ASC Report No. 11/2014

Rate optimality of adaptive algorithms: An axiomatic approach

C. Carstensen, M. Feischl, and D. Praetorius

Institute for Analysis and Scientific Computing Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

10/2014	*N. Zamponi* Analysis of a drift-diffusion model with velocity saturation for spin-polarized transport in semiconductors
9/2014	*M. Feischl, T. Führer, N. Heuer, M. Karkulik, D. Praetorius.* Adaptive boundary element methods: A posteriori error estimators, adaptivity, convergence, and implementation
8/2014	*C. Abert, G. Hrkac, M. Page, D. Praetorius, M. Ruggeri, D. Suess* Spin-polarized transport in ferromagnetic multilayers: An unconditionally convergent FEM integrator
7/2014	*I. Rachunková, S. Staněk, J. Vampolová and E. Weinmüller* On linear ODEs with a time singularity of the first kind and unsmooth inhomogeneity
6/2014	*A. Jüngel, M. Winkler* A degenerate fourth-order parabolic equation modeling Bose-Einstein condensation. Part II: Finite-time blow-up
5/2014	*A. Jüngel, M. Winkler* A degenerate fourth-order parabolic equation modeling Bose-Einstein condensation. Part I: Local existence of solutions
4/2014	*W. Herfort, K. Hofmann, and F. Russo* Locally compact near Abelian groups
3/2014	*H. Woracek* De Branges spaces and growth aspects
2/2014	*J. Gschwindl, I. Rachunková, S. Staněk, and E. Weinmüller* Positive blow-up solutions of nonlinear models from real world dynamics
1/2014	*L. Aceto, C. Magherini, E. Weinmüller* Matrix models for radial Schrödinger eigenproblems defined on a semi-infinite domain

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: `admin@asc.tuwien.ac.at`
WWW: `http://www.asc.tuwien.ac.at`
FAX: `+43-1-58801-10196`

ISBN 978-3-902627-05-6

RATE OPTIMALITY OF ADAPTIVE ALGORITHMS: AN AXIOMATIC APPROACH

Carsten Carstensen[1], **Michael Feischl**[2], **and Dirk Praetorius**[2]

[1] Humboldt-University of Berlin,
Department of Mathematics,
Unter den Linden 6, 10099 Berlin, Germany,
`cc@math.hu-berlin.de`

[2] Vienna University of Technology,
Institute for Analysis and Scientific Computing,
Wiedner Hauptstr. 8-10/E101/4, 1040 Vienna, Austria,
{`Michael.Feischl`, `Dirk.Praetorius`}`@tuwien.ac.at`

Key words: finite element method, boundary element method, a posteriori error estimate, adaptive algorithm, convergence, optimality.

Abstract. Adaptive mesh-refining algorithms dominate the numerical simulations in computational sciences and engineering, because they promise optimal convergence rates in an overwhelming numerical evidence. The mathematical foundation of optimal convergence rates has recently been completed [3] and shall be discussed in this talk. We aim at a simultaneous axiomatic presentation of the proof of optimal convergence rates for adaptive finite elements [10, 4, 2, 5] as well as boundary elements [7, 8] in the spirit of [10]. For this purpose, an overall set of four axioms on the error estimator is sufficient and (partially even) necessary.

Compared to the state of the art in the temporary literature [10, 4, 2, 7, 8, 5], the improvements of [3] can be summarized as follows: First, a general framework is presented which covers the existing literature on rate optimality of adaptive schemes for both, linear as well as nonlinear problems, which is fairly independent of the underlying (conforming, nonconforming, or mixed) finite element or boundary element method. Second, efficiency of the error estimator is not needed. Instead, efficiency exclusively characterizes the approximation classes involved in terms of the bestapproximation error plus data resolution. Third, some general quasi-Galerkin orthogonality is not only sufficient, but also necessary for the R-linear convergence of the error estimator, which is a fundamental ingredient in the current quasi-optimality analysis [10, 4, 2, 7, 8, 5]. Finally, the general analysis allows for various generalizations like equivalent error estimators and inexact solvers as well as different non-homogeneous and mixed boundary conditions.

Acknowledgement. The authors acknowledge support by the Austrian Science Fund (FWF) through the research project P21732 and the doctoral program W1245.

