problem*, *i.e.*,

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega\,,\\ \nabla u\cdot n &= 0 && \text{on } \Gamma_N\,,\\ u &= 0 && \text{on } \Gamma_D\,, \end{aligned} \tag{1.1}$$

where $\Gamma_D,\Gamma_N\subseteq\partial\Omega$ with $\Gamma_D\dot\cup\Gamma_N=\partial\Omega$ and $f\in L^2(\Omega)$, was pointed out in [71] and follows from the celebrated *Prager–Synge identity*, *i.e.*, if $u\in W^{1,2}_D(\Omega)$ is the (weak) solution of the Poisson problem (1.1) and $z=\nabla u\in W^2_N(\mathrm{div};\Omega)$ the solution of the corresponding (Fenchel) dual problem, then for every $v\in W^{1,2}_D(\Omega)$ and $y\in W^2_N(\mathrm{div};\Omega)$ with $\mathrm{div}\,y=-f$ a.e. in Ω, it holds that

$$\tfrac12\|\nabla v-\nabla u\|^2_{2,\Omega}+\tfrac12\|y-z\|^2_{2,\Omega}=\tfrac12\|\nabla v-y\|^2_{2,\Omega}\,. \tag{1.2}$$

The identity (1.2) is an immediate consequence of the L^2-orthogonality of $y-z$ and $\nabla v-\nabla u$ and has the interpretation that the squared L^2-distance between the gradient of a primal approximation $v\in W^{1,2}_D(\Omega)$ and a dual approximation $y\in W^2_N(\mathrm{div};\Omega)$ with $\mathrm{div}\,y=-f$ a.e. in Ω yields an explicit and computable way to determine the sum of the primal and dual approximation errors.

*Email: bartels@mathematik.uni-freiburg.de
†Email: kaltenbach@math.tu-berlin.de

A limitation of using this interpretation practically is that often merely a primal approximation is given and determining an optimal or nearly optimal dual approximation typically is too expensive. For special problems involving two-dimensional Poisson (*cf.* [30]) and obstacle problems (*cf.* [28]), this difficulty has been overcome by constructing a nearly optimal discrete dual vector field via a local post-processing procedure.

1.3 Continuous convex duality

The Prager–Synge identity (1.2) can be generalized to a larger class of convex minimization problems following, *e.g.*, the works [77, 73, 74, 75, 76, 78, 22, 24]: let $\phi\colon \Omega\times\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon \Omega\times\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ be (Lebesgue) measurable functions such that for a.e. $x\in\Omega$, the functions $\phi(x,\cdot)\colon\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and $\psi(x,\cdot)\colon\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ are proper, convex, and lower semi-continuous. Then, the minimization of the functional $I\colon W^{1,p}_D(\Omega)\to\mathbb{R}\cup\{+\infty\}$, for every $v\in W^{1,p}_D(\Omega)$ defined by

$$I(v) \coloneqq \int_\Omega \phi(\cdot,\nabla v)\,\mathrm{d}x + \int_\Omega \psi(\cdot,v)\,\mathrm{d}x\,, \tag{1.3}$$

denotes the *(Fenchel) primal problem*. A corresponding *(Fenchel) dual problem* consists in the maximization of the functional $D\colon W^{p'}_N(\operatorname{div};\Omega)\to\mathbb{R}\cup\{-\infty\}$, for every $y\in W^{p'}_N(\operatorname{div};\Omega)$ defined by

$$D(y) \coloneqq -\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x - \int_\Omega \psi^*(\cdot,\operatorname{div} y)\,\mathrm{d}x\,. \tag{1.4}$$

Here, $\phi^*\colon\Omega\times\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and $\psi^*\colon\Omega\times\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ denote the *Fenchel conjugates* (with respect to the second argument) to $\phi\colon\Omega\times\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon\Omega\times\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$, respectively. *Fenchel–Young inequalities* (*cf.* (2.3)) in combination with an integration-by-parts formula imply a *weak duality relation*, *i.e.*, for every $v\in W^{1,p}_D(\Omega)$ and $y\in W^{p'}_N(\operatorname{div};\Omega)$, it holds that

$$\begin{aligned} I(v) &\geq \int_\Omega \nabla v\cdot y - \phi^*(\cdot,y)\,\mathrm{d}x + \int_\Omega \psi(\cdot,v)\,\mathrm{d}x \\ &= -\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x - \int_\Omega \operatorname{div} y\, v - \psi(\cdot,v)\,\mathrm{d}x \geq D(y)\,. \end{aligned} \tag{1.5}$$

1.4 Continuous strong duality, convex optimality relations, and flux reconstruction

In many cases, *e.g.*, if both $\phi\colon\Omega\times\mathbb{R}^d\to\mathbb{R}$ and $\psi\colon\Omega\times\mathbb{R}\to\mathbb{R}$ are Carathéodory mappings[1] and for every $v\in L^p(\Omega)$ and $y\in L^p(\Omega;\mathbb{R}^d)$, it holds that $\phi(\cdot,y),\psi(\cdot,v)\in L^1(\Omega)$ (*cf.* [48, Thm. 4.1, p. 59, Prop. 1.1, p. 77]), there even holds a *strong duality relation*, *i.e.*, there exists minimizer $u\in W^{1,p}_D(\Omega)$ of (1.3), called *primal solution*, and a maximizer $z\in W^{p'}(\operatorname{div};\Omega)$ of (1.4), called *dual solution*, such that

$$I(u) = D(z)\,. \tag{1.6}$$

The strong duality relation (1.6) is available for a large class of convex minimization problems, *e.g.*, including non-linear Dirichlet problems, obstacle problems, certain non-differentiable problems, and degenerate minimization problems. It is equivalent to *convex optimality relations*, *i.e.*, it holds that

$$z\cdot\nabla u = \phi^*(\cdot,z) + \phi(\cdot,\nabla u) \qquad \text{a.e. in } \Omega\,, \tag{1.7}$$

$$\operatorname{div} z\, u = \psi^*(\cdot,\operatorname{div} z) + \psi(\cdot,u) \qquad \text{a.e. in } \Omega\,. \tag{1.8}$$

The convex optimality relations (1.7), (1.8) characterize equality in the Fenchel–Young inequalities (*cf.* (2.3)) used in the derivation of the dual problem (*cf.* (1.5)). By the *Fenchel–Young identity* (*cf.* (2.4)), the convex optimality relations (1.7), (1.8) are each equivalent to the inclusions

$$z\in\partial_r\phi(\cdot,\nabla u) \qquad \text{a.e. in } \Omega\,, \tag{1.9}$$

$$\operatorname{div} z\in\partial_r\psi(\cdot,u) \qquad \text{a.e. in } \Omega\,, \tag{1.10}$$

where we denote with $\partial_r\phi\colon\Omega\times\mathbb{R}^d\to 2^{\mathbb{R}^d}$ and $\partial_r\psi\colon\Omega\times\mathbb{R}\to 2^{\mathbb{R}}$, the corresponding sub-differentials of $\phi\colon\Omega\times\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon\Omega\times\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ (with respect to the second argument).

[1] A mapping $\Phi\colon\Omega\times\mathbb{R}^\ell\to\mathbb{R}$, $\ell\in\mathbb{N}$, is called *Carathéodory mapping* if for a.e. $x\in\Omega$, the function $\Phi(x,\cdot)\colon\mathbb{R}^\ell\to\mathbb{R}$ is continuous, and for every $r\in\mathbb{R}^\ell$, the function $\Phi(\cdot,r)\colon\Omega\to\mathbb{R}$ is (Lebesgue) measurable.

If $\phi(x,\cdot) \in C^1(\mathbb{R}^d)$ for a.e. $x \in \Omega$ or $\psi(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, then (1.9), (1.10) each become

$$z = D\phi(\cdot, \nabla u) \qquad \text{a.e. in } \Omega\,, \tag{1.11}$$

$$\operatorname{div} z = D\psi(\cdot, u) \qquad \text{a.e. in } \Omega\,, \tag{1.12}$$

where the identity (1.11) has an interpretation as reconstruction formula, in the sense that given a primal solution $u \in W^{1,p}_D(\Omega)$, a dual solution $z \in W^{p'}_N(\operatorname{div};\Omega)$ is immediately available via (1.11).

1.5 *Generalized Prager–Synge identity*

The convex and concave functionals $I\colon W^{1,p}_D(\Omega) \to \mathbb{R}\cup\{+\infty\}$ and $D\colon W^{p'}_N(\operatorname{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$ give rise to the definition of the non-negative coercivity measures $\rho^2_I\colon W^{1,p}_D(\Omega)\times W^{1,p}_D(\Omega) \to \mathbb{R}$ and $\rho^2_{-D}\colon W^{p'}_N(\operatorname{div};\Omega)\times W^{p'}_N(\operatorname{div};\Omega) \to \mathbb{R}$, for every $v, \widehat{v} \in W^{1,p}_D(\Omega)$ and $y, \widehat{y} \in W^{p'}_N(\operatorname{div};\Omega)$ defined by

$$\begin{aligned} \rho^2_I(v,\widehat{v}) &\coloneqq I(v) - I(\widehat{v}) - \langle \delta I(\widehat{v}), v-\widehat{v}\rangle_{W^{1,p}_D(\Omega)}\,,\\ \rho^2_{-D}(y,\widehat{y}) &\coloneqq -D(y) + D(\widehat{y}) + \langle \delta D(\widehat{y}), y-\widehat{y}\rangle_{W^{p'}_N(\operatorname{div};\Omega)}\,, \end{aligned} \tag{1.13}$$

in case that $I\colon W^{1,p}_D(\Omega) \to \mathbb{R}\cup\{+\infty\}$ and $D\colon W^{p'}_N(\operatorname{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$ are Fréchet differentiable. The quantities (1.13) for every $v, \widehat{v} \in W^{1,p}_D(\Omega)$ and $y, \widehat{y} \in W^{p'}_N(\operatorname{div};\Omega)$ measure the distance between the functionals I and $-D$ and their tangents at $I(\widehat{v})$ and $-D(\widehat{y})$ evaluated at v and y, respectively. More generally, the variations δI and δD in (1.13) can be replaced by suitable sub-gradients. Since $\delta I(u) = 0$ and $\delta D(z) = 0$ (or $0 \in \partial I(u)$ and $0 \in \partial D(z)$), in case of a strong duality (1.6), we find a *generalized Prager–Synge identity*, *i.e.*, given a primal solution $u \in W^{1,p}_D(\Omega)$ and dual solution $z \in W^{p'}_N(\operatorname{div};\Omega)$, for every $v \in W^{1,p}_D(\Omega)$ with $I(v) < +\infty$ and $y \in W^{p'}_N(\operatorname{div};\Omega)$ with $D(y) > -\infty$, it holds that

$$\begin{aligned} \rho^2_{\mathrm{tot}}(v,y) &\coloneqq \rho^2_I(v,u) + \rho^2_{-D}(y,z)\\ &= I(v) - I(u) + D(z) - D(y)\\ &= I(v) - D(y)\\ &\eqqcolon \eta^2_{\mathrm{gap}}(v,y)\,, \end{aligned} \tag{1.14}$$

i.e., the inaccessible *total error* $\rho^2_{\mathrm{tot}}(v,y)$ equals the accessible *primal-dual gap estimator* $\eta^2_{\mathrm{gap}}(v,y)$. For the Poisson problem (1.1), we have that $\phi(r) = \phi^*(r) = \frac{1}{2}|r|^2$ for all $r \in \mathbb{R}^d$, $\psi(x,s) = -f(x)s$ for a.e. $x \in \Omega$ and all $s \in \mathbb{R}$, and $\psi^*(x,s) = I^{\Omega}_{\{-f(x)\}}(s)$ for a.e. $x \in \Omega$ and all $s \in \mathbb{R}$, where the latter indicator functional enforces the constraint $\operatorname{div} y = -f$ a.e. in Ω. Hence, for every $v \in W^{1,p}_D(\Omega)$ and $y \in W^{p'}_N(\operatorname{div};\Omega)$ with $\operatorname{div} y = -f$ a.e. in Ω, with an integration-by-parts formula, we find that

$$\begin{aligned} \eta^2_{\mathrm{gap}}(v,y) &= \tfrac{1}{2}\|\nabla v\|^2_{2,\Omega} - (y, \nabla v)_\Omega + \tfrac{1}{2}\|y\|^2_{2,\Omega}\\ &= \tfrac{1}{2}\|\nabla v - y\|^2_{2,\Omega}\,. \end{aligned} \tag{1.15}$$

It is remarkable that the employed global relations lead to an error representation as an integral of a non-negative function. Hence, the right-hand side can be decomposed using partitions and thereby provides meaningful local refinement indicators. This observation can equally be generalized: using the definitions (1.3), (1.4) and an integration-by-parts formula, for every $v \in W^{1,p}_D(\Omega)$ with $I(v) < +\infty$ and $y \in W^{p'}_N(\operatorname{div};\Omega)$ with $D(y) > -\infty$, we arrive at the general representation

$$\begin{aligned} \eta^2_{\mathrm{gap}}(v,y) = &\int_\Omega \phi(\cdot,\nabla v) - y\cdot\nabla v + \phi^*(\cdot, y)\,\mathrm{d}x\\ &+ \int_\Omega \psi(\cdot, v) - \operatorname{div} y\, v + \psi^*(\cdot, \operatorname{div} y)\,\mathrm{d}x\,. \end{aligned} \tag{1.16}$$

Both integrands on the right-hand side of (1.16), by the Fenchel–Young inequality (*cf.* (2.3)), are point-wise non-negative and vanish if and only if convex optimality relations (1.7), (1.8) are satisfied. Hence, the primal dual-gap estimator measures the validity of the convex optimality relations (1.7), (1.8). Note that the convex optimality relations (1.7), (1.8) do not require any regularity of $\phi\colon \Omega\times\mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ and $\psi\colon \Omega\times\mathbb{R} \to \mathbb{R}\cup\{+\infty\}$, which makes the primal-dual gap estimator a predestined a posteriori error estimator for non-differentiable convex minimization problems.

1.6 Discrete reconstruction formula

As in the case of the Poisson problem (1.1), dual problems typically involve constraints (*e.g.*, $\operatorname{div} y = -f$ a.e. in Ω (*cf.* (1.2))) and are, thus, significantly harder to solve than primal problems, which often lead to monotone operators for which efficient iterative numerical solution procedures are available. It is, therefore, fundamental to avoid the explicit numerical solution of dual problems. For a Crouzeix–Raviart approximation of the Poisson problem (1.1), seeking for $u_h^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ such that for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, it holds that

$$(\nabla_h u_h^{cr}, \nabla_h v_h)_\Omega = (f_h, \Pi_h v_h)_\Omega \,, \tag{1.17}$$

where $f_h \in \mathcal{L}^0(\mathcal{T}_h)$ is an element-wise constant approximation of $f \in L^2(\Omega)$, using elementary but fundamental relations between the Crouzeix–Raviart and the Raviart–Thomas element, a remarkable relation, called *Marini formula*, between $u_h^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and a Raviart–Thomas solution $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ of the corresponding dual formulation to (1.17) has been established in [63, 9, 7], *i.e.*,

$$z_h^{rt} = \nabla_h u_h^{cr} - \frac{f_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega \,. \tag{1.18}$$

In (1.17), $\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ denotes the *Crouzeix–Raviart finite element space*, $\mathcal{RT}_N^0(\mathcal{T}_h)$ the *Raviart–Thomas finite element space*, ∇_h the element-wise application of the gradient operator ∇, and Π_h the (local) L^2-projection operator onto the space of element-wise constant functions $\mathcal{L}^0(\mathcal{T}_h)$. Due to $\mathcal{RT}_N^0(\mathcal{T}_h) \subseteq W_N^2(\operatorname{div};\Omega)$ and $\operatorname{div} z_h^{rt} = -f_h$ a.e. in Ω, if $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ given via (1.18) is admissible in (1.2). In other words, in the case of the Poisson problem (1.1), the discrete reconstruction formula (1.18) enables to approximate the primal and the dual problem simultaneously using only approximation (1.17).

1.7 Discrete convex duality

It is possible to construct a discrete primal problem that induces the same (discrete) convex duality relations like the continuous primal problem (1.3), *i.e.*, a corresponding discrete dual problem, a discrete weak duality relation, the equivalence of a discrete strong duality relation to discrete convex optimality relations and, most important, a discrete analogue of the reconstruction formula (1.11) or generalization of the discrete reconstruction formula (1.18), respectively. To this end, we need to perform three non-conforming modifications on the primal energy functional (1.3):

- we replace the energy densities by $\phi_h \colon \Omega \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and $\psi_h \colon \Omega \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, which are, again, are (Lebesgue) measurable functions such that for a.e. $x \in \Omega$, the functions $\phi_h(x,\cdot) \colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and $\psi_h(x,\cdot) \colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ are proper, convex, and lower semi-continuous. In addition, for every $r \in \mathbb{R}^d$ and $s \in \mathbb{R}$, the functions $\phi_h(\cdot, r) \colon \Omega \to \mathbb{R} \cup \{+\infty\}$ and $\psi_h(\cdot, s) \colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ are element-wise constant;
- we incorporate a (local) L^2-projection Π_h onto $\mathcal{L}^0(\mathcal{T}_h)$ into the lower-order term;
- we replace the weak gradient ∇ with the element-wise gradient ∇_h.

With these three non-conforming modifications the *discrete primal problem* is given via the minimization of the functional $I_h^{cr} \colon \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathbb{R} \cup \{+\infty\}$, for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) \coloneqq \int_\Omega \phi_h(\cdot, \nabla_h v_h) \,\mathrm{d}x + \int_\Omega \psi_h(\cdot, \Pi_h v_h) \,\mathrm{d}x \,. \tag{1.19}$$

A corresponding *discrete dual problem* to the minimization of (1.19) consists in the maximization of the functional $D_h^{rt} \colon \mathcal{RT}_N^0(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, for every $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h) \coloneqq -\int_\Omega \phi_h^*(\cdot, \Pi_h y_h) \,\mathrm{d}x + \int_\Omega \psi_h^*(\cdot, \operatorname{div} y_h) \,\mathrm{d}x \,. \tag{1.20}$$

Note that the three non-conforming modifications (more precisely, the first two modifications), in particular, ensure that the integrands in (1.19) and (1.20), respectively, are element-wise constant.

This is crucial for establishing a strong duality relation between the discrete primal problem (1.19) and the discrete dual problem (1.20). As above, Fenchel–Young inequalities imply that a *discrete weak duality relation* applies, *i.e.*, for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$, it holds that

$$\begin{aligned} I_h^{cr}(v_h) &\geq \int_\Omega \nabla_h v_h \cdot \Pi_h y_h - \phi_h^*(\cdot, \Pi_h y_h)\,\mathrm{d}x + \int_\Omega \psi_h(\cdot, \Pi_h v_h)\,\mathrm{d}x \\ &= -\int_\Omega \phi_h^*(\cdot, \Pi_h y_h)\,\mathrm{d}x - \int_\Omega \operatorname{div} y_h v_h - \psi_h(\cdot, v_h)\,\mathrm{d}x \geq D_h^{rt}(y_h)\,. \end{aligned} \tag{1.21}$$

In (1.21), we used a *discrete integration-by-parts formula* without contributions from element sides, *i.e.*, for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}_N^0(\mathcal{T}_h)$, it holds that

$$(\nabla_h v_h, \Pi_h y_h)_\Omega = -(\operatorname{div} y_h, \Pi_h v_h)_\Omega\,. \tag{1.22}$$

The use of the (local) L^2-projection in (1.22) is optional, but crucial or establishing a strong duality relation between the discrete primal problem (1.19) and the discrete dual problem (1.20). There holds a *discrete strong duality relation*, *i.e.*, it holds that

$$I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})\,, \tag{1.23}$$

if and only if there hold *discrete convex optimality relations*, *i.e.*, it holds that

$$\Pi_h z_h^{rt} \cdot \nabla_h u_h^{cr} = \phi_h^*(\cdot, \Pi_h z_h^{rt}) + \phi_h(\cdot, \nabla_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,, \tag{1.24}$$

$$\operatorname{div} z_h^{rt}\, \Pi_h u_h^{cr} = \psi_h^*(\cdot, \operatorname{div} z_h^{rt}) + \psi_h(\cdot, \Pi_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,. \tag{1.25}$$

By the Fenchel–Young identity (*cf.* (2.4)), the discrete convex optimality relations (1.24), (1.25) are each equivalent to the inclusions

$$\Pi_h z_h^{rt} \in \partial_r \phi_h(\cdot, \nabla_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,, \tag{1.26}$$

$$\operatorname{div} z_h^{rt} \in \partial_r \psi_h(\cdot, \Pi_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,. \tag{1.27}$$

If $\phi_h(x,\cdot) \in C^1(\mathbb{R}^d)$ for a.e. $x \in \Omega$ or $\psi_h(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, then (1.26), (1.27) each become

$$\Pi_h z_h^{rt} = D\phi_h(\cdot, \nabla_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,, \tag{1.28}$$

$$\operatorname{div} z_h^{rt} = D\psi_h(\cdot, u_h^{cr}) \qquad \text{a.e. in } \Omega\,. \tag{1.29}$$

Note that, different from the continuous reconstruction formula (1.11), the discrete convex optimality relation (1.28) does not give full information about a discrete dual solution $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$. However, using the additional information provided by the discrete convex optimality relation (1.29), the surjectivity of the divergence operator $\operatorname{div}\colon \mathcal{RT}_N^0(\mathcal{T}_h) \to \mathcal{L}^0(\mathcal{T}_h)$ if $\Gamma_D \neq \emptyset$ (and surjectivity of $\operatorname{div}\colon \mathcal{RT}_N^0(\mathcal{T}_h) \to \mathcal{L}^0(\mathcal{T}_h)/\mathbb{R}$ if $\Gamma_D = \emptyset$), and the *discrete Helmholtz decomposition*

$$(\mathcal{L}^0(\mathcal{T}_h))^d = \operatorname{div}(\ker|_{\mathcal{RT}_N^0(\mathcal{T}_h)}) \oplus \nabla_h(\mathcal{S}_D^{1,cr}(\mathcal{T}_h))\,, \tag{1.30}$$

if $\phi_h(x,\cdot) \in C^1(\mathbb{R}^d)$ and $\psi_h(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, it is possible to establish a discrete strong duality relation (1.23) and a *generalized Marini formula* (*cf.* [16, 38, 59]), *i.e.*, it holds that

$$z_h^{rt} = D\phi_h(\cdot, \nabla_h u_h^{cr}) + \frac{D\psi_h(\cdot, \Pi_h u_h^{cr})}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \qquad \text{a.e. in } \Omega\,. \tag{1.31}$$

Similar to the equivalence of (1.10) to $u \in \partial_r \psi(\cdot, \operatorname{div} z)$ a.e. in Ω, that is $u = D\psi(\cdot, \operatorname{div} z)$ a.e. in Ω if $\psi(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, if instead $\phi_h^*(x,\cdot) \in C^1(\mathbb{R}^d)$ and $\psi_h^*(x,\cdot) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, there exists an inversion of the reconstruction formula (1.31) (*cf.* [24]), *i.e.*, it holds that

$$u_h^{cr} = D\psi_h^*(\cdot, \operatorname{div} z_h^{rt}) + D\phi_h^*(\cdot, \Pi_h z_h^{rt}) \cdot (\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \qquad \text{a.e. in } \Omega\,. \tag{1.32}$$

In many cases, we have that $\psi_h = \psi$, *e.g.*, if $\psi(x,s) = -f(x)s$ for a.e. $x \in \Omega$ and all $s \in \mathbb{R}$ with an element-wise constant function $f \in \mathcal{L}^0(\mathcal{T}_h)$, and, in this case, by Jensen's inequality, we have that $-D_h^{rt}(z_h^{rt}) \leq -D(z_h^{rt})$. Hence, in general, an inconsistent or non-conforming discretization of the primal problem is necessary to ensure strong discrete duality and to obtain a reconstruction formula.

1.8 A priori error estimates

The continuous strong duality relation (1.6) and the discrete strong duality relation (1.23) are useful in deriving a priori error estimates for the primal problem (1.3) and the dual problem (1.4). Letting $v_h \coloneqq \Pi_h^{cr} u \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $y_h \coloneqq \Pi_h^{rt} z \in \mathcal{RT}_N^0(\mathcal{T}_h)$ be the Crouzeix–Raviart and Raviart–Thomas quasi-interpolants of a primal solution $u \in W_D^{1,p}(\Omega)$ and a dual solution $z \in W_N^{p'}(\operatorname{div};\Omega)$, respectively, if a strong duality relation (*i.e.*, (1.6)) applies, due to the discrete weak duality relation (1.21), we have that

$$\begin{aligned} e_{\mathrm{tot},h}^2(v_h, y_h) &\coloneqq \rho_{I_h^{cr}}^2(v_h, u_h^{cr}) + \rho_{-D_h^{rt}}^2(y_h, z_h^{rt}) \\ &= [I_h^{cr}(v_h) - I_h^{cr}(u_h^{cr})] + [D_h^{rt}(z_h^{rt}) - D_h^{rt}(y_h)] \\ &\le I_h^{cr}(v_h) - D_h^{rt}(y_h) \\ &= [I_h^{cr}(v_h) - I(u)] + [D(z) - D_h^{rt}(z_h^{rt})]\,, \end{aligned} \tag{1.33}$$

i.e., the primal approximation error between the discrete primal solution $u_h^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and the interpolant of the primal solution $u \in W_D^{1,p}(\Omega)$ plus the dual approximation error between the discrete dual solution $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ and the interpolant of the dual solution $z \in W_N^{p'}(\operatorname{div};\Omega)$ is bounded in terms of primal and dual interpolation errors. For the Poisson problem (1.1), from (1.33), using that $\nabla_h v_h = \Pi_h \nabla u$ a.e. in Ω together with Jensen's inequality, $\operatorname{div} y_h = \Pi_h \operatorname{div} z$ a.e. in Ω together with $\operatorname{div} z = -f$ a.e. in Ω, an integration-by-parts formula, and the second binomial formula, it follows that

$$\begin{aligned} \tfrac12\|\nabla_h v_h - \nabla_h u_h^{cr}\|_{2,\Omega}^2 + \tfrac12\|y_h - z_h^{rt}\|_{2,\Omega}^2 &\le \tfrac12\|\nabla_h v_h\|_{2,\Omega}^2 - \tfrac12\|\nabla u\|_{2,\Omega}^2 \\ &\quad - (f, v_h - u)_\Omega - \tfrac12\|z\|_{2,\Omega}^2 + \tfrac12\|y_h\|_{2,\Omega}^2 \\ &\le (z, z - y_h)_\Omega - \tfrac12\|z\|_{2,\Omega}^2 + \tfrac12\|y_h\|_{2,\Omega}^2 \\ &= \tfrac12\|z - y_h\|_{2,\Omega}^2\,. \end{aligned} \tag{1.34}$$

If $u \in W^{1+s,2}(\Omega)$ (*i.e.*, $z \in W^{s,2}(\Omega;\mathbb{R}^d)$), $s \in (0,1)$, the right-hand side in (1.34) is of order $\mathcal{O}(h^{2s})$. This short proof can be generalized to a large class of variational problems (*cf.* [16, 58, 18]) and avoids the usage of Strang lemmas to control the effect of the non-conformity of the discretization. It has recently been observed that Crouzeix–Raviart discretizations lead to higher convergence rates than classical conforming methods for certain non-differentiable problems (*cf.* [41, 17]).

1.9 Data approximation and inexact solution

1.9.1 Data approximation

In the case of a linear lower-order term $\psi\colon \Omega \times \mathbb{R} \to \mathbb{R}$, *i.e.*, $\psi(x,s) = -f(x)s$ for a.e. $x \in \Omega$ and all $s \in \mathbb{R}$ for some $f \in L^{p'}(\Omega)$, the discrete dual solution $z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ given via (1.31) is admissible in the continuous dual problem, *i.e.*, satisfies $D(z_h^{rt}) > -\infty$, if and only if $\operatorname{div} y_h = -f$ a.e. in Ω, *i.e.*, if and only if $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$ is element-wise constant. If this is not the case, then we introduce a modified functional $I^{(h)}\colon W_D^{1,p}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W_D^{1,p}(\Omega)$ defined by

$$I^{(h)}(v) \coloneqq \int_\Omega \phi_h(\cdot, \nabla v)\,\mathrm{d}x + \int_\Omega \psi_h(\cdot, \Pi_h v)\,\mathrm{d}x\,. \tag{1.35}$$

Then, for a minimizer $u^{(h)} \in W_D^{1,p}(\Omega)$ of (1.35), due to $I^{(h)}(u^{(h)}) \le I^{(h)}(u)$, we have that

$$\begin{aligned} \rho_I^2(u^{(h)}, u) &= I(u^{(h)}) - I(u) \\ &\le I(u^{(h)}) - I^{(h)}(u^{(h)}) + I^{(h)}(u) - I(u) \\ &\le \|\phi(\cdot, \nabla u^{(h)}) - \phi_h(\cdot, \nabla u^{(h)})\|_{1,\Omega} + \|\psi(\cdot, u^{(h)}) - \psi_h(\cdot, \Pi_h u^{(h)})\|_{1,\Omega} \\ &\quad + \|\phi_h(\cdot, \nabla u) - \phi(\cdot, \nabla u)\|_{1,\Omega} + \|\psi_h(\cdot, \Pi_h u) - \psi(\cdot, u)\|_{1,\Omega}\,. \end{aligned}$$

For autonomous higher-order term $\phi\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ and linear lower-order term $\psi\colon \Omega\times\mathbb{R}\to\mathbb{R}$, *i.e.*, $\psi(x,s) = -f(x)s$ for a.e. $x\in\Omega$ and all $s\in\mathbb{R}$ for some $f\in L^{p'}(\Omega)$, with discretization $\psi_h\colon \Omega\times\mathbb{R}\to\mathbb{R}$, given via $\psi_h(x,s) = -f_h(x)s$ for a.e. $x\in\Omega$ and all $s\in\mathbb{R}$, where $f_h \coloneqq \Pi_h f \in \mathcal{L}^0(\mathcal{T}_h)$, due to $f - f_h \perp \Pi_h(u-u^{(h)})$ in $L^2(\Omega)$, using element-wise Poincaré's inequality, we find that

$$\begin{aligned}\rho_I^2(u^{(h)},u) &= (f-f_h, u-u^{(h)}-\Pi_h(u-u^{(h)}))_\Omega\\ &\le c_P\,\|h_{\mathcal{T}}(f-f_h)\|_{p',\Omega}\|\nabla u-\nabla u^{(h)}\|_{p,\Omega}\,.\end{aligned}$$

Then, the arguments explained above apply to the modified functional (1.35).

1.9.2 Inexact solution

In the case $\phi_h(x,\cdot)\in C^1(\mathbb{R}^d)$ and $\psi_h(x,\cdot)\in C^1(\mathbb{R})$ for a.e. $x\in\Omega$, we can incorporate errors resulting from the inexact iterative solution of the discrete primal problem via discrete residuals. More precisely, if $\widetilde{u}_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ is an inexact approximation of the discrete primal problem, *i.e.*, quasi-minimizer of (1.19), we represent the residual in the discrete $W^{1,2}$-semi-norm. More precisely, we choose $\widetilde{r}_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ such that for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, it holds that

$$(\nabla_h\widetilde{r}_h,\nabla_h v_h)_\Omega = (D\phi_h(\cdot,\nabla_h u_h^{cr}),\nabla_h v_h)_\Omega + (D\psi_h(\cdot,u_h^{cr}),\Pi_h v_h)_\Omega\,.$$

Then, $\widetilde{u}_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ is a minimizer of the functional $\widetilde{I}_h^{cr}\colon\mathcal{S}_D^{1,cr}(\mathcal{T}_h)\to\mathbb{R}$, for every $v_h\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ defined by

$$\widetilde{I}_h^{cr}(v) \coloneqq \int_\Omega \widetilde{\phi}_h(\cdot,\nabla_h v_h)\,\mathrm{d}x + \int_\Omega \psi_h(\cdot,v_h)\,\mathrm{d}x\,.$$

where $\widetilde{\phi}_h\colon\Omega\times\mathbb{R}^d\to\mathbb{R}$ is defined by $\widetilde{\phi}_h(x,r)\coloneqq\phi_h(x,r)-\nabla_h\widetilde{r}_h(x)\cdot r$ for a.e. $x\in\Omega$ and all $r\in\mathbb{R}^d$. The identities and estimates derived above now hold with ϕ_h replaced with $\widetilde{\phi}_h$.

1.10 Properties of the primal-dual gap estimator

In general, the discrete primal solution $u_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ is not admissible in the continuous primal problem since $u_h^{cr}\notin W_D^{1,p}(\Omega)$ and, thus, cannot be inserted in the primal-dual gap estimator, while the discrete dual solution $z_h^{rt}\in\mathcal{RT}_N^0(\mathcal{T}_h)\subseteq W_N^{p'}(\mathrm{div};\Omega)$, up to data approximation terms, is admissible in the continuous dual problem. An admissible approximation $\overline{u}_h^{cr}\in W_D^{1,p}(\Omega)$, *e.g.*, can be obtained via a cheap node-averaging procedure. The generalized Prager–Synge identity (1.14) imposes no restrictions about optimality of the arguments, so that with admissible approximations $\overline{u}_h^{cr}\in W_D^{1,p}(\Omega)$ with $I(\overline{u}_h^{cr})<+\infty$ and $\overline{z}_h^{rt}\in W_N^{p'}(\mathrm{div};\Omega)$ with $D(\overline{z}_h^{rt})>-\infty$, we still have that

$$\rho_{\mathrm{tot}}^2(\overline{u}_h^{cr},\overline{z}_h^{rt}) = \eta_{\mathrm{gap}}^2(\overline{u}_h^{cr},\overline{z}_h^{rt})\,. \tag{1.36}$$

The primal-dual gap error estimator has some remarkable features:

- It is obtained by a simple post-processing procedure of the discrete primal problem (*cf.* (1.31));
- It is reliable and efficient with constants one (*cf.* (1.14));
- It does not require an exact solution of the possibly non-linear primal problem (*cf.* (1.36));
- It is globally equivalent to residual type estimators for many model problems;
- Its integrands are point-wise non-negative and, thus, it is suitable for local mesh-refinement;
- It applies to non-linear, non-differentiable, degenerate, scalar, and vectorial problems and does not require the development of a particular error analysis.

Besides these positive features, some relevant properties are desirable but are not yet established:

- Is it possible to prove optimal convergence of adaptive methods based on the local mesh refinement indicators given by the local contributions to the primal-dual gap estimator?
- Can one devise a general strategy to ensure admissibility of the reconstructed flux in the dual problem?

