AN ABSTRACT ANALYSIS OF OPTIMAL GOAL-ORIENTED ADAPTIVITY

MICHAEL FEISCHL, DIRK PRAETORIUS, AND KRISTOFFER GEORGE VAN DER ZEE

Abstract. We provide an abstract framework for optimal goal-oriented adaptivity for finite element methods and boundary element methods in the spirit of [13]. We prove that this framework covers standard discretizations of general second-order linear elliptic PDEs and hence generalizes available results [7, 33] beyond the Poisson equation.

1. Introduction

1.1. State of the art & contributions. Standard adaptivity aims to approximate some unknown exact solution u at optimal rate in the energy norm; see, e.g., [15, 20, 37] for adaptive finite element methods (FEM), [18, 19, 21, 23] for adaptive boundary element methods (BEM), and [13] for an overview on available results. Instead, goal-oriented adaptivity aims to approximate, at optimal rate, only the functional value $g(u)$ (also called *quantity of interest* in the literature). Goal-oriented adaptivity is usually more important in practice than standard adaptivity. It has therefore attracted much interest also in the mathematical literature; see, e.g., [6, 8, 9, 16, 24, 27, 35] for some prominent contributions. However, as far as convergence and quasi-optimality of goal-oriented adaptivity is concerned, earlier results are only [7, 33] which are concerned with FEM for the Poisson model problem, the work [25] which considers FEM for more general second-order linear elliptic PDEs, but is concerned with convergence only, and the work [17] which considers point errors in adaptive BEM computations. We note that the analytical arguments of [7, 33] are tailored to the Poisson equation and do not directly transfer to the more general setting of [25], and that [17] relies on the symmetry of the variational formulation, so that the quasi-optimality analysis for goal-oriented adaptivity has also been named as an important open problem in the recent work [12].

This work considers the simultaneous adaptive control of two error estimators $\eta_{u,\star}$ and $\eta_{z,\star}$ which satisfy certain abstract axioms from Section 2.4, below. As in [7, 25, 33], the estimator product $\eta_{u,\star}\eta_{z,\star}$ is designed to control the error in goal-oriented adaptivity. This is discussed in Section 1.2 and demonstrated in Section 4–6 for various model problems and FEM resp. BEM. We analyze two adaptive mesh-refining algorithms: While Algorithm A is a variant of the algorithms from [33, 25], Algorithm B has been proposed in [7]. Both algorithms are proved to be linearly convergent with optimal rates in the sense of certain nonlinear approximation classes. Overall, the contributions and advances of the present work can be summarized as follows:

Date: March 9, 2016.

2010 *Mathematics Subject Classification.* 65N30, 65N50, 65Y20, 41A25.

Key words and phrases. adaptivity, goal-oriented algorithm, quantity of interest, convergence, optimal convergence rates, finite element method, boundary element method.

Acknowledgement. MF and DP are supported by the Austrian Science Fund (FWF) under grant P27005 *Optimal adaptivity for BEM and FEM-BEM coupling* as well as through the FWF doctoral school *Dissipation and Dispersion in Nonlinear PDEs*, funded under grant W1245. KVDZ acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) grant EP/I036427/1.

- We give an abstract analysis for optimal goal-oriented adaptivity which applies to general (non-symmetric) second-order linear elliptic PDEs in the spirit of [20] which even extends the problem class of [25].
- While linear convergence of Algorithm A–B holds for all marking parameters $0 < \theta \leq 1$ (Theorem 12), optimal convergence rates are asymptotically guaranteed for $0 < \theta < \theta_{\rm opt}$ (Algorithm A) resp. $0 < \theta < \theta_{\rm opt}/2$ (Algorithm B) for some *a priori* bound $0 < \theta_{\rm opt} < 1$ which depends on the given problem (Theorem 13, 16). Note that such restrictions also apply to the available results for standard adaptivitiy [13, 15, 20, 37].
- The analysis avoids any (discrete) efficiency estimate and thus allows for simple newest vertex bisection, while [7, 33] follow [37] and require local bisec5-refinement. As firstly observed in [3] and later used in [20, 13], the convergence and quasi-optimality analysis relies essentially on reliability of the error estimator, while efficiency is only used to characterize the estimator-based approximation classes in terms of the so-called total error, i.e., error plus data oscillations (Lemma 19). For the Poisson model problem, we thus obtain, in particular, the same result as [33], but under weaker requirements.
- Unlike [7], our proofs avoid any assumption on the resolution of the given data as, e.g., a saturation assumption [7, eq. (4.4)]. In particular, we give the first general quasi-optimality proof for the algorithm from [7], even for the Poisson model problem.
- Unlike [33, 7, 17], we do not require the symmetry of the weak formulation. Instead, we generalize the quasi-orthogonality property from [13]. In particular and unlike [25], our analysis does not enforce the condition that the initial triangulation is sufficiently fine, since we do not exploit the regularity of the dual solution.
- Finally and inspired by [13], our approach is *a priori* independent of the model problems and covers general linear second-order elliptic PDEs in the frame of the Lax-Milgram lemma, discretized by FEM resp. BEM with fixed order polynomials.

Although we shall verify the mentioned estimator axioms only for standard FEM and BEM discretizations, we expect that they can also be verified for discretizations in the frame of isogeometric analysis; see, e.g., [30] for some goal-oriented adaptive IGAFEM.

1.2. Goal-oriented adaptivity in the framework of the Lax-Milgram lemma. The following introduction covers the main application of the abstract theory, we have in mind. Let $\mathcal{X}$ be a Hilbert space with norm $\|\cdot\|_{\mathcal{X}}$, and let $a(\cdot\,,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a continuous and elliptic bilinear form on $\mathcal{X}$. For given continuous linear functionals $f, g \in \mathcal{X}^*$, we aim to approximate $g(u)$, where $u \in \mathcal{X}$ is the unique solution of

$$a(u\,, v) = f(v) \quad \text{for all } v \in \mathcal{X}. \tag{1}$$

Let $\mathcal{X}_\star \subset \mathcal{X}$ be a finite dimensional subspace associated with some triangulation $\mathcal{T}_\star$ of the problem related domain $\Omega \subset \mathbb{R}^d$. Let $U_\star \in \mathcal{X}_\star$ be the unique Galerkin solution to

$$a(U_\star\,, V_\star) = f(V_\star) \quad \text{for all } V_\star \in \mathcal{X}_\star. \tag{2}$$

Furthermore, let $z \in \mathcal{X}$ be the unique solution to the so-called dual problem

$$a(v\,, z) = g(v) \quad \text{for all } v \in \mathcal{X}. \tag{3}$$

Let $Z_\star \in \mathcal{X}_\star$ be the corresponding Galerkin solution to

$$a(V_\star\,, Z_\star) = g(V_\star) \quad \text{for all } V_\star \in \mathcal{X}_\star. \tag{4}$$

Then, it follows

$$|g(u) - g(U_\star)| = |a(u - U_\star\,, z)| = |a(u - U_\star\,, z - Z_\star)| \lesssim \|u - U_\star\|_{\mathcal{X}}\, \|z - Z_\star\|_{\mathcal{X}}. \tag{5}$$

Here and throughout, $\lesssim$ abbreviates $\le$ up to some generic multiplicative factor $C > 0$ which is clear from the context. Finally, suppose that the Galerkin errors on the right-hand side of (5) can be controlled by computable *a posteriori* error estimators, i.e.,

$$\|u - U_\star\|_{\mathcal{X}} \lesssim \eta_{u,\star} \quad \text{and} \quad \|z - Z_\star\|_{\mathcal{X}} \lesssim \eta_{z,\star}. \tag{6}$$

Under these assumptions, we are altogether led to

$$|g(u) - g(U_\star)| \lesssim \eta_{u,\star}\,\eta_{z,\star}. \tag{7}$$

Overall, we thus aim for some adaptive algorithm which drives the computable upper bound on the right-hand side of (7) to zero with optimal rate.

1.3. Outline. In Section 2, we propose two algorithms and outline the main result. Moreover, we provide the abstract framework in terms of four axioms for the estimators. Section 3 proves optimal convergence rates for each algorithm. In Section 4, we apply the abstract theory to conforming goal-oriented FEM for second-order elliptic PDEs. Section 5 covers goal-oriented FEM for the evaluation of some weighted boundary flux, whereas Section 6 considers goal-oriented adaptivity for BEM.

2. Adaptive Algorithms for the Estimator Product

We suppose that each admissible triangulation $\mathcal{T}_\star$ (see Section 2.2 below) allows for the computation of the error estimators $\eta_{w,\star}$, $w \in \{u, z\}$, with local contributions $\eta_{w,\star}(T) \in \mathbb{R}$ for all $T \in \mathcal{T}_\star$. To abbreviate notation, we shall write

$$\eta_{w,\star} := \eta_{w,\star}(\mathcal{T}_\star), \quad \eta_{w,\star}(\mathcal{U}_\star) := \Big(\sum_{T\in\mathcal{U}_\star} \eta_{w,\star}(T)^2\Big)^{1/2} \quad \text{for } w \in \{u, z\} \text{ and all } \mathcal{U}_\star \subseteq \mathcal{T}_\star.$$

We consider two adaptive strategies (Algorithm A–B) which only differ on how elements are marked refinement in Step (II):

Adaptive algorithm. Input: Initial triangulation $\mathcal{T}_0$, marking strategy (fixed below).
Loop: For all $\ell = 0, 1, 2, 3, \dots$ do (I)–(III):

(I) Compute refinement indicators $\eta_{u,\ell}(T)$ and $\eta_{z,\ell}(T)$ for all $T \in \mathcal{T}_\ell$.
(II) Determine a set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of marked elements.
(III) Let $\mathcal{T}_{\ell+1} := \operatorname{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ be the coarsest refinement of $\mathcal{T}_\ell$ such that all marked elements $T \in \mathcal{M}_\ell$ have been refined.

Output: Sequence of successively refined triangulations $\mathcal{T}_\ell$ and corresponding error estimators $\eta_{u,\ell}, \eta_{z,\ell}$ for all $\ell \in \mathbb{N}_0$. ■

Remark 1. *In the frame of Section 1.2, the computation of $\eta_{u,\ell}$ and $\eta_{z,\ell}$ in Step* (I) *usually requires to solve the primal* (2) *and the dual problem* (4) *to obtain U_ℓ resp. Z_ℓ.* ■

The following marking strategies are designed to drive the estimator product $\eta_{u,\star}\eta_{z,\star}$ to zero with optimal rate. This includes, in particular, the problem class from Section 1.2, but also covers point errors in adaptive BEM computations; see the recent own work [17].

2.1. Marking Stategies. First, we propose a modified version of the marking strategy from [33] which allows for more aggressive marking, i.e., less adaptive steps.

Algorithm A. Parameters: $0 < \theta \le 1$, $C_{\rm mark}, C'_{\rm mark} \ge 1$.
Marking: *For all $\ell = 0, 1, 2, 3, \dots$, Step* (II) *of the adaptive algorithm reads as follows:*

(i) *Determine sets $\mathcal{M}_{u,\ell} \subseteq \mathcal{T}_\ell$ and $\mathcal{M}_{z,\ell} \subseteq \mathcal{T}_\ell$ of up to the multiplicative factor $C_{\rm mark}$ minimal cardinality such that*

$$\theta\, \eta_{u,\ell}^2 \le \eta_{u,\ell}(\mathcal{M}_{u,\ell})^2 \quad \textit{and} \quad \theta\, \eta_{z,\ell}^2 \le \eta_{z,\ell}(\mathcal{M}_{z,\ell})^2. \tag{8}$$

(ii) *Choose* $\widetilde{\mathcal{M}}_\ell \in \{\mathcal{M}_{u,\ell}, \mathcal{M}_{z,\ell}\}$ *to be the set of minimal cardinality and choose* $\mathcal{M}_\ell \subseteq \mathcal{M}_{u,\ell} \cup \mathcal{M}_{z,\ell}$ *such that* $\widetilde{\mathcal{M}}_\ell \subseteq \mathcal{M}_\ell$ *and* $\#\mathcal{M}_\ell \le C'_{\rm mark}\#\widetilde{\mathcal{M}}_\ell$. ■

Remark 2. *In our numerical experiments below, we choose* $\mathcal{M}_\ell$ *as follows: Having picked* $\widetilde{\mathcal{M}}_\ell$ *to be the minimal set amongst* $\mathcal{M}_{u,\ell}$ *and* $\mathcal{M}_{z,\ell}$, *we enlarge* $\widetilde{\mathcal{M}}_\ell$ *by adding the largest* $\#\widetilde{\mathcal{M}}_\ell$ *elements of the other set, e.g., if* $\#\mathcal{M}_{u,\ell} \le \#\mathcal{M}_{z,\ell}$, *then* $\mathcal{M}_\ell$ *consists of* $\mathcal{M}_{u,\ell}$ *plus the* $\#\mathcal{M}_{u,\ell}$ *largest contributions of* $\mathcal{M}_{z,\ell}$. *This yields* $C'_{\rm mark} = 2$. ■

Remark 3. *For* $C'_{\rm mark} = 1$ *and hence* $\mathcal{M}_\ell = \widetilde{\mathcal{M}}_\ell$, *the marking strategy of Algorithm A coincides with that of* [33]. *In various numerical experiments, we observed, however, that the described variant with* $C'_{\rm mark} = 2$ *leads to improved results.* ■

Remark 4. *In* [25], *the authors consider Algorithm A, but define* $\mathcal{M}_\ell := \mathcal{M}_{u,\ell} \cup \mathcal{M}_{z,\ell}$ *in step* (ii). *While this also leads to linear convergence in the sense of Theorem 12,* [25] *only proves suboptimal convergence rates* $\min\{s,t\}$ *instead of the optimal rate* $s+t$ *in Theorem 13; see* [25, Section 4]. *We note that the strategy of* [25] *leads to linear convergence* $\eta_{u,\ell+n} \le Cq^n\eta_{u,\ell}$ *and* $\eta_{z,\ell+n} \le Cq^n\eta_{z,\ell}$ *for either estimator and all* $\ell, n \in \mathbb{N}_0$, *where* $C > 0$ *and* $0 < q < 1$ *are independent constants, while the optimal strategies considered in this work only enforce* $\eta_{u,\ell+n}\eta_{z,\ell+n} \le Cq^n\eta_{u,\ell}\eta_{z,\ell}$ *for the product.* ■

Second, the following algorithm has been proposed in [7] for goal-oriented adaptive FEM for the Poisson problem. We note that [7] requires a saturation assumption for the related data oscillation terms in the case of non-polynomial volume forces (see [7, eq. (4.4)] and [7, Theorem 4.1]) which is proved unnecessary by our analysis.