The second aspect arises in model problems with non-differentiable function $\phi\colon\Omega\times\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ and some remedies have been proposed (*cf.* [41, 21, 24]). The first aspect primarily relates to the lack of local efficiency estimates. Due to the global equivalence of primal-dual gap estimators to residual type estimators, for which convergence theories are available, convergence is expected. The verification of this equivalence follows closely the derivation of estimators of that type. In fact, the discrete primal solution $u_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ acts as a substitute of the primal solution $u\in W_D^{1,p}(\Omega)$. In the case of the Poisson problem (1.1), the conforming $P1$-approximation $u_h^{p1}\in\mathcal{S}_D^1(\mathcal{T}_h)$, and $f=f_h$, due to $\Pi_h z_h^{rt}=\nabla_h u_h^{cr}$ a.e. in Ω and $\nabla u_h^{p1}-\nabla_h u_h^{cr}\perp\frac{1}{d}f_h(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})$ in $L^2(\Omega;\mathbb{R}^d)$, we have that

$$\eta_{\mathrm{gap}}^2(u_h^{p1},z_h^{rt})=\tfrac{1}{2}\|\nabla u_h^{p1}-\nabla_h u_h^{cr}\|_{2,\Omega}^2+\tfrac{1}{2d^2}\|f_h(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})\|_{2,\Omega}^2\,.$$

Letting $\delta_h:=u_h^{p1}-u_h^{cr}\in\mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $\Pi_h^{av}\delta_h\in\mathcal{S}_D^1(\mathcal{T}_h)$ be its node-averaging quasi-interpolant, so that, due to $\nabla_h\delta_h\perp\nabla_h(\mathcal{S}_D^1(\mathcal{T}_h))$ in $L^2(\Omega;\mathbb{R}^d)$, via element-wise integration-by-parts, we obtain

$$\begin{aligned}\|\nabla u_h^{p1}-\nabla_h u_h^{cr}\|_{2,\Omega}^2&=(\nabla_h\delta_h,\nabla_h\delta_h-\nabla_h\Pi_h^{av}\delta_h)_\Omega\\&=([\![\nabla u_h^{p1}\cdot n]\!],\{\delta_h-\Pi_h^{av}\delta_h\})_{\mathcal{S}_h}+(f_h,\delta_h-\Pi_h^{av}\delta_h)_\Omega\\&\le c_{av}\,\eta_{\mathrm{res},h}^2(u_h^{p1})\,\|\nabla u_h^{p1}-\nabla_h u_h^{cr}\|_{2,\Omega}\,.\end{aligned}$$

The converse estimate uses typical bubble function arguments (*cf.* [36, 58]).

1.11 Recent related results and open problems

The concepts described above can be generalized to other pairs of finite element methods:

- In [21], a discrete convex duality theory for a first-order Discontinuous Galerkin (DG) method was derived. More precisely, in [21], a discrete primal problem is given via the minimization of the functional $I_h^{dg}\colon\mathcal{L}^1(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$, for every $v_h\in\mathcal{L}^1(\mathcal{T}_h)$ defined by

$$\begin{aligned}I_h^{dg}(v_h):=&\int_\Omega\phi_h(\cdot,\nabla_h v_h)\,\mathrm{d}x+\int_\Omega\psi_h(\cdot,\Pi_h v_h)\,\mathrm{d}x\\&+\sum_{S\in\mathcal{S}_h}\tfrac{\alpha_S}{2}\|[\![v_h]\!]_S(x_S)\|_{2,S}^2+\sum_{S\in\mathcal{S}_h}\tfrac{\beta_S}{2}\|\{v_h\}_S(x_S)\|_{2,S}^2\,,\end{aligned}$$

  and a corresponding (Fenchel) dual problem is given via the maximization of the functional $D_h^{dg}\colon\mathcal{RT}^{0,dg}(\mathcal{T}_h):=(\mathcal{L}^0(\mathcal{T}_h))^d+(\mathrm{id}_{\mathbb{R}^d}-\Pi_h\mathrm{id}_{\mathbb{R}^d})\mathcal{L}^0(\mathcal{T}_h)\to\mathbb{R}\cup\{-\infty\}$, for every $y_h\in\mathcal{RT}^{0,dg}(\mathcal{T}_h)$ defined by

$$\begin{aligned}D_h^{dg}(y_h):=&-\int_\Omega\phi_h^*(\cdot,\Pi_h y_h)\,\mathrm{d}x-\int_\Omega\psi_h^*(\cdot,\operatorname{div}y_h)\,\mathrm{d}x\\&-\sum_{S\in\mathcal{S}_h}\tfrac{1}{2\alpha_S}\|[\![y_h\cdot n]\!]_S\|_{2,S}^2+\sum_{S\in\mathcal{S}_h}\tfrac{1}{2\beta_S}\|\{y_h\cdot n\}_S\|_{2,S}^2\,.\end{aligned}$$

  A discrete strong duality applies, *i.e.*, $I_h^{dg}(u_h^{dg})=D_h^{dg}(z_h^{dg})$ for some $u_h^{dg}\in\mathcal{L}^1(\mathcal{T}_h)$ and $z_h^{dg}\in\mathcal{RT}^{0,dg}(\mathcal{T}_h)$, provided that the parameters $\alpha_S,\beta_S>0$, $S\in\mathcal{S}_h$, are appropriately chosen;
- In [82], a discrete convex duality theory for a Hybrid High-Order (HHO) method was derived, thus, representing the first step towards higher-order element methods.

1.12 Outline of the article

The article is organized as follows. In Section 2, we introduce the employed notation and the relevant function and finite element spaces. In Section 3, we propose a general approach for explicit a posteriori error representation based on convex duality relations. In Section 4, we apply the general concepts of Section 3 to typical model problems including the non-linear Dirichlet problem, the obstacle problem, the Signorini problem, the Rudin–Osher–Fatemi image de-noising problem, a minimization problem jumping coefficients, the Navier–Lamé problem, and the Stokes problem. In Section 5, in the case of the non-linear Dirichlet problem, we establish the global equivalence of the primal-dual gap estimator to a residual type estimator. In Section 6, we establish that the node-averaging quasi-interpolation operator locally preserves approximation capabilities. In Section 7, we review the practical relevance of the theoretical investigations of Section 4.

2. PRELIMINARIES

2.1 Convex analysis

For a (real) Banach space X, which is equipped with the norm $\|\cdot\|_X \colon X \to \mathbb{R}_{\geq 0}$, we denote its corresponding (continuous) dual space by X^* equipped with the dual norm $\|\cdot\|_{X^*} \colon X^* \to \mathbb{R}_{\geq 0}$, defined by $\|x^*\|_{X^*} \coloneqq \sup_{\|x\|_X \leq 1} \langle x^*, x\rangle_X$ for every $x^* \in X^*$, where $\langle \cdot, \cdot \rangle_X \colon X^* \times X \to \mathbb{R}$, defined by $\langle x^*, x\rangle_X \coloneqq x^*(x)$ for every $x^* \in X^*$ and $x \in X$, denotes the duality pairing. A functional $F \colon X \to \mathbb{R} \cup \{+\infty\}$ is called *sub-differentiable* in $x \in X$, if $F(x) < \infty$ and if there exists $x^* \in X^*$, called *sub-gradient*, such that for every $y \in X$, it holds that

$$\langle x^*, y - x\rangle_X \leq F(y) - F(x)\,. \tag{2.1}$$

The *sub-differential* $\partial F \colon X \to 2^{X^*}$ of a functional $F \colon X \to \mathbb{R} \cup \{+\infty\}$ for every $x \in X$ is defined by $(\partial F)(x) \coloneqq \{x^* \in X^* \mid (2.1) \text{ holds for } x^*\}$ if $F(x) < \infty$ and $(\partial F)(x) \coloneqq \emptyset$ else.

For a given functional $F \colon X \to \mathbb{R} \cup \{\pm\infty\}$, we denote its corresponding *(Fenchel) conjugate* by $F^* \colon X^* \to \mathbb{R} \cup \{\pm\infty\}$, which for every $x^* \in X^*$ is defined by

$$F^*(x^*) \coloneqq \sup_{x \in X} \langle x^*, x\rangle_X - F(x)\,. \tag{2.2}$$

If $F \colon X \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex, and lower semi-continuous functional, then also its (Fenchel) conjugate $F^* \colon X^* \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex, and lower semi-continuous functional (*cf.* [48, p. 17]). Furthermore, for every $x^* \in X^*$ and $x \in X$ such that $F^*(x^*) + F(x)$ is well-defined, *i.e.*, critical cancellations $\infty - \infty$ do not occur, the *Fenchel–Young inequality*

$$\langle x^*, x\rangle_X \leq F^*(x^*) + F(x) \tag{2.3}$$

applies. In particular, for every $x^* \in X^*$ and $x \in X$, it holds the *Fenchel–Young identity*

$$x^* \in (\partial F)(x) \quad \Leftrightarrow \quad \langle x^*, x\rangle_X = F^*(x^*) + F(x)\,. \tag{2.4}$$

The following convexity measures for functionals play an important role in the derivation of an explicit a posteriori error representation for convex minimization problems in Section 3; for further information, we refer the reader to [33, 68, 69, 20].

Definition 2.1 (Brègman distance and symmetric Brègman distance)**.** *Let X be a (real) Banach space and $F \colon X \to \mathbb{R} \cup \{+\infty\}$ proper, i.e., $D(F) \coloneqq \{x \in X \mid F(x) < \infty\} \neq \emptyset$.*

(i) The Brègman distance $\sigma_F^2 \colon D(F) \times X \to [0, +\infty]$ *for every $x \in D(F)$ and $y \in X$ is defined by*

$$\sigma_F^2(y, x) \coloneqq F(y) - F(x) - \sup_{x^* \in (\partial F)(x)} \langle x^*, y - x\rangle_X\,,$$

where we use the convention $\sup(\emptyset) \coloneqq -\infty$.

(ii) The symmetric Brègman distance $\sigma_{F,s}^2 \colon D(F)^2 \to [0, +\infty]$ *for every $x, y \in D(F)$ is defined by*

$$\sigma_{F,s}^2(y, x) \coloneqq \sigma_F^2(y, x) + \sigma_F^2(x, y) = \inf_{x^* \in (\partial F)(x);\, y^* \in (\partial F)(y)} \langle x^* - y^*, x - y\rangle_X\,,$$

where we use the convention $\inf(\emptyset) \coloneqq +\infty$.

Definition 2.2 (Optimal convexity measure at a minimizer)**.** *Let X be a (real) Banach space and $F \colon X \to \mathbb{R} \cup \{+\infty\}$ proper. Moreover, let $x \in X$ be minimal for $F \colon X \to \mathbb{R} \cup \{+\infty\}$. Then, the* optimal convexity measure $\rho_F^2 \colon X^2 \to [0, +\infty]$ at $x \in X$ *for every $y \in X$ is defined by*

$$\rho_F^2(y, x) \coloneqq F(y) - F(x) \geq 0\,.$$

Remark 2.3. *Let X be a (real) Banach space and $F \colon X \to \mathbb{R} \cup \{+\infty\}$ proper. Moreover, let $x \in X$ be minimal for $F \colon X \to \mathbb{R} \cup \{+\infty\}$. Then, due to $0 \in (\partial F)(x)$, for every $y \in X$, it holds that*

$$\sigma_F^2(y, x) \leq \rho_F^2(y, x)\,.$$

2.2 Function spaces

Throughout the article, unless otherwise specified, we denote by $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, a bounded polyhedral Lipschitz domain, whose (topological) boundary is disjointly divided into a closed Dirichlet part Γ_D and an open Neumann part Γ_N, *i.e.*, $\Gamma_D, \Gamma_N \subseteq \partial\Omega$ and $\Gamma_D \dot\cup \Gamma_N = \partial\Omega$.

For $\ell \in \mathbb{N}$ and $p \in [1, +\infty)$, we employ the notation

$$U^p_\ell(\Omega) := \left\{ v \in L^p(\Omega;\mathbb{R}^\ell) \mid \nabla v \in L^p(\Omega;\mathbb{R}^{\ell\times d}) \right\},$$
$$Z^p_\ell(\Omega) := \left\{ y \in L^{p'}(\Omega;\mathbb{R}^{\ell\times d}) \mid \operatorname{div} y \in L^{p'}(\Omega;\mathbb{R}^\ell) \right\},$$

where the divergence needs to be understood row-wise , *i.e.*, if $y = (y_{ij})_{i\in\{1,\dots,\ell\},j\in\{1,\dots,d\}} \in Z^p_\ell(\Omega)$, then $(\operatorname{div} y)_i := \sum_{j=1}^d \partial_j y_{ij}$ for all $i = 1,\dots,\ell$. In the special case $\ell = 1$, we employ the standard notation $L^p(\Omega) := L^p(\Omega;\mathbb{R}^1)$, $W^{1,p}(\Omega) := U^p_1(\Omega)$, and $W^{p'}(\operatorname{div};\Omega) := Z^p_1(\Omega)$.

For $\ell \in \mathbb{N}$, a (Lebesgue) measurable set $M \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, and (Lebesgue) measurable functions, vector or tensor fields $u, v\colon M \to \mathbb{R}^\ell$, we employ the inner product

$$(u, v)_M := \int_M u \odot v \,\mathrm{d}x\,,$$

whenever the right-hand side is well-defined, where $\odot\colon \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}$ either denotes scalar multiplication, the Euclidean inner product, or the Frobenius inner product. For $\ell \in \mathbb{N}$, $p \in [1, +\infty]$, and a (Lebesgue) measurable set $M \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, we employ the notation $\|\cdot\|_{p,M} := \|\cdot\|_{L^p(M;\mathbb{R}^\ell)}$.

Denote by $\operatorname{tr}(\cdot)\colon U^p_\ell(\Omega) \to W^{1-\frac{1}{p},p}(\partial\Omega;\mathbb{R}^\ell)$ the trace and by $\operatorname{tr}(\cdot)n\colon Z^p_\ell(\Omega) \to W^{-\frac{1}{p'},p'}(\partial\Omega;\mathbb{R}^\ell)$[2] the normal trace operator. Then, for every $v \in U^p_\ell(\Omega)$ and $y \in Z^p_\ell(\Omega)$, it holds that

$$(\nabla v, y)_\Omega + (v, \operatorname{div} y)_\Omega = \langle \operatorname{tr}(y)n, \operatorname{tr}(v)\rangle_{\partial\Omega}\,, \tag{2.5}$$

where we abbreviate $\langle \operatorname{tr}(y)n, \operatorname{tr}(v)\rangle_\gamma := \langle \operatorname{tr}(y)n, \operatorname{tr}(v)\rangle_{W^{1-\frac{1}{p},p}(\gamma;\mathbb{R}^\ell)}$ for all $y \in W^{-\frac{1}{p'},p'}(\gamma;\mathbb{R}^\ell)$, $v \in W^{1-\frac{1}{p},p}(\gamma;\mathbb{R}^\ell)$, and $\gamma \in \{\Gamma_N, \partial\Omega\}$. Then, for $\ell \in \mathbb{N}$ and $p \in [1, +\infty]$, we employ the notation

$$U^p_{\ell,D}(\Omega) := \left\{ v \in U^p_\ell(\Omega) \mid \operatorname{tr}(v) = 0 \text{ a.e. on } \Gamma_D \right\},$$
$$Z^p_{\ell,N}(\Omega) := \left\{ y \in Z^p_\ell(\Omega) \mid \operatorname{tr}(y)n = 0 \text{ in } W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell) \right\}.$$

In what follows, we omit writing both $\operatorname{tr}(\cdot)$ and $\operatorname{tr}(\cdot)n$ in this context. For $\ell \in \mathbb{N}$ and $p \in [1, +\infty)$, we employ the notation $U^p_{\ell,0}(\Omega) := U^p_{\ell,D}(\Omega)$ if $\Gamma_D = \partial\Omega$ as well as $Z^p_{\ell,0}(\Omega) := Z^p_{\ell,N}(\Omega)$ if $\Gamma_N = \partial\Omega$. In the special case $\ell = 1$, we employ the standard notation $W^{1,p}_D(\Omega) := U^p_{1,D}(\Omega)$, $W^{1,p}_0(\Omega) := U^p_{1,0}(\Omega)$, $W^{p'}_N(\operatorname{div};\Omega) := Z^p_{1,N}(\Omega)$, and $W^{p'}_0(\operatorname{div};\Omega) := Z^p_{1,0}(\Omega)$.

2.3 Triangulations

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.* [49]). Here, $h > 0$ refers to the *average mesh-size*, *i.e.*, if we set $h_T := \operatorname{diam}(T)$ for all $T \in \mathcal{T}_h$, then, we have that $h = \frac{1}{\operatorname{card}(\mathcal{T}_h)} \sum_{T\in\mathcal{T}_h} h_T$. 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 $\{\mathcal{T}_h\}_{h>0}$. We define

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

where for every $M \subseteq \mathbb{R}^d$, we denote by $\dim_{\mathscr{H}}(M) := \inf\{d' \ge 0 \mid \mathscr{H}^{d'}(M) = 0\}$, the Hausdorff dimension. The set $\mathcal{N}_h$ contains the vertices of $\mathcal{T}_h$.

[2]Here, $W^{-\frac{1}{p'},p'}(\gamma;\mathbb{R}^\ell) := (W^{1-\frac{1}{p},p}(\gamma;\mathbb{R}^\ell))^*$ for all $\gamma \in \{\Gamma_N, \partial\Omega\}$.

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

$$\mathcal{L}^k(\mathcal{T}_h) \coloneqq \left\{ v_h \in L^\infty(\Omega) \mid v_h|_T \in \mathbb{P}^k(T) \text{ for all } T \in \mathcal{T}_h \right\},$$
$$\mathcal{S}^k(\mathcal{T}_h) \coloneqq \mathcal{L}^k(\mathcal{T}_h) \cap C^0(\overline{\Omega}).$$

In addition, we set $\mathcal{S}^k_D(\mathcal{T}_h) \coloneqq \mathcal{S}^k(\mathcal{T}_h) \cap W^{1,p}_D(\Omega)$ and $\mathcal{S}^k_0(\mathcal{T}_h) \coloneqq \mathcal{S}^k(\mathcal{T}_h) \cap W^{1,p}_0(\Omega)$. The (local) L^2-projection $\Pi_h \colon L^1(\Omega;\mathbb{R}^\ell) \to (\mathcal{L}^0(\mathcal{T}_h))^\ell$ onto element-wise constant functions, vector or tensor fields, respectively, for every $v \in L^1(\Omega;\mathbb{R}^\ell)$ is defined by

$$\Pi_h v|_T \coloneqq \fint_T v \,\mathrm{d}x \quad \text{for all } T \in \mathcal{T}_h.$$

The element-wise gradient $\nabla_h \colon (\mathcal{L}^1(\mathcal{T}_h))^\ell \to (\mathcal{L}^0(\mathcal{T}_h))^{\ell \times d}$ is given via the element-wise application of the gradient operator, *i.e.*, for every $v_h \in (\mathcal{L}^1(\mathcal{T}_h))^\ell$, we have that $\nabla_h v_h|_T \coloneqq \nabla(v_h|_T)$ for all $T \in \mathcal{T}_h$.

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

$$\mathcal{L}^m(\mathcal{M}_h) \coloneqq \left\{ v_h \in L^\infty(\cup\mathcal{M}_h) \mid v_h|_T \in \mathbb{P}^m(S) \text{ for all } S \in \mathcal{M}_h \right\}.$$

The (local) L^2-projection $\pi_h \colon L^1(\cup\mathcal{S}_h;\mathbb{R}^\ell) \to (\mathcal{L}^0(\mathcal{S}_h))^\ell$ onto side-wise constant functions, vector, or tensor fields, respectively, for every $v \in L^1(\cup\mathcal{S}_h;\mathbb{R}^\ell)$ is defined by

$$\pi_h v|_S \coloneqq \fint_S v \,\mathrm{d}s \quad \text{for all } S \in \mathcal{S}_h.$$

2.3.1 Crouzeix–Raviart element

The *Crouzeix–Raviart finite element space* (*cf.* [44]) is defined as the space of element-wise affine functions that are continuous in the barycenters of inner element sides, *i.e.*,[3]

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

The Crouzeix–Raviart finite element space with homogeneous Dirichlet boundary condition on Γ_D is defined as the space of Crouzeix–Raviart finite element functions that vanish in the barycenters of boundary element sides that belong to Γ_D, *i.e.*,

$$\mathcal{S}^{1,cr}_D(\mathcal{T}_h) \coloneqq \left\{ v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h) \mid \pi_h [\![v_h]\!]_S = 0 \text{ in } S \text{ for all } S \in \mathcal{S}_h^{\Gamma_D} \right\}.$$

We employ the notation $\mathcal{S}^{1,cr}_0(\mathcal{T}_h) = \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ if $\Gamma_D = \partial\Omega$. The functions $\varphi_S \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, that satisfy the Kronecker property $\varphi_S(x_{S'}) = \delta_{S,S'}$ for all $S, S' \in \mathcal{S}_h$, form a basis of $\mathcal{S}^{1,cr}(\mathcal{T}_h)$. Then, the functions $\varphi_S \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$, $S \in \mathcal{S}_h \setminus \mathcal{S}_h^{\Gamma_D}$, form a basis of $\mathcal{S}^{1,cr}_D(\mathcal{T}_h)$. For $\ell \in \mathbb{N}$, we employ the notation

$$U^{cr}_\ell(\mathcal{T}_h) \coloneqq (\mathcal{S}^{1,cr}(\mathcal{T}_h))^\ell, \qquad U^{cr}_{\ell,D}(\mathcal{T}_h) \coloneqq (\mathcal{S}^{1,cr}_D(\mathcal{T}_h))^\ell, \qquad U^{cr}_{\ell,0}(\mathcal{T}_h) \coloneqq (\mathcal{S}^{1,cr}_0(\mathcal{T}_h))^\ell.$$

The (Fortin) quasi-interpolation operator $\Pi^{cr}_h \colon U^p_\ell(\Omega) \to U^{cr}_\ell(\mathcal{T}_h)$, for every $v \in U^p_\ell(\Omega)$ is defined by

$$\Pi^{cr}_h v \coloneqq \sum_{S \in \mathcal{S}_h} v_S \, \varphi_S, \quad \text{where } v_S \coloneqq \fint_S v \,\mathrm{d}s, \tag{2.6}$$

preserves averages of gradients and moments, *i.e.*, for every $v \in U^p_\ell(\Omega)$, it holds that

$$\nabla_h \Pi^{cr}_h v = \Pi_h \nabla v \quad (\mathcal{L}^0(\mathcal{T}_h))^{\ell \times d}, \tag{2.7}$$
$$\pi_h \Pi^{cr}_h v = \pi_h v \qquad (\mathcal{L}^0(\mathcal{S}_h))^\ell. \tag{2.8}$$

In particular, from (2.8), it follows that $\Pi^{cr}_h(U^p_{\ell,D}(\Omega)) \subseteq U^{cr}_{\ell,D}(\mathcal{T}_h)$ and $\Pi^{cr}_h(U^p_{\ell,0}(\Omega)) \subseteq U^{cr}_{\ell,0}(\mathcal{T}_h)$.

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

2.3.2 Raviart–Thomas element

The *(lowest order) Raviart–Thomas finite element space* (*cf.* [72]) is defined as the space of element-wise affine vector fields that have continuous constant normal components on inner elements sides, *i.e.*,[4]

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

The Raviart–Thomas finite element space with homogeneous slip boundary condition on Γ_N is defined as the space of Raviart–Thomas vector fields whose normal components vanish on Γ_N, *i.e.*,

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

We employ the notation $\mathcal{RT}^0_0(\mathcal{T}_h) := \mathcal{RT}^0_N(\mathcal{T}_h)$ if $\Gamma_N = \partial\Omega$. The vector fields $\psi_S \in \mathcal{RT}^0(\mathcal{T}_h)$, $S \in \mathcal{S}_h$, that satisfy 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 $T_+ \cap T_- = S \in \mathcal{S}_h$, form a basis of $\mathcal{RT}^0(\mathcal{T}_h)$. Then, the vector fields $\psi_S \in \mathcal{RT}^0_N(\mathcal{T}_h)$, $S \in \mathcal{S}_h \setminus \Gamma_N$, form a basis of $\mathcal{RT}^0_N(\mathcal{T}_h)$. For $\ell \in \mathbb{N}$, we employ the notations

$$Z^{rt}_\ell(\mathcal{T}_h) := \big\{ y = (y_{ij})_{i\in\{1,\dots,\ell\},j\in\{1,\dots,d\}} \,|\, (y_{ij})_{j\in\{1,\dots,d\}} \in \mathcal{RT}^0(\mathcal{T}_h) \text{ for all } i = 1,\dots,\ell \big\}\,,$$
$$Z^{rt}_{\ell,N}(\mathcal{T}_h) := \big\{ y = (y_{ij})_{i\in\{1,\dots,\ell\},j\in\{1,\dots,d\}} \,|\, (y_{ij})_{j\in\{1,\dots,d\}} \in \mathcal{RT}^0_N(\mathcal{T}_h) \text{ for all } i = 1,\dots,\ell \big\}\,,$$
$$Z^{rt}_{\ell,0}(\mathcal{T}_h) := \big\{ y = (y_{ij})_{i\in\{1,\dots,\ell\},j\in\{1,\dots,d\}} \,|\, (y_{ij})_{j\in\{1,\dots,d\}} \in \mathcal{RT}^0_0(\mathcal{T}_h) \text{ for all } i = 1,\dots,\ell \big\}\,.$$

The corresponding (Fortin) quasi-interpolation operator $\Pi^{rt}_h \colon W^{1,1}(\Omega;\mathbb{R}^{\ell\times d}) \to Z^{rt}_\ell(\mathcal{T}_h)$, for every $y \in W^{1,1}(\Omega;\mathbb{R}^{\ell\times d})$ is defined by

$$\Pi^{rt}_h y := \sum_{S\in\mathcal{S}_h} y_S\, \psi_S\,, \qquad \text{where } y_S := \fint_S y n_S \,\mathrm{d}s\,, \tag{2.9}$$

preserves averages of divergences and normal traces, *i.e.*, for every $y \in W^{1,1}(\Omega;\mathbb{R}^{\ell\times d})$, it holds that

$$\operatorname{div} \Pi^{rt}_h y = \Pi_h \operatorname{div} y \quad (\mathcal{L}^0(\mathcal{T}_h))^\ell\,, \tag{2.10}$$
$$[\Pi^{rt}_h y] n = [\pi_h y] n \quad (\mathcal{L}^0(\mathcal{S}_h))^\ell\,. \tag{2.11}$$

In particular, from (2.11), it follows that $\Pi^{rt}_h(Z^p_{\ell,N}(\Omega)) \subseteq Z^{rt}_{\ell,N}(\mathcal{T}_h)$ and $\Pi^{rt}_h(Z^p_{\ell,0}(\Omega)) \subseteq Z^{rt}_{\ell,0}(\mathcal{T}_h)$.

2.3.3 Discrete integration-by-parts formula

For every $v_h \in U^{cr}_\ell(\mathcal{T}_h)$ and $y_h \in Z^{rt}_\ell(\mathcal{T}_h)$, there holds the *discrete integration-by-parts formula*

$$(\nabla_h v_h, \Pi_h y_h)_\Omega + (\Pi_h v_h,\, \operatorname{div} y_h)_\Omega = (\pi_h v_h, y_h n)_{\partial\Omega}\,. \tag{2.12}$$

which follows from the fact that for every $y_h \in Z^{rt}_\ell(\mathcal{T}_h)$, it holds that $y_h|_T n_T = \text{const}$ on ∂T for all $T \in \mathcal{T}_h$ and $[\![y_h n]\!]_S = 0$ on S for all $S \in \mathcal{S}_h^i$, and for every $v_h \in U^{cr}_\ell(\mathcal{T}_h)$, it holds that $\pi_h [\![v_h]\!]_S = 0$ for all $S \in \mathcal{S}_h^i$. In [38, 59, 41, 21, 16, 22], the discrete integration-by-parts formula (2.12) formed a cornerstone in the derivation of a discrete convex duality theory and, as such, plays a central role in the derivation of the results presented below. Appealing to [19, Sec. 2.4] and [24, Subsec. 2.3.3], there hold the *discrete Helmholtz decompositions*

$$(\mathcal{L}^0(\mathcal{T}_h))^{\ell\times d} = \ker(\operatorname{div}|_{Z^{rt}_{\ell,N}(\mathcal{T}_h)}) \oplus \nabla_h(U^{cr}_{\ell,D}(\mathcal{T}_h))\,, \tag{2.13}$$
$$(\mathcal{L}^0(\mathcal{T}_h))^{\ell} = \ker(\nabla_h|_{U^{cr}_{\ell,D}(\mathcal{T}_h)}) \oplus \operatorname{div}(Z^{rt}_{\ell,N}(\mathcal{T}_h))\,. \tag{2.14}$$

[4] For every inner side $S \in \mathcal{S}_h^i$, the normal jump is defined by $[\![y_h n]\!]_S := y_h|_{T_+} n_{T_+} + y_h|_{T_-} n_{T_-}$ 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 boundary side $S \in \mathcal{S}_h \cap \partial\Omega$, the normal jump is defined by $[\![y_h n]\!]_S := y_h|_T n$ on S, where $T \in \mathcal{T}_h$ satisfies $S \subseteq \partial T$.

3. EXACT A POSTERIORI ERROR ESTIMATION FOR CONVEX MINIMIZATION PROBLEMS

3.1 Continuous convex duality

Primal problem. Let $\phi\colon \Omega \times \mathbb{R}^{\ell\times d} \to \mathbb{R}\cup\{+\infty\}$ and $\psi\colon \Omega\times\mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$ be (Lebesgue) measurable functions such that for a.e. $x\in\Omega$, the functions $\phi(x,\cdot)\colon \mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi(x,\cdot)\colon \mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$ are proper, convex, and lower semi-continuous such that for every $y\in L^p(\Omega;\mathbb{R}^{\ell\times d})$ and $v\in L^p(\Omega;\mathbb{R}^{\ell})$, the following integrals exist and are finite or infinity, *i.e.*,

$$\int_\Omega \phi(\cdot,y)\,\mathrm{d}x, \int_\Omega \psi(\cdot,v)\,\mathrm{d}x \in \mathbb{R}\cup\{+\infty\}\,.$$

Moreover, let $g\in W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$ be given Neumann boundary data and let $u_D\in W^{1-\frac{1}{p},p}(\Gamma_D;\mathbb{R}^\ell)$ be given Dirichlet boundary data. We examine the minimization problem of the functional $I\colon U_\ell^p(\Omega)\to\mathbb{R}\cup\{+\infty\}$, for every $v\in U_\ell^p(\Omega)$ defined by

$$I(v) := \int_\Omega \phi(\cdot,\nabla v)\,\mathrm{d}x + \int_\Omega \psi(\cdot,v)\,\mathrm{d}x - \langle g,v\rangle_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(v)\,, \tag{3.1}$$

where $I^{\Gamma_D}_{\{u_D\}}\colon W^{1-\frac{1}{p},p}(\Gamma_D;\mathbb{R}^\ell)\to\mathbb{R}\cup\{+\infty\}$ for every $\widehat{v}\in W^{1-\frac{1}{p},p}(\Gamma_D;\mathbb{R}^\ell)$ is defined by

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

In what follows, we refer to the minimization of $I\colon U_\ell^p(\Omega)\to\mathbb{R}\cup\{+\infty\}$ as the *primal problem.*

Dual problem. Let $\phi^*\colon \Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi^*\colon\Omega\times\mathbb{R}^\ell\to\mathbb{R}\cup\{+\infty\}$ be the Fenchel conjugates to $\phi\colon\Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon\Omega\times\mathbb{R}^\ell\to\mathbb{R}\cup\{+\infty\}$, respectively, with respect to the second argument and assume that for every $y\in L^{p'}(\Omega;\mathbb{R}^\ell)$ and $v\in L^{p'}(\Omega)$, the following integrals exist and are finite or infinity, *i.e.*,

$$\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x, \int_\Omega \psi^*(\cdot,v)\,\mathrm{d}x \in \mathbb{R}\cup\{+\infty\}\,.$$

Then, a *(Fenchel) dual problem* to the minimization of (3.1) is given via the maximization of the functional $D\colon Z_\ell^p(\Omega)\to\mathbb{R}\cup\{-\infty\}$, for every $y\in Z_\ell^p(\Omega)$ defined by

$$D(y) := -\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x - \int_\Omega \psi^*(\cdot,\operatorname{div} y)\,\mathrm{d}x + \langle yn,u_D\rangle_{\Gamma_D} - I^{\Gamma_N}_{\{g\}}(yn)\,, \tag{3.2}$$

where $I^{\Gamma_N}_{\{g\}}\colon W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)\to\mathbb{R}\cup\{+\infty\}$ for every $\widehat{y}\in W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$ is defined by

$$I^{\Gamma_N}_{\{g\}}(\widehat{y}) := \begin{cases} 0 & \text{if } \widehat{y}=g \text{ in } W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)\,,\\ +\infty & \text{else}\,.\end{cases}$$

We always assume that $\phi\colon\Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon\Omega\times\mathbb{R}^\ell\to\mathbb{R}\cup\{+\infty\}$ are such that (3.1) admits at least one minimizer $u\in U_\ell^p(\Omega)$, called *primal solution*, and that (3.2) admits at least one maximizer $z\in Z_\ell^p(\Omega)$, called *dual solution.* The derivation of a weak duality relation between of (3.1) and (3.2) can be found in the proof of the following results that also establishes the equivalence of a strong duality relation and convex optimality relations.

Proposition 3.1 (Strong duality and convex duality relations)**.** *The following statements apply:*

(i) A weak duality relation *applies, i.e.,*

$$\inf_{v\in U_\ell^p(\Omega)} I(v) \geq \sup_{y\in Z_\ell^p(\Omega)} D(y)\,. \tag{3.3}$$

(ii) A strong duality relation *applies, i.e.,*

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

if and only if the convex optimality relations *apply, i.e.,*

$$z:\nabla u = \phi^*(\cdot,z)+\phi(\cdot,\nabla u) \qquad \text{a.e. in } \Omega\,, \tag{3.5}$$

$$\operatorname{div} z\cdot u = \psi^*(\cdot,\operatorname{div} z)+\psi(\cdot,u) \qquad \text{a.e. in } \Omega\,. \tag{3.6}$$

Proof. ad (i). For every $v \in U^p_\ell(\Omega)$ and $y \in Z^p_\ell(\Omega)$, by the Fenchel–Young inequality (*cf.* (2.3)), we have that

$$\begin{aligned} \phi(\cdot,\nabla v) &\geq y : \nabla v - \phi^*(\cdot, y) && \text{a.e. in } \Omega\,, \\ \psi(\cdot, v) &\geq \operatorname{div} y \cdot v - \psi^*(\cdot, \operatorname{div} y) && \text{a.e. in } \Omega\,. \end{aligned} \tag{3.7}$$

Adding the two inequalities in (3.7), using the integration-by-parts formula (2.5), for every $v \in U^p_\ell(\Omega)$ with $v = u_D$ a.e. on Γ_D and $y \in Z^p_\ell(\Omega)$ with $yn = g$ in $W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$, we find that

$$\begin{aligned} I(v) &= \int_\Omega \phi(\cdot,\nabla v)\,\mathrm{d}x + \int_\Omega \psi(\cdot, v)\,\mathrm{d}x - \langle g, v\rangle_{\Gamma_N} \\ &\geq -\int_\Omega \phi^*(\cdot, y)\,\mathrm{d}x - \int_\Omega \psi^*(\cdot, \operatorname{div} y)\,\mathrm{d}x + \langle yn, u_D\rangle_{\Gamma_D} = D(y)\,. \end{aligned}$$

On the other hand, for every $v \in U^p_\ell(\Omega)$ such that $I^{\Gamma_D}_{\{u_D\}}(v) = +\infty$, we have that $I(v) = +\infty$, and for every $y \in Z^p_\ell(\Omega)$ such that $I^{\Gamma_N}_{\{g\}}(yn) = +\infty$, we have that $D(z) = -\infty$.

ad (ii). The strong duality relation (3.4) is equivalent to

$$\int_\Omega \phi^*(\cdot, z) - z : \nabla u + \phi(\cdot,\nabla u)\,\mathrm{d}x + \int_\Omega \psi^*(\cdot, \operatorname{div} z) - \operatorname{div} z \cdot u + \psi(\cdot, u)\,\mathrm{d}x = 0\,.$$

Therefore, due to (3.7), we find that the strong duality relation (3.4) is equivalent to the convex optimality relations (3.5), (3.6). □

Remark 3.2 (Equivalent convex optimality relations)**.** *(i) If $\phi(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the convex optimality relation (3.5) is equivalent to*

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

(ii) If $\phi^*(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the convex optimality relation (3.5) is equivalent to

$$\nabla u = D\phi^*(\cdot, z) \quad \text{a.e. in } \Omega\,; \tag{3.9}$$

(iii) If $\psi(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, then, by the Fenchel–Young identity (cf. (2.4)), by the Fenchel–Young identity (cf. (2.4)), the convex optimality relation (3.6) is equivalent to

$$\operatorname{div} z = D\psi(\cdot, u) \quad \text{a.e. in } \Omega\,; \tag{3.10}$$

(iv) If $\psi^*(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, then, by the Fenchel–Young identity (cf. (2.4)), by the Fenchel–Young identity (cf. (2.4)), the convex optimality relation (3.6) is equivalent to

$$u = D\psi^*(\cdot, \operatorname{div} z) \quad \text{a.e. in } \Omega\,. \tag{3.11}$$

The convex duality relations (3.5), (3.6) motivate introducing the *primal-dual gap estimator* $\eta^2_{\mathrm{gap}}\colon U^p_\ell(\Omega) \times Z^p_\ell(\Omega) \to [0,+\infty]$, for every $(v,y)^\top \in U^p_\ell(\Omega) \times Z^p_\ell(\Omega)$ defined by

$$\eta^2_{\mathrm{gap}}(v,y) \coloneqq I(v) - D(y)\,. \tag{3.12}$$

Note that the sign of the estimator (3.12) is a consequence of the weak duality relation (3.3).

Together with the optimal convexity measures (*cf.* Definition 2.2) $\rho^2_I\colon U^p_\ell(\Omega) \to [0,+\infty]$ of (3.1) at a primal solution $u \in U^p_\ell(\Omega)$ and $\rho^2_{-D}\colon Z^p_\ell(\Omega) \to [0,+\infty]$ of the negative of (3.2) at a dual solution $z \in Z^p_\ell(\Omega)$, we arrive at the following *generalized Prager–Synge identity.*

Theorem 3.3 (Generalized Prager–Synge identity)**.** *If the strong duality relation (3.4) applies, then the following statements apply:*

(i) For every $v \in U^p_\ell(\Omega)$ and $y \in Z^p_\ell(\Omega)$, for the total error*, we have that*

$$\rho^2_{\mathrm{tot}}(v,y) \coloneqq \rho^2_I(v,u) + \rho^2_{-D}(y,z) = \eta^2_{\mathrm{gap}}(v,y)\,. \tag{3.13}$$

(ii) For every $v \in U^p_\ell(\Omega)$ with $v = u_D$ a.e. in Γ_D and $y \in Z^p_\ell(\Omega)$ with $yn = g$ in $W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$, we have that

$$\begin{aligned} \eta^2_{\mathrm{gap}}(v,y) = &\int_\Omega \phi(\cdot,\nabla v) - \nabla v : y + \phi^*(\cdot, y)\,\mathrm{d}x \\ &+ \int_\Omega \psi(\cdot, v) - v \cdot \operatorname{div} y + \psi^*(\cdot, \operatorname{div} y)\,\mathrm{d}x\,. \end{aligned} \tag{3.14}$$

Proof. ad (i). Due to (3.4), Definition 2.2, and (3.12), for every $v \in U_\ell^p(\Omega)$ and $y \in Z_\ell^p(\Omega)$, we have that

$$\rho_{\mathrm{tot}}^2(v,y) = \rho_I^2(v,u) + \rho_{-D}^2(y,z) = I(v) - I(u) + D(z) - D(y) = \eta_{\mathrm{gap}}^2(v,y)\,.$$

ad (ii). The identitiy (3.14) follows from (3.1), (3.2), the integration-by-parts formula (2.5), and that for every $v \in U_\ell^p(\Omega)$ with $v = u_D$ a.e. in Γ_D and $y \in Z_\ell^p(\Omega)$ with $yn = g$ in $W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$, it holds that

$$\langle yn, v\rangle_{\partial\Omega} = \langle g, v\rangle_{\Gamma_N} + \langle yn, u_D\rangle_{\Gamma_D}\,. \qquad \square$$

Remark 3.4. *(i) By the Fenchel–Young inequality (cf. (2.3)), the integrands in the representation (3.14) are non-negative, i.e., for every $v \in U_\ell^p(\Omega)$ and $y \in Z_\ell^p(\Omega)$, we have that*

$$\begin{aligned} \phi(\cdot,\nabla v) - \nabla v : y + \phi^*(\cdot, y) &\geq 0 \qquad \text{a.e. in } \Omega\,,\\ \psi(\cdot, v) - v \cdot \operatorname{div} y + \psi^*(\cdot, \operatorname{div} y) &\geq 0 \qquad \text{a.e. in } \Omega\,. \end{aligned}$$

and, thus, are suitable as local refinement indicators in an adaptive refinement procedure. Apart from that, due to Proposition 3.1(ii), for every $v \in U_\ell^p(\Omega)$ with $v = u_D$ a.e. in Γ_D and $y \in Z_\ell^p(\Omega)$ with $yn = g$ in $W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$, we have that

$$\begin{aligned} \phi(\cdot,\nabla v) - \nabla v : y + \phi^*(\cdot, y) &= 0 \qquad \text{a.e. in } \Omega\,,\\ \psi(\cdot, v) - v \cdot \operatorname{div} y + \psi^*(\cdot, \operatorname{div} y) &= 0 \qquad \text{a.e. in } \Omega\,, \end{aligned}$$

if and only if

$$I(v) = D(y)\,,$$

i.e., if $v \in U_\ell^p(\Omega)$ is minimal for (3.1) and $y \in Z_\ell^p(\Omega)$ is maximal for (3.2). In other words, given the strong duality relation (3.4), the primal-dual gap estimator measures how well approximations $v \in U_\ell^p(\Omega)$ and $y \in Z_\ell^p(\Omega)$ satisfy the convex optimality relations (3.5), (3.6). This is an advantage compared to residual type a posteriori error estimators which traditionally measure how well an approximation $v \in U_\ell^p(\Omega)$ satisfies the strong formulation of optimality conditions and, consequently, employ classical derivatives of the energy densities. The convex optimality relations (3.5), (3.6), however, do not require any regularity of the energy densities. This makes the primal-dual gap estimator a predestined a posteriori error estimator, in particular, for non-differentiable convex minimization problems.

(ii) Due to Remark 2.3, from Theorem 3.3(i), for every $v \in U_\ell^p(\Omega)$ and $y \in Z_\ell^p(\Omega)$, it follows that

$$\sigma_I^2(v,u) + \sigma_{-D}^2(y,z) \leq \eta_{\mathrm{gap}}^2(v,y)\,.$$

Since the dual problem to the minimization of the negative of (3.2), in turn, consists in the maximization of the negative of (3.1), the roles of the primal problem and the dual problem may be interchanged. An advantage of Theorem 3.3 consists in the fact that it yields reliable and efficient a posteriori error estimators for both the primal problem and the dual problem.

Remark 3.5. *Theorem 3.3 shows that for every $y \in Z_\ell^p(\Omega)$, the estimator $\eta_{I,y}^2 \coloneqq (v \mapsto \eta_{\mathrm{gap}}^2(v,y)) \colon U_\ell^p(\Omega) \to [0,+\infty]$ satisfies*

$$\rho_{\mathrm{tot}}^2(v,y) = \eta_{I,y}^2(v) \quad \text{for all } v \in U_\ell^p(\Omega)\,, \tag{3.15}$$

and for each $v \in U_\ell^p(\Omega)$, the estimator $\eta_{-D,v}^2 \coloneqq (y \mapsto \eta_{\mathrm{gap}}^2(v,y)) \colon Z_\ell^p(\Omega) \to [0,+\infty]$ satisfies

$$\rho_{\mathrm{tot}}^2(v,y) = \eta_{-D,v}^2(y) \quad \text{for all } y \in Z_\ell^p(\Omega)\,. \tag{3.16}$$

For the a posteriori error estimators (3.15) and (3.16) for being numerically practicable, it is necessary to have a computationally cheap way to obtain sufficiently accurate approximations of the dual solution (for (3.15)) and/or of the primal solution (for (3.16)), respectively. In Section 3.2, resorting to (discrete) convex duality relations between a non-conforming Crouzeix–Raviart approximation of the primal problem and a Raviart–Thomas approximation of the dual problem, we arrive at discrete reconstruction formulas, called *generalized Marini formula* (*cf.* [62, 16]).

3.2 Discrete convex minimization problem and discrete convex duality

Discrete primal problem. Let $\phi_h\colon \Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi_h\colon \Omega\times\mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$ be approximations of $\phi\colon \Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi\colon \Omega\times\mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$, respectively, such that for a.e. $x\in\Omega$, the functions $\phi_h(x,\cdot)\colon \mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi_h(x,\cdot)\colon \mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$ are proper, convex, and lower semi-continuous and $\phi_h(\cdot,r),\psi_h(\cdot,s)\in\mathcal{L}^0(\mathcal{T}_h)$ for all $r\in\mathbb{R}^d$ and $s\in\mathbb{R}$. Moreover, let $g_h\in(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}))^\ell$ and $u_D^h\in(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_D}))^\ell$ be approximations of the Neumann boundary data $g\in W^{-\frac{1}{p'},p'}(\Gamma_N;\mathbb{R}^\ell)$ and the Dirichlet boundary data $u_D\in W^{1-\frac{1}{p},p}(\Gamma_D;\mathbb{R}^\ell)$, respectively. We examine the minimization of the functional $I_h^{cr}\colon U_\ell^{cr}(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$, for every $v_h\in U_\ell^{cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) \coloneqq \int_\Omega \phi_h(\cdot,\nabla_h v_h)\,\mathrm{d}x + \int_\Omega \psi_h(\cdot,\Pi_h v_h)\,\mathrm{d}x - (g_h,\pi_h v_h)_{\Gamma_N} + I^{\Gamma_D}_{\{u_D^h\}}(\pi_h v_h)\,, \tag{3.17}$$

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

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

In what follows, we refer to the minimization of $I_h^{cr}\colon U_\ell^{cr}(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$ as the *discrete primal problem.*

Discrete dual problem. A corresponding *discrete (Fenchel) dual problem* to the minimization of (3.17) is found to be given via the maximization of the functional $D_h^{rt}\colon Z_\ell^{rt}(\mathcal{T}_h)\to\mathbb{R}\cup\{-\infty\}$, for every $y_h\in Z_\ell^{rt}(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h) \coloneqq -\int_\Omega \phi_h^*(\cdot,\Pi_h y_h)\,\mathrm{d}x - \int_\Omega \psi_h^*(\cdot,\operatorname{div} y_h)\,\mathrm{d}x + (y_h n,u_D^h)_{\Gamma_D} - I^{\Gamma_N}_{\{g_h\}}(y_h n)\,, \tag{3.18}$$

where $I^{\Gamma_N}_{\{g_h\}}\colon(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}))^\ell\to\mathbb{R}\cup\{+\infty\}$ for every $\widehat{y}_h\in(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}))^\ell$ is defined by

$$I^{\Gamma_N}_{\{g_h\}}(\widehat{y}_h) \coloneqq \begin{cases} 0 & \text{if } \widehat{y}_h = g_h \text{ a.e. in } \Gamma_N\,,\\ +\infty & \text{else}\,. \end{cases}$$

We will always assume that $\phi_h\colon\Omega\times\mathbb{R}^{\ell\times d}\to\mathbb{R}\cup\{+\infty\}$ and $\psi_h\colon\Omega\times\mathbb{R}^{\ell}\to\mathbb{R}\cup\{+\infty\}$ are such that (3.17) admits at least one minimizer $u_h^{cr}\in U_\ell^{cr}(\mathcal{T}_h)$, called *discrete primal solution*, and that (3.18) admits at least one maximizer $z_h^{rt}\in Z_\ell^{rt}(\mathcal{T}_h)$, called *discrete dual solution.* The derivation of the discrete dual problem (3.18) can be found in the proof of the following proposition that also establishes the equivalence of a discrete strong duality relation and discrete convex optimality relations.

Proposition 3.6 (Strong duality and convex duality relations)**.** *The following statements apply:*
(i) A discrete weak duality relation *applies, i.e.,*

$$\inf_{v_h\in U_\ell^{cr}(\mathcal{T}_h)} I_h^{cr}(v_h) \geq \sup_{y_h\in Z_\ell^{rt}(\mathcal{T}_h)} D_h^{rt}(y_h)\,. \tag{3.19}$$

(ii) A discrete strong duality relation *applies, i.e.,*

$$I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})\,, \tag{3.20}$$

if and only if discrete convex optimality relations *apply, i.e.,*

$$\Pi_h z_h^{rt} : \nabla_h u_h^{cr} = \phi_h^*(\cdot,\Pi_h z_h^{rt}) + \phi_h(\cdot,\nabla_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,, \tag{3.21}$$

$$\operatorname{div} z_h^{rt}\cdot\Pi_h u_h^{cr} = \psi_h^*(\cdot,\operatorname{div} z_h^{rt}) + \psi_h(\cdot,\Pi_h u_h^{cr}) \qquad \text{a.e. in } \Omega\,. \tag{3.22}$$

Proof. ad (i). For every $v_h\in U_\ell^{cr}(\mathcal{T}_h)$ and $y_h\in Z_\ell^{rt}(\mathcal{T}_h)$, by the Fenchel–Young inequality (*cf.* (2.3)), we have that

$$\begin{aligned} \phi_h(\cdot,\nabla_h v_h) &\geq \Pi_h y_h : \nabla_h v_h - \phi_h^*(\cdot,\Pi_h y_h) \qquad \text{a.e. in } \Omega\,,\\ \psi_h(\cdot,\Pi_h v_h) &\geq \operatorname{div} y_h\cdot\Pi_h v_h - \psi_h^*(\cdot,\operatorname{div} y_h) \qquad \text{a.e. in } \Omega\,. \end{aligned} \tag{3.23}$$

Adding the two inequalities in (3.23), using the discrete integration-by-parts formula (2.12), for every $v_h \in U_\ell^{cr}(\mathcal{T}_h)$ with $\pi_h v_h = u_D^h$ a.e. on Γ_D and $y_h \in Z_\ell^{rt}(\mathcal{T}_h)$ with $y_h n = g_h$ a.e. on Γ_N, we find that

$$\begin{aligned} I_h^{cr}(v_h) &= \int_\Omega \phi_h(\cdot, \nabla_h v_h)\,\mathrm{d}x + \int_\Omega \psi_h(\cdot, \Pi_h v_h)\,\mathrm{d}x - (g_h, \pi_h v_h)_{\Gamma_N} \\ &\geq -\int_\Omega \phi_h^*(\cdot, \Pi_h y_h)\,\mathrm{d}x - \int_\Omega \psi_h^*(\cdot, \operatorname{div} y_h)\,\mathrm{d}x + (y_h n, u_D^h)_{\Gamma_D} = D_h^{rt}(y_h)\,. \end{aligned}$$

On the other hand, for every $v_h \in U_\ell^{cr}(\mathcal{T}_h)$ such that $I^{\Gamma_D}_{\{u_D^h\}}(\pi_h v_h) = +\infty$, we have that $I_h^{cr}(v_h) = +\infty$, and for every $y_h \in Z_\ell^{rt}(\mathcal{T}_h)$ such that $I^{\Gamma_N}_{\{g_h\}}(y_h n) = +\infty$, we have that $D_h^{rt}(z_h^{rt}) = -\infty$.

ad (ii). The discrete strong duality relation (3.20) is equivalent to

$$\begin{aligned} &\int_\Omega \phi_h^*(\cdot, \Pi_h z_h^{rt}) - \Pi_h z_h^{rt} : \nabla_h u_h^{cr} + \phi_h(\cdot, \nabla_h u_h^{cr})\,\mathrm{d}x \\ &\quad + \int_\Omega \psi_h^*(\cdot, \operatorname{div} z_h^{rt}) - \operatorname{div} z_h^{rt} \cdot \Pi_h u_h^{cr} + \psi_h(\cdot, \Pi_h u_h^{cr})\,\mathrm{d}x = 0\,. \end{aligned}$$

Therefore, due to (3.23), we find that the discrete strong duality relation (3.20) is equivalent to the discrete convex optimality relations (3.21), (3.22). □

Remark 3.7 (Equivalent discrete convex optimality relations)**.** *(i) If $\phi_h(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the discrete convex optimality relation (3.21) is equivalent to*

$$\Pi_h z_h^{rt} = D\phi_h(\cdot, \nabla_h u_h^{cr}) \quad \text{a.e. in } \Omega\,; \tag{3.24}$$

(ii) If $\phi_h^*(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the discrete convex optimality relation (3.22) is equivalent to

$$\nabla_h u_h^{cr} = D\phi_h^*(\cdot, \Pi_h z_h^{rt}) \quad \text{a.e. in } \Omega\,; \tag{3.25}$$

(iii) If $\psi_h(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the discrete convex optimality relation (3.22) is equivalent to

$$\operatorname{div} z_h^{rt} = D\psi_h(\cdot, \Pi_h u_h^{cr}) \quad \text{a.e. in } \Omega\,; \tag{3.26}$$

(iv) If $\psi_h^*(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, by the Fenchel–Young identity (cf. (2.4)), the discrete convex optimality relation (3.22) is equivalent to

$$\Pi_h u_h^{cr} = D\psi_h^*(\cdot, \operatorname{div} z_h^{rt}) \quad \text{a.e. in } \Omega\,. \tag{3.27}$$

The relations (3.24)–(3.27) motivate the following discrete reconstruction formulas for a discrete dual solution $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ from a discrete primal solution $u_h^{cr} \in U_\ell^{cr}(\mathcal{T}_h)$ and vice versa, called *generalized Marini formulas* (*cf.* [62, 16]).

Proposition 3.8 (Generalized Marini formulas)**.** *The following statements apply:*

(i) If $\phi_h(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ and $\psi_h(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, then, given a minimizer $u_h^{cr} \in U_\ell^{cr}(\mathcal{T}_h)$ of (3.17), a maximizer $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ of (3.18) is given by

$$z_h^{rt} = D\phi_h(\cdot, \nabla_h u_h^{cr}) + \frac{D\psi_h(\cdot, \Pi_h u_h^{cr})}{d} \otimes (\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,, \tag{3.28}$$

and a discrete strong duality relation, i.e., (3.20), applies.

(ii) If $\phi_h^*(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ and $\psi_h^*(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, then, given a maximizer $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ of (3.18), a minimizer $u_h^{cr} \in U_\ell^{cr}(\mathcal{T}_h)$ of (3.17) is given by

$$u_h^{cr} = D\psi_h^*(\cdot, \operatorname{div} z_h^{rt}) + D\phi_h^*(\cdot, \Pi_h z_h^{rt})(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,, \tag{3.29}$$

and a discrete strong duality relation, i.e., (3.20), applies.

Proof.

ad (i). By definition, it holds that $z_h^{rt} \in (\mathcal{L}^1(\mathcal{T}_h))^{\ell\times d}$ and the discrete convex optimality relation (3.24) is satisfied. Since $u_h^{rt} \in U_\ell^{cr}(\mathcal{T}_h)$ is minimal for (3.17) as well as $\phi_h(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$ and $\psi_h(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, for every $v_h \in U_{\ell,D}^{cr}(\mathcal{T}_h)$, we have that

$$(D\phi_h(\cdot,\nabla_h u_h^{cr}),\nabla_h v_h)_\Omega + (D\psi_h(\cdot,\Pi_h u_h^{cr}),\Pi_h v_h)_\Omega - (g_h,\pi_h v_h)_{\Gamma_N} = 0\,. \tag{3.30}$$

In particular, (3.30) implies that $D\phi_h(\cdot,\nabla_h u_h^{cr}) \in (\nabla_h(\ker(\Pi_h|_{U_{\ell,0}^{cr}(\mathcal{T}_h)})))^\perp$. Due to [19, Thm. 3.1], it holds that $(\nabla_h(\ker(\Pi_h|_{U_{\ell,0}^{cr}(\mathcal{T}_h)})))^\perp = \Pi_h(Z_\ell^{rt}(\mathcal{T}_h))$. Thus, there exists $y_h \in Z_\ell^{rt}(\mathcal{T}_h)$ such that

$$\Pi_h y_h = D\phi_h(\cdot,\nabla_h u_h^{cr}) \quad \text{a.e. in } \Omega\,. \tag{3.31}$$

As a result, resorting to the discrete integration-by-parts formula (2.12) and to (3.31), (3.30), and (3.24), for every $v_h \in U_{\ell,0}^{cr}(\mathcal{T}_h)$, we find that

$$(y_h - z_h^{rt},\nabla_h v_h)_\Omega = (D\phi_h(\cdot,\nabla_h u_h^{cr}),\nabla_h v_h)_\Omega + (D\psi_h(\cdot,\Pi_h u_h^{cr}),\Pi_h v_h)_\Omega = 0\,. \tag{3.32}$$

On the other hand, we have that $\operatorname{div}(y_h - z_h^{rt}) = 0$ a.e. in Ω, *i.e.*, $y_h - z_h^{rt} \in (\mathcal{L}^0(\mathcal{T}_h))^{\ell\times d}$. Therefore, (3.32) in conjunction with (2.13) implies that $y_h - z_h^{rt} \in (\nabla_h(U_{\ell,0}^{cr}(\mathcal{T}_h)))^\perp = \ker(\operatorname{div}|_{Z_\ell^{rt}(\mathcal{T}_h)})$. As a result, due to $y_h \in Z_\ell^{rt}(\mathcal{T}_h)$, we conclude that $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ with

$$\begin{aligned} \Pi_h z_h^{rt} &= D\phi_h(\cdot,\nabla_h u_h^{cr}) && \text{a.e. in } \Omega\,,\\ \operatorname{div} z_h^{rt} &= D\psi_h(\cdot,\Pi_h u_h^{cr}) && \text{a.e. in } \Omega\,,\\ z_h^{rt} n &= g_h && \text{a.e. on } \Gamma_N\,. \end{aligned} \tag{3.33}$$

By the Fenchel–Young identity (*cf.* (2.4)), $(3.33)_{1,2}$ are equivalent to

$$\begin{aligned} \Pi_h z_h^{rt} : \nabla_h u_h^{cr} &= \phi_h^*(\cdot,\Pi_h z_h^{rt}) + \phi_h(\cdot,\nabla_h u_h^{cr}) && \text{a.e. in } \Omega\,,\\ \operatorname{div} z_h^{rt} \cdot \Pi_h u_h^{cr} &= \psi_h^*(\cdot,\operatorname{div} z_h^{rt}) + \psi_h(\cdot,\Pi_h u_h^{cr}) && \text{a.e. in } \Omega\,. \end{aligned} \tag{3.34}$$

Adding $(3.34)_1$ and $(3.34)_2$, subsequently, integrating with respect to $x \in \Omega$, using the discrete integration-by-parts formula (2.12), and using the definitions (3.17) and (3.18), we arrive at $I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})$, which, by the discrete weak duality relation (3.19), implies that $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ is maximal for (3.18).

ad (ii). By definition, it holds that $u_h^{cr} \in (\mathcal{L}^1(\mathcal{T}_h))^\ell$ and the discrete convex optimality relation (3.27) is satisfied. Since $z_h^{rt} \in Z_\ell^{rt}(\mathcal{T}_h)$ is maximal for (3.18) as well as $\phi_h^*(x,\cdot) \in C^1(\mathbb{R}^{\ell\times d})$ for a.e. $x \in \Omega$ and $\psi_h^*(x,\cdot) \in C^1(\mathbb{R}^\ell)$ for a.e. $x \in \Omega$, for every $y_h \in Z_{\ell,N}^{rt}(\mathcal{T}_h)$, we have that

$$(D\phi_h^*(\cdot,\Pi_h z_h^{rt}),\Pi_h y_h)_\Omega + (D\psi_h^*(\cdot,\operatorname{div} z_h^{rt}),\operatorname{div} y_h)_\Omega + (y_h n, u_D^h)_{\Gamma_D} = 0\,. \tag{3.35}$$

In particular, (3.35) implies that $D\phi_h^*(\cdot,\Pi_h z_h^{rt}) \in (\ker(\operatorname{div}|_{Z_{\ell,0}^{rt}(\mathcal{T}_h)}))^\perp$. Due to (2.13), it holds that $(\ker(\operatorname{div}|_{Z_{\ell,0}^{rt}(\mathcal{T}_h)}))^\perp = \nabla_h(U_\ell^{cr}(\mathcal{T}_h))$. Therefore, there exists $v_h \in U_\ell^{cr}(\mathcal{T}_h)$ such that

$$\nabla_h v_h = D\phi_h^*(\cdot,\Pi_h z_h^{rt}) \quad \text{a.e. in } \Omega\,. \tag{3.36}$$

As a result, resorting to the discrete integration-by-parts formula (2.12) and to (3.36), (3.35), and (3.27), for every $y_h \in Z_{\ell,0}^{rt}(\mathcal{T}_h)$, we find that

$$(v_h - u_h^{cr},\operatorname{div} y_h)_\Omega = -(D\phi_h^*(\cdot,\Pi_h z_h^{rt}),\Pi_h y_h)_\Omega - (D\psi_h^*(\cdot,\operatorname{div} z_h^{rt}),\operatorname{div} y_h)_\Omega = 0\,. \tag{3.37}$$

On the other hand, we have that $\nabla_h(v_h - u_h^{cr}) = 0$ a.e. in Ω, *i.e.*, $v_h - u_h^{cr} \in (\mathcal{L}^0(\mathcal{T}_h))^\ell$. Therefore, (3.37) in conjunction with (2.14) implies that $v_h - u_h^{cr} \in (\operatorname{div}(Z_{\ell,0}^{rt}(\mathcal{T}_h)))^\perp = \ker(\nabla_h|_{U_\ell^{cr}(\mathcal{T}_h)})$. As a result, due to $v_h \in U_\ell^{cr}(\mathcal{T}_h)$, we conclude that $u_h^{cr} \in U_\ell^{cr}(\mathcal{T}_h)$ with

$$\begin{aligned} \nabla_h u_h^{cr} &= D\phi_h^*(\cdot,\Pi_h z_h^{rt}) && \text{a.e. in } \Omega\,,\\ \Pi_h u_h^{cr} &= D\psi_h^*(\cdot,\operatorname{div} z_h^{rt}) && \text{a.e. in } \Omega\,,\\ \pi_h u_h^{cr} &= u_D^h && \text{a.e. on } \Gamma_D\,. \end{aligned} \tag{3.38}$$

By the Fenchel–Young identity (*cf.* (2.4)), $(3.38)_{1,2}$ is equivalent to (3.34). Adding $(3.34)_1$ and $(3.34)_2$, subsequently, integrating with respect to $x \in \Omega$, using the discrete integration-by-parts formula (2.12), and using the definitions (3.17) and (3.18), we arrive at $I_h^{cr}(u_h^{cr}) = D_h^{rt}(z_h^{rt})$, which, by the discrete weak duality relation (3.19), implies that $u_h^{cr} \in U_\ell^{cr}(\mathcal{T}_h)$ is minimal for (3.17). □

4. MODEL PROBLEMS

In this section, we illustrate the procedure described above by addressing model problems.

4.1 Non-linear Dirichlet problem

4.1.1 Continuous problem

A class of variational problems that includes the non-linear Laplace operator (*cf.* [38, 59, 58, 13]) involves a function $\phi\colon \Omega \times \mathbb{R}^d \to \mathbb{R}$ that satisfies

(A.1) $\phi\colon \Omega \times \mathbb{R}^d \to \mathbb{R}$ is a Carathéodory mapping and $\phi(x,\cdot)\colon \mathbb{R}^d \to \mathbb{R}$ is convex for a.e. $x \in \Omega$;

(A.2) There exist constants $\alpha_m, \alpha_M > 0$, functions $\beta_m, \beta_M \in L^1(\Omega)$, and a variable exponent $p \in L^\infty(\Omega)$ with $p^- := \operatorname{ess\,sup}_{x\in\Omega} p(x)$ such that for a.e. $x \in \Omega$ and $r \in \mathbb{R}^d$, it holds that

$$\beta_m(x) + \alpha_m\,|r|^{p(x)} \leq \phi(x,r) \leq \beta_M(x) + \alpha_M\,|r|^{p(x)}\,;$$

and a function $\psi\colon \Omega\times\mathbb{R} \to \mathbb{R}$ defined by $\psi(x,s) := -f(x)s$ for a.e. $x \in \Omega$ and all $s \in \mathbb{R}$, where $f \in L^{p'(\cdot)}(\Omega)$[5] and $p'(x) := \frac{p(x)}{p(x)-1}$ for a.e. $x \in \Omega$. Then, for $g \in L^{(p^-)'}(\Gamma_N)$ and $u_D \in W^{1-\frac{1}{p^-},p^-}(\Gamma_D)$, the *non-linear Dirichlet problem* is given via the optimality condition of the minimization of the functional $I\colon W^{1,p(\cdot)}(\Omega) \to \mathbb{R}\cup\{+\infty\}$, for every $v \in W^{1,p(\cdot)}(\Omega)$ defined by

$$I(v) := \int_\Omega \phi(\cdot,\nabla v)\,\mathrm{d}x - (f,v)_\Omega - \langle g, v\rangle_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(v)\,, \tag{4.1}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon W^{p'(\cdot)}(\operatorname{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$, for every $y \in W^{p'(\cdot)}(\operatorname{div};\Omega)$ defined by

$$D(y) := -\int_\Omega \phi^*(\cdot,y)\,\mathrm{d}x + \langle y\cdot n, u_D\rangle_{\Gamma_D} - I^\Omega_{\{-f\}}(\operatorname{div} y) - I^{\Gamma_N}_{\{g\}}(y\cdot n)\,. \tag{4.2}$$

As a result, given the two functionals (4.1) and (4.2), the corresponding primal-dual gap estimator $\eta^2_{\mathrm{gap}}\colon W^{1,p(\cdot)}(\Omega) \times W^{p'(\cdot)}(\operatorname{div};\Omega) \to \mathbb{R}$, for every $v \in W^{1,p(\cdot)}(\Omega)$ with $v = u_D$ a.e. in Γ_D and $y \in W^{p'(\cdot)}(\operatorname{div};\Omega)$ with $\operatorname{div} y = -f$ a.e. in Ω and $y\cdot n = g$ a.e. on Γ_N, is given via

$$\eta^2_{\mathrm{gap}}(v,y) := \int_\Omega \phi(\cdot,\nabla v) - \nabla v\cdot y + \phi^*(\cdot,y)\,\mathrm{d}x\,. \tag{4.3}$$

The integrand of (4.3), by the Fenchel–Young inequality (*cf.* (2.3)), is point-wise non-negative and, by the Fenchel–Young identity (*cf.* (2.4)), vanishes if and only if $y \in \partial_r\phi(\cdot,\nabla v)$ a.e. in Ω.

4.1.2 Discrete problem

Let $\phi_h\colon \Omega\times\mathbb{R}^d \to \mathbb{R}$ be an approximation $\phi\colon \Omega\times\mathbb{R}^d \to \mathbb{R}$ satisfying (A.1) and $\phi_h(\cdot,r) \in \mathcal{L}^0(\mathcal{T}_h)$ for all $r \in \mathbb{R}^d$. Moreover, let $f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $g_h \in \mathcal{L}^0(\mathcal{S}^{\Gamma_N}_h)$, and $u^h_D \in \mathcal{L}^0(\mathcal{S}^{\Gamma_D}_h)$ be approximations of $f \in L^{p'(\cdot)}(\Omega)$, $g \in L^{(p^-)'}(\Gamma_N)$, and $u_D \in W^{1-\frac{1}{p^-},p^-}(\Gamma_D)$. Then, the *discrete non-linear Dirichlet problem* is given via the minimization of the functional $I^{cr}_h\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defined by

$$I^{cr}_h(v_h) := \int_\Omega \phi_h(\cdot,\nabla_h v_h)\,\mathrm{d}x - (f_h,\Pi_h v_h)_\Omega - (g_h,\pi_h v_h)_{\Gamma_N} + I^{\Gamma_D}_{\{u^h_D\}}(\pi_h v_h)\,. \tag{4.4}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D^{rt}_h\colon \mathcal{RT}^0(\mathcal{T}_h) \to \mathbb{R}\cup\{-\infty\}$, for every $y_h \in \mathcal{RT}^0(\mathcal{T}_h)$ defined by

$$D^{rt}_h(y_h) := -\int_\Omega \phi^*_h(\cdot,\Pi_h y_h)\,\mathrm{d}x + (y_h\cdot n, u^h_D)_{\Gamma_D} - I^\Omega_{\{-f_h\}}(\operatorname{div} y_h) - I^{\Gamma_N}_{\{g_h\}}(y_h\cdot n)\,. \tag{4.5}$$

If $\phi_h(x,\cdot) \in C^1(\mathbb{R}^d)$ for a.e. $x \in \Omega$, then given a discrete primal solution $u^{cr}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, a discrete dual solution $z^{rt}_h \in \mathcal{RT}^0(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z^{rt}_h = D\phi_h(\cdot,\nabla_h u^{cr}_h) - \frac{f_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,. \tag{4.6}$$

[5] $L^{p'(\cdot)}(\Omega) := \{v \in L^1(\Omega) \mid |v|^{p'(\cdot)} \in L^1(\Omega)\}$, $W^{1,p(\cdot)}(\Omega) := \{v \in W^{1,1}(\Omega) \mid |v|^{p(\cdot)}, |\nabla v|^{p(\cdot)} \in L^1(\Omega)\}$, $W^{p'(\cdot)}(\operatorname{div};\Omega) := \{y \in W^1(\operatorname{div};\Omega) \mid |y|^{p'(\cdot)}, |\operatorname{div} y|^{p'(\cdot)} \in L^1(\Omega)\}$.