Algorithm B. PARAMETERS: $0 < \theta \le 1$, $C_{\rm mark} \ge 1$.
MARKING: *For all* $\ell = 0, 1, 2, 3, \ldots$, *Step* (II) *of the adaptive algorithm reads as follows:*
(i) *Assemble refinement indicators* $\rho_\ell(T)^2 := \eta_{u,\ell}(T)^2\eta_{z,\ell}^2 + \eta_{u,\ell}^2\eta_{z,\ell}(T)^2$ *for all* $T \in \mathcal{T}_\ell$.
(ii) *Determine a set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\rm mark}$ *minimal cardinality such that*

$$\theta\,\rho_\ell^2 \le \rho_\ell(\mathcal{M}_\ell)^2. \tag{9}$$ ■

2.2. Mesh-refinement. We suppose that the mesh-refinement is a deterministic and fixed strategy, e.g., newest vertex bisection [38]. For each triangulation $\mathcal{T}$ and marked elements $\mathcal{M} \subseteq \mathcal{T}$, we let $\mathcal{T}' := \mathrm{refine}(\mathcal{T}, \mathcal{M})$ be the coarsest triangulation, where all elements $T \in \mathcal{M}$ have been refined, i.e., $\mathcal{M} \subseteq \mathcal{T}\backslash\mathcal{T}'$. We write $\mathcal{T}' \in \mathrm{refine}(\mathcal{T})$, if there exist finitely many triangulations $\mathcal{T}^{(0)}, \ldots, \mathcal{T}^{(n)}$ and sets $\mathcal{M}^{(j)} \subseteq \mathcal{T}^{(j)}$ such that $\mathcal{T} = \mathcal{T}^{(0)}$, $\mathcal{T}' = \mathcal{T}^{(n)}$ and $\mathcal{T}^{(j)} = \mathrm{refine}(\mathcal{T}^{(j-1)}, \mathcal{M}^{(j-1)})$ for all $j = 1, \ldots, n$, where we formally allow $n = 0$, i.e., $\mathcal{T} = \mathcal{T}^{(0)} \in \mathrm{refine}(\mathcal{T})$. To abbreviate notation, let $\mathbb{T} := \mathrm{refine}(\mathcal{T}_0)$, where $\mathcal{T}_0$ is the given initial triangulation of Algorithms A–B.

2.3. Main result. Let $\mathbb{T}_N := \{\mathcal{T} \in \mathbb{T} : \#\mathcal{T} - \#\mathcal{T}_0 \le N\}$ denote the (finite) set of all refinements of $\mathcal{T}_0$ which have at most N elements more than $\mathcal{T}_0$. For $s > 0$ and $w \in \{u, z\}$, we write $w \in \mathbb{A}_s$ if

$$\|w\|_{\mathbb{A}_s} := \sup_{N\in\mathbb{N}_0}\Big((N+1)^s \min_{\mathcal{T}_\star\in\mathbb{T}_N} \eta_{w,\star}\Big) < \infty,$$

where $\eta_{w,\star}$ is the error estimator associated with the optimal triangulation $\mathcal{T}_\star \in \mathbb{T}_N$. In explicit terms, $\|w\|_{\mathbb{A}_s} < \infty$ means that an algebraic convergence rate $\mathcal{O}(N^{-s})$ for the error estimator is possible, if the optimal triangulations are chosen.

For either algorithm, our main result is twofold: First, we prove linear convergence (Section 3.1): For each $0 < q < 1$, there exists some n such that for all $\ell \in \mathbb{N}$, it holds

$\eta_{u,\ell+n}\,\eta_{z,\ell+n} \le q\,\eta_{u,\ell}\,\eta_{z,\ell}$. Second, we prove optimal convergence behavior (Section 3.3): With respect to the number of elements $N \simeq \#\mathcal{T}_\ell - \#\mathcal{T}_0$, the product $\eta_{u,\ell}\,\eta_{z,\ell}$ decays with order $\mathcal{O}(N^{-(s+t)})$ for each possible algebraic rate $s+t>0$, i.e., $\|u\|_{\mathbb{A}_s} + \|z\|_{\mathbb{A}_t} < \infty$.

Remark 5. *Since our analysis works with the estimator instead of the error, it avoids the use of any (discrete) efficiency bound. Unlike* [7, 33], *this allows to use simple newest vertex bisection. Moreover, Lemma 19 below states that for standard FEM our approximation classes $\mathbb{A}_s$ coincide with those of* [7, 15, 33] *which are defined through the so-called total error (i.e., error plus data oscillations).* ■

2.4. Axioms of Adaptivity. Recall the notation of Section 2.2. Let $\mathbb{d}_w(\cdot\,,\cdot) : \mathbb{T}\times\mathbb{T} \to \mathbb{R}_{\ge 0}$ denote a distance function on the set of admissible triangulations which satisfies

$$C_{\rm dist}^{-1}\mathbb{d}_w(\mathcal{T}\,,\mathcal{T}'') \le \mathbb{d}_w(\mathcal{T}\,,\mathcal{T}') + \mathbb{d}_w(\mathcal{T}'\,,\mathcal{T}'') \quad \text{for all } \mathcal{T},\mathcal{T}',\mathcal{T}'' \in \mathbb{T},$$
$$\mathbb{d}_w(\mathcal{T}\,,\mathcal{T}') \le C_{\rm dist}\mathbb{d}_w(\mathcal{T}'\,,\mathcal{T}) \quad \text{for all } \mathcal{T},\mathcal{T}' \in \mathbb{T},$$

with some uniform constant $C_{\rm dist} > 0$; see also Remark 8 below.

The convergence and optimality analysis of the adaptive algorithms requires the following four *axioms of adaptivity* [13], where (A4) is relaxed when compared to [13]:

(A1) *Stability on non-refined elements*: There exists $C_{\rm stb} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\bullet)$ the corresponding error estimators satisfy

$$|\eta_{w,\star}(\mathcal{T}_\bullet \cap \mathcal{T}_\star) - \eta_{w,\bullet}(\mathcal{T}_\bullet \cap \mathcal{T}_\star)| \le C_{\rm stb}\,\mathbb{d}_w(\mathcal{T}_\bullet\,,\mathcal{T}_\star).$$

(A2) *Reduction on refined elements*: There exist $0 < q_{\rm red} < 1$ and $C_{\rm red} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\bullet)$ the corresponding error estimators satisfy

$$\eta_{w,\star}(\mathcal{T}_\star \backslash \mathcal{T}_\bullet)^2 \le q_{\rm red}\,\eta_{w,\bullet}(\mathcal{T}_\bullet \backslash \mathcal{T}_\star)^2 + C_{\rm red}\,\mathbb{d}_w(\mathcal{T}_\bullet\,,\mathcal{T}_\star)^2.$$

(A3) *Discrete reliability*: There exists $C_{\rm rel} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\bullet)$, there exists $\mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star) \subseteq \mathcal{T}_\bullet$ with $\mathcal{T}_\bullet \backslash \mathcal{T}_\star \subseteq \mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star)$ such that

$$\mathbb{d}_w(\mathcal{T}_\star\,,\mathcal{T}_\bullet) \le C_{\rm rel}\,\eta_{w,\ell}(\mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star)) \quad \text{and} \quad \#\mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star) \le C_{\rm rel}\,\#(\mathcal{T}_\bullet \backslash \mathcal{T}_\star).$$

(A4) *Quasi-orthogonality*: Let $\mathcal{T}_{\ell_n}$ be the (possibly finite) subsequence of triangulations $\mathcal{T}_\ell$ generated by Algorithm A or B which satisfy

$$\theta\,\eta_{w,\ell_n}^2 \le \eta_{w,\ell_n}(\mathcal{T}_{\ell_n} \backslash \mathcal{T}_{\ell_n+1})^2. \tag{10}$$

Then, for all $\varepsilon > 0$, there exists $C_{\rm orth}(\varepsilon) > 0$ such that for all $n \le N$, for which $\mathcal{T}_{\ell_n}, \dots, \mathcal{T}_{\ell_N}$ are well-defined, it holds

$$\sum_{j=n}^{N} \big(\mathbb{d}_w(\mathcal{T}_{\ell_{j+1}}\,,\mathcal{T}_{\ell_j})^2 - \varepsilon\,\eta_{w,\ell_j}^2\big) \le C_{\rm orth}(\varepsilon)\,\eta_{w,\ell_n}^2.$$

We recall some observations of [13].

Lemma 6 (quasi-monotonicity of estimator [13, Lemma 3.5])**.** *There exists $C_{\rm mon} > 0$ which depends only on* (A1)–(A3), *such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\bullet)$, it holds $\eta_{w,\star}^2 \le C_{\rm mon}\,\eta_{w,\bullet}^2$.* ■

Lemma 7 (optimality of Dörfler marking [13, Proposition 4.12])**.** *Suppose stability* (A1) *and discrete reliability* (A3). *For all $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2)^{-1}$, there exists some $0 < \kappa_{\rm opt} < 1$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\bullet)$, it holds*

$$\eta_{w,\star}^2 \le \kappa_{\rm opt}\,\eta_{w,\bullet}^2 \quad \Longrightarrow \quad \theta\,\eta_{w,\bullet}^2 \le \eta_{w,\bullet}(\mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star))^2, \tag{11}$$

where $\mathcal{R}_w(\mathcal{T}_\bullet, \mathcal{T}_\star)$ is the set of refined elements from (A3). ■

Remark 8. (i) *In the setting of Section 1.2, let* $w \in \{u, z\}$ *with* $W_\star \in \{U_\star, Z_\star\}$ *being the corresponding Galerkin solution for* $\mathcal{T}_\star \in \mathbb{T}$*. The abstract distance is then usually defined by* $\mathrm{d\!l}_w(\mathcal{T}_\bullet, \mathcal{T}_\star) := a(W_\star - W_\bullet, W_\star - W_\bullet)^{1/2} \simeq \|W_\star - W_\bullet\|_{\mathcal{X}}$*; see Section 4–6 below.*

(ii) *Suppose that the bilinear form* $a(\cdot,\cdot)$ *is additionally symmetric, and let* $|\!|\!|v|\!|\!| := a(v,v)^{1/2}$ *denote the equivalent energy norm on* $\mathcal{X}$*. Then, nestedness* $\mathcal{X}_n \subseteq \mathcal{X}_m \subseteq \mathcal{X}_k$ *of the discrete spaces for all* $k \geq m \geq n$ *implies the Galerkin orthogonality*

$$|\!|\!|W_k - W_m|\!|\!|^2 + |\!|\!|W_m - W_n|\!|\!|^2 = |\!|\!|W_k - W_n|\!|\!|^2 \quad \text{for all } k \geq m \geq n.$$

This and (A3) *imply*

$$\sum_{j=n}^{N} \mathrm{d\!l}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 = \sum_{j=n}^{N} \big(|\!|\!|W_{\ell_{j_{N+1}}} - W_{\ell_j}|\!|\!|^2 - |\!|\!|W_{\ell_{j_{N+1}}} - W_{\ell_{j+1}}|\!|\!|^2\big)$$

$$\leq |\!|\!|W_{\ell_{j_{N+1}}} - W_{\ell_n}|\!|\!|^2 \overset{\text{(A3)}}{\lesssim} \eta^2_{w,\ell_n}.$$

This shows the quasi-orthogonality (A4) *with* $\varepsilon = 0$ *and* $C_{\mathrm{orth}}(\varepsilon) = C^2_{\mathrm{rel}}$*.* ■

2.5. Generalized linear convergence. The following estimator reduction is first found in [15] for $\mathcal{T}_\star = \mathcal{T}_{\ell+1}$ and, e.g., proved along the lines of [13, Lemma 4.7].

Lemma 9 (generalized estimator reduction)**.** *Let* $0 < \theta \leq 1$*. Let* $\mathcal{T}_\ell \in \mathbb{T}$ *and* $\mathcal{T}_{\ell+1} \in \mathrm{refine}(\mathcal{T}_\ell)$*. Suppose that the refined elements satisfy the Dörfler marking*

$$\theta\, \eta^2_{w,\ell} \leq \eta_{w,\ell}(\mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1})^2. \tag{12}$$

Then, there exist constants $0 < q_{\mathrm{est}} < 1$ *and* $C_{\mathrm{est}} > 0$ *which depend only on* (A1)–(A2) *and* θ*, such that for all* $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_{\ell+1})$*, it holds*

$$\eta^2_{w,\star} \leq q_{\mathrm{est}}\, \eta^2_{w,\ell} + C_{\mathrm{est}}\, \mathrm{d\!l}_w(\mathcal{T}_\star, \mathcal{T}_\ell)^2. \tag{13}$$

■

The following result generalizes [13, Proposition 4.10] to the present setting. We note that (A3) enters only through the quasi-monotonicity of the estimator (Lemma 6).