4.2 Obstacle problem

4.2.1 Continuous problem

Non-differentiable lower-order contributions arise in formulating an obstacle problem as a variation problem (*cf.* [53, 37]): for an external force $f \in L^2(\Omega)$, Dirichlet boundary data $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$, and an obstacle $\chi \in W^{1,2}(\Omega)$ with $\chi \le u_D$ a.e. on Γ_D, the *obstacle problem* is given via the minimization of the functional $I\colon W^{1,2}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\nabla v\|^2_{2,\Omega} - (f,v)_\Omega + I^\Omega_+(v-\chi) + I^{\Gamma_D}_{\{u_D\}}(v)\,, \tag{4.7}$$

where the indicator functional $I^\Omega_+\colon L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$ for every $\widehat{v} \in L^2(\Omega)$ is defined by

$$I^\Omega_+(\widehat{v}) \coloneqq \begin{cases} 0 & \text{if } \widehat{v} \ge 0 \text{ a.e. } \Omega\,,\\ +\infty & \text{else}\,. \end{cases}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon L^2(\Omega;\mathbb{R}^d) \to \mathbb{R} \cup \{-\infty\}$, for every $y \in W^2_N(\operatorname{div};\Omega)$ defined by

$$D(y) \coloneqq -\tfrac{1}{2}\|y\|^2_{2,\Omega} - (\operatorname{div} y + f, \chi)_\Omega + \langle y\cdot n, u_D\rangle_{\Gamma_D} - I^\Omega_-(f + \operatorname{div} y)\,, \tag{4.8}$$

where the indicator functional $I^\Omega_-\colon L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$ for every $\widehat{y} \in L^2(\Omega)$ is defined by

$$I^\Omega_-(\widehat{y}) \coloneqq \begin{cases} 0 & \text{if } \widehat{y} \le 0 \text{ a.e. } \Omega\,,\\ +\infty & \text{else}\,. \end{cases}$$

The dual problem is, in general, ill-posed in $W^2_N(\operatorname{div};\Omega)$. A maximizing vector field $z \in L^2(\Omega;\mathbb{R}^d)$ with distributional divergence defines a non-negative distribution $\lambda \coloneqq -\operatorname{div} y - f \in (W^{1,2}_D(\Omega))^*$. As a result, given the two functionals (4.7) and (4.8), the corresponding primal-dual gap estimator $\eta^2_{\mathrm{gap}}\colon W^{1,2}(\Omega) \times W^2_N(\operatorname{div};\Omega) \to \mathbb{R}$, for every $v \in W^{1,2}(\Omega)$ with $v = u_D$ a.e. on Γ_D and $v \ge \chi$ a.e. in Ω and $y \in W^2_N(\operatorname{div};\Omega)$ and $\operatorname{div} y \le -f$ a.e. in Ω, using the integration-by-parts formula (2.5), is given via

$$\eta^2_{\mathrm{gap}}(v,y) = \tfrac{1}{2}\|\nabla v - y\|^2_{2,\Omega} + (-\operatorname{div} y - f, v-\chi)_\Omega\,. \tag{4.9}$$

The first part of (4.9) measures a violation of the optimality condition $y = \nabla v$ a.e. in Ω and the second part of (4.9) measures a violation of the complementarity property $(\operatorname{div} y + f, v-\chi)_\Omega = 0$.

4.2.2 Discrete problem

Let $f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $u^h_D \in \mathcal{L}^0(\mathcal{S}^{\Gamma_D}_h)$, and $\chi_h \in \mathcal{L}^0(\mathcal{T}_h)$ be approximations of $f \in L^2(\Omega)$, $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$, and $\chi \in W^{1,2}(\Omega)$. Then, the *discrete obstacle problem* is given via the minimization of the functional $I^{cr}_h\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R} \cup \{+\infty\}$, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defined by

$$I^{cr}_h(v_h) \coloneqq \tfrac{1}{2}\|\nabla_h v_h\|^2_{2,\Omega} - (f_h, \Pi_h v_h)_\Omega + I^\Omega_+(\Pi_h v_h - \chi_h) + I^{\Gamma_D}_{\{u^h_D\}}(\pi_h v_h)\,. \tag{4.10}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D^{rt}_h\colon \mathcal{RT}^0_N(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, for every $y_h \in \mathcal{RT}^0_N(\mathcal{T}_h)$ defined by

$$D^{rt}_h(y_h) \coloneqq -\tfrac{1}{2}\|\Pi_h y_h\|^2_{2,\Omega} - (\operatorname{div} y_h + f_h, \chi_h)_\Omega + (y_h\cdot n, u^h_D)_{\Gamma_D} - I^\Omega_-(f_h + \operatorname{div} y_h)\,. \tag{4.11}$$

Given a discrete primal solution $u^{cr}_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and a Lagrange multiplier $\overline{\lambda}^{cr}_h \in \Pi_h(\mathcal{S}^{1,cr}_D(\mathcal{T}_h))$, for every $v_h \in \mathcal{S}^{1,cr}_D(\mathcal{T}_h)$ satisfying

$$(\overline{\lambda}^{cr}_h, \Pi_h v_h)_\Omega = (f_h, \Pi_h v_h)_\Omega - (\nabla_h u^{cr}_h, \nabla_h v_h)_\Omega\,, \tag{4.12}$$

proceeding as in the proof Proposition 3.8(i), a discrete dual solution $z^{rt}_h \in \mathcal{RT}^0(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z^{rt}_h = \nabla_h u^{cr}_h - \frac{f_h - \overline{\lambda}^{cr}_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,. \tag{4.13}$$

4.3 Scalar Signorini problem

4.3.1 Continuous problem

A scalar variant of elastic contact problems leads to a variational problem with an inequality constraint on a subset Γ_C of the boundary on which penetration of an obstacle is prevented (*cf.* [80]): for $\Gamma_D, \Gamma_N, \Gamma_C \subseteq \partial\Omega$ with $\Gamma_D \dot\cup \Gamma_N \dot\cup \Gamma_C = \partial\Omega$, $f \in L^2(\Omega)$, $g \in W^{-\frac{1}{2},2}(\Gamma_N)$, $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$, and $\chi \in W^{1,2}(\Omega)$ with $\chi = u_D$ a.e. on Γ_D, the *scalar Signorini problem* is given via the minimization of the functional $I\colon W^{1,2}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,2}_D(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\nabla v\|^2_{2,\Omega} - (f,v)_\Omega - \langle g, v\rangle_{\Gamma_N} + I_+^{\Gamma_C}(v-\chi) + I^{\Gamma_D}_{\{u_D\}}(v)\,, \tag{4.14}$$

where the indicator functional $I_+^{\Gamma_C}\colon W^{\frac{1}{2},2}(\Gamma_C) \to \mathbb{R}\cup\{+\infty\}$ for every $\widehat{v} \in W^{\frac{1}{2},2}(\Gamma_C)$ is defined by

$$I_+^{\Gamma_C}(\widehat{v}) \coloneqq \begin{cases} 0 & \text{if } \widehat{v} \geq 0 \text{ a.e. on } \Gamma_C\,,\\ +\infty & \text{else}\,. \end{cases}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon W^2(\mathrm{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$, for every $y \in W^2(\mathrm{div};\Omega)$ defined by

$$\begin{aligned} D(y) \coloneqq -\tfrac{1}{2}\|y\|^2_{2,\Omega} + \langle y\cdot n, \chi\rangle_{\Gamma_D\cup\Gamma_C} - I^{\Omega}_{\{-f\}}(\mathrm{div}\, y)\\ - I^{\Gamma_N}_{\{g\}}(y\cdot n) - I_+^{\Gamma_C}(y\cdot n)\,. \end{aligned} \tag{4.15}$$

As a result, given the two functionals (4.14) and (4.15), the corresponding primal-dual gap estimator $\eta^2_{\mathrm{gap}}\colon W^{1,2}(\Omega)\times W^2(\mathrm{div};\Omega)\to\mathbb{R}$, for every $v \in W^{1,2}(\Omega)$ with $v = u_D$ a.e. on Γ_D and $v \geq \chi$ a.e. on Γ_C and $y \in W^2(\mathrm{div};\Omega)$ with $y\cdot n = g$ a.e. on Γ_N and $y\cdot n \geq 0$ in $W^{-\frac{1}{2},2}(\Gamma_N)$, using the integration-by-parts formula (2.5), is given via

$$\eta^2_{\mathrm{gap}}(v,y) = \tfrac{1}{2}\|\nabla v - y\|^2_{2,\Omega} + \langle y\cdot n, v-\chi\rangle_{\Gamma_C}\,. \tag{4.16}$$

The first part of (4.16) measures a violation of the optimality relation $y = \nabla v$ a.e. in Ω and the second part of (4.16) measures a violation of the complementarity property $\langle y\cdot n, v-\chi\rangle_{\Gamma_C} = 0$.

4.3.2 Discrete problem

Let $f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $g_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, $u_D^h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$, and $\chi_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_C})$ be approximations of $f \in L^2(\Omega)$, $g \in W^{-\frac{1}{2},2}(\Gamma_N)$, $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$, and $\chi|_{\Gamma_C} \in W^{\frac{1}{2},2}(\Gamma_C)$. Then, the *discrete scalar Signorini problem* is given 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

$$\begin{aligned} I_h^{cr}(v_h) \coloneqq \tfrac{1}{2}\|\nabla_h v_h\|^2_{2,\Omega} - (f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N}\\ + I^{\Gamma_D}_{\{u_D^h\}}(\pi_h v_h) + I_+^{\Gamma_C}(\pi_h v_h - \chi_h)\,, \end{aligned} \tag{4.17}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D_h^{rt}\colon \mathcal{RT}^0(\mathcal{T}_h) \to \mathbb{R}\cup\{-\infty\}$, for every $y_h \in \mathcal{RT}^0(\mathcal{T}_h)$ defined by

$$\begin{aligned} D_h^{rt}(y_h) \coloneqq -\tfrac{1}{2}\|\Pi_h y_h\|^2_{2,\Omega} + (y_h\cdot n, \pi_h u_D^h)_{\Gamma_D\cup\Gamma_C}\\ - I^{\Omega}_{\{-f_h\}}(\mathrm{div}\, y_h) - I^{\Gamma_N}_{\{g_h\}}(y_h\cdot n) - I_+^{\Gamma_C}(y_h\cdot n)\,. \end{aligned} \tag{4.18}$$

Appealing to Proposition 3.8(i), given a discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, a discrete dual solution $z_h^{rt} \in \mathcal{RT}^0(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z_h^{rt} = \nabla_h u_h^{cr} - \frac{f_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,.$$

For an element-wise affine obstacle $\chi \in \mathcal{S}^1(\mathcal{T}_h)$, an admissible approximation $\overline{u}_h^{cr} \in W^{1,2}(\Omega)$, *e.g.*, can be obtained via node-averaging combined with setting $\overline{u}_h^{cr}(\nu) \coloneqq \max\{\chi(\nu), (\Pi_h^{av} u_h^{cr})(\nu)\}$ for all boundary nodes $\nu \in \mathcal{N}_h$ belonging to the contact boundary Γ_C.

4.4 Rudin–Osher–Fatemi image de-noising model

4.4.1 Continuous problem

A model problem that requires the usage of spaces of functions with bounded variation is the *Rudin–Osher–Fatemi image de-noising model* (*cf.* [79, 15]): for a noisy image $g \in L^2(\Omega)$ and a fidelity parameter $\alpha > 0$, it is given via the minimization of the functional $I\colon BV(\Omega) \cap L^2(\Omega) \to \mathbb{R}$, for every $v \in BV(\Omega) \cap L^2(\Omega)$ defined by

$$I(v) \coloneqq |\mathrm{D}v|(\Omega) + \tfrac{\alpha}{2}\|v - g\|_{2,\Omega}^2 \,. \tag{4.19}$$

Here, $|\mathrm{D}(\cdot)|(\Omega)\colon L^1_{\mathrm{loc}}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in L^1_{\mathrm{loc}}(\Omega)$ defined by

$$|\mathrm{D}v|(\Omega) \coloneqq \sup_{\varphi \in C_c^\infty(\Omega;\mathbb{R}^d)\colon \|\varphi\|_{\infty,\Omega} \leq 1} (v, \operatorname{div}\varphi)_\Omega \,,$$

denotes the total variation functional and $BV(\Omega) \coloneqq \{v \in L^1(\Omega) \mid |\mathrm{D}v|(\Omega) < +\infty\}$ the space of functions with bounded variation. The total variation functional can be seen as an extension of the semi-norm in $W^{1,1}(\Omega)$ and allows for discontinuous minimizers. For every $v \in L^2(\Omega)$, it can be characterized via the Fenchel duality

$$|\mathrm{D}v|(\Omega) = \sup_{y \in W_0^2(\operatorname{div};\Omega)} -(v, \operatorname{div} y)_\Omega - I^\Omega_{K_1(0)}(y) \,, \tag{4.20}$$

where the indicator functional $I^\Omega_{K_1(0)}\colon W_0^2(\operatorname{div};\Omega) \to \mathbb{R} \cup \{+\infty\}$ for every $\widehat{y} \in W_0^2(\operatorname{div};\Omega)$ is defined by

$$I^\Omega_{K_1(0)}(\widehat{y}) \coloneqq \begin{cases} 0 & \text{if } |\widehat{y}| \leq 1 \text{ a.e. in } \Omega \,, \\ +\infty & \text{else} \,. \end{cases}$$

By means of the relation (4.20), one finds that (*cf.* [56]) the (Fenchel) (pre-)dual problem is given via the maximization of $D\colon W_0^2(\operatorname{div};\Omega) \to \mathbb{R} \cup \{-\infty\}$, for every $y \in W_0^2(\operatorname{div};\Omega)$ defined by

$$D(y) \coloneqq -I^\Omega_{K_1(0)}(y) - \tfrac{1}{2\alpha}\|\operatorname{div} y + \alpha\, g\|_{2,\Omega}^2 + \tfrac{\alpha}{2}\|g\|_{2,\Omega}^2 \,. \tag{4.21}$$

As a result, given the two functionals (4.19) and (4.21), the corresponding primal-dual gap estimator $\eta_{\mathrm{gap}}^2\colon BV(\Omega) \cap L^2(\Omega) \times W_0^2(\operatorname{div};\Omega) \to \mathbb{R}$, for every $v \in BV(\Omega) \cap L^2(\Omega)$ and $y \in W_0^2(\operatorname{div};\Omega)$ with $|y| \leq 1$ a.e. in Ω, is given via

$$\eta_{\mathrm{gap}}^2(v, y) = |\mathrm{D}v|(\Omega) + (v, \operatorname{div} y)_\Omega + \tfrac{1}{2\alpha}\|\operatorname{div} y + \alpha\,(v - g)\|_{2,\Omega}^2 \,. \tag{4.22}$$

The first two parts of (4.22) measure a violation of the optimality condition $|\mathrm{D}v|(\Omega) = -(v, \operatorname{div} y)_\Omega$ and the third part of (4.22) measures a violation of the optimality condition $\operatorname{div} y = \alpha(v - g)$ a.e. in Ω.

4.4.2 Discrete problem

Let $\phi_h \in C^1(\mathbb{R}^d)$ and $g_h \in \mathcal{L}^0(\mathcal{T}_h)$ be approximations of the Euclidean length $|\cdot|$ and $g \in L^2(\Omega)$. Then, the *discrete Rudin–Osher–Fatemi image de-noising model* is given via the minimization of the functional $I_h^{cr}\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}$, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) = \int_\Omega \phi_h(\nabla_h v_h)\,\mathrm{d}x + \tfrac{\alpha}{2}\|\Pi_h v_h - g_h\|_{2,\Omega}^2 \,. \tag{4.23}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D_h^{rt}\colon \mathcal{RT}_0^0(\mathcal{T}_h) \to \mathbb{R} \cup \{-\infty\}$, for every $y_h \in \mathcal{RT}_0^0(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h) \coloneqq -\int_\Omega \phi_h^*(\Pi_h y_h)\,\mathrm{d}x - \tfrac{1}{2\alpha}\|\operatorname{div} y_h + \alpha\, g_h\|_{2,\Omega}^2 + \tfrac{\alpha}{2}\|g_h\|_{2,\Omega}^2 \,. \tag{4.24}$$

Appealing to Proposition 3.8(i), given a discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, a discrete dual solution $z_h^{rt} \in \mathcal{RT}_0^0(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z_h^{rt} = D\phi_h(\nabla_h u_h^{cr}) + \frac{\alpha(\Pi_h u_h^{cr} - g_h)}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega \,. \tag{4.25}$$

4.5 Jumping coefficients

4.5.1 Continuous problem

The derivation of sharp a posteriori error estimates is particularly challenging if a partial differential equation involves coefficients whose minimal and maximal values are not comparable: for a right-hand side $f \in L^2(\Omega)$, Neumann boundary data $g \in W^{-\frac{1}{2},2}(\Gamma_N)$, and Dirichlet boundary data $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$, a related model problem is given via the minimization of the functional $I\colon W^{1,2}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, for every $v \in W^{1,2}(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|A^{\frac{1}{2}}(\cdot)\nabla v\|_{2,\Omega}^2 - (f,v)_\Omega - \langle g, v\rangle_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(v)\,, \tag{4.26}$$

where $A\colon \Omega \to \mathbb{R}^{d\times d}$ is a tensor-valued mapping has the following properties:

(B.1) $A\colon \Omega \to \mathbb{R}^{d\times d}$ is (Lebesgue) measurable;

(B.2) For a.e. $x \in \Omega$, the tensor $A(x) \in \mathbb{R}^{d\times d}$ is symmetric and positive definite;

(B.3) There exist constants $\alpha_m, \alpha_M > 0$ such that for every $r \in \mathbb{R}^d$ and a.e. $x \in \Omega$, it holds that

$$\alpha_m|r|^2 \le |A^{\frac{1}{2}}(x)r|^2 = A(x)r\cdot r \le \alpha_M|r|^2\,.$$

Note that, due to (B.2), for a.e. $x \in \Omega$, the tensor $A(x) \in \mathbb{R}^{d\times d}$ admits a root $A(x)^{\frac{1}{2}} \in \mathbb{R}^{d\times d}$. The tensor-valued mapping $A^{\frac{1}{2}}\colon \Omega \to \mathbb{R}^{d\times d}$ in (4.26) is defined by $A^{\frac{1}{2}}(x) \coloneqq A(x)^{\frac{1}{2}}$ for a.e. $x \in \Omega$. Since for a.e. $x \in \Omega$, the tensor $A(x)^{\frac{1}{2}} \in \mathbb{R}^{d\times d}$ is symmetric and positive definite, it is invertible. The tensor-valued mapping $A^{-\frac{1}{2}}\colon \Omega \to \mathbb{R}^{d\times d}$ is defined by $A^{-\frac{1}{2}}(x) \coloneqq A(x)^{-\frac{1}{2}}$ for a.e. $x \in \Omega$. If $\phi\colon \Omega\times\mathbb{R}^d \to \mathbb{R}$ is defined by $\phi(x,r) \coloneqq \frac{1}{2}|A^{\frac{1}{2}}(x)r|^2$ for a.e. $x \in \Omega$ and all $r \in \mathbb{R}^d$, then $\phi^*\colon \Omega\times\mathbb{R}^d \to \mathbb{R}$ for a.e. $x \in \Omega$ and every $s \in \mathbb{R}^d$ is found to be given via

$$\phi^*(x,s) = |A^{-\frac{1}{2}}(x)s|^2\,.$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon W^2(\operatorname{div};\Omega) \to \mathbb{R}\cup\{-\infty\}$, for every $y \in W^2(\operatorname{div};\Omega)$ defined by

$$D(y) \coloneqq -\tfrac{1}{2}\|A^{-\frac{1}{2}}(\cdot)y\|_{2,\Omega}^2 + \langle y\cdot n, u_D\rangle_{\Gamma_D} - I^{\Omega}_{\{-f\}}(\operatorname{div} y) - I^{\Gamma_N}_{\{g\}}(y\cdot n)\,. \tag{4.27}$$

As a result, given the two functionals (4.26) and (4.27), the corresponding primal-dual gap estimator $\eta^2_{\mathrm{gap}}\colon W^{1,2}(\Omega)\times W^2(\operatorname{div};\Omega) \to \mathbb{R}$, for every $v \in W^{1,2}(\Omega)$ with $v = u_D$ a.e. on Γ_D and $y \in W^2(\operatorname{div};\Omega)$ with $y\cdot n = g$ in $W^{-\frac{1}{2},2}(\Gamma_N)$ and $\operatorname{div} y = -f$ a.e. in Ω, using the integration-by-parts formula (2.5), is given via

$$\eta^2_{\mathrm{gap}}(v,y) = \tfrac{1}{2}\|A^{\frac{1}{2}}(\cdot)\nabla v - A^{-\frac{1}{2}}(\cdot)y\|_{2,\Omega}^2\,, \tag{4.28}$$

measuring the violation of the optimality condition $y = A(\cdot)\nabla v$ a.e. in Ω.

4.5.2 Discrete problem

Let $A_h \in (\mathcal{L}^0(\mathcal{T}_h))^{d\times d}$ be an approximation of $A\colon \Omega \to \mathbb{R}^{d\times d}$ satisfying (B.2). Moreover, let $f_h \in \mathcal{L}^0(\mathcal{T}_h)$, $g_h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_N})$, and $u_D^h \in \mathcal{L}^0(\mathcal{S}_h^{\Gamma_D})$ be approximations of $f \in L^2(\Omega)$, $g \in W^{-\frac{1}{2},2}(\Gamma_N)$, and $u_D \in W^{\frac{1}{2},2}(\Gamma_D)$. Then, the discrete problem is given via the minimization of the functional $I_h^{cr}\colon \mathcal{S}^{1,cr}(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) \coloneqq \tfrac{1}{2}\|A_h^{\frac{1}{2}}(\cdot)\nabla_h v_h\|_{2,\Omega}^2 - (f_h, \Pi_h v_h)_\Omega - (g_h, \pi_h v_h)_{\Gamma_N} + I^{\Gamma_D}_{\{u_D^h\}}(\pi_h v_h)\,. \tag{4.29}$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D_h^{rt}\colon \mathcal{RT}^0(\mathcal{T}_h) \to \mathbb{R}\cup\{-\infty\}$, for every $y_h \in \mathcal{RT}^0(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h) \coloneqq -\tfrac{1}{2}\|A_h^{-\frac{1}{2}}(\cdot)y_h\|_{2,\Omega}^2 + (y_h\cdot n, u_D^h)_{\Gamma_D} - I^{\Omega}_{\{-f_h\}}(\operatorname{div} y_h) - I^{\Gamma_N}_{\{g_h\}}(y_h\cdot n)\,. \tag{4.30}$$

Appealing to Proposition 3.8(i), given a discrete primal solution $u_h^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$, a discrete dual solution $z_h^{rt} \in \mathcal{RT}^0(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z_h^{rt} = A_h(\cdot)\nabla_h u_h^{cr} - \frac{f_h}{d}(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,. \tag{4.31}$$

4.6 Navier–Lamé problem

4.6.1 Continuous problem

Small deformations of elastic bodies are modeled by the Navier–Lamé equations (*cf.* [43]): for an external force $f \in L^2(\Omega;\mathbb{R}^d)$, Neumann boundary data $g \in W^{-\frac{1}{2},2}(\Gamma_N;\mathbb{R}^d)$, and Dirichlet boundary data $u_D \in W^{\frac{1}{2},2}(\Gamma_D;\mathbb{R}^d)$, the *Navier–Lamé minimization problem* is given via optimality condition of the minimization of the functional $I : U_d^2(\Omega) \to \mathbb{R}\cup\{+\infty\}$, for every $v \in U_d^2(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\mathbb{C}^{\frac{1}{2}}\varepsilon(v)\|_{2,\Omega}^2 - (f,v)_\Omega - \langle g, v\rangle_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(v)\,. \tag{4.32}$$

Here, the symmetric gradient $\varepsilon\colon U_d^2(\Omega) \to L^2(\Omega;\mathbb{R}^{d\times d})$ is defined by $\varepsilon(v) \coloneqq \frac{1}{2}(\nabla v + \nabla v^\top)$ a.e. in Ω for all $v \in U_d^2(\Omega)$ and the positive definite linear operator $\mathbb{C}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is defined by $\mathbb{C}R \coloneqq 2\,\mu\,R + \lambda\,(\operatorname{tr} R)\,\mathrm{I}_{d\times d}$ for all $R \in \mathbb{R}^{d\times d}$, where $\lambda, \mu > 0$ denote the so-called Lamé constants. If $\phi\colon \mathbb{R}^{d\times d} \to \mathbb{R}_{\geq 0}$ is defined by $\phi(R) \coloneqq \frac{1}{2}|\mathbb{C}^{\frac{1}{2}}R|^2 = \mathbb{C}R : R$ for all $R \in \mathbb{R}^{d\times d}$, then $\phi^*\colon \mathbb{R}^{d\times d} \to \mathbb{R}_{\geq 0}$ for every $S \in \mathbb{R}^{d\times d}$ is found to be given via

$$\phi^*(S) = \tfrac{1}{2}|\mathbb{C}^{-\frac{1}{2}}S|^2 = \mathbb{C}^{-1}S : S\,,$$

where $\mathbb{C}^{-1}\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ is given via $\mathbb{C}^{-1}S = \frac{1}{d^2\lambda+2d\mu}\,(\operatorname{tr} S)\,\mathrm{I}_{d\times d} + \frac{1}{2\mu}\operatorname{dev} S$ for all $S \in \mathbb{R}^{d\times d}$, where $\operatorname{dev} S \coloneqq S - \frac{1}{d}\,(\operatorname{tr} S)\,\mathrm{I}_{d\times d}$ is the deviatoric part. Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon Z_d^2(\Omega) \to \mathbb{R}\cup\{-\infty\}$, for every $y \in Z_d^2(\Omega)$ defined by

$$D(y) \coloneqq -\tfrac{1}{2}\|y\|_{2,\Omega}^2 + \langle yn, u_D\rangle_{\Gamma_N} - I^{\Omega}_{\{-f\}}(\operatorname{div} y) - I^{\Gamma_N}_{\{g\}}(yn)\,. \tag{4.33}$$

As a result, given the two functionals (4.32) and (4.33), the corresponding primal-dual gap estimator $\eta_{\mathrm{gap}}^2\colon U_d^2(\Omega)\times Z_d^2(\Omega) \to \mathbb{R}$, for every $v \in U_d^2(\Omega)$ with $v = u_D$ a.e. on Γ_D and $y \in Z_d^2(\Omega)$ with $yn = g$ in $W^{-\frac{1}{2},2}(\Gamma_N;\mathbb{R}^d)$ and $\operatorname{div} y = -f$ a.e. in Ω, using the integration-by-parts formula (2.5), is given via

$$\eta_{\mathrm{gap}}^2(v,y) \coloneqq \tfrac{1}{2}\|\mathbb{C}^{\frac{1}{2}}\varepsilon(v) - \mathbb{C}^{-\frac{1}{2}}y\|_{2,\Omega}^2\,. \tag{4.34}$$

4.6.2 Discrete problem

The canonical discretization of (4.32) with a Crouzeix–Raviart method is unstable due to the lack of a discrete (non-conforming variant of) Korn's inequality, *i.e.*, in general, there exist non-trivial vector fields $v_h \in U_d^{cr}(\mathcal{T}_h)\setminus\{0\}$ with $\varepsilon_h(v_h) = 0$ a.e. in Ω, where the element-wise symmetric gradient $\varepsilon_h : U_d^{cr}(\mathcal{T}_h) \to (\mathcal{L}^0(\mathcal{T}_h))^{d\times d}$ is defined by $\varepsilon_h(v_h)|_T \coloneqq \varepsilon(v_h|_T)$ for all $T \in \mathcal{T}_h$ and $v_h \in U_d^{cr}(\mathcal{T}_h)$. Hence, a stabilization is required, for $f_h \in (\mathcal{L}^0(\mathcal{T}_h))^d$, $g_h \in (\mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}))^d$, and $u_D^h \in (\mathcal{L}^0(\mathcal{S}_h^{\Gamma_D}))^d$, leading to a functional $I_h^{cr,\mathrm{stab}}\colon U_d^{cr}(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, for every $v_h \in U_d^{cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr,\mathrm{stab}}(v_h) \coloneqq \tfrac{1}{2}\|\mathbb{C}^{\frac{1}{2}}\varepsilon_h(v_h)\|_{2,\Omega}^2 + s_h(v_h,v_h) - (f_h,\Pi_h v_h)_\Omega - (g_h,\pi_h v_h)_{\Gamma_N} + I^{\Gamma_D}_{\{u_D\}}(\pi_h v_h)\,,$$

where $s_h : U_d^{cr}(\mathcal{T}_h)\times U_d^{cr}(\mathcal{T}_h) \to \mathbb{R}$ is a symmetric bilinear form, so that the problem is well-posed (*cf.* [29]). Given the minimizer $u_h^{cr} \in U_d^{cr}(\mathcal{T}_h)$ of $I_h^{cr,\mathrm{stab}}\colon U_d^{cr}(\mathcal{T}_h) \to \mathbb{R}\cup\{+\infty\}$, we choose a residual given via element-wise gradient of a Crouzeix–Raviart vector field, *i.e.*, we choose $r_h \in U_{d,D}^{cr}(\mathcal{T}_h)$ such that for every $v_h \in U_{d,D}^{cr}(\mathcal{T}_h)$, it holds that

$$(\nabla_h r_h, \nabla_h v_h)_\Omega = s_h(u_h,v_h) = (f_h,\Pi_h v_h)_\Omega + (g_h,\pi_h v_h)_{\Gamma_N} - (\mathbb{C}\varepsilon_h(u_h^{cr}),\nabla_h v_h)_\Omega\,. \tag{4.35}$$

Then, from (4.35), proceeding as in the proof of Proposition 3.8(i), we find that a tensor field $z_h^{rt} \in Z_d^{rt}(\mathcal{T}_h)$ with $\Pi_h z_h^{rt} = \mathbb{C}\varepsilon_h(u_h^{cr}) + \nabla_h r_h$ a.e. in Ω, $\operatorname{div} z_h^{rt} = -f_h$ a.e. in Ω, and $z_h^{rt} n = g_h$ a.e. in Γ_N is given via the generalized Marini formula

$$z_h^{rt} = \mathbb{C}\varepsilon_h(u_h^{cr}) + \nabla_h r_h - \frac{1}{d} f_h \otimes (\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,. \tag{4.36}$$

The possible asymmetry of (4.36) can be seen as part of the discretization error. In general, it is not true that (4.36) is optimal for a discrete dual problem. It is, however, admissible in the continuous dual problem (4.33) if $f = f_h$ and $g = g_h$.

4.7 Stokes' problem

4.7.1 Continuous problem

Stokes' problem can be formulated as a minimization problem over divergence-free velocity fields (*cf.* [27]): for an external force $f \in L^2(\Omega;\mathbb{R}^d)$, Neumann boundary data $g \in W^{-\frac{1}{2},2}(\Gamma_N;\mathbb{R}^d)$, and Dirichlet boundary data $u_D \in W^{\frac{1}{2},2}(\Gamma_D;\mathbb{R}^d)$, the *Stokes minimization problem* is defined via the minimization of the functional $I\colon U_d^2(\Omega) \to \mathbb{R}\cup\{+\infty\}$, for every $v \in U_d^2(\Omega)$ defined by

$$I(v) \coloneqq \tfrac{1}{2}\|\nabla v\|_{2,\Omega}^2 + I_{\{0\}}^{\Omega}(\operatorname{tr}\nabla v) - (f,v)_{\Omega} - \langle g, v\rangle_{\Gamma_N} + I_{\{u_D\}}^{\Gamma_D}(v)\,. \tag{4.37}$$

where the indicator functional $I_{\{0\}}^{\Omega}\colon L^2(\Omega)\to\mathbb{R}\cup\{+\infty\}$ for every $\widehat{v}\in L^2(\Omega)$ is defined by

$$I_{\{0\}}^{\Omega}(\widehat{v}) \coloneqq \begin{cases} 0 & \text{if } \widehat{v}=0 \text{ a.e. in } \Omega\,,\\ +\infty & \text{else}\,. \end{cases}$$

If the function $\phi\colon \mathbb{R}^{d\times d}\to\mathbb{R}\cup\{+\infty\}$ is defined by $\phi(R)\coloneqq \frac{1}{2}|R|^2 - I_{\{0\}}^{\Omega}(\operatorname{tr} R)$ for all $R\in\mathbb{R}^{d\times d}$, then $\phi^*\colon \mathbb{R}^{d\times d}\to\mathbb{R}\cup\{+\infty\}$ for every $S\in\mathbb{R}^{d\times d}$ is found to be given via

$$\phi^*(x,S) = \tfrac{1}{2}|\operatorname{dev} S|^2\,.$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D\colon Z_d^2(\Omega)\to\mathbb{R}\cup\{-\infty\}$, for every $y\in Z_d^2(\Omega)$ defined by

$$D(y) \coloneqq -\tfrac{1}{2}\|\operatorname{dev} y\|_{2,\Omega}^2 + \langle yn, u_D\rangle_{\Gamma_N} - I_{\{-f\}}^{\Omega}(\operatorname{div} y) - I_{\{g\}}^{\Gamma_N}(yn)\,. \tag{4.38}$$

As a result, given the two functionals (4.37) and (4.38), the corresponding primal-dual gap estimator $\eta_{\mathrm{gap}}^2\colon U_d^2(\Omega)\times Z_d^2(\Omega)\to\mathbb{R}$, for every $v\in U_d^2(\Omega)$ with $v=u_D$ a.e. on Γ_D and $\operatorname{tr}\nabla v = 0$ a.e. in Ω and $y\in Z_d^2(\Omega)$ with $\operatorname{div} y = -f$ a.e. in Ω and $yn = g$ in $W^{-\frac{1}{2},2}(\Gamma_N;\mathbb{R}^d)$, using the integration-by-parts formula (2.5) and that $(\operatorname{dev} y, \nabla v)_{\Omega} = (y,\nabla v)_{\Omega}$, is given via

$$\eta_{\mathrm{gap}}^2(v,y) = \tfrac{1}{2}\|\nabla v - y\|_{2,\Omega}^2\,,$$

measuring the violation of the optimality condition $y=\nabla v$ a.e. in Ω.

4.7.2 Discrete problem

Let $f_h\in(\mathcal{L}^0(\mathcal{T}_h))^d$, $g_h\in(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_N}))^d$, and $u_D^h\in(\mathcal{L}^0(\mathcal{S}_h^{\Gamma_D}))^d$ be approximations of $f\in L^2(\Omega;\mathbb{R}^d)$, $g\in W^{-\frac{1}{2},2}(\Gamma_N;\mathbb{R}^d)$, and $u_D\in W^{\frac{1}{2},2}(\Gamma_D;\mathbb{R}^d)$. Then, the *discrete Stokes minimization problem* is given via the minimization of the functional $I_h^{cr}\colon U_d^{cr}(\mathcal{T}_h)\to\mathbb{R}\cup\{+\infty\}$, for every $v_h\in U_d^{cr}(\mathcal{T}_h)$ defined by

$$I_h^{cr}(v_h) \coloneqq \tfrac{1}{2}\|\nabla_h v_h\|_{2,\Omega}^2 + I_{\{0\}}^{\Omega}(\operatorname{tr}\nabla_h v_h) - (f_h,\Pi_h v_h)_{\Omega} - (g_h,\pi_h v_h)_{\Gamma_N} + I_{\{u_D^h\}}^{\Omega}(\pi_h v_h)\,.$$

Then, a corresponding (Fenchel) dual problem is given via the maximization of the functional $D_h^{rt}\colon Z_d^{rt}(\mathcal{T}_h)\to\mathbb{R}\cup\{-\infty\}$, for every $y_h\in Z_d^{rt}(\mathcal{T}_h)$ defined by

$$D_h^{rt}(y_h) \coloneqq -\tfrac{1}{2}\|\Pi_h\operatorname{dev} y_h\|_{2,\Omega}^2 + (y_h n, u_D^h)_{\Gamma_N} - I_{\{-f_h\}}^{\Omega}(\operatorname{div} y_h) - I_{\{g_h\}}^{\Gamma_N}(y_h n)\,.$$

Appealing to Proposition 3.8(i), given a discrete primal solution $u_h^{cr}\in U_d^{cr}(\mathcal{T}_h)$, a discrete dual solution $z_h^{rt}\in Z_d^{rt}(\mathcal{T}_h)$ is immediately available via the generalized Marini formula

$$z_h^{rt} = \nabla_h u_h^{cr} - \frac{1}{d} f_h\otimes(\mathrm{id}_{\mathbb{R}^d} - \Pi_h \mathrm{id}_{\mathbb{R}^d}) \quad \text{a.e. in } \Omega\,.$$

Note that, different from the previous model problem, an admissible approximation $\overline{u}_h^{cr}\in U_d^2(\Omega)$, *i.e.*, $\overline{u}_h^{cr} = u_D$ a.e. on Γ_D and $\operatorname{tr}\nabla\overline{u}_h^{cr} = 0$ a.e. in Ω, cannot be obtained via simple node-averaging since, in general, we have that $|\{\operatorname{tr}\nabla\Pi_h^{av}u_h^{cr}\neq 0\}|>0$, although, by construction, we have that $\operatorname{tr}\nabla_h u_h^{cr}=0$ a.e. in Ω. Instead an approximation $\overline{u}_h^{cr}\in U_d^2(\Omega)$ can be obtained via node-averaging combined with a local divergence-correction procedure (*cf.* [85]), in which one solves local discrete Stokes problems in finite element spaces with higher polynomial degree. Since these local problems can be solved in parallel, the overall cost of the local divergence-correction procedure is moderate.

5. Equivalence to residual type error estimators

In the case of the *φ-Dirichlet problem*, *i.e.*, if we have that $\phi := \varphi \circ |\cdot| \in C^0(\mathbb{R}^d) \cap C^2(\mathbb{R}^d \setminus \{0\})$, where $\varphi\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is an N-function (*cf.* Appendix A.2) that satisfies the following conditions (*cf.* [46, Assumption 1]):

(C.1) φ satisfies Δ_2-condition (*i.e.*, $\Delta_2(\varphi) < \infty$) and the ∇_2-condition (*i.e.*, $\Delta_2(\varphi^*) < \infty$);

(C.2) $\varphi \in C^2(0,\infty)$ and uniformly with respect to $t \geq 0$, it holds that[6]

$$\varphi'(t) \sim t\,\varphi''(t)\,;$$

and if we have that $\psi(x,\cdot) := (s \mapsto -f(x)s) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$, assuming that $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$ and $g = g_h = 0$, we can relate the primal-dual gap estimator to the residual type estimator in [46], which, in fact, coincides with the standard residual type estimator for the Poisson problem (1.1) (*i.e.*, $\varphi(t) := \frac{1}{2}t^2$ for ll $t \geq 0$). If $u_h^{p1} \in \mathcal{S}_D^1(\mathcal{T}_h)$ is the unique minimizer of $I_h^{p1} := I|_{\mathcal{S}_D^1(\mathcal{T}_h)}\colon \mathcal{S}_D^1(\mathcal{T}_h) \to \mathbb{R}$, the residual type estimator denotes the quantity

$$\eta_{\mathrm{res},h}^2(u_h^{p1}) := \sum_{T \in \mathcal{T}_h} \eta_{\mathrm{res},T}^2(u_h^{p1})\,, \tag{5.1}$$

where, if $h_S := \operatorname{diam}(S)$ for all $S \in \mathcal{S}_h$, for every $T \in \mathcal{T}_h$ and $S \in \mathcal{S}_h^i$ with $S \subseteq \partial T$

$$\begin{aligned} \eta_{\mathrm{res},T}^2(u_h^{p1}) &:= \eta_{E,T}^2(u_h^{p1}) + \sum_{S \in \mathcal{S}_h^i\,:\, S \subseteq \partial T} \eta_{J,S}^2(u_h^{p1})\,, \\ \eta_{E,T}^2(u_h^{p1}) &:= \|(\varphi_{|\nabla u_h^{p1}|})^*(h_{\mathcal{T}}|f_h|)\|_{1,T}\,, \\ \eta_{J,S}^2(u_h^{p1}) &:= \big\|h_S^{\frac{1}{2}}[\![F(\nabla u_h^{p1})]\!]_S\big\|_{2,T}^2\,. \end{aligned} \tag{5.2}$$

In (5.2), for every $a \geq 0$, the function $(\varphi_a)^*\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is the Fenchel conjugate of $\varphi_a\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ (*cf.* Appendix A.2) and the function $F\colon \mathbb{R}^d \to \mathbb{R}^d$ for every $r \in \mathbb{R}^d$ is defined by

$$F(r) := \sqrt{\frac{\varphi'(|r|)}{|r|}}\,r\,.$$

In [46, Lem. 8 & Cor. 11], it has been shown that the error estimator (5.1) is reliable and efficient with respect to the primal approximation error, *i.e.*, there exist constants $c_{\mathrm{rel}}, c_{\mathrm{eff}} > 0$ such that

$$c_{\mathrm{rel}}\,\|F(\nabla u_h^{p1}) - F(\nabla u)\|_{2,\Omega}^2 \leq \eta_{\mathrm{res},h}^2(u_h^{p1}) \leq c_{\mathrm{eff}}\,\|F(\nabla u_h^{p1}) - F(\nabla u)\|_{2,\Omega}^2\,. \tag{5.3}$$

Generalizing the procedure in [35, 54, 55, 22] and resorting to properties of the node-averaging quasi-interpolation operator $\Pi_h^{av}\colon \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathcal{S}_D^1(\mathcal{T}_h)$ (*cf.* Appendix A.1), we are able to establish the global equivalence of the primal-dual gap estimator (4.3) in the case $v = u_h^{p1} \in \mathcal{S}_D^1(\mathcal{T}_h)$ and $y = z_h^{rt} \in \mathcal{RT}_N^0(\mathcal{T}_h)$ to the residual type estimator (5.1).

Theorem 5.1. *Let $\phi = \varphi \circ |\cdot| \in C^0(\mathbb{R}^d) \cap C^2(\mathbb{R}^d \setminus \{0\})$, where $\varphi\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is an N-function satisfying (C.1), (C.2), $\psi(x,\cdot) := (s \mapsto -f(x)s) \in C^1(\mathbb{R})$ for a.e. $x \in \Omega$ for $f = f_h \in \mathcal{L}^0(\mathcal{T}_h)$, and let $g = g_h = 0$. Then, it holds that*

$$\eta_{\mathrm{res},h}^2(u_h^{p1}) \sim \eta_{\mathrm{gap}}^2(u_h^{p1}, z_h^{rt}) = \int_\Omega \phi(\nabla u_h^{p1}) - \nabla u_h^{p1} \cdot z_h^{rt} + \phi^*(z_h^{rt})\,\mathrm{d}x\,, \tag{5.4}$$

where the equivalence $\sim$ depends only on ω_0, $\Delta_2(\varphi)$, and $\nabla_2(\varphi)$.

Remark 5.2. *Theorem 5.1 extends the results in [46] by the aspect that the residual type estimator (5.1) is not only equivalent to the primal approximation (i.e., to $\rho_I^2(u_h^{p1}, u)$), but to the primal approximation plus the dual approximation error (i.e., to $\rho_{\mathrm{tot}}^2(u_h^{p1}, z_h^{rt}) := \rho_I^2(u_h^{p1}, u) + \rho_{-D}^2(z_h^{rt}, z)$). In other words, the residual type estimator (5.1) also provides control of the approximation error of the Raviart–Thomas approximation of the dual problem.*

[6] Here, we employ the notation $f \sim g$ for two (Lebesgue) measurable functions $f, g\colon \Omega \to \mathbb{R}$, if there exists a constant $c > 0$ such that $c^{-1} f \leq g \leq cf$ almost everywhere in Ω.

Proof of (Theorem 5.1). *ad* $\eta^2_{\mathrm{gap}}(u_h^{p1}, z_h^{rt}) \le c\,\eta^2_{\mathrm{res},h}(u_h^{p1})$. Using the discrete optimality relation (3.21) (which is equivalent to (3.24)), we find that

$$\left.\begin{aligned}\phi(\nabla u_h^{p1}) - \nabla u_h^{p1}\cdot\Pi_h z_h^{rt} + \phi^*(z_h^{rt})\\ = \phi(\nabla u_h^{p1}) - D\phi(\nabla_h u_h^{cr})\cdot(\nabla u_h^{p1} - \nabla_h u_h^{cr}) + \phi(\nabla_h u_h^{cr})\\ + \phi^*(z_h^{rt}) - \phi^*(\Pi_h z_h^{rt})\,.\end{aligned}\right\} \quad \text{a.e. in } \Omega\,. \tag{5.5}$$

On the other hand, by the convexity of $\phi, \phi^* \in C^1(\mathbb{R}^d)$, we have that

$$\begin{aligned}-\phi(\nabla_h u_h^{cr}) &\le -\phi(\nabla u_h^{p1}) + D\phi(\nabla u_h^{p1})\cdot(\nabla u_h^{p1} - \nabla_h u_h^{cr}) && \text{a.e. in } \Omega\,,\\ -\phi^*(\Pi_h z_h^{rt}) &\le -\phi^*(z_h^{rt}) + D\phi^*(z_h^{rt})\cdot(z_h^{rt} - \Pi_h z_h^{rt}) && \text{a.e. in } \Omega\,,\end{aligned} \tag{5.6}$$

Therefore, using (5.5) and (5.6) together with $D\phi^*(\Pi_h z_h^{rt}) \perp (z_h^{rt} - \Pi_h z_h^{rt})$ in $L^2(\Omega;\mathbb{R}^d)$, we find that

$$\begin{aligned}\eta^2_{\mathrm{gap}}(u_h^{p1}, z_h^{rt}) &\le (D\phi(\nabla u_h^{p1}) - D\phi(\nabla_h u_h^{cr}), \nabla u_h^{p1} - \nabla_h u_h^{cr})_\Omega\\ &\quad + (D\phi^*(z_h^{rt}) - D\phi^*(\Pi_h z_h^{rt}), z_h^{rt} - \Pi_h z_h^{rt})_\Omega\\ &=: I_h^1 + I_h^2\,.\end{aligned} \tag{5.7}$$

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

ad I_h^1. Abbreviating $e_h := u_h^{p1} - u_h^{cr} \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and using Galerkin orthogonality of the continuous and the discrete primal problem, we find that

$$\begin{aligned}I_h^1 &= (D\phi(\nabla u_h^{p1}), \nabla_h(e_h - \Pi_h^{av} e_h))_\Omega\\ &\quad + (f_h, \Pi_h^{av} e_h - e_h)_\Omega\\ &\quad + (D\phi(\nabla u_h^{p1}) - D\phi(\nabla u), \nabla \Pi_h^{av} e_h)_\Omega\\ &=: I_h^{1,1} + I_h^{1,2} + I_h^{1,3}\,.\end{aligned} \tag{5.8}$$

Let us next estimate $I_h^{1,1}$, $I_h^{1,2}$, and $I_h^{1,3}$:

ad $I_h^{1,1}$. Using that $[\![D\phi(\nabla u_h^{p1})\cdot n(e_h - \Pi_h^{av} e_h)]\!]_S = [\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S \{e_h - \Pi_h^{av} e_h\}_S + \{D\phi(\nabla u_h^{p1})\cdot n\}_S [\![e_h - \Pi_h^{av} e_h]\!]_S$ on S, $\pi_h [\![e_h - \Pi_h^{av} e_h]\!]_S = 0$ and $\{D\phi(\nabla u_h^{p1})\}_S = \mathrm{const}$ on S for all $S \in \mathcal{S}_h^i$, an element-wise integration-by-parts, a discrete trace inequality (*cf.* [49, Lem. 12.8]), and Proposition A.1, denoting for every $S \in \mathcal{S}_h$ by $\omega_S := \bigcup\{T \in \mathcal{T}_h \mid S \subseteq \partial T\}$ the side patch and for every $T \in \mathcal{T}_h$ by $\omega_T := \bigcup\{T' \in \mathcal{T}_h \mid T \cap T' \neq \emptyset\}$ the element patch, we find that

$$\begin{aligned}I_h^1 &= ([\![D\phi(\nabla u_h^{p1})\cdot n]\!], \{e_h - \Pi_h^{av} e_h\})_{\mathcal{S}_h^i}\\ &\le \sum_{S\in\mathcal{S}_h^i} |[\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S| \|\{e_h - \Pi_h^{av} e_h\}_S\|_{1,S}\\ &\le c \sum_{S\in\mathcal{S}_h^i} |[\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S| h_S^{-1} \|e_h - \Pi_h^{av} e_h\|_{1,\omega_S}\\ &\le c \sum_{S\in\mathcal{S}_h^i} \sum_{T\in\mathcal{T}_h : T\subseteq\omega_S} \| |[\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S| \nabla_h e_h\|_{1,\omega_T}\,.\end{aligned} \tag{5.9}$$

Then, for every $T \in \mathcal{T}_h$, using in ω_T, the ε-Young inequality (*cf.* (A.1)) for $\varphi_{|\nabla u_h^{p1}(T)|}\colon \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ and $(\varphi_{|\nabla u_h^{p1}(T)|})^*(|[\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S|) \sim |[\![F(\nabla u_h^{p1})]\!]_S|^2$ for all $S \in \mathcal{S}_h^i$ with $S \subseteq \partial T$ (*cf.* [46, Cor. 6]), where we write $\nabla u_h^{p1}(T)$ to indicate that the shift on ω_T depends only on the value of ∇u_h^{p1} on T, from (5.9), for every $\varepsilon > 0$, we deduce that

$$\begin{aligned}I_h^{1,1} &\le c \sum_{S\in\mathcal{S}_h^i} \sum_{T\in\mathcal{T}_h : T\subseteq\omega_S} c_\varepsilon \|(\varphi_{|\nabla u_h^{p1}(T)|})^*(|[\![D\phi(\nabla u_h^{p1})\cdot n]\!]_S|)\|_{1,\omega_T}\\ &\quad + \varepsilon\, c \sum_{S\in\mathcal{S}_h^i} \sum_{T\in\mathcal{T}_h : T\subseteq\omega_S} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T}\\ &\le c_\varepsilon \sum_{S\in\mathcal{S}_h^i} \eta^2_{J,S}(u_h^{p1}) + \varepsilon\, c \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T}\,.\end{aligned} \tag{5.10}$$

ad $I_h^{1,2}$. Using the ε-Young inequality (*cf.* (A.1)) and Proposition A.1, for every $\varepsilon>0$, we obtain

$$\begin{aligned} I_h^{1,2} &\le c_\varepsilon \sum_{T\in\mathcal{T}_h} \|(\varphi_{|\nabla u_h^{p1}|})^*(h_T|f_h|)\|_{1,T} \\ &\quad + \varepsilon \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}|}(h_T^{-1}|e_h-\Pi_h^{av}e_h|)\|_{1,T} \\ &\le c_\varepsilon \sum_{T\in\mathcal{T}_h} \eta_{E,T}^2(u_h^{p1}) + \varepsilon\, c \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T}\,. \end{aligned} \tag{5.11}$$

ad $I_h^{1,3}$. The ε-Young inequality (*cf.* (A.1)), the equivalence $(\varphi_{|\nabla u_h^{p1}|})^*(|D\phi(\nabla u_h^{p1})-D\phi(\nabla u)|) \sim |F(\nabla u_h^{p1})-F(\nabla u)|^2$ (*cf.* [46, Cor. 6]), and Proposition A.1, for every $\varepsilon>0$, yield that

$$\begin{aligned} I_h^{1,3} &\le c_\varepsilon \sum_{T\in\mathcal{T}_h} \|(\varphi_{|\nabla u_h^{p1}|})^*(|D\phi(\nabla u_h^{p1})-D\phi(\nabla u)|)\|_{1,T} \\ &\quad + \varepsilon \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}|}(|\nabla\Pi_h^{av}e_h|)\|_{1,T} \\ &\le c_\varepsilon \|F(\nabla u_h^{p1})-F(\nabla u)\|_{2,\Omega}^2 + \varepsilon\, c \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T}\,. \end{aligned} \tag{5.12}$$

Combining (5.10)–(5.12) in (5.8), using (5.3) in doing so, for every $\varepsilon>0$, we conclude that

$$I_h^1 \le c_\varepsilon\, \eta_{\mathrm{res},h}^2(u_h^{p1}) + \varepsilon\, c \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T}\,. \tag{5.13}$$

Proceeding as in [46, p. 9 & 10], we find that

$$\sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}(T)|}(|\nabla_h e_h|)\|_{1,\omega_T} \le c \sum_{T\in\mathcal{T}_h} \|\varphi_{|\nabla u_h^{p1}|}(|\nabla_h e_h|)\|_{1,\omega_T} + c \sum_{S\in\mathcal{S}_h^i} \eta_{J,S}^2(u_h^{p1})\,. \tag{5.14}$$

Therefore, using (5.14) in (5.13) together with the equivalence chain $\varphi_{|\nabla u_h^{p1}|}(|\nabla_h e_h|) \sim |F(\nabla u_h^{p1})-F(\nabla_h u_h^{cr})|^2 \sim (D\phi(\nabla u_h^{p1})-D\phi(\nabla_h u_h^{cr}))\cdot(\nabla u_h^{p1}-\nabla_h u_h^{cr})$ a.e. in Ω, for every $\varepsilon>0$, we arrive at

$$I_h^1 \le c_\varepsilon\, \eta_{\mathrm{res},h}^2(u_h^{p1}) + \varepsilon\, c\, I_h^1\,. \tag{5.15}$$

Resorting the reconstruction formula (4.6) and [46, Lem. 3], with $F^*\colon \mathbb{R}^d\to\mathbb{R}^d$ for every $r\in\mathbb{R}^d$ defined by

$$F^*(r) \coloneqq \sqrt{\frac{(\varphi^*)'(|r|)}{|r|}}\, r\,,$$

and the equivalence $|F^*(z_h^{rt})-F^*(\Pi_h z_h^{rt})|^2 \sim (D\phi^*(z_h^{rt})-D\phi^*(\Pi_h z_h^{rt}))\cdot(z_h^{rt}-\Pi_h z_h^{rt})$ a.e. in Ω, a change of shift (*cf.* [46, Cor. 28]), and the equivalence $|F(\nabla u_h^{p1})-F(\nabla_h u_h^{cr})|^2 \sim (D\phi(\nabla u_h^{p1})-D\phi(\nabla_h u_h^{cr}))\cdot(\nabla u_h^{p1}-\nabla_h u_h^{cr})$ a.e. in Ω, for every $\varepsilon>0$, we find that

$$\begin{aligned} I_h^2 &\le c\, \|(\varphi_{|\nabla_h u_h^{cr}|})^*(h_T|f_h|)\|_{1,\Omega} \\ &\le c_\varepsilon\, \eta_{\mathrm{res},h}^2(u_h^{p1}) + \varepsilon\, c\, I_h^1\,. \end{aligned} \tag{5.16}$$

For $\varepsilon>0$ sufficiently small, from (5.15) and (5.16) in (5.7), we conclude that

$$\eta_{\mathrm{gap}}^2(u_h^{p1}, z_h^{rt}) \le c\, \eta_{\mathrm{res},h}^2(u_h^{p1})\,. \tag{5.17}$$

ad $\eta_{\mathrm{res},h}^2(u_h^{p1}) \le c\,\eta_{\mathrm{gap}}^2(u_h^{p1}, z_h^{rt})$. From Theorem 3.3(i), (5.17), and (5.3) together with the equivalence $\rho_I^2(u_h^{p1}, u) \sim \|F(\nabla u_h^{p1})-F(\nabla u)\|_{2,\Omega}^2$, we conclude that

$$\begin{aligned} \rho_{\mathrm{tot}}^2(u_h^{p1}, z_h^{rt}) &= \eta_{\mathrm{gap}}^2(u_h^{p1}, z_h^{rt}) \\ &\le c\, \eta_{\mathrm{res},h}^2(u_h^{p1}) \\ &\le c\, \rho_I^2(u_h^{p1}, u) \\ &\le c\, \rho_{\mathrm{tot}}^2(u_h^{p1}, z_h^{rt})\,. \end{aligned}$$

□

6. QUASI-OPTIMALITY OF NODE-AVERAGING OPERATOR

Since, in general, for a discrete primal solution, we have that $u_h^{cr} \notin U_\ell^p(\Omega)$, it is necessary to post-process the discrete primal solution to obtain an admissible approximation $\overline{u}_h^{cr} \in U_\ell^p(\Omega)$. Here, it is convenient to enforce admissibility via the node-averaging quasi-interpolation operator $\Pi_h^{av} : \mathcal{S}_D^{1,cr}(\mathcal{T}_h) \to \mathcal{S}_D^1(\mathcal{T}_h)$ (*cf.* Appendix A.1), which satisfies the following local best-approximation result.

Lemma 6.1. *Let $\phi = \varphi \circ |\cdot| \in C^0(\mathbb{R}^d) \cap C^2(\mathbb{R}^d \setminus \{0\})$, where $\varphi\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is an N-function satisfying (C.1), (C.2). Then, there exists a constant $c > 0$ depending only on ω_0, $\Delta_2(\varphi)$, and $\nabla_2(\varphi)$, such that for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$ and $T \in \mathcal{T}_h$, it holds that*

$$\begin{aligned}\|F(\nabla_h v_h) - F(\nabla \Pi_h^{av} v_h)\|_{2,T}^2 \leq c \inf_{v \in W_D^{1,1}(\Omega)\,:\, F(\nabla v) \in L^2(\Omega;\mathbb{R}^d)} \|F(\nabla_h v_h) - F(\nabla v)\|_{2,\omega_T}^2 \\ + c \left\| h_{\mathcal{S}}^{1/2} [\![F(\nabla_h v_h)]\!] \right\|_{2,\mathcal{S}_h^i(T)}^2 ,\end{aligned}$$

where $\mathcal{S}_h^i(T) \coloneqq \{S \in \mathcal{S}_h^i \mid S \cap T \neq \emptyset\}$ for all $T \in \mathcal{T}_h$ and $h_{\mathcal{S}}|_S \coloneqq h_S$ for all $S \in \mathcal{S}_h$.

Proof. Follows along the lines of the proof of [58, Lem. 3.8] up to obvious adjustments. □

Using Lemma 6.1, in turn, we can deduce that the local distance of a node-averaged Crouzeix–Raviart function to a Sobolev function on an element is bounded by the local distance of the same Crouzeix–Raviart function to the same Sobolev function on an element patch plus an additive term quantifying the local fractional higher regularity of the Sobolev function, which justifies the usage of the node-averaging quasi-interpolation operator in local mesh refinement procedures.

In order to express the fractional regularity of functions, we make use of Nikolskiĭ spaces. For given $p \in [1, \infty)$, $\beta \in (0, 1]$, an open set $G \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, and $v \in L^p(G)$, the *Nikolskiĭ semi-norm* is defined by

$$[v]_{N^{\beta,p}(G)} \coloneqq \sup_{h \in \mathbb{R}^d \setminus \{0\}} |h|^{-\beta} \left(\int_{G \cap (G-h)} |v(\cdot + h) - v|^p \,\mathrm{d}x \right)^{\frac{1}{p}} < \infty .$$

Then, for $p \in [1, \infty)$ and $\beta \in (0, 1]$, the *Nikolskiĭ space* is defined by

$$N^{\beta,p}(G) \coloneqq \left\{ v \in L^p(G) \mid [\cdot]_{N^{\beta,p}(G)} < \infty \right\},$$

and the *Nikolskiĭ norm* $\|\cdot\|_{N^{\beta,p}(G)} \coloneqq \|\cdot\|_{p,G} + [\cdot]_{N^{\beta,p}(G)}$ turns $N^{\beta,p}(G)$ into a Banach space.

Proposition 6.2. *Let $\phi = \varphi \circ |\cdot| \in C^0(\mathbb{R}^d) \cap C^2(\mathbb{R}^d \setminus \{0\})$, where $\varphi\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is an N-function satisfying (C.1), (C.2). Then, there exists a constant $c > 0$ depending only on ω_0, $\Delta_2(\varphi)$, and $\nabla_2(\varphi)$, such that for every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, $v \in W_D^{1,1}(\Omega)$ with $F(\nabla v) \in L^2(\Omega;\mathbb{R}^d)$, and $T \in \mathcal{T}_h$, it holds that*

$$\|F(\nabla \Pi_h^{av} v_h) - F(\nabla v)\|_{2,T}^2 \leq c \|F(\nabla_h v_h) - F(\nabla v)\|_{2,\omega_T}^2 + c \inf_{r \in \mathbb{R}^d} \|F(\nabla v) - F(r)\|_{2,\omega_T}^2 , \tag{6.1}$$

and if, in addition, $F(\nabla v) \in N^{\beta,2}(\mathrm{int}(\omega_T);\mathbb{R}^d)$ with $\beta \in (0, 1]$, then it holds that

$$\|F(\nabla \Pi_h^{av} v_h) - F(\nabla v)\|_{2,T}^2 \leq c \|F(\nabla_h v_h) - F(\nabla v)\|_{2,\omega_T}^2 + c\, h_T^{2\beta} [F(\nabla v)]_{N^{\beta,2}(\mathrm{int}(\omega_T))}^2 . \tag{6.2}$$

Proof. ad (6.1). For every $v_h \in \mathcal{S}_D^{1,cr}(\mathcal{T}_h)$, $v \in W_D^{1,1}(\Omega)$ with $F(\nabla v) \in L^2(\Omega;\mathbb{R}^d)$, and $T \in \mathcal{T}_h$, it holds that

$$\begin{aligned}\left\| h_{\mathcal{S}}^{1/2} [\![F(\nabla_h v_h)]\!] \right\|_{2,\mathcal{S}_h^i(T)}^2 &= \inf_{r \in \mathbb{R}^d} \left\| h_{\mathcal{S}}^{1/2} [\![F(\nabla_h v_h) - F(r)]\!] \right\|_{2,\mathcal{S}_h^i(T)}^2 \\ &\leq c \inf_{r \in \mathbb{R}^d} \|F(\nabla_h v_h) - F(r)\|_{2,\omega_T}^2 \\ &\leq c \|F(\nabla_h v_h) - F(\nabla v)\|_{2,\omega_T}^2 + c \inf_{r \in \mathbb{R}^d} \|F(\nabla v) - F(r)\|_{2,\omega_T}^2 ,\end{aligned} \tag{6.3}$$

so that, using Lemma 6.1, we conclude that (6.1) applies.

ad (6.2). If, in addition, $F(\nabla v) \in N^{\beta,2}(\mathrm{int}(\omega_T);\mathbb{R}^d)$ with $\beta \in (0, 1]$, choosing $r = \fint_{\omega_T} \nabla v \,\mathrm{d}x$ for every $T \in \mathcal{T}_h$ in (6.1), by [34, ineqs. (4.6), (4.7)], we find that (6.2) applies. □

7. Numerical experiments

In this section, we review the practical relevance of the theoretical investigations of Section 4. In doing so, we restrict to scalar model problems of Section 4 (*i.e.*, we restrict to the case $\ell = 1$). The vectorial model problems of Section 4 (*cf.* Subsections 4.6, 4.7) will be experimentally investigated in forthcoming articles. All experiments were conducted deploying the finite element software package `FEniCS` (version 2019.1.0, *cf.* [60]). All graphics were generated using the `Matplotlib` library (version 3.5.1, *cf.* [57]) and the `Vedo` library (version 2023.4.4, *cf.* [65]).

7.1 Implementation details regarding the adaptive mesh refinement procedure

The experiments are based on the following generic *adaptive algorithm* (*cf.* [11, 14, 86]):

Algorithm 7.1 (AFEM). *Let $\varepsilon_{\mathrm{STOP}} > 0$, $\theta \in (0,1)$. and $\mathcal{T}_0$ an initial triangulation of Ω. Then, for every $k \in \mathbb{N} \cup \{0\}$:*

('Solve') *Compute a discrete primal solution $u_k^{cr} := u_{h_k}^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ (i.e., a minimizer of (3.17)) and a discrete dual solution $z_k^{rt} := z_{h_k}^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ (i.e., a maximizer of (3.18)). Post-process $u_k^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ and $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ to obtain admissible approximations $\overline{u}_k^{cr} \in W^{1,p}(\Omega)$ with $I(\overline{u}_k^{cr}) < \infty$ and $\overline{z}_k^{rt} \in W^{p'}(\mathrm{div};\Omega)$ with $D(\overline{z}_k^{rt}) > -\infty$;*

('Estimate') *Compute the resulting local refinement primal-dual indicators $\{\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt})\}_{T\in\mathcal{T}_k}$. If $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) \le \varepsilon_{\mathrm{STOP}}$, 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*

$$\sum_{T\in\mathcal{M}_k} \eta_T^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) \ge \theta^2 \sum_{T\in\mathcal{T}_k} \eta_T^2(\overline{u}_k^{cr}, \overline{z}_k^{rt})\,;$$

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

Remark 7.2. *(i) If not otherwise specified, we employ the parameter $\theta = \frac{1}{2}$ in step ('Mark').*
(ii) To find the set $\mathcal{M}_k \subseteq \mathcal{T}_k$ in step ('Mark'), we resort to the Dörfler marking strategy (cf. [47]).
(iii) 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. [86]) for $d = 2$ and by deploying the Plaza–Carey refinement algorithm (cf. [70]) for $d = 3$.

7.2 Inhomogeneuous $p(\cdot)$-Dirichlet problem

In this subsection, we review the theoretical findings of Subsection 4.1.

7.2.1 Implementation details regarding the optimization procedure

Before we present our numerical experiments, we briefly outline implementation details regarding the optimization procedure.

Remark 7.3. *(i) The discrete primal solution $u_k^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ (i.e., minimizer of (4.4)) in step ('Solve') is computed using the Newton line search algorithm of* `PETSc` *(version 3.17.3, cf. [12]) (with an absolute tolerance of $\tau_{\mathrm{abs}} = 1.0\times10^{-8}$ and a relative tolerance of $\tau_{\mathrm{rel}} = 1.0\times10^{-10}$) to the corresponding discrete Euler–Lagrange equations. The linear system emerging in each Newton step is solved using the sparse direct solver from* `MUMPS` *(version 5.5.0, cf. [5]);*
(ii) The reconstruction of the discrete dual solution $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ (i.e., maximizer of (4.5)) in step ('Solve') is based on the generalized Marini formula (4.6);
(iii) As conforming approximations in step ('Solve'), we employ $\overline{u}_k^{cr} := \Pi_{h_k}^{av} u_k^{cr} \in \mathcal{S}^1(\mathcal{T}_k) \subseteq W^{1,p(\cdot)}(\Omega)$ and $\overline{z}_k^{rt} = z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k) \subseteq W^{p'(\cdot)}(\mathrm{div};\Omega)$.
(iv) The local refinement indicators $\{\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt})\}_{T\in\mathcal{T}_k} \subseteq \mathbb{R}_{\ge0}$, for every $T \in \mathcal{T}_h$, are given via

$$\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt}) := \int_T \phi(\cdot, \nabla\overline{u}_k^{cr}) - \nabla\overline{u}_k^{cr}\cdot z_k^{rt} + \phi^*(\cdot, z_k^{rt})\,\mathrm{d}x\,,$$

which follows from restricting $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr}, z_k^{rt})$ (cf. (4.3)) to each element $T \in \mathcal{T}_h$.