Proposition 10 (generalized linear convergence)**.** *Let* $\mathcal{T}_\ell$ *be a sequence of successively refined triangulations, i.e.,* $\mathcal{T}_\ell \in \mathrm{refine}(\mathcal{T}_{\ell-1})$ *for all* $\ell \in \mathbb{N}$*. Let* $0 < \theta \leq 1$*. Then, there exist* $0 < q_{\mathrm{conv}} < 1$ *and* $C_{\mathrm{conv}} > 0$ *which depend only on* (A1)–(A4) *and* θ*, such that the following holds: Let* $\ell, n \in \mathbb{N}_0$ *and suppose that there are at least* $k \leq n$ *indices* $\ell \leq \ell_1 < \ell_2 < \cdots < \ell_k < \ell + n$ *such that*

$$\theta\, \eta^2_{w,\ell_j} \leq \eta_{w,\ell_j}(\mathcal{T}_{\ell_j} \backslash \mathcal{T}_{\ell_j+1})^2 \quad \text{for all } j = 1, \ldots k. \tag{14}$$

Then, the error estimator satisfies

$$\eta^2_{w,\ell+n} \leq C_{\mathrm{conv}}\, q^k_{\mathrm{conv}}\, \eta^2_{w,\ell}. \tag{15}$$

Proof. To abbreviate notation, set $\ell_0 := \ell$. Note that $\mathcal{T}_{\ell_{k+1}} \in \mathrm{refine}(\mathcal{T}_{\ell_k+1})$. Therefore, the estimator reduction (13) shows for all $\varepsilon > 0$ and all $0 \leq j \leq k$

$$\sum_{i=k-j}^{k} \eta^2_{w,\ell_{i+1}} \leq \sum_{i=k-j}^{k} \Big(q_{\mathrm{est}} \eta^2_{w,\ell_i} + C_{\mathrm{est}} \mathrm{d\!l}_w(\mathcal{T}_{\ell_{i+1}}, \mathcal{T}_{\ell_i})^2\Big)$$

$$= \sum_{i=k-j}^{k} \Big((q_{\mathrm{est}} + C_{\mathrm{est}}\varepsilon)\eta^2_{w,\ell_i} + C_{\mathrm{est}}\big(\mathrm{d\!l}_w(\mathcal{T}_{\ell_{i+1}}, \mathcal{T}_{\ell_i})^2 - \varepsilon \eta^2_{w,\ell_i}\big)\Big).$$

Choose $\varepsilon < (1-q_{\rm est})C_{\rm est}^{-1}$ so that $\kappa := 1-(q_{\rm est}+C_{\rm est}\varepsilon) > 0$. For $0 \le j \le k$, (A4) shows

$$\begin{aligned}(16)\qquad \kappa \sum_{i=k-j}^{k} \eta^2_{w,\ell_{i+1}} &\le \eta^2_{w,\ell_{k-j}} + C_{\rm est} \sum_{i=k-j}^{k} \big(\mathbb{d}_w(\mathcal{T}_{\ell_{i+1}}, \mathcal{T}_{\ell_i})^2 - \varepsilon \eta^2_{w,\ell_i}\big)\\ &\le (1+C_{\rm est}C_{\rm orth}(\varepsilon))\eta^2_{w,\ell_{k-j}}.\end{aligned}$$

With $C := (1+C_{\rm est}C_{\rm orth}(\varepsilon))/\kappa > 1$, mathematical induction below shows

$$(17)\qquad \eta^2_{w,\ell_k} \le (1-C^{-1})^j \sum_{i=k-j}^{k} \eta^2_{w,\ell_i} \quad \text{for all } 0 \le j \le k.$$

To see (17), note that the case $j=0$ holds with equality. Suppose that (17) holds for $j<k$. This induction hypothesis and (16) show

$$\begin{aligned}\eta^2_{w,\ell_k} \le (1-C^{-1})^j \sum_{i=k-j}^{j} \eta^2_{w,\ell_i} &= (1-C^{-1})^j \Big(\Big(\sum_{i=k-(j+1)}^{k} \eta^2_{w,\ell_i}\Big) - \eta^2_{w,\ell_{k-(j+1)}}\Big)\\ &\overset{(16)}{\le} (1-C^{-1})^{j+1} \sum_{i=k-(j+1)}^{k} \eta^2_{w,\ell_i},\end{aligned}$$

which proves the validity of the induction step. Hence, the assertion (17) holds for all $j \le k$. By use of Lemma 6, (17) for $j = k-1$, and (16) for $j=k$, we obtain

$$\begin{aligned}C_{\rm mon}^{-1}\eta^2_{w,\ell+n} \le \eta^2_{w,\ell_k} &\overset{(17)}{\le} (1-C^{-1})^{k-1}\sum_{i=1}^{k}\eta^2_{w,\ell_i} \le (1-C^{-1})^{k-1}\sum_{i=0}^{k}\eta^2_{w,\ell_{i+1}}\\ &\overset{(16)}{\le} (1-C^{-1})^{k-1}C\,\eta^2_{w,\ell_0} = (1-C^{-1})^k C/(1-C^{-1})\,\eta^2_{w,\ell}.\end{aligned}$$

This concludes the proof with $C_{\rm conv} = CC_{\rm mon}/(1-C^{-1})$ and $q_{\rm conv} = (1-C^{-1})$. ■

3. Optimal Convergence of Adaptive Algorithms

Throughout this section, we suppose that the error estimators $\eta_{u,\ell}$ and $\eta_{z,\ell}$ satisfy the respective assumptions (A1)–(A4) of Section 2.4. Without loss of generality, we suppose that $\eta_{u,\ell}$ and $\eta_{z,\ell}$ satisfy the axioms (A1)–(A4) with the same constants.

Remark 11. *The axioms* (A1)–(A4) *are designed for weighted-residual error estimators in the frame of FEM and BEM. For optimal adaptivity for the energy error, it is sufficient that for $w \in \{u,z\}$ the error estimator $\eta_{w,\ell}$ used in the adaptive algorithm is locally equivalent to some error estimator $\widetilde{\eta}_{w,\ell}$ which satisfies* (A1)–(A4)*, i.e.,*

$$\eta_{\ell,w}(T) \lesssim \widetilde{\eta}_{\ell,w}(\omega_\ell(T)) \quad \text{and} \quad \widetilde{\eta}_{\ell,w}(T) \lesssim \eta_{\ell,w}(\omega_\ell(T)) \quad \text{for all } T \in \mathcal{T}_\ell,$$

where $\omega_\ell(T)$ denotes a patch of T; see [13, Section 8]*. Then, the convergence (Theorem 12) as well as optimality results (Theorem 13 and 16) remain valid. We leave the details to the reader, but note that this covers averaging-based error estimators, hierarchical error estimators, as well as estimators based on equilibrated fluxes; see* [13, 29]. ■

3.1. Linear convergence.

The following result is independent of $C_{\rm mark}$, and we may formally also choose $C_{\rm mark} = \infty = C'_{\rm mark}$. Discrete reliability (A3) only enters through the quasi-monotonicity of the estimator (Lemma 6). In the frame of the Lax-Milgram lemma from Section 1.2, the quasi-monotonicity already follows from classical reliability (6); see [13, Lemma 3.6].

Theorem 12. *For all $0 < \theta \le 1$, there exist $0 < q_{\rm lin} < 1$ and $C_{\rm lin} > 0$ which depend only on* (A1)–(A4) *and θ, such that Algorithms A–B are linearly convergent in the sense of*

$$\eta_{u,\ell+n}\eta_{z,\ell+n} \le C_{\rm lin} q_{\rm lin}^n \eta_{u,\ell}\eta_{z,\ell} \quad \text{for all } \ell, n \in \mathbb{N}_0. \tag{18}$$

Proof for Algorithm A. In each step of Algorithm A, the set $\widetilde{\mathcal{M}}_j$ satisfies either the Dörfler marking (8) for $\eta_{u,j}$ or for $\eta_{z,j}$. With $\widetilde{\mathcal{M}}_j \subseteq \mathcal{M}_j \subseteq \mathcal{T}_j \backslash \mathcal{T}_{j+1}$, this implies for n successive meshes $\mathcal{T}_j$, $j = \ell, \dots, \ell + n$, that $\mathcal{T}_j \backslash \mathcal{T}_{j+1}$ satisfies k-times the Dörfler marking (14) for $\eta_{u,j}$ and $(n-k)$-times the Dörfler marking for $\eta_{z,j}$. Proposition 10 thus shows

$$\eta_{u,\ell+n}^2 \le C_{\rm conv}\, q_{\rm conv}^k\, \eta_{u,\ell}^2 \quad \text{as well as} \quad \eta_{z,\ell+n}^2 \le C_{\rm conv}\, q_{\rm conv}^{n-k}\, \eta_{z,\ell}^2.$$

Altogether, this proves

$$\eta_{u,\ell+n}^2\, \eta_{z,\ell+n}^2 \le C_{\rm conv}^2\, q_{\rm conv}^k\, \eta_{u,\ell}^2\, \eta_{z,\ell}^2.$$

This concludes (18) with $q_{\rm lin} = q_{\rm conv}^{1/2}$ and $C_{\rm lin} = C_{\rm conv}$. ∎

Proof for Algorithm B. Note that $\rho_\ell^2 = 2\, \eta_{u,\ell}^2 \eta_{z,\ell}^2$. Therefore, (9) becomes

$$2\theta\, \eta_{u,\ell}^2 \eta_{z,\ell}^2 \le \eta_{u,\ell}(\mathcal{M}_\ell)^2\, \eta_{z,\ell}^2 + \eta_{u,\ell}^2\, \eta_{z,\ell}(\mathcal{M}_\ell)^2.$$

In particular, this shows that

$$\theta\, \eta_{u,\ell}^2 \le \eta_{u,\ell}(\mathcal{M}_\ell)^2 \quad \text{or} \quad \theta\, \eta_{z,\ell}^2 \le \eta_{z,\ell}(\mathcal{M}_\ell)^2.$$

Arguing as for Algorithm A, we conclude the proof. ∎

3.2. Fine properties of mesh-refinement. Unlike linear convergence, the proof of optimal convergence rates is more strongly tailored to the mesh-refinement used. First, we suppose that each refined element has at least two sons, i.e.,

$$\#(\mathcal{T} \backslash \mathcal{T}') + \#\mathcal{T} \le \#\mathcal{T}' \quad \text{for all } \mathcal{T} \in \mathbb{T} \text{ and all } \mathcal{T}' \in \text{refine}(\mathcal{T}). \tag{19}$$

Second, we require the mesh-closure estimate

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \le C_{\rm mesh} \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \quad \text{for all } \ell \in \mathbb{N}, \tag{20}$$

where $C_{\rm mesh} > 0$ depends only on $\mathcal{T}_0$. This has first been proved for 2D newest vertex bisection in [10] and has later been generalized to arbitrary dimension $d \ge 2$ in [38]. While both works require an additional admissibility assumption on $\mathcal{T}_0$, this has at least been proved unnecessary for 2D in [28]. Finally, it has been proved in [15, 37] that newest vertex bisection ensures the overlay estimate, i.e., for all triangulations $\mathcal{T}, \mathcal{T}' \in \mathbb{T}$ there exists a common refinement $\mathcal{T} \oplus \mathcal{T}' \in \text{refine}(\mathcal{T}) \cap \text{refine}(\mathcal{T}')$ which satisfies

$$\#(\mathcal{T} \oplus \mathcal{T}') \le \#\mathcal{T} + \#\mathcal{T}' - \#\mathcal{T}_0. \tag{21}$$

We note that for newest vertex bisection, the triangulation $\mathcal{T} \oplus \mathcal{T}'$ is, in fact, the overlay of $\mathcal{T}$ and $\mathcal{T}'$. For 1D bisection (e.g., for 2D BEM computations in Section 6), the algorithm from [2] satisfies (19)–(21) and guarantees that the local mesh-ratio is uniformly bounded. For meshes with first-order hanging nodes, (19)–(21) are analyzed in [11], while T-spline meshes for isogeometric analysis are considered in [34].

3.3. Optimal convergence rates. Our proofs of the following theorems (Theorem 13, 16) follow the ideas of [33] as worked out in [17]. We include it here for the sake of completeness and a self-contained presentation.

Theorem 13. *Suppose that the mesh-refinement satisfies* (19)–(21). *Let* $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2)^{-1}$. *Then, Algorithm A implies the existence of* $C_{\rm opt} > 0$ *which depends only on* θ, $C_{\rm mesh}$, $C_{\rm mark}$, $C'_{\rm mark}$, *and* (A1)–(A4), *such that for all* $s, t > 0$ *the assumption* $(u, z) \in \mathbb{A}_s \times \mathbb{A}_t$ *implies for all* $\ell \in \mathbb{N}_0$

$$\eta_{u,\ell}\eta_{z,\ell} \le \frac{C_{\rm opt}^{1+s+t}}{(1 - q_{\rm lin}^{1/(s+t)})^{s+t}} \|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t} (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-(s+t)} \tag{22}$$

i.e., Algorithm A guarantees that the estimator product decays asymptotically with any possible algebraic rate.

Corollary 14. *Assume that the estimators both have finite optimal convergence rate, i.e.,*

$$s_{\max} := \sup\{s > 0 \,:\, \|u\|_{\mathbb{A}_s} < \infty\} < \infty \quad \text{and} \quad t_{\max} := \sup\{t > 0 \,:\, \|z\|_{\mathbb{A}_t} < \infty\} < \infty.$$

Then, for any $0 < s < s_{\max}$ *and* $0 < t < t_{\max}$, *there exist subsequences such that*

$$\eta_{u,\ell_k} \lesssim (\#\mathcal{T}_{\ell_k} - \#\mathcal{T}_0)^{-s} \text{ for all } k \in \mathbb{N} \quad \text{as well as} \quad \eta_{z,\ell_j} \lesssim (\#\mathcal{T}_{\ell_j} - \#\mathcal{T}_0)^{-t} \text{ for all } j \in \mathbb{N},$$

where the hidden constants additionally depend on $s_{\max} - s > 0$ *resp.* $t_{\max} - t > 0$.

Proof. Let $0 < \widetilde{s} < s_{\max}$. Choose $\varepsilon > 0$ with $s := \widetilde{s} + 2\varepsilon < s_{\max}$ and $t := t_{\max} - \varepsilon > 0$. By choice of $t_{\max}$, it holds $\eta_{z,\ell} \not\lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-(t_{\max}+\varepsilon)}$; see [13, Theorem 4.1(ii)]. Hence,

$$\forall C > 0 \,\forall \ell \in \mathbb{N}\, \exists k \ge \ell \quad \eta_{z,k} > C\, (\#\mathcal{T}_k - \#\mathcal{T}_0)^{-(t_{\max}+\varepsilon)}.$$

Consequently, there exists a subsequence with $\eta_{z,\ell_k} \ge (\#\mathcal{T}_{\ell_k} - \#\mathcal{T}_0)^{-(t_{\max}+\varepsilon)}$. With Theorem 13, the same subsequence satisfies

$$\eta_{u,\ell_k} \le \eta_{u,\ell_k}\eta_{z,\ell_k} (\#\mathcal{T}_{\ell_k} - \#\mathcal{T}_0)^{t_{\max}+\varepsilon} \overset{(22)}{\lesssim} (\#\mathcal{T}_{\ell_k} - \#\mathcal{T}_0)^{-(s+t)+(t_{\max}+\varepsilon)} = (\#\mathcal{T}_{\ell_k} - \#\mathcal{T}_0)^{-\widetilde{s}}.$$

The same argument applies to an appropriate subsequence of $\eta_{z,\ell}$. ■

The heart of the proof of Theorem 13 is the following lemma.

Lemma 15. *For any* $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2)^{-1}$ *and* $\ell \in \mathbb{N}_0$, *there exist* $C_1, C_2 > 0$ *and some* $\mathcal{T}_\star \in \mathrm{refine}(\mathcal{T}_\ell)$ *such that the sets* $\mathcal{R}_u(\mathcal{T}_\ell, \mathcal{T}_\star)$ *and* $\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star)$ *from the discrete reliability* (A3) *satisfy for all* $s, t > 0$ *with* $(u, z) \in \mathbb{A}_s \times \mathbb{A}_t$

$$\max\{\#\mathcal{R}_u(\mathcal{T}_\ell, \mathcal{T}_\star)\,,\, \#\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star)\} \le C_1\, (C_2\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}\, (\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}. \tag{23}$$

Moreover, $\mathcal{R}_u(\mathcal{T}_\ell, \mathcal{T}_\star)$ *or* $\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star)$ *satisfies the Dörfler marking, i.e., it holds*

$$\theta\eta_{u,\ell}^2 \le \eta_{u,\ell}\big(\mathcal{R}_u(\mathcal{T}_\ell, \mathcal{T}_\star)\big)^2 \quad \text{or} \quad \theta\eta_{z,\ell}^2 \le \eta_{z,\ell}\big(\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star)\big)^2. \tag{24}$$

The constants C_1, C_2 *depend only on* θ *and* (A1)–(A3).

Proof. Adopt the notation of Lemma 7. For $\varepsilon := C_{\rm mon}^{-1}\kappa_{\rm opt}\,\eta_{u,\ell}\eta_{z,\ell}$, the quasi-monotonicity of the estimators (Lemma 6) yields $\varepsilon \le \kappa_{\rm opt}\,\eta_{u,0}\eta_{z,0} < \|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t} < \infty$. Choose the minimal $N \in \mathbb{N}_0$ such that $\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t} \le \varepsilon\,(N+1)^{s+t}$. Choose $\mathcal{T}_{\varepsilon_1}, \mathcal{T}_{\varepsilon_2} \in \mathbb{T}_N$ with $\eta_{u,\varepsilon_1} = \min_{\mathcal{T}_\star \in \mathbb{T}_N} \eta_{u,\star}$ and $\eta_{z,\varepsilon_2} = \min_{\mathcal{T}_\star \in \mathbb{T}_N} \eta_{z,\star}$. Define $\mathcal{T}_\varepsilon := \mathcal{T}_{\varepsilon_1} \oplus \mathcal{T}_{\varepsilon_2}$ and $\mathcal{T}_\star := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$. Then, Lemma 6, the definition of the approximation classes, and the choice of N give

$$\eta_{u,\star}\eta_{z,\star} \le C_{\rm mon}\eta_{u,\varepsilon_1}\eta_{z,\varepsilon_2} \le C_{\rm mon}(N+1)^{-(s+t)}\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t} \le C_{\rm mon}\varepsilon = \kappa_{\rm opt}\,\eta_{u,\ell}\eta_{z,\ell}.$$

This implies $\eta_{u,\star}^2 \le \kappa_{\rm opt}\,\eta_{u,\ell}^2$ or $\eta_{z,\star}^2 \le \kappa_{\rm opt}\,\eta_{z,\ell}^2$, and Lemma 7 hence proves (24). It remains to derive (23). First, note that

$$\max\{\#\mathcal{R}_u(\mathcal{T}_\ell, \mathcal{T}_\star)\,,\, \#\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star)\} \overset{(A3)}{\le} C_{\rm rel}\,\#(\mathcal{T}_\ell \backslash \mathcal{T}_\star) \overset{(19)}{\le} C_{\rm rel}(\#\mathcal{T}_\star - \#\mathcal{T}_\ell). \tag{25}$$

Second, minimality of N yields

$$N < (\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}\varepsilon^{-1/(s+t)} = C\,(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}$$

with $C := (\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}(C_{\rm mon}^{-1}\kappa_{\rm opt})^{-1/(s+t)} = (C_{\rm mon}\kappa_{\rm opt}^{-1}\,\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}$. According to the choice of $\mathcal{T}_\star$, the overlay estimate (21) yields

$$\#\mathcal{T}_\star - \#\mathcal{T}_\ell \overset{(21)}{\le} \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \overset{(21)}{\le} \#\mathcal{T}_{\varepsilon_1} + \#\mathcal{T}_{\varepsilon_2} - 2\,\#\mathcal{T}_0 \le 2N < 2C\,(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}. \tag{26}$$

Combining (25)–(26), we conclude (23) with $C_1 = 2C_{\rm rel}$ and $C_2 = C_{\rm mon}/\kappa_{\rm opt}$. ∎

Proof of Theorem 13. According to (24) of Lemma 15 and the marking strategy in Algorithm A, for all $j \in \mathbb{N}_0$, there hold the implications

$$\begin{aligned}
\widetilde{\mathcal{M}}_j = \mathcal{M}_{u,j} \quad &\Longrightarrow \quad \#\mathcal{M}_{u,j} \le C_{\rm mark}\,\#\mathcal{R}_u(\mathcal{T}_j,\mathcal{T}_\star),\\
\widetilde{\mathcal{M}}_j = \mathcal{M}_{z,j} \quad &\Longrightarrow \quad \#\mathcal{M}_{z,j} \le C_{\rm mark}\,\#\mathcal{R}_z(\mathcal{T}_j,\mathcal{T}_\star).
\end{aligned}$$

This yields

$$\begin{aligned}
\frac{1}{C'_{\rm mark}}\,\#\mathcal{M}_j \le \#\widetilde{\mathcal{M}}_j &= \min\{\#\mathcal{M}_{u,j}\,,\,\#\mathcal{M}_{z,j}\}\\
&\le C_{\rm mark}\,\max\{\#\mathcal{R}_u(\mathcal{T}_j,\mathcal{T}_\star)\,,\,\#\mathcal{R}_z(\mathcal{T}_j,\mathcal{T}_\star)\}.
\end{aligned} \tag{27}$$

With the mesh-closure estimate (20) and estimate (23) of Lemma 15, we obtain

$$\begin{aligned}
\#\mathcal{T}_\ell - \#\mathcal{T}_0 &\overset{(20)}{\le} C_{\rm mesh}\sum_{j=0}^{\ell-1}\#\mathcal{M}_j\\
&\overset{(23)}{\le} C_{\rm mesh}C_{\rm mark}C'_{\rm mark}C_1\,(C_2\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}\sum_{j=0}^{\ell-1}(\eta_{u,j}\eta_{z,j})^{-1/(s+t)}.
\end{aligned}$$

Linear convergence (18) implies

$$\eta_{u,\ell}\eta_{z,\ell} \le C_{\rm lin}\,q_{\rm lin}^{\ell-j}\eta_{u,j}\eta_{z,j} \quad \text{for all } 0 \le j \le \ell$$

and hence

$$(\eta_{u,j}\eta_{z,j})^{-1/(s+t)} \le C_{\rm lin}^{1/(s+t)}q_{\rm lin}^{(\ell-j)/(s+t)}(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}.$$

With $0 < q := q_{\rm lin}^{1/(s+t)} < 1$, the geometric series applies and yields

$$\sum_{j=0}^{\ell-1}(\eta_{u,j}\eta_{z,j})^{-1/(s+t)} \le C_{\rm lin}^{1/(s+t)}(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}\sum_{j=0}^{\ell-1}q^{\ell-j} \le \frac{C_{\rm lin}^{1/(s+t)}}{1-q_{\rm lin}^{1/(s+t)}}\,(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}.$$

Combining this with the first estimate, we obtain

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \le \frac{C_{\rm mesh}C_{\rm mark}C'_{\rm mark}C_1}{1-q_{\rm lin}^{1/(s+t)}}\,(C_{\rm lin}C_2\,\|u\|_{\mathbb{A}_s}\|z\|_{\mathbb{A}_t})^{1/(s+t)}\,(\eta_{u,\ell}\eta_{z,\ell})^{-1/(s+t)}.$$

Altogether, we conclude (22) with $C_{\rm opt} = \max\{C_{\rm lin}C_2, C_{\rm mesh}C_{\rm mark}C'_{\rm mark}C_1\}$. ∎

Theorem 16. *Let $\theta_{\rm opt} := (1 + C_{\rm stb}C_{\rm rel})^{-1}$. For any $0 < \theta < \theta_{\rm opt}/2$, Algorithm B guarantees optimal algebraic convergence rates in the sense of Theorem 13 and Corollary 14.*

Proof. Arguing as for Algorithm A, we only need to show that (27) remains valid. Note that $0 < 2\theta < \theta_{\rm opt}$. Therefore, estimate (24) of Lemma 15 yields

$$2\theta\,\eta_{u,j}^2 \le \eta_{u,j}\big(\mathcal{R}_u(\mathcal{T}_j,\mathcal{T}_\star)\big)^2 \quad \text{or} \quad 2\theta\,\eta_{z,j}^2 \le \eta_{z,j}\big(\mathcal{R}_z(\mathcal{T}_j,\mathcal{T}_\star)\big)^2.$$

Either for $\mathcal{R}_j := \mathcal{R}_u(\mathcal{T}_j,\mathcal{T}_\star)$ or for $\mathcal{R}_j := \mathcal{R}_z(\mathcal{T}_j,\mathcal{T}_\star)$ this implies

$$\theta\,\rho_j^2 = 2\theta\,\eta_{u,j}^2\eta_{z,j}^2 \le \eta_{u,j}(\mathcal{R}_j)^2\,\eta_{z,j}^2 + \eta_{u,j}^2\,\eta_{z,j}(\mathcal{R}_j)^2 = \rho_j(\mathcal{R}_j)^2.$$

According to the marking strategy in Algorithm B, we obtain

$$\#\mathcal{M}_j \le C_{\rm mark}\#\mathcal{R}_j \le C_{\rm mark}\,\max\{\#\mathcal{R}_u(\mathcal{T}_j,\mathcal{T}_\star)\,,\,\#\mathcal{R}_z(\mathcal{T}_j,\mathcal{T}_\star)\}$$

which is (27). Therefore, the claim follows with $C_{\rm opt} = \max\{C_{\rm lin}C_2, C_{\rm mesh}C_{\rm mark}C_1\}$. ■

Remark 17. *Our numerical experiments below do not show that Algorithm B leads to suboptimal convergence rates for large θ, where Algorithm A still is optimal. However, this has been observed in* [17] *for the point evaluation in adaptive BEM computations.* ■

4. Goal-Oriented Adaptive FEM for Second-Order Linear Elliptic PDEs

In this section, we prove that our analysis implies convergence and optimality of goal-oriented AFEM for general second-order linear elliptic PDEs. **4.1. Model problem.** Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with polygonal boundary. For given $f_1, g_1 \in L^2(\Omega)$ and $\boldsymbol{f}_2, \boldsymbol{g}_2 \in L^2(\Omega;\mathbb{R}^d)$, define

$$f(v) := \int_\Omega f_1 v - \boldsymbol{f}_2\cdot\nabla v\,dx \quad \text{and} \quad g(v) := \int_\Omega g_1 v - \boldsymbol{g}_2\cdot\nabla v\,dx.$$

We aim to compute $g(u)$, where $u \in H_0^1(\Omega)$ solves the weak formulation

$$a(u,v) := \int_\Omega \big(\boldsymbol{A}\nabla u\cdot\nabla v + \boldsymbol{b}\cdot\nabla u v + cuv\big)\,dx = f(v) \quad \text{for all } v \in \mathcal{X} := H_0^1(\Omega), \tag{28}$$

where $\boldsymbol{A} \in W^{1,\infty}(\Omega;\mathbb{R}^{d\times d}_{\rm sym})$, $\boldsymbol{b} \in W^{1,\infty}(\Omega;\mathbb{R}^d)$, and $c \in L^\infty(\Omega)$. We suppose that $a(\cdot,\cdot)$ is elliptic on $H_0^1(\Omega)$ so that the problem fits in the framework of Section 1.2. To formulate the residual error estimators in (31)–(32) below, we additionally require that $\operatorname{div}\boldsymbol{f}_2, \operatorname{div}\boldsymbol{g}_2$ exist in $L^2(\Omega)$ elementwise on the initial mesh $\mathcal{T}_0$ and that the edge jumps satisfy $[\boldsymbol{f}_2\cdot n], [\boldsymbol{g}_2\cdot n] \in L^2(\partial T)$ for all $T \in \mathcal{T}_0$. (For instance, this is satisfied if $\boldsymbol{f}_2, \boldsymbol{g}_2$ are $\mathcal{T}_0$-piecewise constant.) Note that the corresponding differential operator $\mathcal{L}$ is non-symmetric as

$$\mathcal{L}w := -\operatorname{div}(\boldsymbol{A}\nabla w) + \boldsymbol{b}\cdot\nabla w + cw \neq -\operatorname{div}(\boldsymbol{A}\nabla w) - \boldsymbol{b}\cdot\nabla w + (c - \operatorname{div}\boldsymbol{b})w =: \mathcal{L}^\top w. \tag{29}$$

Remark 18. *For the ease of presentation, we focus on (homogeneous) Dirichlet conditions. We note that the extension to mixed Dirichlet-Neumann-Robin boundary conditions is easily possible; see* [3, 13, 22] *in the frame of standard AFEM. However, our analysis currently requires that the Dirichlet data belong to the coarsest trace space $\mathcal{S}^1(\mathcal{T}_0|_\Gamma)$, so that $u - U_\ell$ resp. $z - Z_\ell$ are admissible test functions. The latter fails for general inhomogeneous Dirichlet conditions. We believe that the rigorous analysis of this problem is beyond the current work and requires further ideas beyond those of standard AFEM* [3, 13, 22]. ■

4.2. Discretization. For a regular triangulation $\mathcal{T}_\star$ of Ω and $p \in \mathbb{N}$, define $\mathcal{P}^p(\mathcal{T}_\star) := \{V \in L^2(\Omega) \,:\, V|_T \text{ is polynomial of degree } \le p \text{ for all } T \in \mathcal{T}_\star\}$. Let $U_\star, Z_\star \in \mathcal{X}_\star := \mathcal{S}_0^p(\mathcal{T}_\star) := \mathcal{P}^p(\mathcal{T}_\star) \cap H_0^1(\Omega)$ be the unique FEM solutions of (2) resp. (4), i.e.,

$$U_\star \in \mathcal{S}_0^p(\mathcal{T}_\star) \quad \text{such that} \quad a(U_\star, V_\star) = f(V_\star) \quad \text{for all } V_\star \in \mathcal{S}_0^p(\mathcal{T}_\star), \tag{30a}$$

$$Z_\star \in \mathcal{S}_0^p(\mathcal{T}_\star) \quad \text{such that} \quad a(V_\star, Z_\star) = g(V_\star) \quad \text{for all } V_\star \in \mathcal{S}_0^p(\mathcal{T}_\star). \tag{30b}$$

4.3. Residual error estimator. For $T \in \mathcal{T}_\star$, let $h_T := |T|^{1/d}$ and $\mathcal{L}|_T$ (resp. $\mathcal{L}^\top|_T$) be the natural restriction of $\mathcal{L}$ (resp. $\mathcal{L}^\top$) to T. Then, the residual error estimators read

$$\eta_{u,\star}(T)^2 := h_T^2 \|\mathcal{L}|_T U_\star - f_1 - \operatorname{div} \boldsymbol{f}_2\|_{L^2(T)}^2 + h_T \|[(\boldsymbol{A}\nabla U_\star + \boldsymbol{f}_2)\cdot n]\|_{L^2(\partial T\cap\Omega)}^2, \tag{31}$$

$$\eta_{z,\star}(T)^2 := h_T^2 \|\mathcal{L}^\top|_T Z_\star - g_1 - \operatorname{div} \boldsymbol{g}_2\|_{L^2(T)}^2 + h_T \|[(\boldsymbol{A}\nabla Z_\star + \boldsymbol{g}_2)\cdot n]\|_{L^2(\partial T\cap\Omega)}^2. \tag{32}$$

There holds reliability (6); see, e.g., [1, 39]. Therefore, Section 1.2 yields

$$|g(u) - g(U_\star)| \lesssim \eta_{u,\star}\eta_{z,\star}. \tag{33}$$

Moreover, efficiency and the Céa lemma prove that $\mathbb{A}_s$ from Section 2.3 coincides with the approximation class based on the total error (see [7, 15, 33]). The following result is proved in [20, Lemma 5.1] for $\boldsymbol{f}_2 = 0 = \boldsymbol{g}_2$, but holds verbatim in the present case.