7.2.2 Example with corner singularity on L-shape domain

For our numerical experiments, we choose $\Omega := (-1,1)^2 \setminus ([0,1]\times[-1,0])$, $\Gamma_D := \partial\Omega$, $\Gamma_N := \emptyset$, $u_D = 0 \in W^{1-\frac{1}{p^-},p^-}(\Gamma_D)$, $p \in C^\infty(\overline{\Omega})$, for every $x \in \overline{\Omega}$ defined by

$$p(x) := p^- + \tfrac{1}{2}|x|^2\,,$$

where $p^- \in \{1.5, 2\}$, $\phi\colon \Omega\times\mathbb{R}^2 \to \mathbb{R}$ (satisfying (A.1), (A.2)), for every $x \in \overline{\Omega}$ and $r \in \mathbb{R}^2$ defined by

$$\phi(x,r) := \tfrac{1}{p(x)-1}|r|^{p(x)-1}r - \tfrac{1}{p(x)(p(x)-1)}|r|^{p(x)}\,,$$

and as manufactured primal solution $u \in W^{1,p(\cdot)}(\Omega)$ (*i.e.*, minimizer of (4.1)), in polar coordinates, for every $(r,\varphi)^\top \in (0,\infty)\times(0,2\pi)$ defined by

$$u(r,\varphi) := (1 - r^2\cos^2(\varphi))(1 - r^2\sin^2(\varphi))r^{\sigma(r)}\sin(\tfrac{2}{3}\varphi)\,,$$

where $\sigma \in C^\infty(0,+\infty)$ with $\sigma(r) \in (0,1)$ for all $r \in (0,1)$, *i.e.*, we choose $f \in L^{p'(\cdot)}(\Omega)$ accordingly.

As approximations, for $k = 0,\dots,20$, we employ $\phi_k := \phi_{h_k}\colon \Omega\times\mathbb{R}^2 \to \mathbb{R}$, for a.e. $x \in \Omega$ and every $r \in \mathbb{R}^2$ defined by

$$\begin{aligned}\phi_{h_k}(x,r) := {}& \tfrac{1}{p_k(x)-1}(h_k^2 + |r|)^{p_k(x)-1}r \\ & - \tfrac{1}{p_k(x)(p_k(x)-1)}((h_k^2+|r|)^{p_k(x)} - h_k^{2p_k(x)})\,,\end{aligned}$$

where $p_k := p_{h_k} \in \mathcal{L}^0(\mathcal{T}_h)$ is defined by

$$p_{h_k}|_T := p(x_T) \quad \text{for all } T \in \mathcal{T}_k\,,$$

where $x_T := \frac{1}{3}\sum_{\nu\in\mathcal{N}_h\,:\,\nu\in T}\nu$ is the barycenter of T for all $T \in \mathcal{T}_k$, and $f_k := f_{h_k} := \Pi_{h_k} f \in \mathcal{L}^0(\mathcal{T}_h)$.

For every $p^- \in \{1.5, 2\}$, the function $\sigma \in C^\infty(0,+\infty)$ for every $r \in (0,\infty)$ is defined by

$$\sigma(r) := 1.01 - \tfrac{1}{p^- + r^2}\,,$$

which precisely yields the fractional regularity

$$F(\cdot,\nabla u) \in \mathcal{N}^{\frac{1}{2},2}(\Omega;\mathbb{R}^2)\,, \tag{7.1}$$

where $F\colon \Omega\times\mathbb{R}^2 \to \mathbb{R}^2$ is defined by $F(x,r) := |r|^{\frac{p(x)-2}{2}}r$ for a.e. $x \in \Omega$ and all $r \in \mathbb{R}^2$.

In the case of uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1), the fractional regularity (7.1) let us expect the reduced convergence rate $h_k \sim N_k^{-\frac{1}{2}}$, $k = 0,\dots,20$, where $N_k := \dim(\mathcal{S}_0^{1,cr}(\mathcal{T}_k))$, for the alternative total error quantity

$$\left.\begin{aligned}\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}_k^{cr}, z_k^{rt}) := {}& \|F(\cdot,\nabla\overline{u}_k^{cr}) - F(\cdot,\nabla u)\|^2_{2,\Omega} \\ & + \|F^*(\cdot,z_k^{rt}) - F^*(\cdot,z)\|^2_{2,\Omega}\end{aligned}\right\}\,, \quad k = 0,\dots,20\,,$$

where $F^*\colon \Omega\times\mathbb{R}^2 \to \mathbb{R}^2$ is defined by $F^*(x,r) := |r|^{\frac{p'(x)-2}{2}}r$ for a.e. $x \in \Omega$ and all $r \in \mathbb{R}^2$, which up to a multiplicative constant is a lower bound for the total error $\rho^2_{\mathrm{tot}}(\overline{u}_k^{cr}, z_k^{rt})$, $k = 0,\dots,20$, *i.e.*, there exists a constant $c_{p(\cdot)} > 0$, depending only on $p^- := \operatorname{ess\,inf}_{x\in\Omega} p(x)$ and $p^+ := \operatorname{ess\,sup}_{x\in\Omega} p(x)$, such that

$$\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}_k^{cr}, z_k^{rt}) \le c_{p(\cdot)}\,\rho^2_{\mathrm{tot}}(\overline{u}_k^{cr}, z_k^{rt})\,, \quad k = 0,\dots,20\,.$$

The coarsest triangulation $\mathcal{T}_0$ of Figure 2 consists of 96 elements and 65 vertices. In Figure 1, for $p^- \in \{1.5, 2\}$, one sees that uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) yields the expected reduced convergence rate $h_k \sim N_k^{-\frac{1}{2}}$, $k = 0,\dots,4$, while adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1) yields the quasi-optimal convergence rate $h_k^2 \sim N_k^{-1}$, $k = 0,\dots,20$. In particular, for every $p^- \in \{1.5, 2\}$ and $k = 0,\dots,20$, when using adaptive mesh refinement, and $k = 0,\dots,4$, when using uniform mesh refinement, the primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}_k^{cr}, z_k^{rt})$ is reliable and efficient with respect to $\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}_k^{cr}, z_k^{rt})$, $k = 0,\dots,20$, although it is an upper bound only up to a constant. This is due to the unspecified constant $c_{p(\cdot)} > 0$.

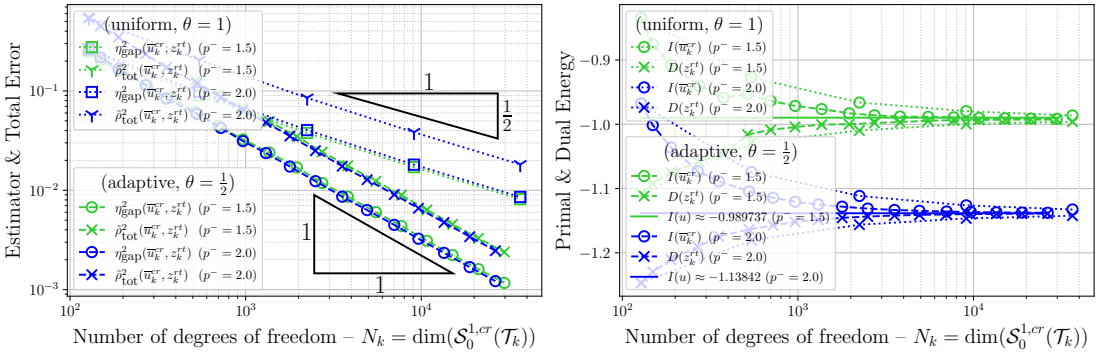

Figure 1: LEFT: primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}^{cr}_k, z^{rt}_k)$ and alternative total error $\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}^{cr}_k, z^{rt}_k)$; RIGHT: primal energy $I(\overline{u}^{cr}_k)$ and dual energy $D(z^{rt}_k)$; each for $p^- \in \{1.5, 2\}$ and $k = 0, \dots, 20$, when using adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1), and for $k = 0, \dots, 4$, when uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1), in the inhomogeneous $p(\cdot)$-Dirichlet problem.

In Figure 2, for $p^- \in \{1.5, 2\}$, one finds that Algorithm 7.1 refines towards the origin, where the gradient of the primal solution $u \in W^{1,p(\cdot)}(\Omega)$ has its singularity. More precisely, Figure 2 displays the triangulations $\mathcal{T}_k$, $k \in \{0, 10, 20\}$, generated by Algorithm 7.1 in the case $p^- = 2$.

This behavior can also be seen in Figure 3, in which the discrete primal solution $u^{cr}_{10} \in \mathcal{S}^{1,cr}_0(\mathcal{T}_{10})$, the node-averaged discrete primal solution $\Pi^{av}_{h_{10}} u^{cr}_{10} \in \mathcal{S}^1_0(\mathcal{T}_{10})$, and the discrete dual solution $z^{rt}_{10} \in \mathcal{RT}^0(\mathcal{T}_{10})$ are plotted, each in the case $p^- = 2$.

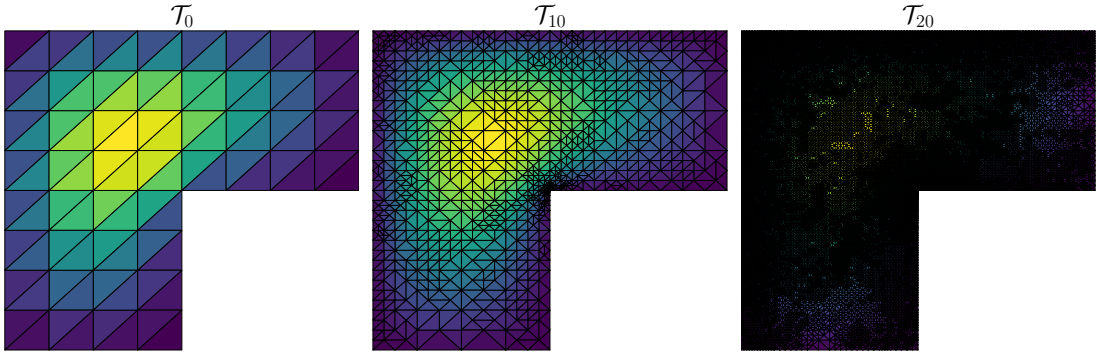

Figure 2: Initial triangulation $\mathcal{T}_0$ and adaptively refined triangulations $\mathcal{T}_k$, $k \in \{10, 20\}$, generated by Algorithm 7.1 for $\theta = \frac{1}{2}$, each in the case $p^- = 2$ in the inhomogeneous $p(\cdot)$-Dirichlet problem.

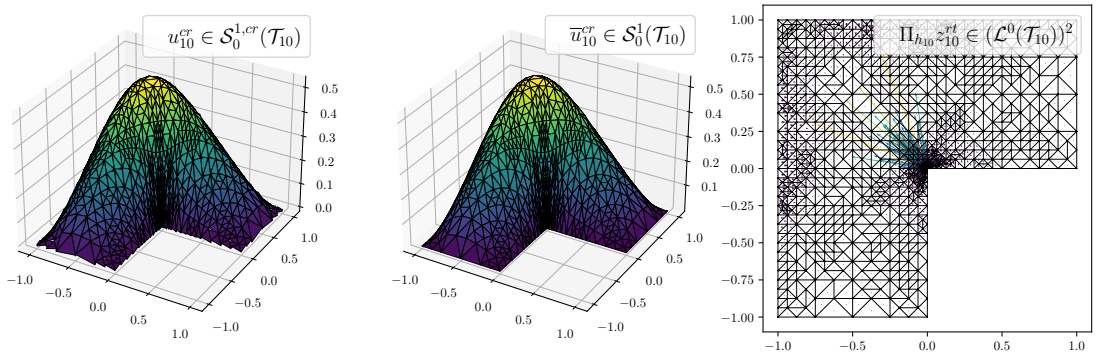

Figure 3: LEFT: discrete primal solution $u^{cr}_{10} \in \mathcal{S}^{1,cr}_0(\mathcal{T}_{10})$; MIDDLE: node-averaged discrete primal solution $\overline{u}^{cr}_{10} \in \mathcal{S}^1_0(\mathcal{T}_{10})$; RIGHT: (local) L^2-projection (onto $(\mathcal{L}^0(\mathcal{T}_{10}))^2$) of discrete dual solution $z^{rt}_{10} \in \mathcal{RT}^0(\mathcal{T}_{10})$, each in the the case $p^- = 2$ in the inhomogeneous $p(\cdot)$-Dirichlet problem.

7.3 Obstacle problem

In this subsection, we review the theoretical findings of Subsection 4.2.

7.3.1 Implementation details regarding the optimization procedure

Before we present our numerical experiments, again, we briefly outline implementation details regarding the optimization procedure.

Remark 7.4. *(i) The discrete primal solution $u_k^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ (i.e., minimizer of (4.10)) and the discrete Lagrange multiplier $\overline{\lambda}_k^{cr} \in \Pi_{h_k}(\mathcal{S}_D^{1,cr}(\mathcal{T}_k))$ (i.e., solution of (4.12)) in step ('Solve') are computed using a primal-dual active set strategy interpreted as a locally super-linear converging semi-smooth Newton method (cf. [18, Alg. 6.1 with $\alpha = 1$]). Since only a finite number of active sets are possible, the algorithm terminates after a finite number of iterations at $(u_k^{cr}, \overline{\lambda}_k^{cr})^\top \in \mathcal{S}^{1,cr}(\mathcal{T}_k) \times \Pi_{h_k}(\mathcal{S}_D^{1,cr}(\mathcal{T}_k))$. The linear system emerging in each semi-smooth Newton step is solved using the sparse direct solver of* `SciPy` *(version 1.8.1, cf. [87]);*

(ii) The reconstruction of the discrete dual solution $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ (i.e., maximizer of (4.11)) in step ('Solve') is based on the generalized Marini formula (4.13);

(iii) As conforming approximations in step ('Solve'), we employ $\overline{u}_k^{cr} := \max\{\Pi_{h_k}^{av} u_k^{cr}, \chi\} \in W^{1,2}(\Omega)$ and $\overline{z}_k^{rt} = z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k) \subseteq W^2(\mathrm{div}; \Omega)$.

(iv) The local refinement indicators $\{\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt})\}_{T \in \mathcal{T}_k} \subseteq \mathbb{R}_{\geq 0}$, for every $T \in \mathcal{T}_h$, are given via

$$\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt}) := \tfrac{1}{2}\|\nabla \overline{u}_k^{cr} - z_k^{rt}\|_{2,T}^2 + (-\mathrm{div}\, z_k^{rt} - f, \overline{u}_k^{cr} - \chi)_T \,,$$

which follows from restricting $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr}, z_k^{rt})$ (cf. (4.9)) to each element $T \in \mathcal{T}_h$.

7.3.2 Example with unknown exact solution

For our numerical experiments, we choose $\Omega := (-\frac{3}{2}, \frac{3}{2})^2$, $\Gamma_D := \partial\Omega$, $\Gamma_N := \emptyset$, $f = 0 \in L^2(\Omega)$, $u_D = 0 \in W^{\frac{1}{2},2}(\partial\Omega)$, and $\chi \in W_0^{1,2}(\Omega)$ (*cf.* Figure 4), for every $x = (x_1, x_2)^\top \in \Omega$ defined by

$$\chi(x_1, x_2) := \max\{0, \min\{\min\{x_1 + 1, \tfrac{1}{2}, 1 - x_1\}, \min\{x_2 + 1, \tfrac{1}{2}, 1 - x_2\}\}\} \,.$$

As approximation, for $k = 0, \dots, 25$, we employ $\chi_k := \chi_{h_k} := \Pi_{h_k} \chi \in \mathcal{L}^0(\mathcal{T}_k)$ (*cf.* Figure 4). The primal solution $u \in W_0^{1,2}(\Omega)$ (*i.e.*, minimizer of (4.7)) is not known and cannot be expected to satisfy $u \in W^{2,2}(\Omega)$ inasmuch as $\chi \notin W^{2,2}(\Omega)$. In consequence, uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) is expected to yield a reduced convergence rate compared to the quasi-optimal convergence rate $h_k^2 \sim N_k^{-1}$, $k = 0, \dots, 25$, where $N_k := \dim(\mathcal{S}_0^{1,cr}(\mathcal{T}_k)) + \dim(\mathcal{L}^0(\mathcal{T}_k))$.

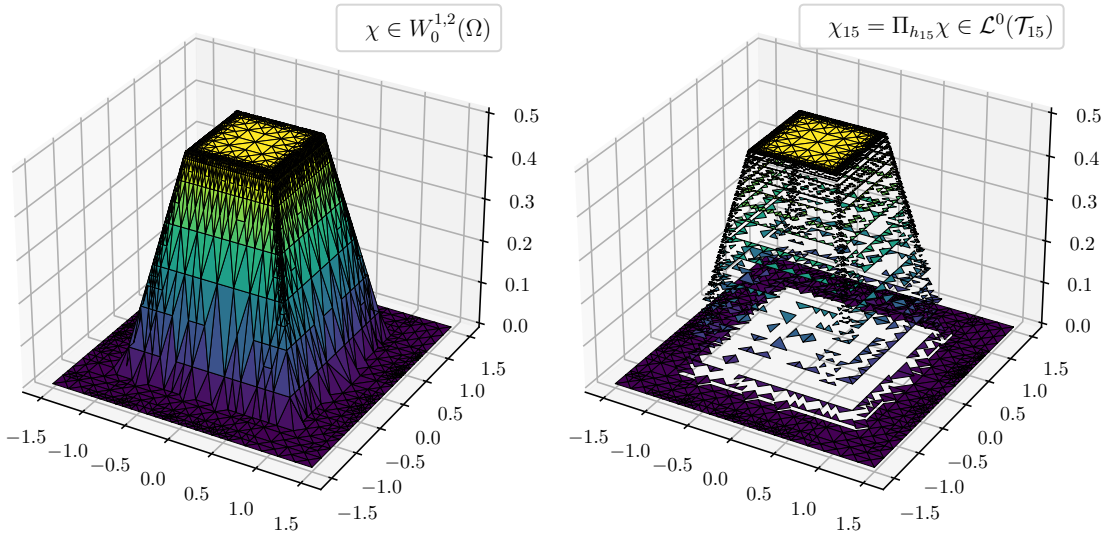

Figure 4: LEFT: nodal interpolant of the obstacle $\chi \in W_0^{1,2}(\Omega)$; RIGHT: (local) L^2-projection (onto $\mathcal{L}^0(\mathcal{T}_h)$) $\chi_{15} := \Pi_{h_{15}} \chi \in \mathcal{L}^0(\mathcal{T}_{15})$ of the obstacle $\chi \in W_0^{1,2}(\Omega)$.

The coarsest triangulation $\mathcal{T}_0$ in Figure 5 consists of 64 elements and 41 vertices. More precisely, Figure 5 displays the triangulations $\mathcal{T}_k$, $k\in\{0,5,10,15,20,25\}$, generated by Algorithm 7.1 for $\theta=\frac{1}{2}$. The discrete contact zones $\mathcal{C}_k^{cr}:=\{\Pi_{h_k}u_k^{cr}=\chi_k\}=\{\overline{\lambda}_k^{cr}<0\}$, $k\in\{0,5,10,15,20,25\}$, are plotted in white in Figure 5 while their complements are shaded.

Algorithm 7.1 refines the triangulations towards the contact zone $\mathcal{C}:=\{u=\chi\}$ (*cf.* Figure 5). The discrete contact zones $\mathcal{C}_k^{cr}$, $k\in\{0,\dots,25\}$, reduce to $\mathcal{C}$. This can also be observed in Figure 7, where the discrete primal solution $u_{15}^{cr}\in\mathcal{S}_0^{1,cr}(\mathcal{T}_{15})$, the node-averaged discrete primal solution $\Pi_{h_{15}}^{av}u_{15}^{cr}\in\mathcal{S}_0^1(\mathcal{T}_{15})$, the discrete Lagrange multiplier $\overline{\lambda}_{15}^{cr}\in\Pi_{h_{15}}(\mathcal{S}_0^{1,cr}(\mathcal{T}_{15}))$, and the discrete dual solution $z_{15}^{rt}\in\mathcal{RT}^0(\mathcal{T}_{15})$ are plotted. In Figure 6, one sees that uniform mesh refinement (*i.e.*, $\theta=1$ in Algorithm 7.1) yields the expected reduced convergence rate $h_k\sim N_k^{-\frac{1}{2}}$, $k=0,\dots,4$, while adaptive mesh refinement (*i.e.*, $\theta=\frac{1}{2}$ in Algorithm 7.1) yields the quasi-optimal convergence rate $h_k^2\sim N_k^{-1}$, $k=0,\dots,25$.

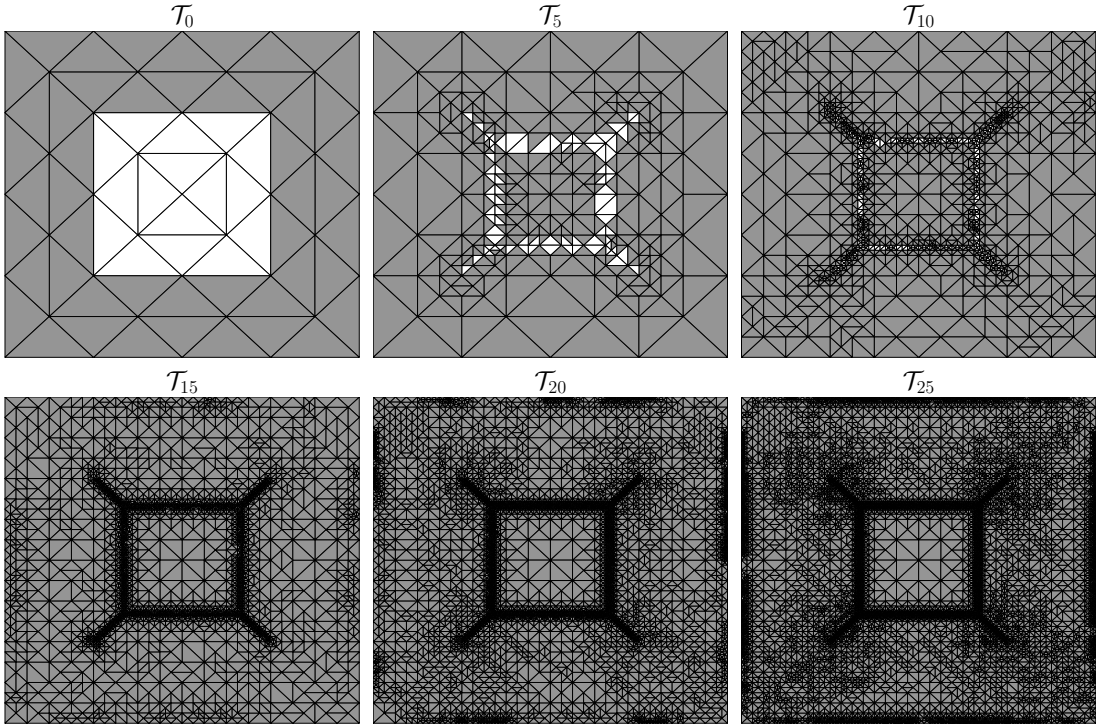

Figure 5: Adaptively refined triangulations $\mathcal{T}_k$, $k\in\{0,5,10,15,20,25\}$, with discrete contact zones $\mathcal{C}_k^{cr}$, $k\in\{0,5,10,15,20,25\}$, shown in white in the obstacle problem.

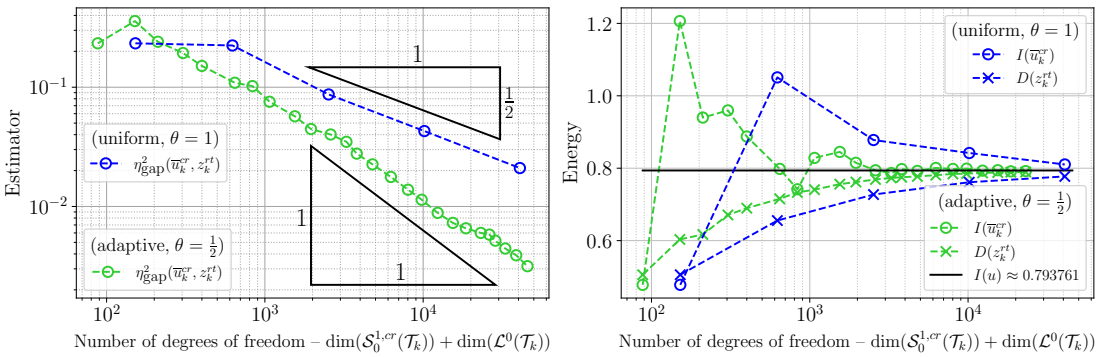

Figure 6: LEFT: primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}_k^{cr},z_k^{rt})$; RIGHT: primal energy $I(\overline{u}_k^{cr})$, dual energy $D(z_k^{rt})$, and primal energy $I(u)$ approximated via Aitken's δ^2-process (*cf.* [3]); each for $k=0,\dots,25$, when using adaptive mesh refinement (*i.e.*, $\theta=\frac{1}{2}$ in Algorithm 7.1), and for $k=0,\dots,4$, when uniform mesh refinement (*i.e.*, $\theta=1$ in Algorithm 7.1), in the obstacle problem.

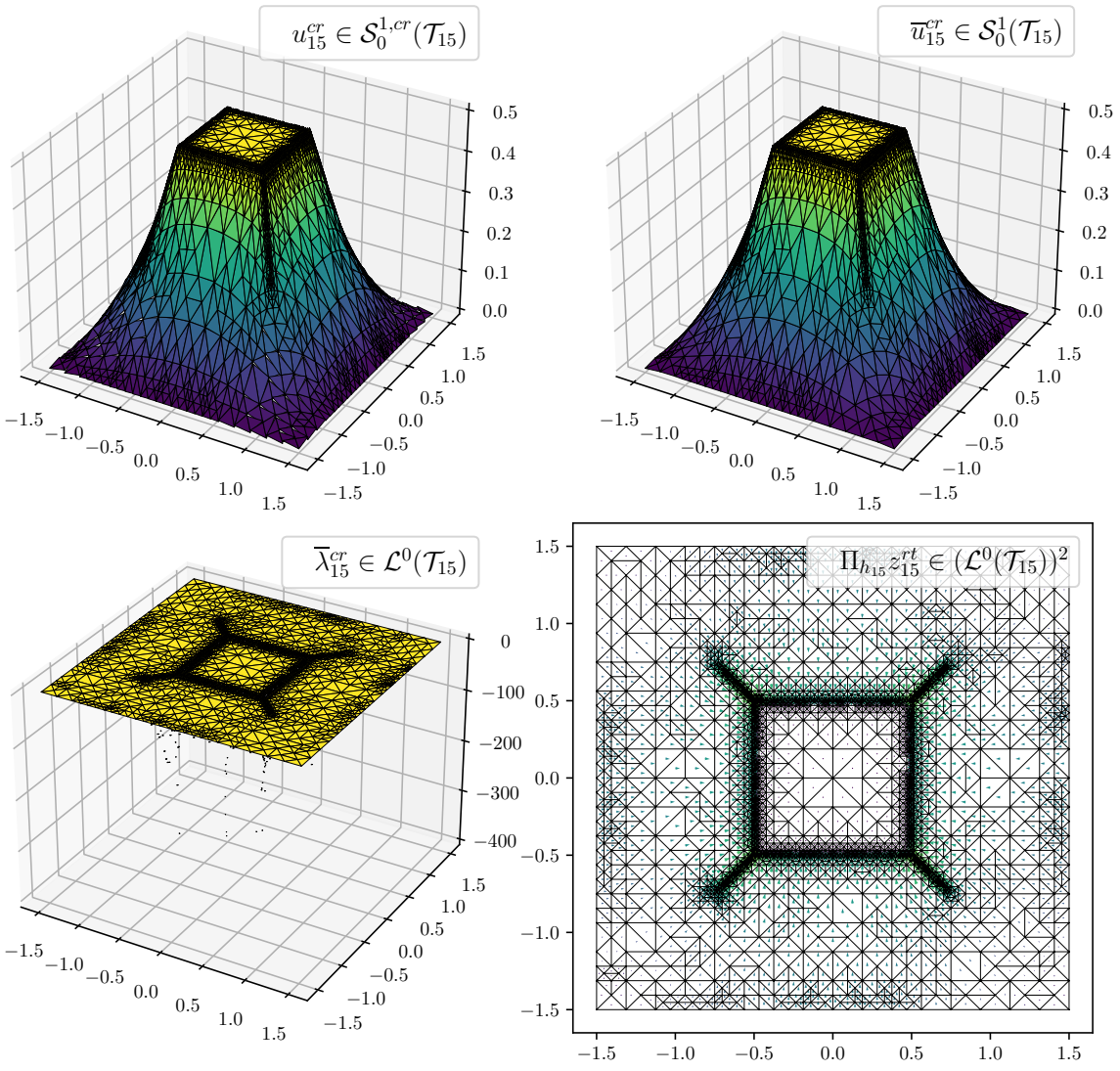

Figure 7: UPPER LEFT: discrete primal solution $u_{15}^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{15})$; UPPER RIGHT: node-averaged discrete primal solution $\overline{u}_{15}^{cr} \in \mathcal{S}_0^{1}(\mathcal{T}_{15})$; LOWER LEFT: discrete Lagrange multiplier $\overline{\lambda}_{15}^{cr} \in \mathcal{L}^0(\mathcal{T}_{15})$; LOWER RIGHT: (local) L^2-projection of the discrete dual solution $z_{15}^{rt} \in \mathcal{RT}^0(\mathcal{T}_{15})$; each on the triangulation $\mathcal{T}_{15}$ in the obstacle problem.

7.4 Rudin–Osher–Fatemi (ROF) image denoising problem

In this subsection, we review the theoretical findings of Subsection 4.4. To compare approximations to an exact solution, we impose homogeneous Dirichlet boundary conditions on $\Gamma_D = \partial\Omega$, even though, then, a corresponding existence theory is difficult to establish, in general. However, the set-up derived in Subsection 4.4 carries over verbatimly with $\Gamma_N = \emptyset$ provided that the existence of a minimizer is a priori guaranteed.

7.4.1 Implementation details regarding the optimization procedure

Before we present our numerical experiments, again, we briefly outline implementation details regarding the optimization procedure.

Remark 7.5. *(i) The discrete primal solution $u_k^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$ in step ('Solve') is computed using a semi-implicit discretized L^2-gradient flow (cf. [24, Alg. 5.1]) for fixed step-size $\tau = 1$, stopping criterion $\varepsilon_{stop}^{h_k} := \frac{h_k}{\sqrt{20}}$, and initial condition $u_k^0 = 0 \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$. Appealing to [24, Prop. 5.2(ii)], [24, Alg. 5.1] is unconditionally strongly stable, so that employing the fixed step-size $\tau = 1$ is a canonical choice. The stopping criterion $\varepsilon_{stop}^{h_k} := \frac{h_k}{\sqrt{20}}$ ensures (cf. the argumentation below [24, Alg. 5.1]) that the final iterate $u_{h_k}^{i^*} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$, $i^* \in \mathbb{N}$, is a sufficiently accurate approximation (in $L^2(\Omega)$) of the discrete primal solution $u_k^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$, in the sense that its accuracy does not violate the best possible convergence rate (cf. [24, Rem. 5.6]). The linear systems emerging in each gradient descent step is solved using the conjugate gradient method of* `PETSc` *(version 3.17.3, cf. [12]) preconditioned with an incomplete LU factorization.*

(ii) The reconstruction of the discrete dual solution $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ (i.e., maximizer of (4.24)) in step ('Solve') is based on the generalized Marini formula (4.25).

(iii) As conforming approximations in step ('Solve'), we employ $\overline{u}_k^{cr} := \Pi_{h_k}^{\partial\Omega} u_k^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$ with $\overline{u}_k^{cr} = 0$ a.e. on $\partial\Omega$, where the operator $\Pi_{h_k}^{\partial\Omega} \colon \mathcal{S}^{1,cr}(\mathcal{T}_k) \to \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$ for every $v_{h_k} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ is defined by

$$\Pi_{h_k}^{\partial\Omega} v_{h_k} := \sum_{S \in \mathcal{S}_{h_k}^i \,:\, S \cap \partial\Omega = \emptyset} v_{h_k}(x_S)\, \varphi_S \,,$$

where $x_S := \frac{1}{d} \sum_{\nu \in \mathcal{N}_h \,:\, \nu \in S} \nu$ denotes the barycenter of S for all $S \in \mathcal{S}_h$, and $\overline{z}_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ with $|\overline{z}_k^{rt}| \leq 1$ a.e. in Ω, defined by

$$\overline{z}_k^{rt} := \frac{z_k^{rt}}{\max\{1, \|z_k^{rt}\|_{\infty,\Omega}\}} \,. \tag{7.2}$$

Note that the post-processing $\overline{u}_k^{cr} := \Pi_{h_k}^{\partial\Omega} u_k^{cr}$ is only due to the imposed homogeneous Dirichlet boundary condition. In the case $\Gamma_D = \emptyset$, the choice $\overline{u}_k^{cr} := u_k^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ is always admissible.

(iv) The local refinement indicators $\{\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt})\}_{T \in \mathcal{T}_k} \subseteq \mathbb{R}_{\geq 0}$, for every $T \in \mathcal{T}_h$, are given via

$$\begin{aligned}
\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt}) &:= \eta_{\mathrm{gap},A,T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) + \eta_{\mathrm{gap},B,T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) \,, \\
\eta_{\mathrm{gap},A,T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) &:= \|\nabla_h \overline{u}_k^{cr}\|_{1,T} - (\nabla_h \overline{u}_k^{cr}, \Pi_h \overline{z}_k^{rt})_T + \sum_{S \in \mathcal{S}_h \,:\, S \subseteq \partial T} \|[\![\overline{u}_k^{cr}]\!]_S\|_{1,S} \,, \\
\eta_{\mathrm{gap},B,T}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt}) &:= \tfrac{1}{2\alpha} \|\operatorname{div} \overline{z}_k^{rt} + \alpha\,(\overline{u}_k^{cr} - g)\|_{2,T}^2 \,,
\end{aligned}$$

where we used in (4.22) the discrete integration-by-parts formula (2.12) and the discrete representation of the total variation (cf. [25]), i.e., for every $v_h \in \mathcal{L}^1(\mathcal{T}_h)$, it holds that

$$|\mathrm{D} v_h|(\Omega) = \|\nabla_h v_h\|_{1,\Omega} + \sum_{S \in \mathcal{S}_h} \|[\![v_h]\!]_S\|_{1,S} \,,$$

to arrive at an alternative representation for (4.22), i.e., for every $v_h \in \mathcal{S}^{1,cr}(\mathcal{T}_h)$ and $y_h \in \mathcal{RT}_0^0(\mathcal{T}_h)$ with $|y_h| \leq 1$ a.e. in Ω, we have that

$$\begin{aligned}
\eta_{\mathrm{gap},T}^2(v_h, y_h) = \|\nabla_h v_h\|_{1,\Omega} - (\nabla_h v_h, \Pi_h y_h)_\Omega + \sum_{S \in \mathcal{S}_h} \|[\![v_h]\!]_S\|_{1,S} \\
+ \tfrac{1}{2\alpha} \|\operatorname{div} y_h + \alpha\,(v_h - g)\|_{2,T}^2 \,.
\end{aligned}$$

7.5 *Example with Lipschitz continuous dual solution*

We examine an example from [23]. In this example, we let $\Omega := (-1,1)^d$, $d \in \{2,3\}$, $\Gamma_D := \partial\Omega$, $r := \frac{1}{2}$, $\alpha := 10$, and $g := \chi_{B_r^d(0)} \in BV(\Omega) \cap L^\infty(\Omega)$. Then, the primal solution $u \in BV(\Omega) \cap L^\infty(\Omega)$ and a dual solution $z \in W^2(\mathrm{div};\Omega) \cap L^\infty(\Omega;\mathbb{R}^d)$, for a.e. $x \in \Omega$ are given via

$$u(x) := (1 - \tfrac{d}{\alpha r})\, g(x)\,, \qquad z(x) := \begin{cases} -\frac{x}{r} & \text{if } |x| < r\,, \\ -\frac{rx}{|x|^d} & \text{if } |x| \geq r\,. \end{cases} \tag{7.3}$$

As approximations, for $k = 0, \ldots, 25$, we employ $\phi_k := \phi_{h_k} \in C^1(\mathbb{R}^d)$, for every $r \in \mathbb{R}^d$ defined by $\phi_{h_k}(r) := (1 - h_k^2)\,(h_k^4 + |r|^2)^{\frac{1}{2}}$, and $g_k := g_{h_k} := \Pi_{h_k} g \in \mathcal{L}^0(\mathcal{T}_k)$.