Lemma 19. *Let $w \in \{u, z\}$. Then, there holds $w \in \mathbb{A}_s$ if and only if*

$$\sup_{N\in\mathbb{N}_0} \Big((N+1)^s \min_{\mathcal{T}_\star\in\mathbb{T}_N} \big(\min_{V_\star\in\mathcal{X}_\star} \|w - V_\star\|_{\mathcal{X}} + \operatorname{osc}_{w,\star}(V_\star)\big)\Big) < \infty,$$

where $\operatorname{osc}_{w,\star}(V_\star)^2 = \sum_{T\in\mathcal{T}_\star} \operatorname{osc}_{w,\star}(T, V_\star)^2$ *and*

$$\begin{aligned}
\operatorname{osc}_{u,\star}^2(T, V_\star) &:= h_T^2 \|(1 - \Pi_T^{2p-2})(\mathcal{L}|_T V_\star - f_1 - \operatorname{div} \boldsymbol{f}_2)\|_{L^2(T)}^2 \\
&\quad + h_T \|(1 - \Pi_{\partial T}^{2p-1})[(\boldsymbol{A}\nabla V_\star + \boldsymbol{f}_2)\cdot n]\|_{L^2(\partial T\cap\Omega)}^2, \\
\operatorname{osc}_{z,\star}^2(T, V_\star) &:= h_T^2 \|(1 - \Pi_T^{2p-2})(\mathcal{L}^\top|_T V_\star - g_1 - \operatorname{div} \boldsymbol{g}_2)\|_{L^2(T)}^2 \\
&\quad + h_T \|(1 - \Pi_{\partial T}^{2p-1})[(\boldsymbol{A}\nabla V_\star + \boldsymbol{g}_2)\cdot n]\|_{L^2(\partial T\cap\Omega)}^2.
\end{aligned}$$

Here, $\Pi_T^q : L^2(T) \to \mathcal{P}^q(T)$ denotes the L^2-orthogonal projection onto polynomials of degree q and $\Pi_{\partial T}^q : L^2(\partial T) \to \mathcal{P}^q(\mathcal{S}_{\partial T})$ denotes the L^2-orthogonal projection onto (discontinuous) piecewise polynomials of degree q on the faces of T. ■

4.4. Verification of axioms. For newest vertex bisection [38], the assumptions of Section 3.2 are satisfied. It remains to verify the axioms (A1)–(A4), where $\mathbb{d}_w(\mathcal{T}_\ell, \mathcal{T}_\star) := a(W_\ell - W_\star, W_\ell - W_\star)^{1/2} \simeq \|W_\ell - W_\star\|_{H^1(\Omega)}$ and W_ℓ resp. $W_\star$ are the corresponding FEM approximations of $w \in \{u, z\}$.

Theorem 20. *The conforming discretization* (30) *of the model problem of Section 4.1 with the residual error estimators* (31)–(32) *satisfies* (A1)–(A4) *for both $w \in \{u, z\}$ with $q_{\rm red} = 2^{-1/d}$ and $\mathcal{R}_w(\mathcal{T}_\ell, \mathcal{T}_\star) = \mathcal{T}_\ell \backslash \mathcal{T}_\star$. Therefore, Algorithm A–B are linearly convergent with optimal rates in the sense of Theorem 12, 13, and 16 for the upper bound in* (33).

Proof of Theorem 20, (A1)–(A3). The work [15] considers some symmetric model problem with $\boldsymbol{b} = 0$ and $c \geq 0$ as well as $\boldsymbol{f}_2 = 0 = \boldsymbol{g}_2$. Stability (A1) and reduction (A2) are essentially part of the proof of [15, Corollary 3.4]. The discrete reliability (A3) is found in [15, Lemma 3.6]. Both proofs transfer verbatim to the present situation. ■

Lemma 21. *In the setting of Theorem 20, there holds convergence*

$$\lim_{\ell\to\infty} \|U_\infty - U_\ell\|_{H^1(\Omega)} = 0 = \lim_{\ell\to\infty} \|Z_\infty - Z_\ell\|_{H^1(\Omega)}, \tag{34}$$

for certain $U_\infty, Z_\infty \in H_0^1(\Omega)$. Moreover, there holds at least $U_\infty = u$ or $Z_\infty = z$.

Proof. Adaptive mesh-refinement guarantees nestedness $\mathcal{X}_\ell \subseteq \mathcal{X}_\star$ for all $\mathcal{T}_\ell \in \mathbb{T}$ and $\mathcal{T}_\star \in \operatorname{refine}(\mathcal{T}_\ell)$. As in [13, Section 3.6] or [5, Lemma 6.1], the Céa lemma thus implies *a priori* convergence, i.e., there exist $U_\infty, Z_\infty \in \mathcal{X}_\infty := \overline{\bigcup_{\ell\in\mathbb{N}_0} \mathcal{X}_\ell} \subseteq H_0^1(\Omega)$ such that

$$\lim_{\ell\to\infty} \|U_\infty - U_\ell\|_{H^1(\Omega)} = 0 = \lim_{\ell\to\infty} \|Z_\infty - Z_\ell\|_{H^1(\Omega)}.$$

This proves (34). For $w \in \{u, z\}$, let $\ell_{w,n}$ denote the subsequences which satisfy

$$\theta\eta_{w,\ell_{w,n}}^2 \le \eta_{w,\ell_{w,n}}(\mathcal{M}_{w,\ell_{w,n}})^2 \quad \text{for all } n \in \mathbb{N}.$$

There holds $\#\{\ell_{w,n} : n \in \mathbb{N}\} = \infty$ for at least one $w \in \{u, z\}$. While this is obvious for Algorithm A, it follows for Algorithm B from the proof of Theorem 12. For this particular w, (34) implies $\mathrm{d}_w(\mathcal{T}_{\ell_{w,n+1}}, \mathcal{T}_{\ell_{w,n}})^2 \to 0$ as $n \to \infty$. Moreover, Lemma 9 states

$$\eta_{w,\ell_{w,n+1}}^2 \le q_{\mathrm{est}}\eta_{w,\ell_{w,n}}^2 + C_{\mathrm{est}}\mathrm{d}_w(\mathcal{T}_{\ell_{w,n+1}}, \mathcal{T}_{\ell_{w,n}})^2 \quad \text{for all } n \in \mathbb{N}.$$

These observations and elementary calculus yield $\eta_{w,\ell_{w,n}} \to 0$; see, e.g., [4, Lemma 2.3]. Reliability (6) of $\eta_{w,\ell}$ concludes $\lim_{n\to\infty} \|w - W_{\ell_{w,n}}\|_{H^1(\Omega)} = 0$, i.e., $w = W_\infty$. ■

Proof of Theorem 20, (A4). With Lemma 21, the proof of [20, Lemma 3.5] shows the weak convergence in $H_0^1(\Omega)$ for $W_\infty \in \{U_\infty, Z_\infty\}$

$$\frac{W_\infty - W_{\ell_n}}{\|W_\infty - W_{\ell_n}\|_{H^1(\Omega)}} \rightharpoonup 0 \quad \text{and} \quad \frac{W_{\ell_{n+1}} - W_{\ell_n}}{\|W_{\ell_{n+1}} - W_{\ell_n}\|_{H^1(\Omega)}} \rightharpoonup 0 \quad \text{as } \ell \to \infty.$$

Define $\mathrm{d}_w(\mathcal{T}_\infty, \cdot) := a(W_\infty - (\cdot), W_\infty - (\cdot))^{1/2}$. With this, [20, Proposition 3.6] applies for the primal as well as the dual problem and shows that given any $0 < \delta < 1$, there exists $j_\delta \in \mathbb{N}$ such that all $j \ge j_\delta$ satisfy

$$\mathrm{d}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 \le \frac{1}{1-\delta}\mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_j})^2 - \mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_{j+1}})^2. \tag{35}$$

The discrete reliability (A3) and the convergence (34) yield

$$\mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_j}) = \lim_{k\to\infty} \mathrm{d}_w(\mathcal{T}_{\ell_k}, \mathcal{T}_{\ell_j}) \le C_{\mathrm{rel}}\eta_{w,\ell_j}. \tag{36}$$

With (35)–(36), the quasi-monotonicity from Lemma 6 (since (A1)–(A3) have already been verified) implies for $\delta = 1 - 1/(1 + \varepsilon C_{\mathrm{rel}}^{-2})$ and hence $1/(1-\delta) = 1 + \varepsilon C_{\mathrm{rel}}^{-2}$ that

$$\begin{aligned}
&\sum_{j=n}^{N} \big(\mathrm{d}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 - \varepsilon C_{\mathrm{rel}}^{-2}\mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_j})^2\big) \\
&\quad \overset{(35)}{\le} \sum_{j=j_\delta}^{N} \big((\frac{1}{1-\delta} - \varepsilon C_{\mathrm{rel}}^{-2})\mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_j})^2 - \mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_{j+1}})^2\big) + \sum_{j=n}^{j_\delta-1} \mathrm{d}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 \\
&\quad \le \mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_{j_\delta}})^2 + C_{\mathrm{rel}}^2 \sum_{j=n}^{j_\delta-1} \eta_{w,\ell_j}^2 \overset{(36)}{\le} (1 + j_\delta)C_{\mathrm{rel}}^2 C_{\mathrm{mon}}\eta_{w,\ell_n}^2.
\end{aligned} \tag{37}$$

Another application of the reliability (36) shows

$$\begin{aligned}
\sum_{j=n}^{N} \big(\mathrm{d}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 - \varepsilon\eta_{w,\ell_j}^2\big) &\overset{(36)}{\le} \sum_{j=n}^{N} \big(\mathrm{d}_w(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j})^2 - \varepsilon C_{\mathrm{rel}}^{-2}\mathrm{d}_w(\mathcal{T}_\infty, \mathcal{T}_{\ell_j})^2\big) \\
&\overset{(37)}{\le} (1 + j_\delta)C_{\mathrm{rel}}^2 C_{\mathrm{mon}}\eta_{w,\ell_n}^2.
\end{aligned}$$

This proves (A4) with $C_{\mathrm{orth}}(\varepsilon) := (1 + j_\delta)C_{\mathrm{rel}}^2 C_{\mathrm{mon}}$. ■

4.5. Numerical experiment I: Goal oriented FEM for the Poisson equation. As proposed in [33, Example 7.3], we consider the Poisson model problem (i.e., $\boldsymbol{A} = \boldsymbol{I}$, $\boldsymbol{b} = \boldsymbol{0}$, and $c = 0$) on the unit cube $\Omega = (0,1)^2 \subset \mathbb{R}^2$, while a nonsymmetric second-order elliptic operator is considered in Section 5.5. Figure 1 (left) shows the initial mesh $\mathcal{T}_0$ together with the triangles $T_f := \mathrm{conv}\{(0,0), (\frac{1}{2}, 0), (0, \frac{1}{2})\}$ and $T_g := \mathrm{conv}\{(1,1), (\frac{1}{2}, 1), (1, \frac{1}{2})\}$.

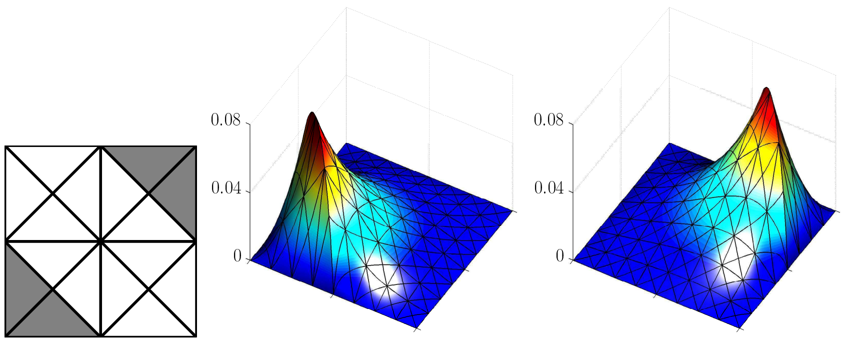

FIGURE 1. Example from Section 4.5: The initial mesh $\mathcal{T}_0$ (left) and the triangles T_f (bottom left) and T_g (top right) indicated in gray. An approximation to the primal (middle) and dual solution (right) on a uniform mesh with 256 elements, where the singularities of both are clearly visible.

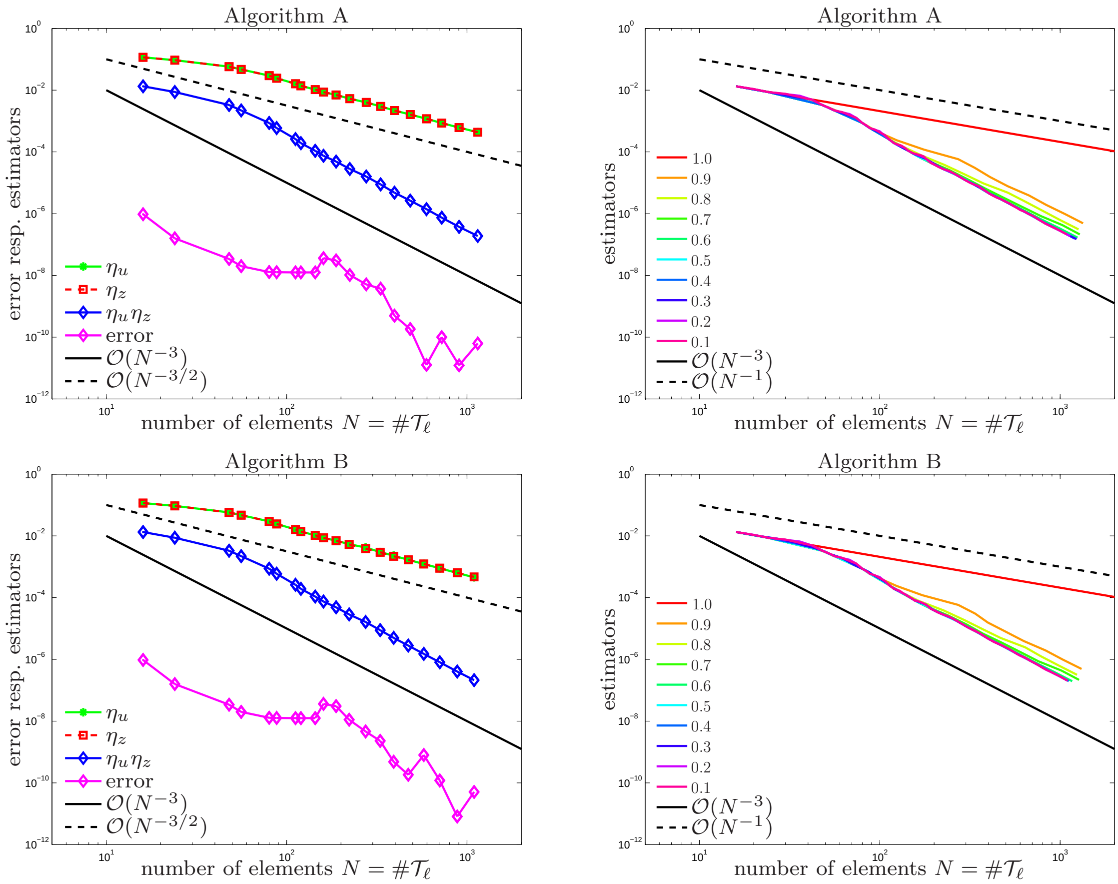

FIGURE 2. Example from Section 4.5: Estimators $\eta_{u,\ell}$ and $\eta_{z,\ell}$, estimator product $\eta_{u,\ell}\eta_{z,\ell}$, as well as goal error $|g(u) - g(U_\ell)|$ as output of Algorithm A–B with $\theta = 0.5$ (left) resp. estimator product for various $\theta \in \{0.1, \dots, 0.9\}$ as well as for $\theta = 1.0$, i.e., uniform mesh-refinement.

Choosing $f_1 = 0$, $\boldsymbol{f}_2 = (\chi_{T_f}, 0)$, $g_1 = 0$, $\boldsymbol{g}_2 = (\chi_{T_g}, 0)$, where χ_ω for $\omega \subset \mathbb{R}^2$ denotes the characteristic function, the right-hand sides of the primal (1) and dual problem (3) are

$$f(v) = -\int_{T_f} \frac{\partial v}{\partial x_1}\,dx \quad \text{resp.} \quad g(u) = -\int_{T_g} \frac{\partial u}{\partial x_1}\,dx.$$

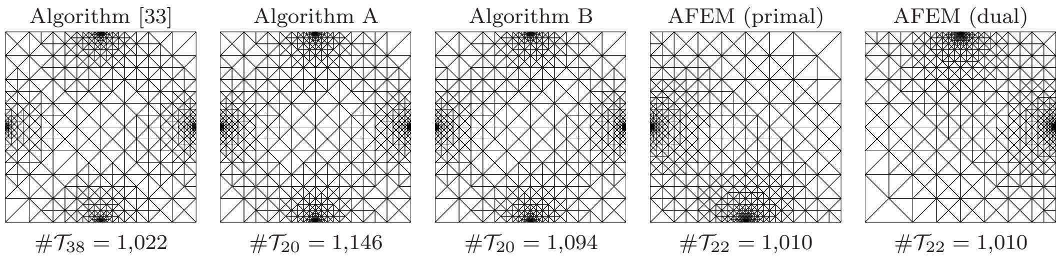

FIGURE 3. Example from Section 4.5: Meshes generated by goal-oriented algorithms as well as standard (non-goal-oriented) AFEM driven by the primal error estimator resp. the dual error estimator for $\theta = 0.5$.

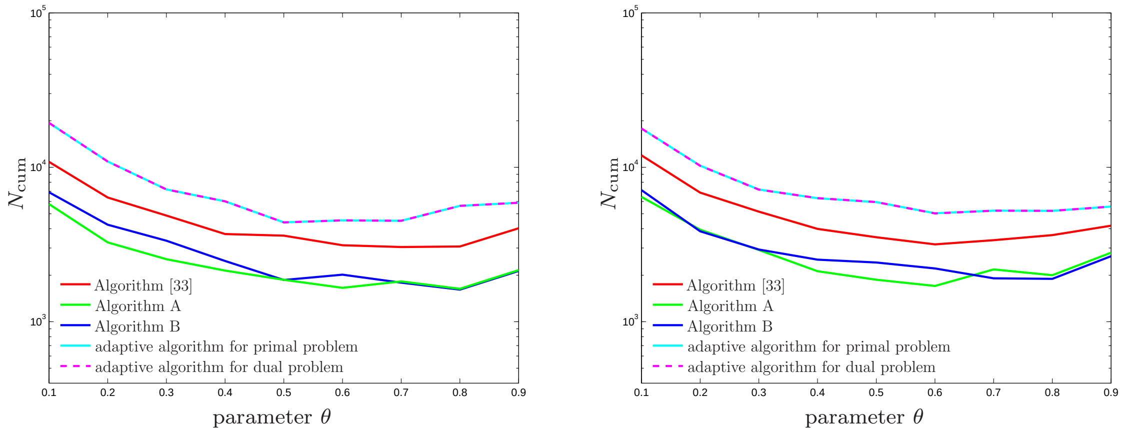

FIGURE 4. Example from Section 4.5: To compare the adaptive strategies, we plot the cumulative number of elements $N_{\rm cum} := \sum_{j=0}^{\ell} \#\mathcal{T}_j$ necessary to reach a prescribed accuracy $\eta_{u,\ell}\eta_{z,\ell} \le \mathsf{tol}$ over $\theta \in \{0.1, \dots, 0.9\}$ for $p = 3$ and $\mathsf{tol} = 10^{-5}$ (left) resp. $p = 2$ and $\mathsf{tol} = 10^{-4}$ (right).

Figure 1 also shows some approximations of the primal and dual solution, where the singularities of u along $\operatorname{conv}\{(\frac{1}{2},0),(0,\frac{1}{2})\}$ resp. z along $\operatorname{conv}\{(\frac{1}{2},1),(1,\frac{1}{2})\}$ are clearly visible.

We consider and compare five adaptive mesh-refining strategies:

- the goal-oriented algorithm from [33], i.e., Algorithm A with $C'_{\rm mark} = 1$,
- Algorithm A with $C'_{\rm mark} = 2$ as described in Remark 2,
- Algorithm B originally proposed in [7],
- standard adaptivity for the primal problem, i.e., Algorithm A with $\mathcal{M}_\ell := \mathcal{M}_{u,\ell}$,
- standard adaptivity for the dual problem, i.e., Algorithm A with $\mathcal{M}_\ell := \mathcal{M}_{z,\ell}$.

To compare these strategies, we compute the cumulative number of elements

$$N_{\rm cum} := \sum_{j=0}^{\ell} \#\mathcal{T}_j, \tag{38}$$

which is necessary to reach a prescribed accuracy of $\eta_{u,\ell}\eta_{z,\ell} \le \mathsf{tol}$. Since the overall runtime depends on the entire history of adaptively generated meshes, the definition of $N_{\rm cum}$ reflects the total amount of work in the adaptive process.

Overall, we find that the goal-oriented adaptive algorithms lead to optimal convergence behavior $\eta_{u,\ell}\eta_{z,\ell} = \mathcal{O}(N^{-3})$ for $p = 3$ (see Figure 2), while standard adaptivity for the primal or dual problem only leads to $\eta_{u,\ell}\eta_{z,\ell} = \mathcal{O}(N^{-2})$ for $p = 3$ (not displayed). This is also reflected in Figure 4, where we plot $N_{\rm cum}$ over the marking paraemter $0.1 \le \theta \le 0.9$:

For $\mathsf{tol} = 10^{-5}$ and $p = 3$, $N_{\rm cum}$ is smallest for Algorithm A–B and $\theta = 0.8$. For $\mathsf{tol} = 10^{-4}$ and $p = 2$, $N_{\rm cum}$ is smallest for Algorithm A and $\theta = 0.6$.

5. Goal-Oriented Adaptive FEM for Flux Evaluation

5.1. Model problem. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with polygonal boundary $\Gamma := \partial\Omega$. Given $f_1 \in L^2(\Omega)$ and $\boldsymbol{f_2} = 0$, let $u \in H^1_0(\Omega)$ be the solution to (28). For $\Lambda \in H^{1/2}(\Gamma)$, we aim to evaluate the weighted boundary flux

$$g(u) := \int_\Gamma (\boldsymbol{A}\nabla u) \cdot n \, \Lambda \, ds. \tag{39a}$$

For smooth u, $g(u)$ can be rewritten as

$$g(u) = \int_\Omega \operatorname{div}(\boldsymbol{A}\nabla u) z \, dx + \int_\Omega \boldsymbol{A}\nabla u \cdot \nabla z = a(u, z) - f(z) =: N_z(u) \tag{39b}$$

for all $z \in H^1(\Omega)$ with $z|_\Gamma = \Lambda$. Since the right-hand side is well-defined for $u \in H^1_0(\Omega)$, this is a valid generalization of the flux [24, Section 7]. Let z be the unique solution of the following inhomogeneous Dirichlet problem:

$$z \in H^1(\Omega) \text{ with } z|_\Gamma = \Lambda \quad \text{such that} \quad a(v, z) = 0 \quad \text{for all } v \in H^1_0(\Omega).$$

Then, it holds $N_z(u) = -f(z)$.

5.2. Discretization. With the notation of Section 4.2, consider $\mathcal{S}^p(\mathcal{T}_\star) := \mathcal{P}^p(\mathcal{T}_\star) \cap H^1(\Omega)$ and $\mathcal{S}^p_0(\mathcal{T}_\star) := \mathcal{P}^p(\mathcal{T}_\star) \cap H^1_0(\Omega)$. Let $U_\star$ be the unique FEM solution of

$$U_\star \in \mathcal{S}^p_0(\mathcal{T}_\star) \quad \text{such that} \quad a(U_\star, V_\star) = f(V_\star) \quad \text{for all } V_\star \in \mathcal{S}^p_0(\mathcal{T}_\star). \tag{40a}$$

Suppose that $\Lambda \in \mathcal{S}^p(\mathcal{T}_0|_\Gamma) := \{V_0|_\Gamma \,:\, V_0 \in \mathcal{S}^p(\mathcal{T}_0)\}$ belongs to the discrete trace space with respect to the initial mesh $\mathcal{T}_0$. Let $Z_\star$ be the unique FEM solution of

$$Z_\star \in \mathcal{S}^p(\mathcal{T}_\star) \text{ with } Z_\star|_\Gamma = \Lambda \quad \text{such that} \quad a(V_\star, Z_\star) = 0 \quad \text{for all } V_\star \in \mathcal{S}^p_0(\mathcal{T}_\star). \tag{40b}$$

To approximate $N_z(u)$ from (39), define

$$N_{z,\star}(U_\star) = -f(Z_\star). \tag{41}$$

Lemma 22. *There holds*

$$|N_z(u) - N_{z,\star}(U_\star)| \le C_{\rm flux} \|u - U_\star\|_{H^1(\Omega)} \|z - Z_\star\|_{H^1(\Omega)},$$

where $C_{\rm flux} > 0$ depends only on $a(\cdot, \cdot)$.

Proof. Since $z - Z_\star \in H^1_0(\Omega)$, there holds

$$\begin{aligned} |N_z(u) - N_{z,\star}(U_\star)| &= |f(z) - f(Z_\star)| = |f(z - Z_\star)| = |a(u, z - Z_\star)| \\ &= |a(u - U_\star, z - Z_\star)| \lesssim \|u - U_\star\|_{H^1(\Omega)} \|z - Z_\star\|_{H^1(\Omega)}, \end{aligned}$$

where we used the definition of z and $Z_\star$. ∎

5.3. Residual error estimator. With $\Lambda \in \mathcal{S}^p(\mathcal{T}_0|_\Gamma)$, the residual error estimators remain the same as in (31)–(32) with $g_1 = 0$ and $\boldsymbol{f}_2 = 0 = \boldsymbol{g}_2$, i.e.,

$$\eta_{u,\star}(T)^2 := h_T^2 \|\mathcal{L}|_T U_\star - f_1\|^2_{L^2(T)} + h_T \|[\boldsymbol{A}\nabla U_\star \cdot n]\|^2_{L^2(\partial T \cap \Omega)}, \tag{42}$$

$$\eta_{z,\star}(T)^2 := h_T^2 \|\mathcal{L}^\top|_T Z_\star\|^2_{L^2(T)} + h_T \|[\boldsymbol{A}\nabla Z_\star \cdot n]\|^2_{L^2(\partial T \cap \Omega)}. \tag{43}$$

Lemma 22 together with the reliability of $\eta_{w,\star}$ for $w \in \{u, z\}$ (see, e.g., [3, Proposition 3] for the inhomogeneous Dirichlet problem for z) implies

$$|N_z(u) - N_{z,\star}(U_\star)| \lesssim \eta_{u,\star} \eta_{z,\star}. \tag{44}$$

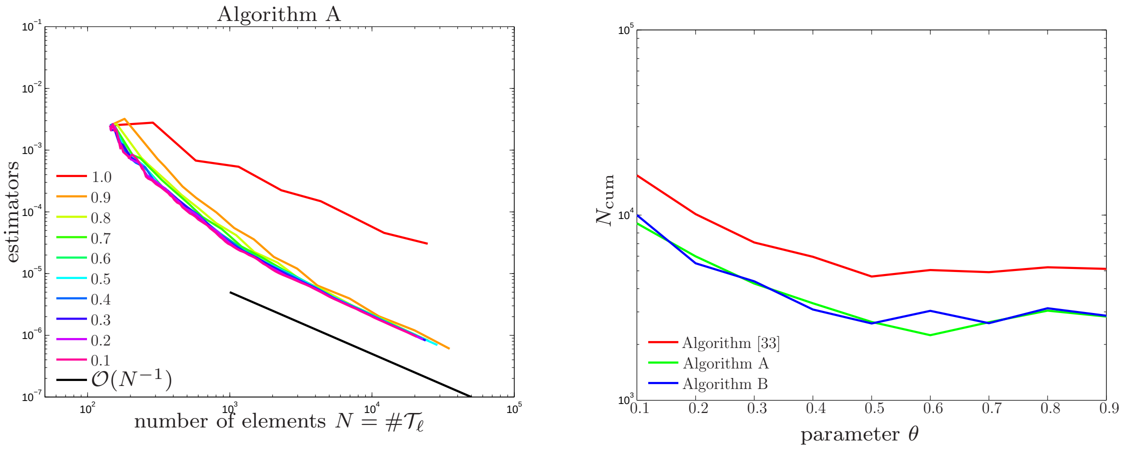

FIGURE 5. Example from Section 5.5 for $p = 1$ and $\nu = 10^{-3}$: Estimator product as output of Algorithm A for various $\theta \in \{0.1, \dots, 0.9\}$ as well as for $\theta = 1.0$, i.e., uniform refinement (left) and cumulative number of elements $N_{\text{cum}} := \sum_{j=0}^{\ell} \#\mathcal{T}_j$ necessary to reach a prescribed accuracy $\eta_{u,\ell}\eta_{z,\ell} \leq 10^{-4}$ over $\theta \in \{0.1, \dots, 0.9\}$.