2D Case. The coarsest triangulation $\mathcal{T}_0$ of Figure 8 consists of 32 elements and 25 vertices. More precisely, Figure 8 displays the triangulations $\mathcal{T}_k$, $k \in \{0, 15, 25\}$, generated by Algorithm 7.1. A refinement towards $\partial B_r^2(0)$, *i.e.*, the jump set J_u of the primal solution $u \in BV(\Omega) \cap L^\infty(\Omega)$ (*cf.* (7.3)) is reported. This behavior is also seen in Figure 10, in which the discrete primal solution $u_{15}^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{15})$, the (local) L^2-projection (onto $\mathcal{L}^0(\mathcal{T}_{15})$) $\Pi_{h_{15}} u_{15}^{cr} \in \mathcal{L}^0(\mathcal{T}_{15})$, the (local) L^2-projections (onto $\mathcal{L}^1(\mathcal{T}_{15})$) of the modulus of the discrete dual solution $z_{15}^{rt} \in \mathcal{RT}^0(\mathcal{T}_{15})$, and the (local) L^2-projections (onto $\mathcal{L}^1(\mathcal{T}_{15})$) of the unit-length scaled discrete dual solution $\overline{z}_{15}^{rt} \in \mathcal{RT}^0(\mathcal{T}_{15})$ (*cf.* (7.2)) are plotted. In Figure 9, one sees that uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) yields the expected reduced the convergence rate $h_k \sim N_k^{-\frac{1}{2}}$, $k = 0, \ldots, 25$, (predicted by [41, 16]), while adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1) yields the quasi-optimal convergence rate $h_k^2 \sim N_k^{-1}$, $k = 0, \ldots, 25$. In addition, Figure 9 indicates the primal-dual gap estimator is reliable and efficient respect to the alternative total error quantity

$$\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}_k^{cr}, \overline{z}_k^{rt}) := \tfrac{\alpha}{2} \|\overline{u}_k^{cr} - u\|^2_{2,\Omega} + \tfrac{1}{2\alpha} \|\mathrm{div}\, \overline{z}_k^{rt} - \mathrm{div}\, z\|^2_{2,\Omega}\,, \quad k = 0, \ldots, 25\,, \tag{7.4}$$

which is a lower bound for the total error $\rho^2_{\mathrm{tot}}(\overline{u}_k^{cr}, \overline{z}_k^{rt}) = \eta^2_{\mathrm{gap}}(\overline{u}_k^{cr}, \overline{z}_k^{rt})$, $k = 0, \ldots, 25$.

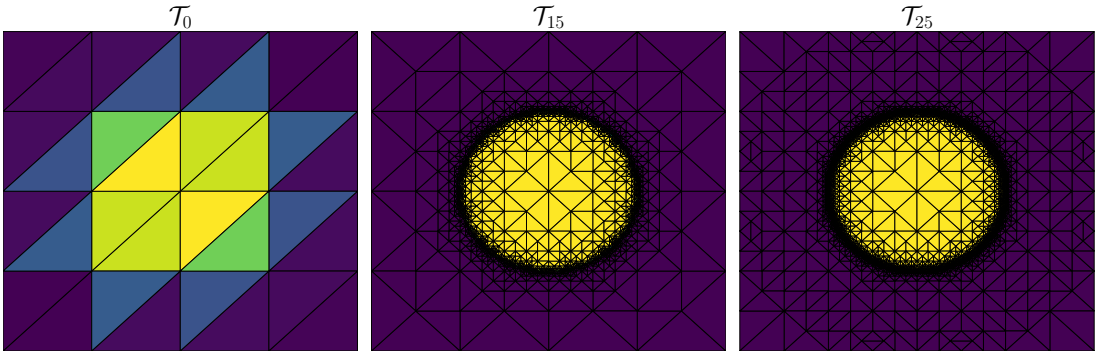

Figure 8: Initial triangulation $\mathcal{T}_0$ and adaptively refined triangulations $\mathcal{T}_k$, $k \in \{15, 25\}$, generated by Algorithm 7.1 for $\theta = \frac{1}{2}$ in the Rudin–Osher–Fatemi image de-noising problem.

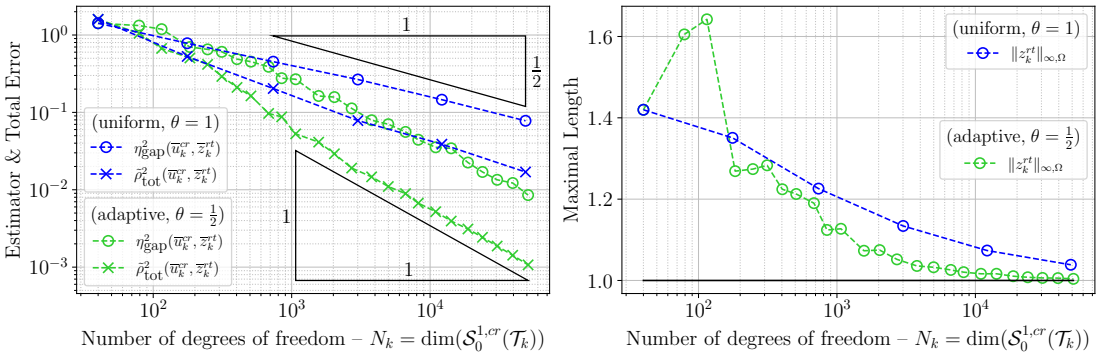

Figure 9: LEFT: primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}_k^{cr}, \overline{z}_k^{rt})$ and alternative total error $\tilde{\rho}^2(\overline{u}_k^{cr}, \overline{z}_k^{rt})$; RIGHT: maximal length of discrete dual solution $\|z_k^{rt}\|_{\infty,\Omega}$; each for $k = 0, \ldots, 25$, when using adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1), and for $k = 0, \ldots, 5$, when using uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1), in the Rudin–Osher–Fatemi image de-noising problem.

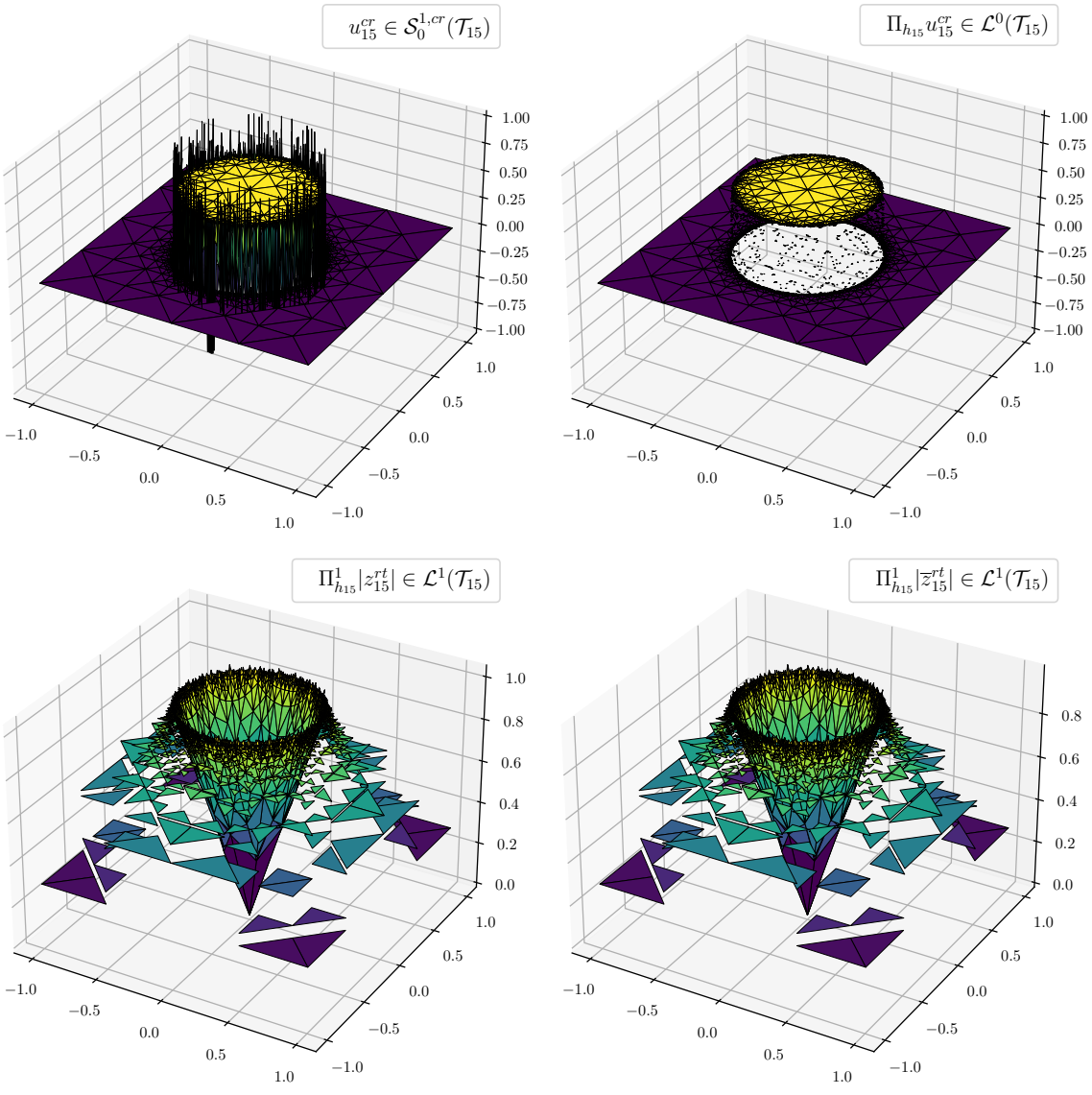

Figure 10: UPPER LEFT: discrete primal solution $u_{15}^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{15})$, UPPER RIGHT: (local) L^2-projection of discrete primal solution $\Pi_{h_{15}} u_{15}^{cr} \in \mathcal{L}^0(\mathcal{T}_{15})$; LOWER LEFT: (local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_{15})$) of discrete dual solution $\Pi_{h_{15}}^1 |z_{15}^{rt}| \in \mathcal{L}^1(\mathcal{T}_{15})$; LOWER RIGHT: (local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_{15})$) of unit-length scaled discrete dual solution $\Pi_{h_{15}}^1 |\overline{z}_{15}^{rt}| \in \mathcal{L}^1(\mathcal{T}_{15})$ (*cf.* (7.2)); each on the triangulation $\mathcal{T}_{15}$ in the Rudin–Osher–Fatemi image de-noising problem.

3D Case. The initial triangulation $\mathcal{T}_0$ of Algorithm 7.1 consists of 384 elements and 125 vertices. Algorithm 7.1 refines the triangulations towards $\partial B_r^3(0)$, *i.e.*, the jump set J_u of the exact solution $u \in BV(\Omega) \cap L^\infty(\Omega)$ (*cf.* (7.3)), which can be observed in Figure 11, in which, more concretely, for $k \in \{0, 5, 9\}$, the discrete primal solution $u_k^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$ and the (local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_k)$) of the modulus of the discrete dual solution $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ are plotted. In Figure 12, one sees that uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) yields the expected reduced convergence rate $h_k \sim N_k^{-\frac{1}{3}}$, $k = 0, \dots, 3$, (predicted by [41, 16]), while adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1) yields the improved convergence rate $h_k^2 \sim N_k^{-\frac{2}{3}}$, $k = 0, \dots, 9$.

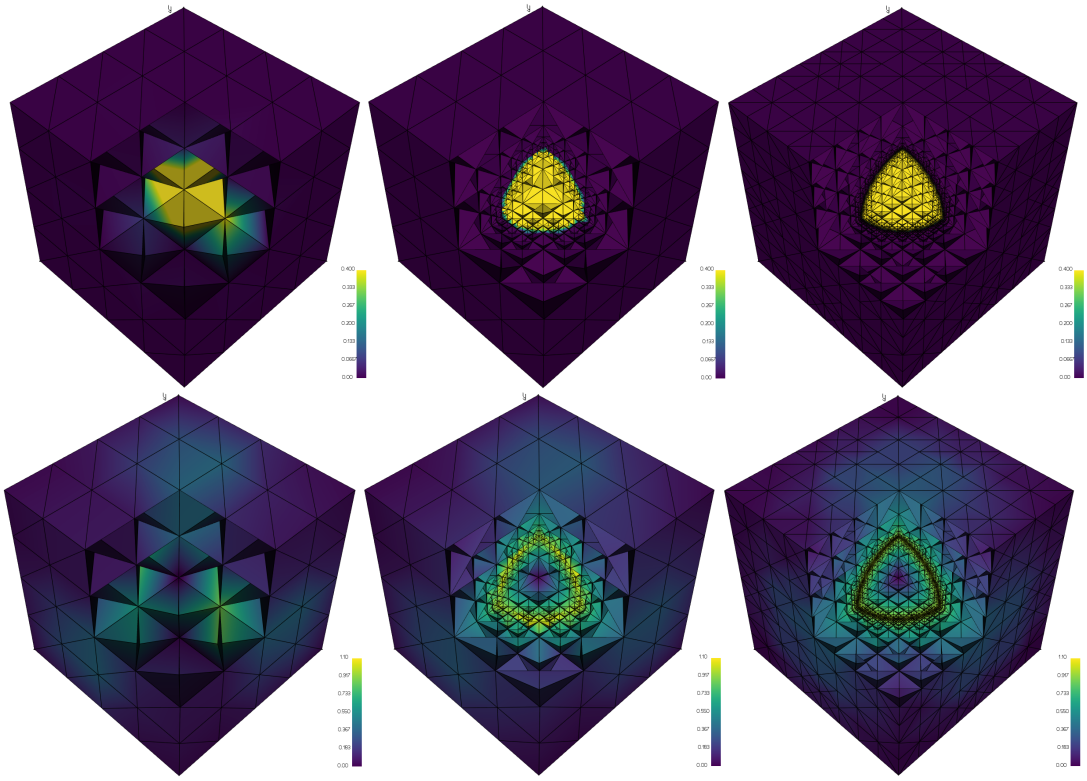

Figure 11: TOP: discrete primal solutions $u_k^{cr} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k)$, $k \in \{0, 5, 9\}$; BOTTOM: (local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_k)$) of the moduli of the discrete dual solutions $\Pi_{h_k}^1 |z_k^{rt}| \in \mathcal{L}^1(\mathcal{T}_k)$, $k \in \{0, 5, 9\}$; each on the triangulation $\mathcal{T}_k$, $k \in \{0, 5, 9\}$, in the Rudin–Osher–Fatemi image de-noising problem.

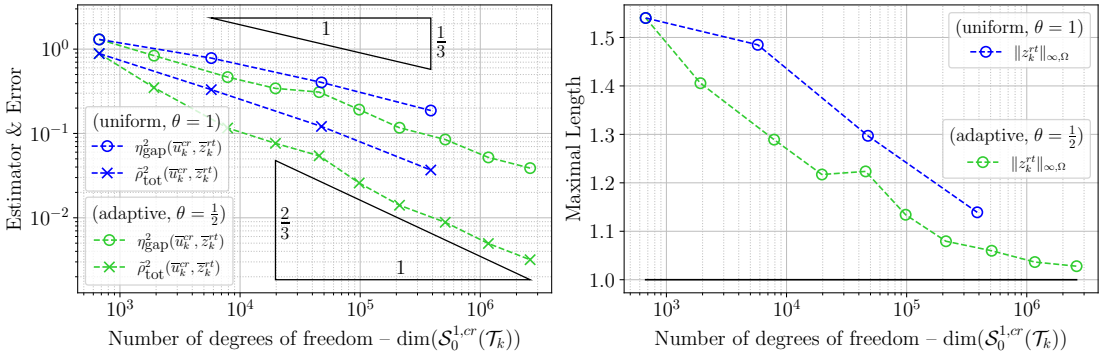

Figure 12: LEFT: primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}_k^{cr}, \overline{z}_k^{rt})$ and alternative total error $\tilde{\rho}^2_{\mathrm{tot}}(\overline{u}_k^{cr}, \overline{z}_k^{rt})$; RIGHT: maximal length of discrete dual solution $\|z_k^{rt}\|_{\infty,\Omega}$; each for $k = 0, \dots, 9$, when using adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1), and for $k = 0, \dots, 3$, when using uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1), in the Rudin–Osher–Fatemi image de-noising problem.

7.6 *Jumping coefficients*

In this subsection, we review the theoretical findings of Subsection 4.5.

7.6.1 *Implementation details regarding the optimization procedure*

Before we present our numerical experiments, again, we briefly outline implementation details regarding the optimization procedure.

Remark 7.6. *(i) The discrete primal solution $u_k^{cr} \in \mathcal{S}^{1,cr}(\mathcal{T}_k)$ (i.e., minimizer of (4.29)) in step ('Solve') is computed using the sparse direct solver from* MUMPS *(version 5.5.0, cf. [5]) applied to the corresponding discrete Euler–Lagrange equation;*

(ii) The reconstruction of the discrete dual solution $z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k)$ (i.e., maximizer of (4.30)) in step ('Solve') is based on the generalized Marini formula (4.31);

(iii) As conforming approximations in ('Solve'), we employ $\overline{u}_k^{cr} := \Pi_{h_k}^{av} u_k^{cr} \in \mathcal{S}^1(\mathcal{T}_k) \subseteq W^{1,2}(\Omega)$ and $\overline{z}_k^{rt} = z_k^{rt} \in \mathcal{RT}^0(\mathcal{T}_k) \subseteq W^2(\mathrm{div};\Omega)$.

(iv) The local refinement indicators $\{\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt})\}_{T\in\mathcal{T}_k} \subseteq \mathbb{R}_{\geq 0}$, for every $T \in \mathcal{T}_h$, are given via

$$\eta_{\mathrm{gap},T}^2(\overline{u}_k^{cr}, z_k^{rt}) := \tfrac{1}{2}\|A^{\frac{1}{2}}(\cdot)\nabla\overline{u}_k^{cr} - A^{-\frac{1}{2}}(\cdot)z_k^{rt}\|_{2,T}^2\,,$$

which follows from restricting $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr}, z_k^{rt})$ (cf. (4.28)) to each element $T \in \mathcal{T}_h$.

7.6.2 *Example with unknown exact solution*

For our numerical experiments, we choose $\Omega := (-1,1)^2$, $\Gamma_D := \partial\Omega$, $\Gamma_N := \emptyset$, $u_D = 0 \in W^{\frac{1}{2},2}(\Gamma_D)$, $f = 1 \in L^2(\Omega)$, and as jumping coefficient matrix $A^\varepsilon\colon \Omega \to \mathbb{R}^{2\times 2}$, for every $\varepsilon \in \{16, 32, 64\}$ and $x \in \Omega$ defined by

$$A^\varepsilon(x) := \begin{cases} \varepsilon \mathrm{I}_{2\times 2} & \text{if } |x + \mathrm{e}_1| < 1\,, \\ \frac{1}{\varepsilon}\mathrm{I}_{2\times 2} & \text{else}\,, \end{cases} \tag{7.5}$$

where $\mathrm{e}_1 := (1,0)^\top \in \mathbb{R}^2$.

As approximation, for $k = 0, \dots, 40$ and $\varepsilon \in \{16, 32, 64\}$, we employ $A_k^\varepsilon := A_{h_k}^\varepsilon := \Pi_{h_k}^\varepsilon A^\varepsilon \in (\mathcal{L}^0(\mathcal{T}_k^\varepsilon))^{2\times 2}$. For every $\varepsilon \in \{16, 32, 64\}$, the primal solution $u^\varepsilon \in W_0^{1,2}(\Omega)$ (*i.e.* minimizer of (4.26) with $A = A^\varepsilon$) is not known and cannot be expected to satisfy $u^\varepsilon \in W^{2,2}(\Omega)$ since $A^\varepsilon \notin C^0(\overline{\Omega}; \mathbb{R}^{2\times 2})$. As a consequence, uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) is expected to yield a reduced convergence rate compared to the quasi-optimal convergence rate $(h_k^\varepsilon)^2 \sim (N_k^\varepsilon)^{-1}$, where $N_k^\varepsilon := \dim(\mathcal{S}_0^{1,cr}(\mathcal{T}_k^\varepsilon))$.

The coarsest triangulation $\mathcal{T}_0$ consists of 32 halved elements and 25 vertices. In Figure 13, for every $\varepsilon \in \{16, 32, 64\}$, the final triangulations $\mathcal{T}_{40}^\varepsilon$ generated by Algorithm 7.1 are displayed. In it, a refinement towards $\partial B_1^2(\mathrm{e}_1) \cap \Omega$, *i.e.*, the jump set J_{A^ε} of the discontinuous coefficient matrix $A^\varepsilon \in L^\infty(\Omega; \mathbb{R}^{2\times 2})$ (*cf.* (7.5)) is reported.

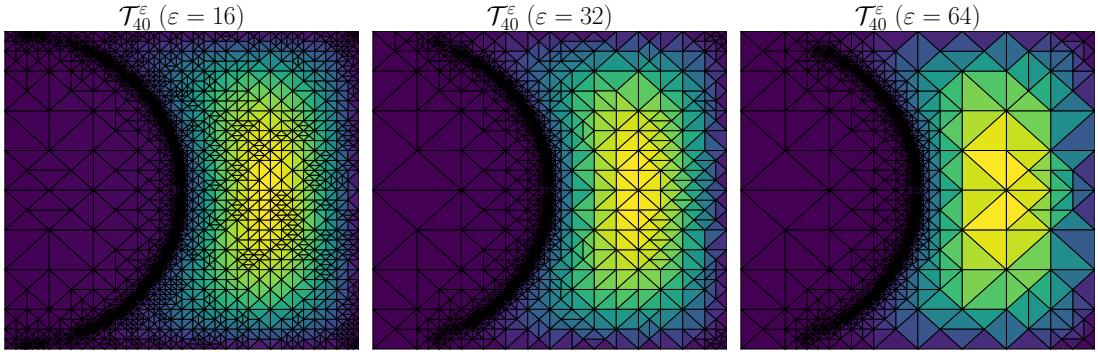

Figure 13: Final triangulations $\mathcal{T}_{40}^\varepsilon$, $\varepsilon \in \{16, 32, 64\}$, generated by Algorithm 7.1 for $\theta = \frac{1}{2}$ in the jumping coefficient problem.

This behavior is also seen in Figure 14, in which, for every $\varepsilon \in \{16, 32, 64\}$, the discrete primal solution $u_{30}^{cr,\varepsilon} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{30}^{\varepsilon})$ (*i.e.*, minimizer of (4.29) with $A_{h_k} = A_{h_k}^{\varepsilon}$) and the node-averaged discrete primal solution $\overline{u}_{30}^{cr,\varepsilon} \in \mathcal{S}_0^{1}(\mathcal{T}_{30}^{\varepsilon})$ are plotted. In addition, Figures 13, 14 show that for increasing value of $\varepsilon \in \{16, 32, 64\}$, the refinement is more concentrated at the jump set $J_{A^\varepsilon} = \partial B_1^2(\mathrm{e}_1) \cap \Omega$.

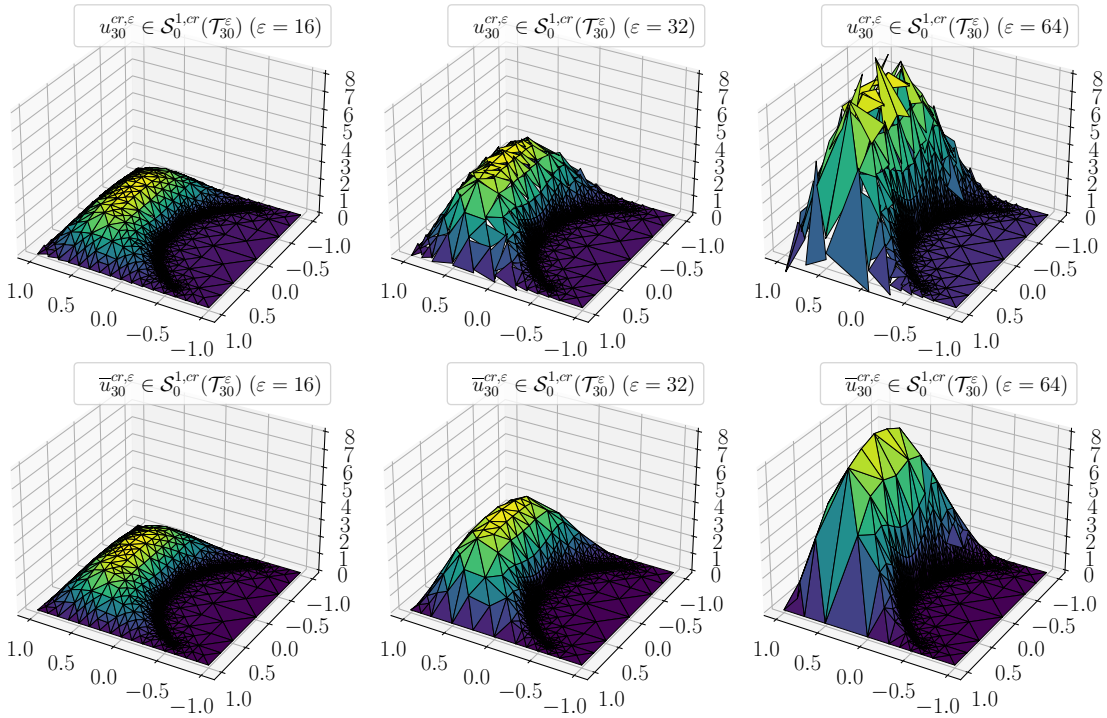

Figure 14: TOP: discrete primal solutions $u_{30}^{cr,\varepsilon} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{30}^{\varepsilon})$, $\varepsilon \in \{16, 32, 64\}$; BOTTOM: node-averaged discrete primal solution $\overline{u}_{30}^{cr,\varepsilon} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{30}^{\varepsilon})$, $\varepsilon \in \{16, 32, 64\}$; each on the triangulation $\mathcal{T}_{30}^{\varepsilon}$ in the jumping coefficient problem.

In Figure 15, for every $\varepsilon \in \{16, 32, 64\}$, one sees that uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1) yields the expected reduced convergence rate $(h_k^{\varepsilon})^{0.7} \sim (N_k^{\varepsilon})^{-0.35}$, $k = 0, \ldots, 5$, while adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1) yields the quasi-optimal convergence rate $(h_k^{\varepsilon})^2 \sim (N_k^{\varepsilon})^{-1}$, $k = 0, \ldots, 40$.

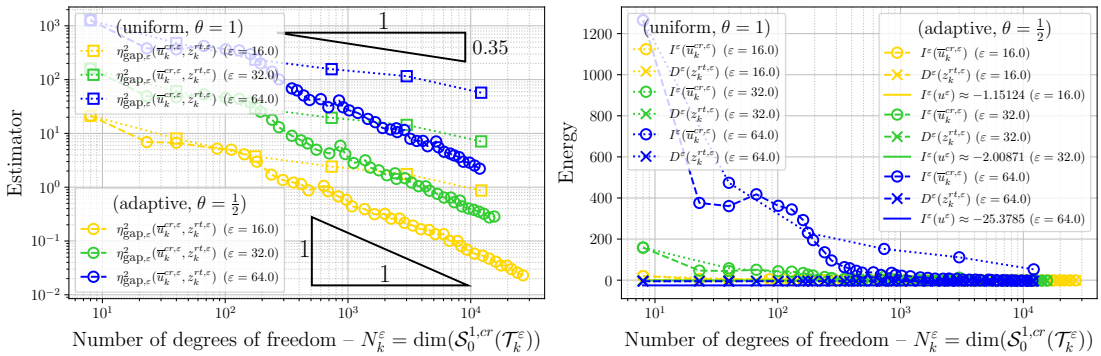

Figure 15: LEFT: primal-dual gap estimator $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr,\varepsilon}, z_k^{rt,\varepsilon})$; RIGHT: primal energy $I^{\varepsilon}(\overline{u}_k^{cr,\varepsilon})$, dual energy $D^{\varepsilon}(z_k^{rt,\varepsilon})$, and primal energy $I^{\varepsilon}(u^{\varepsilon})$ approximated via Aitken's δ^2-process (*cf.* [3]); each for $\varepsilon \in \{16, 32, 64\}$ and $k = 0, \ldots, 40$, when using adaptive mesh refinement (*i.e.*, $\theta = \frac{1}{2}$ in Algorithm 7.1), for $k = 0, \ldots, 40$ and for $k = 0, \ldots, 5$, when uniform mesh refinement (*i.e.*, $\theta = 1$ in Algorithm 7.1), in the jumping coefficient problem.

7.7 Anisotropic mesh refinement

In this subsection, we examine the behavior of the primal-dual gap estimator with respect to anisotropic mesh refinement. For an extensive examination of anisotropic mesh refinement, we refer the reader to [6, 67]. We restrict to the Poisson problem (1.1), *i.e.*, we employ the implementation of Subsection 4.5, but in the case $A(x) \coloneqq \mathrm{I}_{d\times d}$ for a.e. $x \in \Omega$.

7.7.1 Example with unknown exact solution

For our numerical experiments, we choose $\Omega \coloneqq (-1,1)^d \setminus ([0,1]\times[-1,0]^{d-1})$, $d \in \{2,3\}$, $\Gamma_D \coloneqq \partial\Omega$, $\Gamma_N \coloneqq \emptyset$, $u_D = 0 \in W^{\frac{1}{2},2}(\Gamma_D)$, and $f = 1 \in L^2(\Omega)$. For different grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$ and $k = 0, \dots, 40$, we anisotropically refine the triangulations $\mathcal{T}_k^\beta$ towards the origin, where the gradient of the unknown primal solution $u \in W_0^{1,2}(\Omega)$ (*i.e.*, minimizer of (4.26) with $A \coloneqq I_{d\times d}$) is expected to have a singularity. Note that the grading strength $\beta = 1$ precisely corresponds to uniform mesh refinement.

2D Case. For $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, the coarsest triangulation $\mathcal{T}_0^\beta$ consists of 6 element and 8 vertices. Figure 16 depicts the anisotropically refined triangulation $\mathcal{T}_{20}^\beta$ for grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$. In it, one observes that for increasing value grading strength $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, the potential singularity of the gradient of the unknown primal solution $u \in W_0^{1,2}(\Omega)$ at the origin is better resolved. The same behavior can be seen in Figure 18, in which for different grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, the discrete primal solution $u_{20}^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{20}^\beta)$ (*i.e.*, minimizer of (4.26) with $A_{h_{20}} = \mathrm{I}_{2\times 2}$) and the node-averaged discrete primal solution $\overline{u}_{20}^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{20}^\beta)$ are depicted. In Figure 17, one sees that the grading strengths $\beta = \frac{1}{2}$ and $\beta = 1$ yield the reduced convergence rates $(h_k^\beta)^{-0.8} \sim (N_k^\beta)^{-0.4}$ (for $\beta = \frac{1}{2}$) and $(h_k^\beta)^{-1.5} \sim (N_k^\beta)^{-0.75}$ (for $\beta = 1$), respectively, where $N_k^\beta = \dim(\mathcal{S}_0^{1,cr}(\mathcal{T}_k^\beta))$, while the grading strength $\beta = \frac{3}{2}$ yields the quasi-optimal convergence rate $(h_k^\beta)^2 \sim (N_k^\beta)^{-1}$. Overall, we find that the primal-dual gap estimator is robust with respect to the choice of grading strengths.

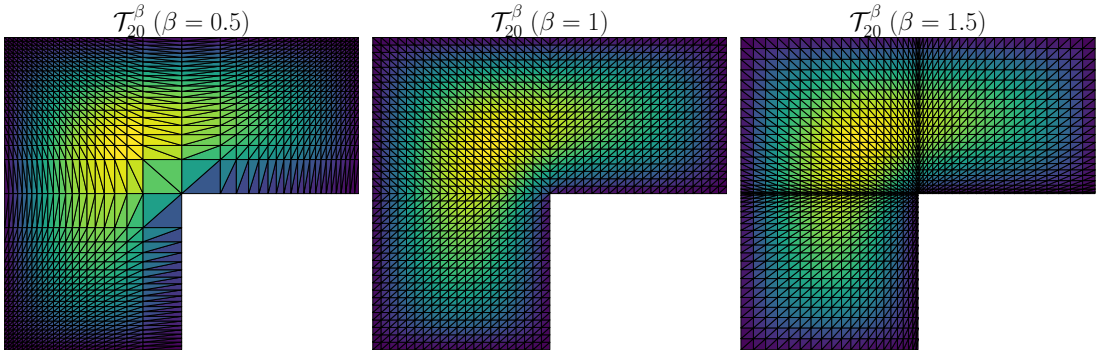

Figure 16: Anisotropically refined triangulations $\mathcal{T}_{20}^\beta$ for different grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$.

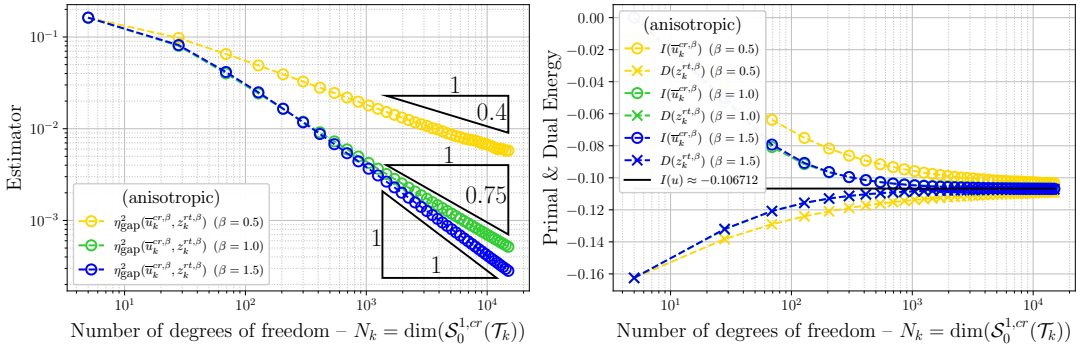

Figure 17: LEFT: primal-dual gap estimator $\eta_{\mathrm{gap}}^2(\overline{u}_k^{cr,\varepsilon}, z_k^{rt,\varepsilon})$; RIGHT: primal energy $I(\overline{u}_k^{cr,\beta})$, dual energy $D(z_k^{rt,\beta})$, and primal energy $I(u)$ approximated via Aitken's δ^2-process (*cf.* [3]); each for grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$ and $k = 0, \dots, 40$.

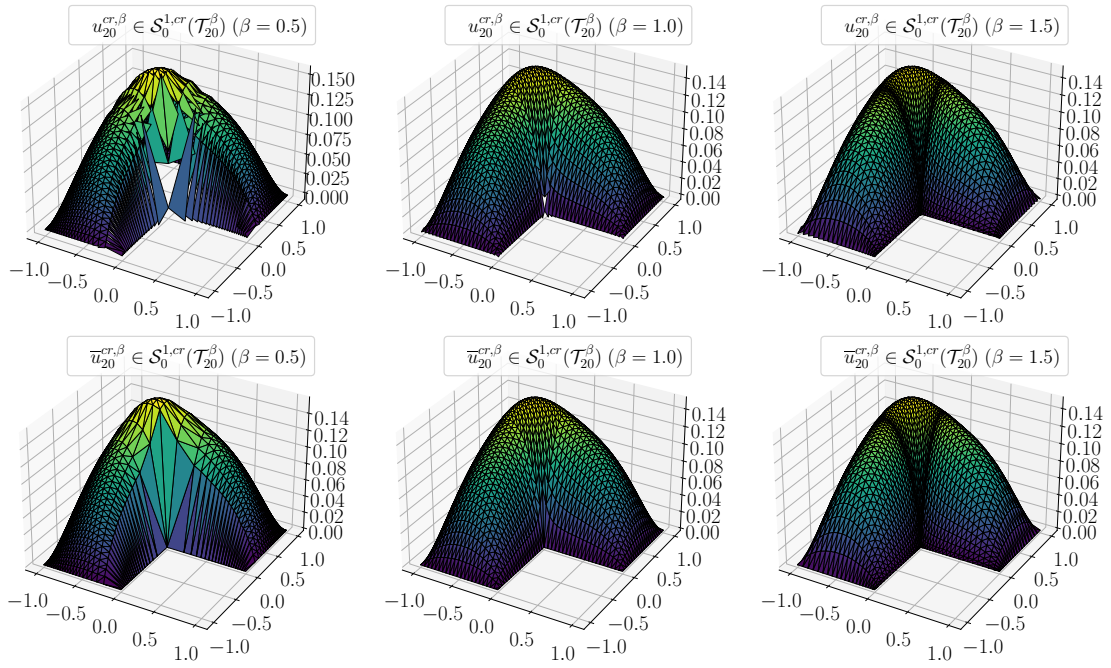

Figure 18: TOP: discrete primal solutions $u_{20}^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{20}^\beta)$, $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$; BOTTOM: node-averaged discrete primal solution $\overline{u}_{20}^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_{20}^\beta)$, $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$; each on the anisotropically refined triangulation $\mathcal{T}_{30}^\beta$ for grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$.

3D Case. For $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, the coarsest triangulation $\mathcal{T}_0^\beta$ consists of 336 elements and 117 vertices. In Figure 20 and Figure 21, for $k \in \{0, 2, 4\}$ and different grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, the discrete primal solution $u_k^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k^\beta)$ (*i.e.*, minimizer of (4.26) with $A_{h_k} = \mathrm{I}_{2\times 2}$) and the (local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_k)$) of the modulus of discrete dual solution $z_k^{rt,\beta} \in \mathcal{RT}^0(\mathcal{T}_k^\beta)$ are depicted. In Figure 19, one sees that the grading strengths $\beta = \frac{1}{2}$ and $\beta = 1$ yield the reduced convergence rates $(h_k^\beta)^{-\frac{3}{4}} \sim (N_k^\beta)^{-\frac{1}{4}}$ (for $\beta = \frac{1}{2}$) and $(h_k^\beta)^{-\frac{3}{2}} \sim (N_k^\beta)^{-\frac{1}{2}}$ (for $\beta = 1$), where $N_k^\beta = \dim(\mathcal{S}_0^{1,cr}(\mathcal{T}_k^\beta))$, while the grading strength $\beta = \frac{3}{2}$ yield the improved convergence rate $(h_k^\beta)^2 \sim (N_k^\beta)^{-\frac{2}{3}}$.

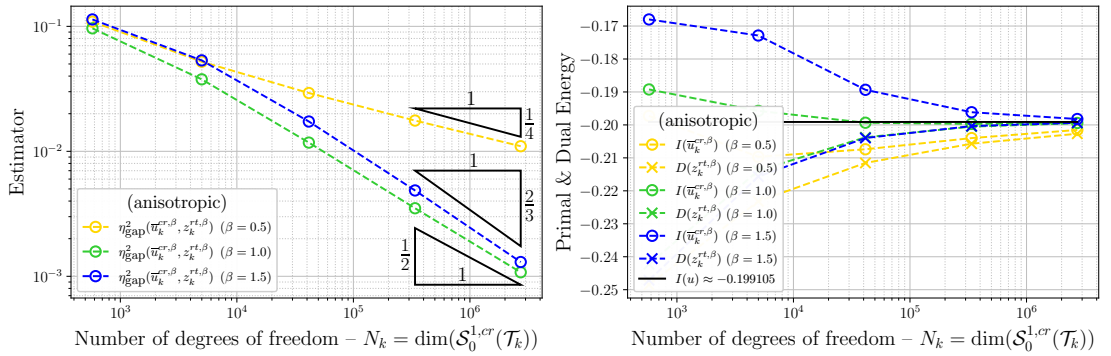

Figure 19: LEFT: primal-dual gap estimator $\eta^2_{\mathrm{gap}}(\overline{u}_k^{cr,\varepsilon}, z_k^{rt,\varepsilon})$; RIGHT: primal energy $I(\overline{u}_k^{cr,\beta})$, dual energy $D(z_k^{rt,\beta})$, and primal energy $I(u)$ approximated via Aitken's δ^2-process (*cf.* [3]); each for grading strengths $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$ and $k = 0, \ldots, 3$.

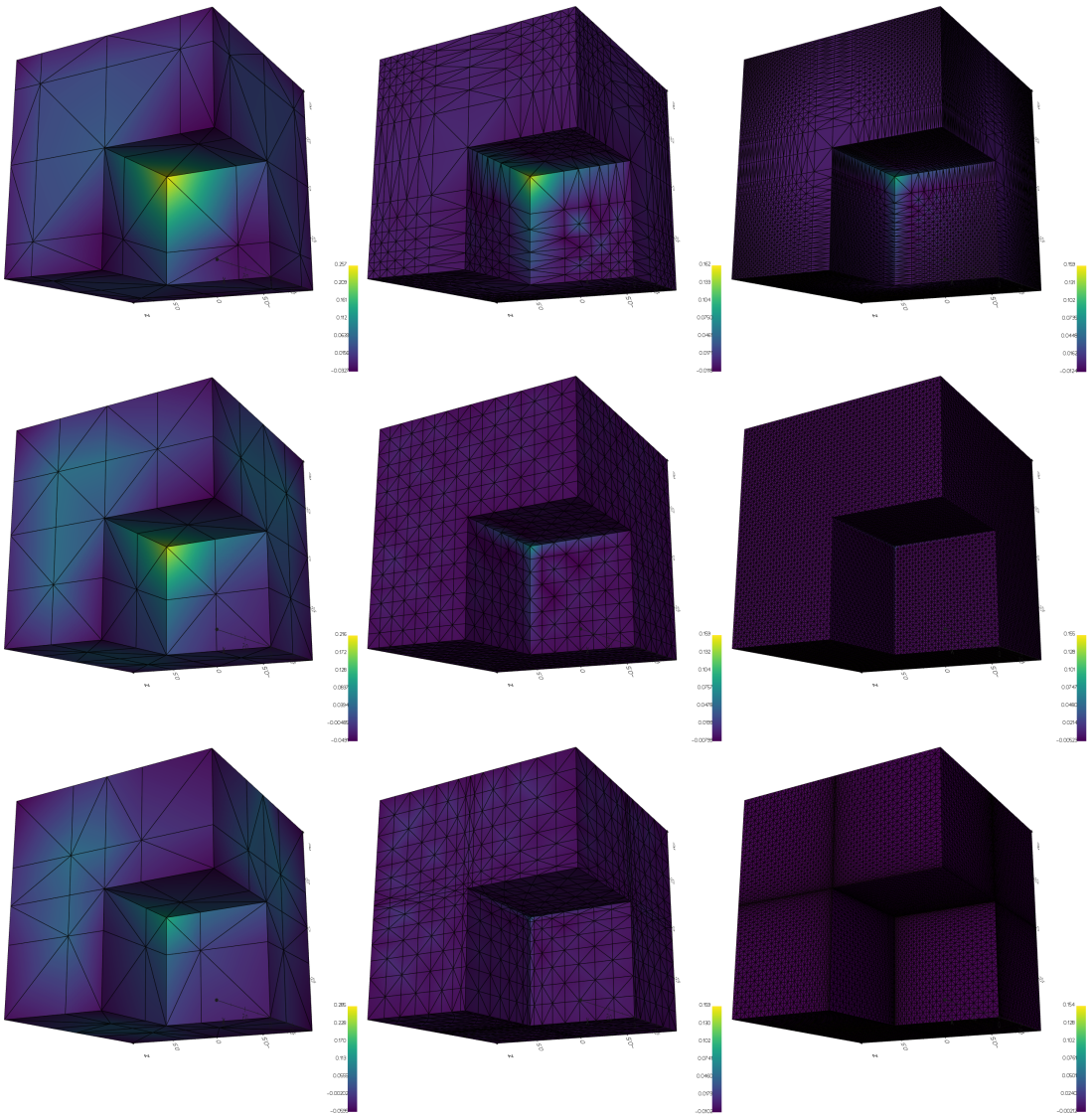

Figure 20: Discrete primal solution $u_k^{cr,\beta} \in \mathcal{S}_0^{1,cr}(\mathcal{T}_k^\beta)$ on triangulations$\mathcal{T}_k^\beta$ obtained using anisotropic mesh refinement for $k \in \{0, 2, 4\}$ (from left two right) and $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$ (from top to bottom).

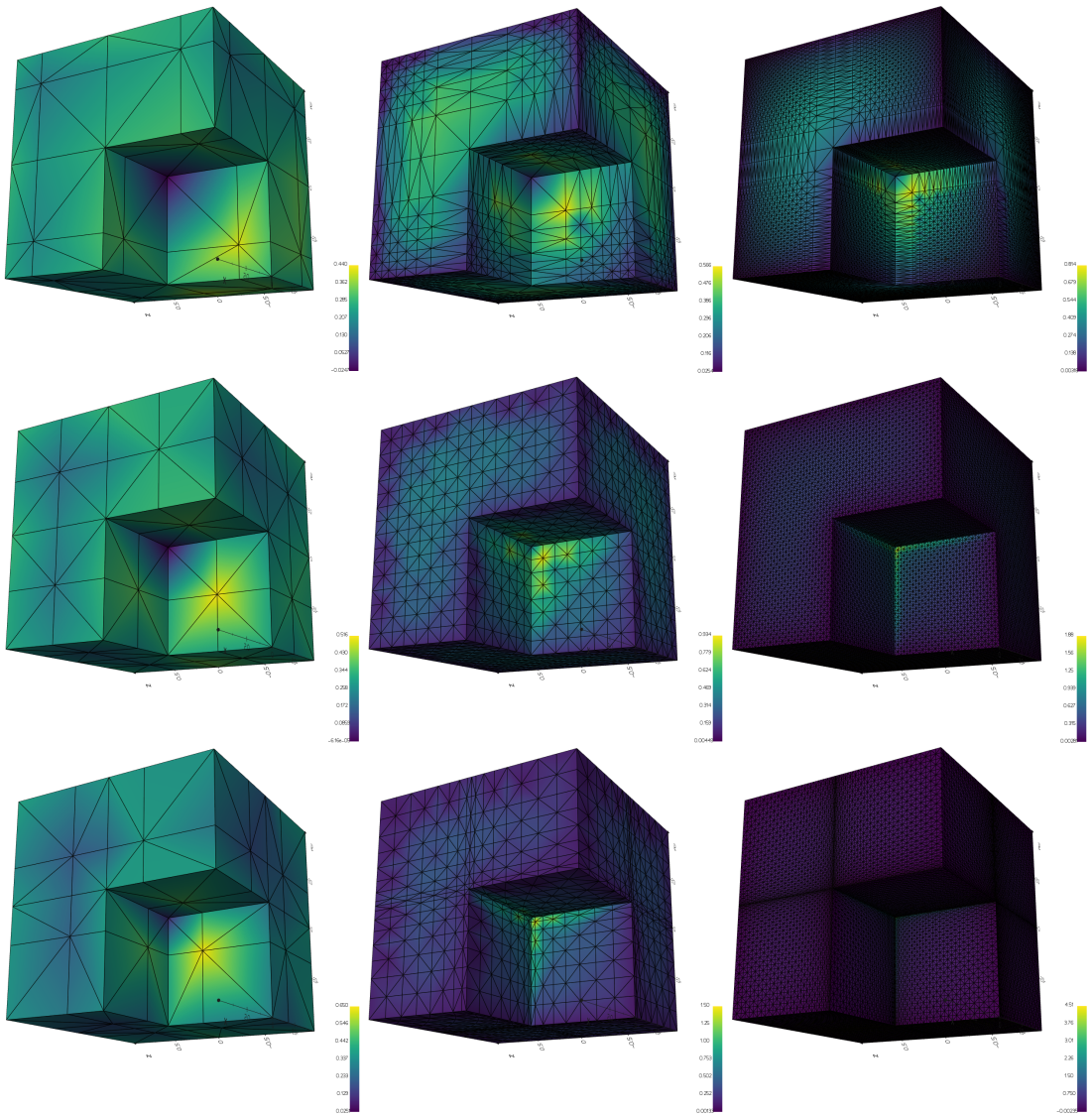

Figure 21: (Local) L^2-projection (onto $\mathcal{L}^1(\mathcal{T}_k^\beta)$) of modulus of discrete dual solution $\Pi^1_{h_k^\beta}|z_k^{cr,\beta}| \in \mathcal{L}^1(\mathcal{T}_k^\beta)$ on triangulation $\mathcal{T}_k^\beta$ obtained using anisotropic mesh refinement for $k \in \{0, 2, 4\}$ (from left two right) and $\beta \in \{\frac{1}{2}, 1, \frac{3}{2}\}$ (from top to bottom).

A. Node-averaging quasi-interpolation operator and interpolation error estimate in terms of (shifted) N-functions

In this appendix, we recall the definition of node-averaging operator $\Pi_h^{av}\colon \mathcal{L}^1(\mathcal{T}_h) \to \mathcal{S}^1_D(\mathcal{T}_h)$ and an interpolation error estimate in terms of (shifted) N-functions.

A.1 Node-averaging quasi-interpolation operator

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

$$\Pi_h^{av} v_h \coloneqq \sum_{\nu\in\mathcal{N}_h} \langle v_h\rangle_\nu \varphi_\nu\,, \qquad \langle v_h\rangle_\nu \coloneqq \begin{cases} \frac{1}{\operatorname{card}(\mathcal{T}_h(\nu))}\sum_{T\in\mathcal{T}_h(\nu)} (v_h|_T)(\nu) & \text{if } \nu \in \Omega\cup\Gamma_N\,,\\ 0 & \text{if } \nu\in\Gamma_D\,,\end{cases}$$

where we denote by $(\varphi_\nu)_{\nu\in\mathcal{N}_h}$ the nodal basis of $\mathcal{S}^1(\mathcal{T}_h)$.

A.2 Interpolation error estimate in terms of (shifted) N-functions

A convex function $\varphi\colon \mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ is said to be an *N-function* if and only if $\varphi(0)=0$, $\varphi(t)>0$ for all $t>0$, $\lim_{t\to 0}\varphi(t)/t=0$, and $\lim_{t\to\infty}\varphi(t)/t=\infty$. Then, there exists a right-derivative $\varphi'\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{>0}$, which is non-decreasing and satisfies $\varphi'(0)=0$, $\varphi'(t)>0$ for all $t>0$, and $\lim_{t\to\infty}\varphi'(t)=\infty$. In addition, an N-function $\varphi\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ is said to satisfy the *Δ_2-condition* (in short, $\varphi\in\Delta_2$) if and only if there exists a constant $c>0$ such that $\varphi(2t)\leq c\,\varphi(t)$ for all $t\geq 0$. We denote the smallest such constant by $\Delta_2(\varphi)$. An N-function $\varphi\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ is said to satisfy the *∇_2-condition* (in short, $\varphi\in\nabla_2$), if its Fenchel conjugate $\varphi^*\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ is an N-function satisfying the Δ_2-condition. If $\varphi\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ satisfies the Δ_2- and the ∇_2-condition (in short, $\varphi\in\Delta_2\cap\nabla_2$), then we define the corresponding family of *shifted N-functions* $\varphi_a\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$, $a\geq 0$, for every $t\geq 0$ by

$$\varphi_a(t) \coloneqq \int_0^t \varphi_a'(s)\,\mathrm{d}s\,, \qquad \text{where } \varphi_a'(s)\coloneqq \frac{\varphi'(a+s)}{a+s}s \text{ for all } s\geq 0\,.$$

Appealing to [46, Lem. 22], it holds that $c_\varphi \coloneqq \sup_{a\geq 0}\Delta_2(\varphi_a)<\infty$. In particular, for every $\varepsilon>0$, there exists a constant $c_\varepsilon>0$, not depending on $a\geq 0$, such that for every $t,s\geq 0$ and $a\geq 0$, there holds the following *ε-Young inequality*:

$$s\,t \leq c_\varepsilon\,(\varphi_a)^*(s) + \varepsilon\,\varphi_a(t)\,. \tag{A.1}$$

Proposition A.1. *Let $\varphi\colon\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ be an N-function such that $\varphi\in\Delta_2\cap\nabla_2$. Then, for every $v_h\in\mathcal{S}^{1,cr}(\mathcal{T}_h)$, $m\in\{0,1,2\}$, $a\geq 0$ and $T\in\mathcal{T}_h$, we have that*

$$\int_T \varphi_a(h_T^m|\nabla_h^m(v_h-\Pi_h^{av}v_h)|)\,\mathrm{d}x \leq c_{av}\int_{\omega_T}\varphi_a(h_T|\nabla_h v_h|)\,\mathrm{d}x\,.$$

where $c_{av}>0$ depends only on c_φ and ω_0.

Proof. See [58, Cor. A.2]. □

References

[1] M. Ainsworth, Robust a posteriori error estimation for nonconforming finite element approximation, *SIAM J. Numer. Anal.* **42** no. 6 (2005), 2320–2341. doi:10.1137/S0036142903425112.

[2] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, *Pure and Applied Mathematics (New York)*, Wiley-Interscience [John Wiley & Sons], New York, 2000.doi:10.1002/9781118032824.

[3] 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.
[4] A. ALONSO, Error estimators for a mixed method, *Numer. Math.* **74** no. 4 (1996), 385–395. doi:10.1007/s002110050222.
[5] P. R. AMESTOY, I. S. DUFF, J. KOSTER, and J.-Y. L'EXCELLENT, A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM Journal on Matrix Analysis and Applications* **23** no. 1 (2001), 15–41.
[6] T. APEL, *Anisotropic finite elements: local estimates and applications*, *Advances in Numerical Mathematics*, B. G. Teubner, Stuttgart, 1999.
[7] T. ARBOGAST and Z. CHEN, On the implementation of mixed methods as nonconforming methods for second-order elliptic problems, *Math. Comp.* **64** no. 211 (1995), 943–972. doi:10.2307/2153478.
[8] D. N. ARNOLD and F. BREZZI, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates, *RAIRO Modél. Math. Anal. Numér.* **19** no. 1 (1985), 7–32.doi:10.1051/m2an/1985190100071.
[9] D. N. ARNOLD and F. BREZZI, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates, *Mathematical Modelling and Numerical Analysis* **19** (1985), 7–32.
[10] I. BABUŠKA, R. DURÁN, and R. RODRÍGUEZ, Analysis of the efficiency of an a posteriori error estimator for linear triangular finite elements, *SIAM J. Numer. Anal.* **29** no. 4 (1992), 947–964.doi:10.1137/0729058.
[11] I. BABUŠKA and T. STROUBOULIS, *The finite element method and its reliability*, *Numerical Mathematics and Scientific Computation*, The Clarendon Press, Oxford University Press, New York, 2001.
[12] S. BALAY ET AL., PETSc Web page, https://www.mcs.anl.gov/petsc, 2019. Available at https://www.mcs.anl.gov/petsc.
[13] A. K. BALCI and A. KALTENBACH, Error analysis for a crouzeix-raviart approximation of the variable exponent dirichlet problem, 2023.
[14] W. BANGERTH and R. RANNACHER, *Adaptive finite element methods for differential equations*, *Lectures in Mathematics ETH Zürich*, Birkhäuser Verlag, Basel, 2003. doi:10.1007/978-3-0348-7605-6.
[15] 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.
[16] 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.
[17] S. BARTELS and A. KALTENBACH, *Error estimates for total-variation regularized minimization problems with singular dual solutions*, 2022. Available at http://arxiv.org/abs/1905.13591.
[18] S. BARTELS and A. KALTENBACH, Error analysis for a Crouzeix–Raviart approximation of the obstacle problem, 2023. doi:10.48550/ARXIV.2302.01646.
[19] 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.
[20] S. BARTELS, Error control and adaptivity for a variational model problem defined on functions of bounded variation, *Math. Comp.* **84** no. 293 (2015), 1217–1240.doi:10.1090/S0025-5718-2014-02893-7.
[21] 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.
[22] S. BARTELS and A. KALTENBACH, Explicit and efficient error estimation for convex minimization problems, *Math. Comp.* **92** no. 343 (2023), 2247–2279. doi:10.1090/mcom/3821.
[23] S. BARTELS, R. TOVEY, and F. WASSMER, Singular solutions, graded meshes,and adaptivity for total-variation regularized minimization problems, *ESAIM Math. Model. Numer. Anal.* **56** no. 6 (2022), 1871–1888. doi:10.1051/m2an/2022056.
[24] S. BARTELS and A. KALTENBACH, Explicit a posteriori error representation for variational problems and application to TV-minimization, 2023.
[25] L. BAUMGÄRTNER, R. BERGMANN, R. HERZOG, S. SCHMIDT, and J. VIDAL-NÚNEZ, Total generalized variation for piecewise constant functions on triangular meshes with applications in imaging, *SIAM J. Imaging Sci.* **16** no. 1 (2023), 313–339. doi:10.1137/22M1505281.
[26] R. BECKER, D. CAPATINA, and R. LUCE, Local flux reconstructions for standard finite element methods on triangular meshes, *SIAM J. Numer. Anal.* **54** no. 4 (2016), 2684–2706.

doi:10.1137/16M1064817.
[27] D. Boffi, F. Brezzi, and M. Fortin, *Finite Elements for the Stokes Problem*, pp. 45–100, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. doi:10.1007/978-3-540-78319-0_2.
[28] D. Braess, R. H. W. 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.
[29] D. Braess and J. Schöberl, Equilibrated residual error estimator for edge elements, *Math. Comp.* **77** no. 262 (2008), 651–672. doi:10.1090/S0025-5718-07-02080-7.
[30] D. Braess, *Finite elements*, third ed., Cambridge University Press, Cambridge, 2007, Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker. doi:10.1017/CBO9780511618635.
[31] D. Braess, An a posteriori error estimate and a comparison theorem for the nonconforming P_1 element, *Calcolo* **46** no. 2 (2009), 149–155. doi:10.1007/s10092-009-0003-z.
[32] D. Braess and J. Schöberl, Equilibrated residual error estimator for edge elements, *Math. Comp.* **77** no. 262 (2008), 651–672.doi:10.1090/S0025-5718-07-02080-7.
[33] L. Brégman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Computational Mathematics and Mathematical Physics* **7** no. 3 (1967), 200–217. doi:https://doi.org/10.1016/0041-5553(67)90040-7.
[34] D. Breit, L. Diening, J. Storn, and J. Wichmann, The parabolic p-Laplacian with fractional differentiability, *IMA J. Numer. Anal.* **41** no. 3 (2021), 2110–2138. doi:10.1093/imanum/draa081.
[35] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, third ed., *Texts in Applied Mathematics* **15**, Springer, New York, 2008.doi:10.1007/978-0-387-75934-0.
[36] S. C. Brenner, Forty years of the Crouzeix-Raviart element, *Numer. Methods Partial Differential Equations* **31** no. 2 (2015), 367–396.doi:10.1002/num.21892.
[37] L. A. Caffarelli, The obstacle problem revisited, *J. Fourier Anal. Appl.* **4** no. 4-5 (1998), 383–402. doi:10.1007/BF02498216.
[38] C. Carstensen and D. J. Liu, Nonconforming FEMs for an optimal design problem, *SIAM J. Numer. Anal.* **53** no. 2 (2015), 874–894. doi:10.1137/130927103.
[39] C. Carstensen and S. A. Funken, Fully reliable localized error control in the FEM, *SIAM J. Sci. Comput.* **21** no. 4 (1999/00), 1465–1484. doi:10.1137/S1064827597327486.
[40] C. Carstensen, J. Gedicke, and D. Rim, Explicit error estimates for Courant, Crouzeix-Raviart and Raviart-Thomas finite element methods, *J. Comput. Math.* **30** no. 4 (2012), 337–353. doi:10.4208/jcm.1108-m3677.
[41] 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.
[42] L. Chamoin and F. Legoll, An introductory review on a posteriori error estimation in finite element computations, *SIAM Rev.* **65** no. 4 (2023), 963–1028. doi:10.1137/21M1464841.
[43] P. G. Ciarlet, *Mathematical elasticity. Volume I. Three-dimensional elasticity*, *Classics in Applied Mathematics* **84**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, [2022] ©2022, Reprint of the 1988 edition [0936420].
[44] 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.
[45] E. Dari, R. Duran, C. Padra, and V. Vampa, A posteriori error estimators for nonconforming finite element methods, *RAIRO Modél. Math. Anal. Numér.* **30** no. 4 (1996), 385–400. doi:10.1051/m2an/1996300403851.
[46] 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.
[47] W. Dörfler, A convergent adaptive algorithm for Poisson's equation, *SIAM J. Numer. Anal.* **33** no. 3 (1996), 1106–1124.doi:10.1137/0733054.
[48] 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.
[49] 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.
[50] A. Ern, A. F. Stephansen, and M. Vohralík, Guaranteed and robust discontinuous Galerkin a

posteriori error estimates for convection-diffusion-reaction problems, *J. Comput. Appl. Math.* **234** no. 1 (2010), 114–130. doi:10.1016/j.cam.2009.12.009.

[51] A. ERN and M. VOHRALÍK, Four closely related equilibrated flux reconstructions for nonconforming finite elements, *C. R. Math. Acad. Sci. Paris* **351** no. 1-2 (2013), 77–80. doi:10.1016/j.crma.2013.01.001.

[52] A. ERN and M. VOHRALÍK, Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations, *SIAM J. Numer. Anal.* **53** no. 2 (2015), 1058–1081.doi:10.1137/130950100.

[53] A. FRIEDMAN, *Variational principles and free-boundary problems*, *A Wiley-Interscience Publication*, John Wiley & Sons, Inc., New York, 1982, Pure and Applied Mathematics.

[54] T. GUDI, A new error analysis for discontinuous finite element methods for linear elliptic problems, *Math. Comp.* **79** no. 272 (2010), 2169–2189. doi:10.1090/S0025-5718-10-02360-4.

[55] T. GUDI, Some nonstandard error analysis of discontinuous Galerkin methods for elliptic problems, *Calcolo* **47** no. 4 (2010), 239–261.doi:10.1007/s10092-010-0022-9.

[56] M. HINTERMÜLLER and K. KUNISCH, Total bounded variation regularization as a bilaterally constrained optimization problem, *SIAM J. Appl. Math.* **64** no. 4 (2004), 1311–1333.doi:10.1137/S0036139903422784.

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

[58] A. KALTENBACH, Error analysis for a crouzeix–raviart approximation of the p-dirichlet problem, *Journal of Numerical Mathematics* **0** no. 0 (2023). doi:doi:10.1515/jnma-2022-0106.

[59] D. J. LIU, A. Q. LI, and Z. R. CHEN, Nonconforming FEMs for the p-Laplace problem, *Adv. Appl. Math. Mech.* **10** no. 6 (2018), 1365–1383.doi:10.4208/aamm.

[60] 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.

[61] R. LUCE and B. I. WOHLMUTH, A local a posteriori error estimator based on equilibrated fluxes, *SIAM J. Numer. Anal.* **42** no. 4 (2004), 1394–1414. doi:10.1137/S0036142903433790.

[62] L. D. MARINI, An inexpensive method for the evaluation of the solution of the lowest order Raviart-Thomas mixed method, *SIAM J. Numer. Anal.* **22** no. 3 (1985), 493–496. doi:10.1137/0722029.

[63] L. D. MARINI, An inexpensive method for the evaluation of the solution of the lowest order Raviart-Thomas mixed method, *SIAM J. Numer. Anal.* **22** no. 3 (1985), 493–496. doi:10.1137/0722029.

[64] P. MORIN, R. H. NOCHETTO, and K. G. SIEBERT, Local problems on stars: a posteriori error estimators, convergence, and performance, *Math. Comp.* **72** no. 243 (2003), 1067–1097. doi:10.1090/S0025-5718-02-01463-1.

[65] M. T. MUSY, marcomusy/vedo: 2023.4.4, March 2023. doi:10.5281/zenodo.7734756.

[66] P. NEITTAANMÄKI and S. REPIN, *Reliable methods for computer simulation*, *Studies in Mathematics and its Applications* **33**, Elsevier Science B.V., Amsterdam, 2004, Error control and a posteriori estimates.

[67] S. NICAISE and E. CREUSÉ, Isotropic and anisotropic a posteriori error estimation of the mixed finite element method for second order operators in divergence form, *Electron. Trans. Numer. Anal.* **23** (2006), 38–62.

[68] R. H. NOCHETTO, G. SAVARÉ, and C. VERDI, A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations, *Communications on Pure and Applied Mathematics* **53** no. 5 (2000), 525–589. doi:https://doi.org/10.1002/(SICI)1097-0312(200005)53:5<525::AID-CPA1>3.0.CO;2-M.

[69] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, and W. YIN, An iterative regularization method for total variation-based image restoration, *Multiscale Modeling & Simulation* **4** no. 2 (2005), 460–489. doi:10.1137/040605412.

[70] A. PLAZA and G. F. CAREY, Local refinement of simplicial grids based on the skeleton, *Appl. Numer. Math.* **32** no. 2 (2000), 195–218 (English). doi:10.1016/S0168-9274(99)00022-7.

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

[72] 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.

[73] S. Repin, *A posteriori estimates for partial differential equations*, *Radon Series on Computational and Applied Mathematics* **4**, Walter de Gruyter GmbH & Co. KG, Berlin, 2008. doi:10.1515/9783110203042.

[74] S. Repin and J. Valdman, Error identities for variational problems with obstacles, *ZAMM Z. Angew. Math. Mech.* **98** no. 4 (2018), 635–658. doi:10.1002/zamm.201700105.

[75] S. I. Repin, A posteriori error estimates for approximate solutions of variational problems with power growth functionals, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **249** no. Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funkts. 29 (1997), 244–255, 317. doi:10.1007/BF02680150.

[76] S. I. Repin, A posteriori error estimation for nonlinear variational problems by duality theory, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **243** no. Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funktsiĭ. 28 (1997), 201–214, 342.doi:10.1007/BF02673600.

[77] S. I. Repin, A posteriori error estimates for approximate solutions to variational problems with strongly convex functionals, *J. Math. Sci. (New York)* **97** no. 4 (1999), 4311–4328, Problems of mathematical physics and function theory. doi:10.1007/BF02365047.

[78] S. I. Repin, A posteriori error estimation for variational problems with uniformly convex functionals, *Math. Comp.* **69** no. 230 (2000), 481–500.doi:10.1090/S0025-5718-99-01190-4.

[79] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, **60**, 1992, Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991), pp. 259–268. doi:10.1016/0167-2789(92)90242-F.

[80] A. Signorini, Questioni di elasticità non linearizzata e semilinearizzata, *Rend. Mat. e Appl. (5)* **18** (1959), 95–139.

[81] I. Smears and M. Vohralík, Simple and robust equilibrated flux *a posteriori* estimates for singularly perturbed reaction-diffusion problems, *ESAIM Math. Model. Numer. Anal.* **54** no. 6 (2020), 1951–1973. doi:10.1051/m2an/2020034.

[82] N. T. Tran, Discrete weak duality of hybrid high-order methods for convex minimization problems, 2023.

[83] A. Veeser and R. Verfürth, Explicit upper bounds for dual norms of residuals, *SIAM J. Numer. Anal.* **47** no. 3 (2009), 2387–2405. doi:10.1137/080738283.

[84] R. Verfürth, A note on constant-free a posteriori error estimates, *SIAM J. Numer. Anal.* **47** no. 4 (2009), 3180–3194. doi:10.1137/080726239.

[85] R. Verfürth and P. Zanotti, A quasi-optimal Crouzeix-Raviart discretization of the Stokes equations, *SIAM J. Numer. Anal.* **57** no. 3 (2019), 1082–1099. doi:10.1137/18M1177688.

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

[87] 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.

[88] M. Vohralík, Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients, *J. Sci. Comput.* **46** no. 3 (2011), 397–438.doi:10.1007/s10915-010-9410-1.