5.4. Verification of axioms. For newest vertex bisection, the assumptions of Section 3.2 are satisfied. It remains to verify the axioms (A1)–(A4), where $\mathrm{d\!l}_w(\mathcal{T}_\ell, \mathcal{T}_\star) := a(W_\ell - W_\star, W_\ell - W_\star)^{1/2} \simeq \|W_\ell - W_\star\|_{H^1(\Omega)}$.

Theorem 23. *The conforming discretization* (40) *of the model problem of Section 5.1 with the residual error estimators* (42)–(43) *satisfies* (A1)–(A4) *for both $w \in \{u, z\}$ with $q_{\text{red}} = 2^{-1/d}$ and $\mathcal{R}_w(\mathcal{T}_\ell, \mathcal{T}_\star) = \mathcal{T}_\ell \backslash \mathcal{T}_\star$. Therefore, Algorithm A–B are linearly convergent with optimal rates in the sense of Theorem 12, 13, and 16 for the upper bound in* (44).

Proof. For the primal problem, (A1)–(A4) follow from Theorem 20. For the dual problem, (A1)–(A2) follow from Theorem 20, since the estimator did not change. The discrete reliability (A3) is proved in [3] for general $\Lambda \in H^1(\Gamma)$. For $\Lambda \in \mathcal{S}^p(\mathcal{T}_0|_\Gamma)$, the proof simplifies vastly and shows $\mathcal{R}_z(\mathcal{T}_\ell, \mathcal{T}_\star) = \mathcal{T}_\ell \backslash \mathcal{T}_\star$. To see the quasi-orthogonality (A4), choose a discrete extension $\widehat{\Lambda} \in \mathcal{S}^1(\mathcal{T}_0)$ with $\widehat{\Lambda}|_\Gamma = \Lambda$. Consider the solution $Z_\star^0 \in \mathcal{S}_0^p(\mathcal{T}_\star)$ of

$$a(V_\star, Z_\star^0) = -a(V_\star, \widehat{\Lambda}) \quad \text{for all } V_\star \in \mathcal{S}_0^p(\mathcal{T}_\star).$$

Then, there holds $Z_\star = Z_\star^0 + \widehat{\Lambda}$ and consequently $\mathrm{d\!l}_z(\mathcal{T}_{\ell_{j+1}}, \mathcal{T}_{\ell_j}) \simeq \|Z_{\ell_{j+1}} - Z_{\ell_j}\|_{H^1(\Omega)} = \|Z_{\ell_{j+1}}^0 - Z_{\ell_j}^0\|_{H^1(\Omega)}$. Since $Z_\star^0$ is the solution to a homogeneous Dirichlet problem, the proof of (A4) follows analogously to that of Theorem 20. ∎

5.5. Numerical experiment II: Flux-oriented adaptive FEM for convection–diffusion. We consider a numerical experiment similar to [32, Section 5.3] for some convection-diffusion problem in 2D. Throughout, we use lowest-order FEM, i.e., $p = 1$. Let $\Omega = (0,1)^2 \subset \mathbb{R}^2$. Set $\boldsymbol{A} = \nu \boldsymbol{I}$, with $\nu > 0$ the diffusion coefficient, $\boldsymbol{b} = (y, \frac{1}{2} - x)$, which is a rotating convective field around $(\frac{1}{2}, 0)$, and $c = 0$. With $\operatorname{div} \boldsymbol{b} = 0$, it holds

$$\mathcal{L} = -\nu\Delta + \boldsymbol{b} \cdot \nabla \qquad \text{and} \qquad \mathcal{L}^\top = -\nu\Delta - \boldsymbol{b} \cdot \nabla\,.$$

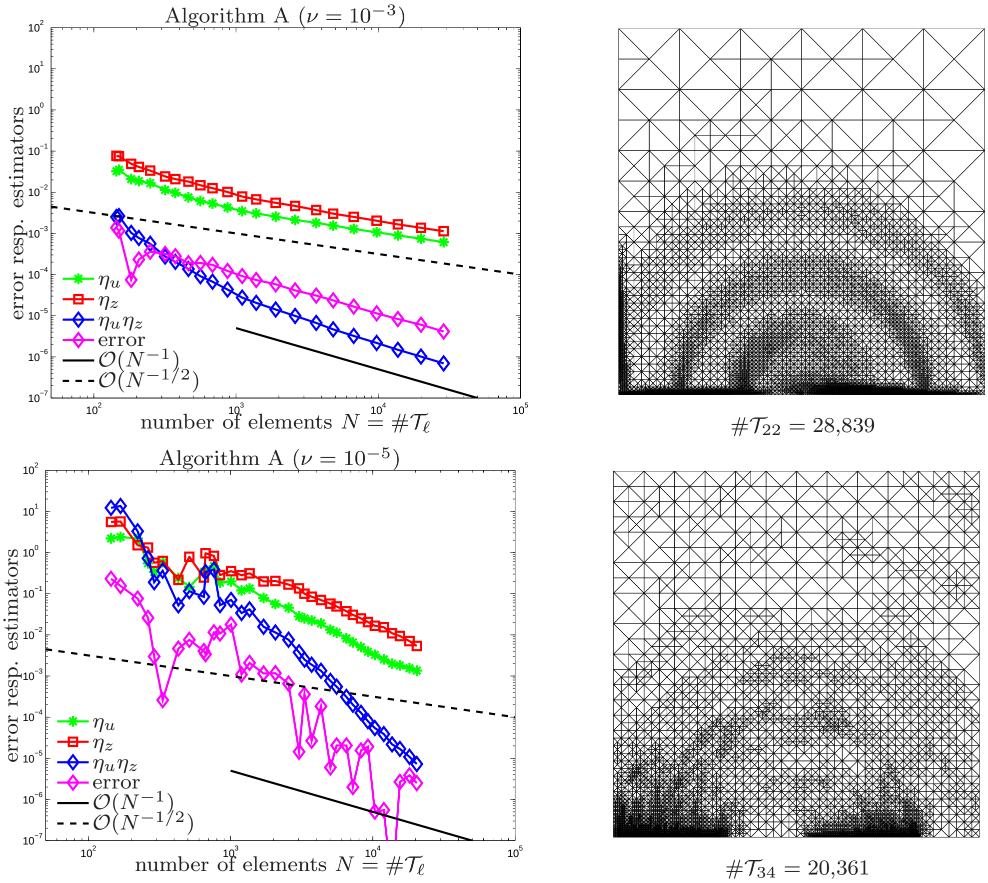

FIGURE 6. Example from Section 5.5: To study the robustness of the goal-oriented algorithm with respect to the diffusion coefficient $\nu = 10^{-3}$ (top) and $\nu = 10^{-5}$ (bottom), we plot $\eta_{u,\ell}$, $\eta_{z,\ell}$, and $\eta_{u,\ell}\eta_{z,\ell}$, as well as the goal error $|N_z(u) - N_{z,\ell}(U_\ell)|$ as output of Algorithm A with $\theta = 0.6$ over the numbers of elements $\#\mathcal{T}_\ell$ (left). We show some related discrete meshes with $> 20{,}000$ elements (right).

We set $f(v) = 0$ and consider non-homogeneous Dirichlet data on $\partial\Omega$ for the primal problem, a pulse, defined by the continuous piecewise linear function

$$u_{\mathrm{Dir}}(x,y) = \begin{cases} 6(x - \frac{1}{6}), & \text{if } \frac{1}{6} \le x < \frac{1}{3}\,,\ y = 0, \\ 6(\frac{1}{2} - x), & \text{if } \frac{1}{3} \le x < \frac{1}{2}\,,\ y = 0, \\ 0, & \text{otherwise}\,. \end{cases}$$

Note that u_{Dir} trivially extends to some discrete function $u_{\mathrm{Dir}} \in \mathcal{S}^1(\mathcal{T}_0)$ if $\mathcal{T}_0$ is chosen appropriately. Therefore, we can rewrite the problem into a homogeneous Dirichlet problem. To that end, write $u = u_0 + u_{\mathrm{Dir}}$ with $u_0 \in H_0^1(\Omega)$ and solve

$$a(u_0, v) = f(v) - a(u_{\mathrm{Dir}}, v) \quad \text{for all } v \in H_0^1(\Omega).$$

Note that the additional term on the right-hand side is of the form $\mathrm{div}\boldsymbol{\lambda} + \lambda$ for some $\mathcal{T}_0$-element wise constant $\boldsymbol{\lambda}$ and some $\lambda \in L^2(\Omega)$. A direct computation shows that the weighted-residual error estimator with respect to u_0 coincides with $\eta_{u,\ell}$. Arguing as in the proof of Theorem 23, we see that the estimator satisfies the axioms (A1)–(A4). Altogether, the problem thus fits in the frame of our analysis.

The primal solution corresponds to the clockwise convection–diffusion of this pulse. We choose the boundary weight function $\Lambda : \partial\Omega \to \mathbb{R}$ as the shifted pulse

$$\Lambda(x,y) = \begin{cases} 6(x-\frac{2}{3}), & \text{if } \frac{2}{3} \le x < \frac{5}{6}\,,\ y=0, \\ 6(1-x), & \text{if } \frac{5}{6} \le x < 1\,,\ y=0, \\ 0, & \text{otherwise}\,. \end{cases}$$

The dual solution corresponds to the counter-clockwise convection–diffusion of this pulse. For small ν, the (primal and dual) pulses are transported from $\partial\Omega$ into Ω and eventually back to $\partial\Omega$ where a boundary layer develops. The uniform initial triangulation $\mathcal{T}_0$ ensures that the (primal and dual) Dirichlet data belong to the discrete trace space $\mathcal{S}^1(\mathcal{T}_0|_\Gamma)$.

For $\nu = 10^{-3}$ and a large range of values of $\theta \in \{0.1, \dots, 0.9\}$, Figure 5 (left) shows that Algorithm A yields the optimal convergence rate $\mathcal{O}(N^{-1})$ for the flux quantity of interest and lowest-order elements $p = 1$, while uniform mesh-refinement appears to be slightly suboptimal. Algorithm B leads to similar results (not displayed).

To compare the overall performance of the different algorithms, Figure 5 (right) visualizes the cumulative number of elements $N_{\rm cum}$ (see (38)) which is necessary to reach a prescribed accuracy of $\eta_{u,\ell}\eta_{z,\ell} \le 10^{-4}$. We observe that $N_{\rm cum}$ is smallest for relatively large values $\theta \ge 0.5$, with Algorithm [33] being less efficient than Algorithm A and B. Overall, Algorithm A with $\theta = 0.6$ seems to be the best choice.

Figure 6 illustrates the effect of varying $\nu \in \{10^{-3}, 10^{-5}\}$. Because ν is relatively small, both the primal and the dual solution have significant boundary layers. The optimal convergence rate of the estimator product is observed for the indicated values of ν, however, the pre-asymptotic regime is longer for smaller values of ν. This is to be expected, as the hidden constant in (44) depends on the reliability constants for the estimators, which in turn depend on ν.

6. Goal oriented BEM

In this section, we extend ideas from [21] and prove that our abstract frame of convergence and optimality of goal-oriented adaptivity applies also to the BEM.

6.1. Model problem. Let $\Gamma \subseteq \partial\Omega$ denote some relatively open boundary part of the Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$. Given $F, \Lambda \in H^1(\Gamma)$, we aim to compute

$$g(u) := \int_\Gamma \Lambda u \, ds, \tag{45}$$

where u solves the weakly-singular integral equation

$$\mathcal{V}u(x) := \int_\Gamma G(x,y)u(y)\,dy = F(x) \quad \text{almost everywhere on } \Gamma. \tag{46}$$

Here, $G : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ denotes the Newton kernel

$$G(x,y) := \begin{cases} -\frac{1}{2\pi}\log|x-y| & \text{for } d=2, \\ \frac{1}{4\pi|x-y|} & \text{for } d=3. \end{cases}$$

The single-layer operator extends to a linear and continuous operator $\mathcal{V} : \widetilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$, where $H^{1/2}(\Gamma) := \{\widehat{v}|_\Gamma \,:\, \widehat{v} \in H^1(\Omega)\}$ is the trace space of $H^1(\Omega)$ and $\widetilde{H}^{-1/2}(\Gamma)$ denotes its dual space; see, e.g., [31, 26, 36] for the functional analytic setting. For $d = 3$ as well as supposed that $\operatorname{diam}(\Omega) < 1$ for $d = 2$, the induced bilinear form

$$a(u,v) := \langle \mathcal{V}u\,,\, v\rangle := \int_\Gamma (\mathcal{V}u)(x)v(x)\,dx \quad \text{for } u,v \in \mathcal{X} := \widetilde{H}^{-1/2}(\Gamma)$$

is continuous, symmetric, and $\widetilde{H}^{-1/2}(\Gamma)$-elliptic. In particular, $\|\!|v|\!\|^2 := a(v,v)$ defines an equivalent norm on $\widetilde{H}^{-1/2}(\Gamma)$. The problem fits in the frame of Section 1.2. More precisely and according to the Hahn-Banach theorem, (46) is equivalent to (1), where the right-hand side of (1) reads $f(v) := \int_\Gamma Fv\,dx$. Moreover, the goal functional from (45) satisfies $g \in \widetilde{H}^{-1/2}(\Gamma)^* = H^{1/2}(\Gamma)$, where the integral is understood as the duality pairing between $\widetilde{H}^{-1/2}(\Gamma)$ and its dual $H^{1/2}(\Gamma)$.

6.2. Discretization. Let $\mathcal{T}_\star$ be a regular triangulation of Γ into affine line segments for $d = 2$ resp. flat surface triangles for $d = 3$. For each element $T \in \mathcal{T}_\star$, let $\gamma_T : T_{\rm ref} \to T$ be an affine bijection, where the reference element is $T_{\rm ref} = [0,1]$ for $d = 2$ resp. $T_{\rm ref} = \operatorname{conv}\{(0,0),(0,1),(1,0)\}$ for $d = 3$. For some polynomial degree $p \geq 1$, define

$$\mathcal{X}_\star := \mathcal{P}^p(\mathcal{T}_\star) := \{V_\star : \Gamma \to \mathbb{R} \,:\, V_\star \circ \gamma_T \in \mathcal{P}^p(T_{\rm ref}) \text{ for all } T \in \mathcal{T}_\star\},$$

where $\mathcal{P}^p(T_{\rm ref}) := \{q \in L^2(T_{\rm ref}) \,:\, q \text{ is polynomial of degree} \leq p \text{ on } T_{\rm ref}\}$. Let $U_\star, Z_\star$ be the unique BEM solutions of (2) resp. (4), i.e.,

$$\text{(47a)} \qquad U_\star \in \mathcal{P}^p(\mathcal{T}_\star) \quad \text{such that} \quad a(U_\star, V_\star) = f(V_\star) \quad \text{for all } V_\star \in \mathcal{P}^p(\mathcal{T}_\star),$$

$$\text{(47b)} \qquad Z_\star \in \mathcal{P}^p(\mathcal{T}_\star) \quad \text{such that} \quad a(V_\star, Z_\star) = g(V_\star) \quad \text{for all } V_\star \in \mathcal{P}^p(\mathcal{T}_\star).$$

6.3. Residual error estimator. The residual error estimators from [14] for the discrete primal problem (2) and the discrete dual problem (4) read

$$\text{(48)} \qquad \eta_{u,\star}(T)^2 := h_T\|\nabla(\mathcal{V}U_\star - F)\|^2_{L^2(T)} \quad \text{and} \quad \eta_{z,\star}(T)^2 := h_T\|\nabla(\mathcal{V}Z_\star - \Lambda)\|^2_{L^2(T)}.$$

The error estimators satisfy reliability (6); see, e.g., [14]. The abstract analysis of Section 1.2 thus results in

$$\text{(49)} \qquad |g(u) - g(U_\star)| \lesssim \eta_{u,\star}\eta_{z,\star}.$$

6.4. Verification of axioms. With 2D newest vertex bisection [38] for $d = 3$ resp. the extended 1D bisection from [2] for $d = 2$, the assumptions of Section 3.2 are satisfied. It remains to verify (A1)–(A4), where $\mathrm{d\!l}_w(\mathcal{T}_\ell, \mathcal{T}_\star) := \|\!|W_\ell - W_\star|\!\| \simeq \|W_\ell - W_\star\|_{\widetilde{H}^{-1/2}(\Gamma)}$.

Theorem 24. *The conforming discretization* (47) *of the model problem of Section 6.1 with the residual error estimators* (48) *satisfies* (A1)–(A4) *for both $w \in \{u, z\}$ with $q_{\rm red} = 2^{-1/(d-1)}$ and $\mathcal{R}_w(\mathcal{T}_\ell, \mathcal{T}_\star) = \{T \in \mathcal{T}_\ell \,:\, \exists T' \in \mathcal{T}_\ell \backslash \mathcal{T}_\star \quad T \cap T' \neq \emptyset\}$, i.e., refined elements plus one additional layer of elements. Therefore, Algorithm A–B are linearly convergent with optimal rates in the sense of Theorem 12, 13, and 16 for the upper bound in* (49).

Proof. The assumptions (A1)–(A2) and (A3) are proved in [21, Proposition 4.2, Proposition 5.3] for the lowest-order case. The general case is proved in [18]. The quasi-orthogonality (A4) follows from symmetry of $a(\cdot,\cdot)$ and (A3); see Remark 8. ∎

6.5. Numerical experiment with conforming weight function. Let $\Omega \subset \mathbb{R}^2$ with $\operatorname{diam}(\Omega) = 1/\sqrt{2}$ be the L-shaped domain from Figure 7. On the boundary $\Gamma := \partial\Omega$, consider $\phi(x) := r^{2/3}\cos(2\alpha/3)$ for polar coordinates $r(x), \alpha(x)$ with origin $(0,0)$. Let $\mathcal{K} : H^{1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma)$, for all $-1/2 \leq s \leq 1/2$, be the double-layer potential which is formally defined as (n_y denotes the outer unit normal on Γ at y)

$$\mathcal{K}\phi(x) := -\frac{1}{2\pi}\int_\Gamma \frac{(x-y)\cdot n_y}{|x-y|^2}\phi(y)\,dy.$$

Consider the model problem (46) with

$$F := (\mathcal{K} + 1/2)\phi.$$

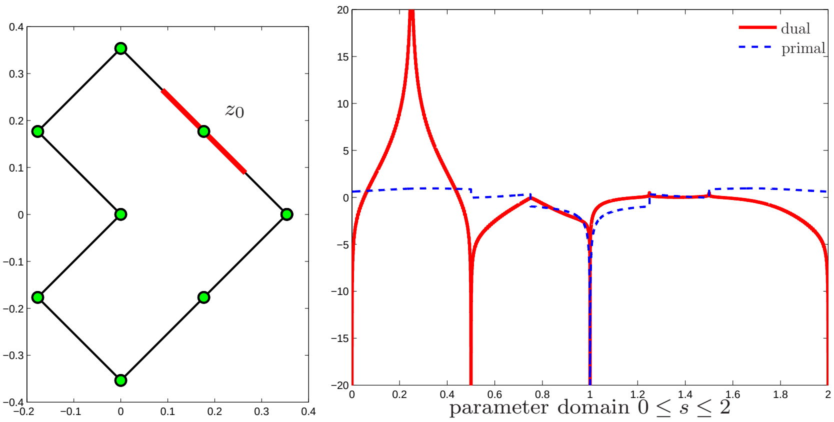

FIGURE 7. Example from Section 6.5: Domain Ω with initial triangulation $\mathcal{T}_0$ (left) and primal and dual solution plotted over the arc-length (right), where $s = 1$ (resp. $s = 0.25$) corresponds to the reentrant corner (resp. z_0).

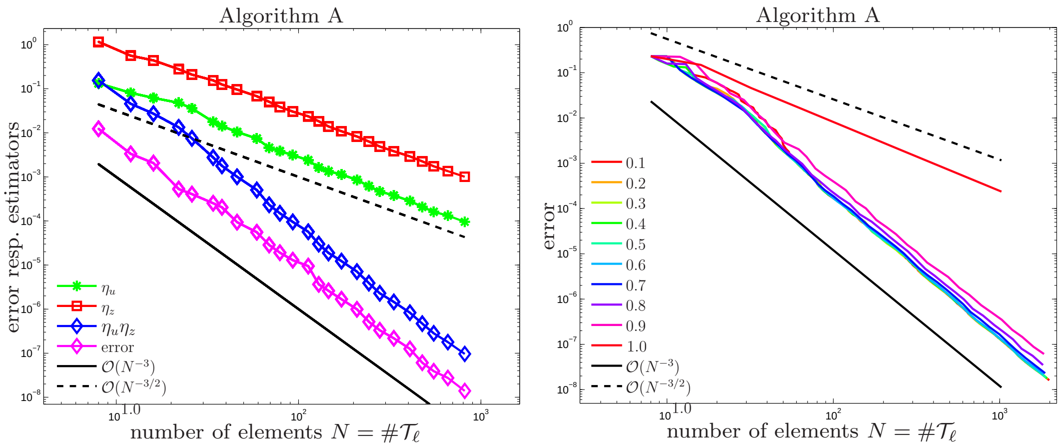

FIGURE 8. Example from Section 6.5: Estimators and goal error $|g(u) - g(U_\ell)|$ as output of Algorithm A for $\theta = 0.5$ (left) resp. estimator product $\eta_{u,\ell}\eta_{z,\ell}$ for various $\theta \in \{0.1, \dots, 0.9\}$ as well as for $\theta = 1.0$, i.e., uniform refinement (right).

It is known [26, 31, 36] that (46) is equivalent to the Laplace-Dirichlet problem

$$\Delta P = 0 \text{ in } \Omega \text{ subject to Dirichlet boundary conditions } P = \phi \text{ on } \Gamma,$$

and the exact solution of (46) is the normal derivative $u = \partial_n P$ of P. The initial mesh $\mathcal{T}_0$ is shown in Figure 7. As weight function $\Lambda \in \mathcal{S}^1(\mathcal{T}_0)$, we consider the hat function defined by $\Lambda(z_0) = 1$ and $\Lambda(z) = 0$ for all other nodes z of $\mathcal{T}_0$ (the node z_0 is indicated in Figure 7).

For the lowest-order case $p = 0$ and $\theta = 0.5$ in Algorithm A, Figure 8 shows the convergence rates of the error estimators η_u, η_z, their product $\eta_u\eta_z$, and the error in the goal functional $|g(u) - g(U_\ell)|$. Moreover, we compare the convergence rate of the estimator product for different values of $\theta \in \{0.1, \dots, 0.9\}$. For either choice of θ, we observe the optimal convergence rate $(\#\mathcal{T}_\ell)^{-3/2}$ for the respective error estimators as well as $(\#\mathcal{T}_\ell)^{-3}$ for the error in the goal functional.

References

[1] Mark Ainsworth and J. Tinsley Oden, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience, New York, 2000.

[2] Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, and Dirk Praetorius, *Efficiency and optimality of some weighted-residual error estimator for adaptive 2D boundary element methods*, Comput. Methods Appl. Math., 13 (2013), pp. 305–332.

[3] Markus Aurada, Michael Feischl, Josef Kemetmüller, Markus Page, and Dirk Praetorius, *Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 1207–1235.

[4] Markus Aurada, Samuel Ferraz-Leite, and Dirk Praetorius, *Estimator reduction and convergence of adaptive bem*, Appl. Numer. Math., 62 (2012), pp. 787–801.

[5] Ivo Babuska and Michael Vogelius, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numer. Math., 44 (1984), pp. 75–102.

[6] Wolfgang Bangerth and Rolf Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics, ETH Zürich, Birkhäuser, 2003.

[7] Roland Becker, Elodie Estecahandy, and David Trujillo, *Weighted marking for goal-oriented adaptive finite element methods*, SIAM J. Numer. Anal., 49 (2011), pp. 2451–2469.

[8] Roland Becker and Rolf Rannacher, *A feed-back approach to error control in finite element methods: basic analysis and examples*, East-West J. Numer. Math., 4 (1996), pp. 237–264.

[9] Roland Becker and Rolf Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica, 10 (2001), pp. 1–102.

[10] Peter Binev, Wolfgang Dahmen, and Ron DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., 97 (2004), pp. 219–268.

[11] Andrea Bonito and Ricardo H. Nochetto, *Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method*, SIAM J. Numer. Anal., 48 (2010), pp. 734–771.

[12] Markus Bürg and Murtazo Nazarov, *Goal-oriented adaptive finite element methods for elliptic problems revisited*, J. Comput. Appl. Math., 287 (2015), pp. 125–147.

[13] Carsten Carstensen, Michael Feischl, Marcus Page, and Dirk Praetorius, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253.

[14] Carsten Carstensen, Matthias Maischak, and Ernst P. Stephan, *A posteriori error estimate and h-adaptive algorithm on surfaces for Symm's integral equation*, Numer. Math., 90 (2001), pp. 197–213.

[15] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550.

[16] Kenneth Eriksson, Don Estep, Peter Hansbo, and Claes Johnson, *Introduction to adaptive methods for differential equations*, in Acta numerica, 1995, Acta Numer., Cambridge Univ. Press, Cambridge, 1995, pp. 105–158.

[17] Michael Feischl, Thomas Führer, Gregor Gantner, Alexander Haberl, and Dirk Praetorius, *Adaptive boundary element methods for optimal convergence of point errors*, Numer. Math., 132 (2016), pp. 541–567.

[18] Michael Feischl, Thomas Führer, Michael Karkulik, Jens Markus Melenk, and Dirk Praetorius, *Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, Part I: Weakly-singular integral equation*, Calcolo, 51 (2014), pp. 531–562.

[19] ——, *Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, Part II: Hyper-singular integral equation*, Electron. Trans. Numer. Anal., 44 (2015), pp. 153–176.

[20] Michael Feischl, Thomas Führer, and Dirk Praetorius, *Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems*, SIAM J. Numer. Anal., 52 (2014), pp. 601–625.

[21] Michael Feischl, Michael Karkulik, Jens Markus Melenk, and Dirk Praetorius, *Quasi-optimal convergence rate for an adaptive boundary element method*, SIAM J. Numer. Anal., 51 (2013), pp. 1327–1348.

[22] Michael Feischl, Marcus Page, and Dirk Praetorius, *Convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data*, J. Comput. Appl. Math., 255 (2014), pp. 481–501.

[23] Tsogtgerel Gantumur, *Adaptive boundary element methods with convergence rates*, Numer. Math., 124 (2013), pp. 471–516.
[24] Michael B. Giles and Endre Süli, *Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality*, Acta Numer., 11 (2002), pp. 145–236.
[25] Michael Holst and Sara Pollock, *Convergence of goal-oriented adaptive finite element methods for nonsymmetric operators*, Numer. Methods Partial Differential Equations, published online first (2015). DOI: `10.1002/num.22002`.
[26] George C. Hsiao and Wolfgang L. Wendland, *Boundary integral equations*, vol. 164 of Applied Mathematical Sciences, Springer-Verlag, Berlin, 2008.
[27] Claes Johnson and Anders Szepessy, *Adaptive finite element methods for conservation laws based on a posteriori error estimates*, Comm. Pure Appl. Math., 48 (1995), pp. 199–234.
[28] Michael Karkulik, David Pavlicek, and Dirk Praetorius, *On 2D newest vertex bisection: optimality of mesh-closure and H^1-stability of L_2-projection*, Constr. Approx., 38 (2013), pp. 213–234.
[29] Christian Kreuzer and Kunibert G. Siebert, *Decay rates of adaptive finite elements with Dörfler marking*, Numer. Math., 117 (2011), pp. 679–716.
[30] Gokturk Kuru, Clemens V. Verhoosel, Kristoffer G. van der Zee, and E. Harald van Brummelen, *Goal-adaptive isogeometric analysis with hierarchical splines*, Comput. Methods Appl. Mech. Engrg., 270 (2014), pp. 270–292.
[31] William McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.
[32] Khamron Mekchay and Ricardo H. Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal., 43 (2005), pp. 1803–1827.
[33] Mario S. Mommer and Rob Stevenson, *A goal-oriented adaptive finite element method with convergence rates*, SIAM J. Numer. Anal., 47 (2009), pp. 861–886.
[34] Philipp Morgenstern and Daniel Peterseim, *Analysis-suitable adaptive T-mesh refinement with linear complexity*, Comput. Aided Geom. Design, 34 (2015), pp. 50–66.
[35] Serge Prudhomme and J. Tinsley Oden, *On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors*, Comput. Methods Appl. Mech. Engrg., 176 (1999), pp. 313–331. New advances in computational methods (Cachan, 1997).
[36] Stefan A. Sauter and Christoph Schwab, *Boundary element methods*, vol. 39 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2011.
[37] Rob Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
[38] ——, *The completion of locally refined simplicial partitions created by bisection*, Math. Comp., 77 (2008), pp. 227–241.
[39] Rüdiger Verfürth, *A posteriori error estimation techniques for finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.

School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
E-mail address: `M.Feischl@unsw.edu.au`

Vienna University of Technology, Institute for Analysis and Scientific Computing, Wiedner Hauptstr. 8–10, 1040 Vienna, Austria
E-mail address: `Dirk.Praetorius@tuwien.ac.at`

University of Nottingham, School of Mathematical Sciences, University Park, Nottingham NG7 2RD, UK
E-mail address: `KG.vanderZee@nottingham.ac.uk`