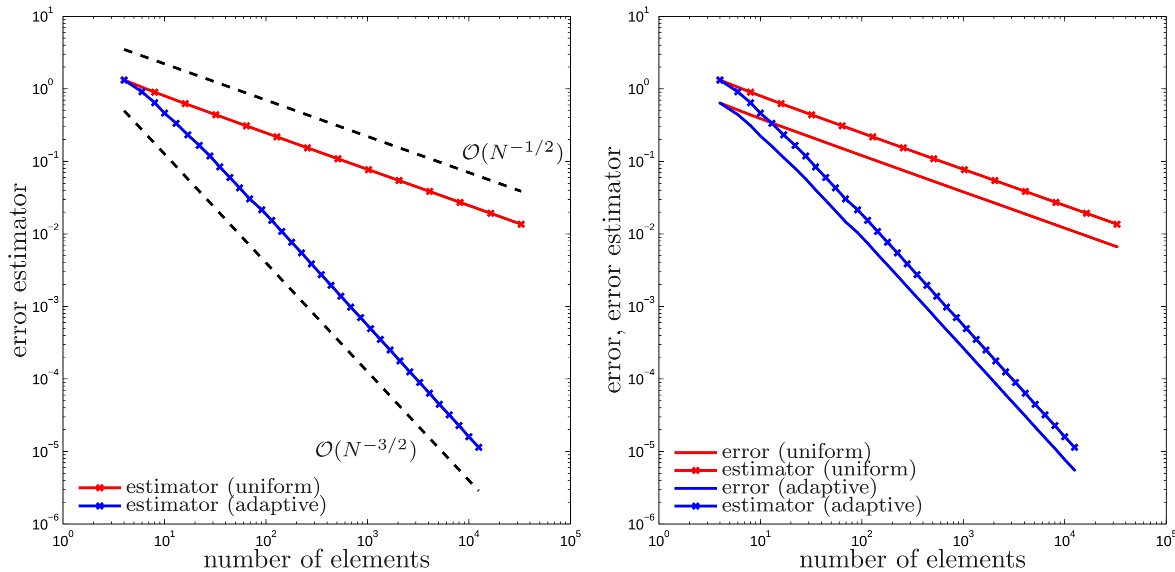

Figure 1: Numerical results of Algorithm 1 with $\theta = 0.25$ for lowest-order 2D BEM for weakly-singular integral equation $\mathcal{V}u = f$ on the slit $\Gamma = (-1,1) \times \{0\}$ with $f(x,0) = -x$ (see Section 5). Adaptivity is driven by the weighted-residual error estimator defined in (25) below. The energy space is $\widetilde{H}^{-1/2}(\Gamma)$. For lowest-order BEM, the optimal order of convergence thus is $\mathcal{O}(N^{-3/2})$, while singularities of ϕ reduce the order down to $\mathcal{O}(N^{-1/2})$ for uniform mesh-refinement (left). The error estimator provides a lower bound (efficiency (7)) and an upper bound (reliability (6)) for the (in general) unknown error (right).

1 INTRODUCTION

The ultimate goal of adaptive mesh-refining algorithms is to compute a discrete solution with error below a prescribed tolerance at the expense of, up to a multiplicative constant, the minimal computational cost. Although adaptive strategies have been successfully employed since the eighties and empirically led to optimal convergence rates in an exhaustive number of numerical experiments, the mathematical understanding of convergence and optimal rates has been a long standing issue.

In practice, the adaptive algorithm iterates the loop

$$\boxed{\texttt{solve}} \longrightarrow \boxed{\texttt{estimate}} \longrightarrow \boxed{\texttt{mark}} \longrightarrow \boxed{\texttt{refine}} \tag{1}$$

and so provides a sequence of discrete solutions U_ℓ which approximate some target u, as well as the corresponding error estimators η_ℓ which measure the approximation error $\|u - U_\ell\|$.

Figure 1 shows the typical outcome of an adaptive algorithm for some 2D boundary element computation and illustrates the following main results, we aim to explain and analyze below. Precise statements of these main results are given in Section 6.

Main Result 1 *The adaptive algorithm guarantees R-linear convergence in the sense of*

$$\eta_{\ell+n} \leq C q^n \, \eta_\ell \quad \textit{for all } \ell, n \in \mathbb{N}_0, \tag{2}$$

with certain constants $C > 0$ and $0 < q < 1$. ■

Main Result 2 *The adaptive algorithm is quasi-optimal in the sense that it recovers the optimal algebraic convergence rate after a possible preasymptotic phase.* ■

The remainder of this work is organized as follows: Section 2 provides our formulation of the adaptive loop (1) and states the four axioms (A1)–(A4) of adaptivity from [3]. Section 3 collects some remarks on the abstract frame and important consequences of these axioms. In many important situations, which are discussed in Section 4, the quasi-orthogonality (A3) is automatically satisfied. Section 5 gives three prominent examples and comments on the verification of the axioms (A1)–(A4). We conclude this work with the thorough mathematical formulation of Main Result 1–2 in Section 6.

2 ADAPTIVE ALGORITHM AND ABSTRACT AXIOMS

Throughout, we consider conforming triangulations which consist of simplices. For local mesh-refinement, we employ *newest vertex bisection* (NVB). For the precise mesh-refinement rules, we refer to, e.g., [11]. Together with the given initial triangulation $\mathcal{T}_0$, this fixes the set $\mathbb{T}$ of all conforming triangulations which can theoretically be obtained.

For $\mathcal{T} \in \mathbb{T}$ and $\mathcal{M} \subseteq \mathcal{T}$, we write $\mathcal{T}' = \texttt{refine}(\mathcal{T}, \mathcal{M}) \in \mathbb{T}$ for the one-level refinement, i.e., the coarsest conforming refinement of $\mathcal{T}$ such that all *marked simplices* $T \in \mathcal{M}$ have been bisected. Moreover, we write $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$, if $\mathcal{T}' \in \mathbb{T}$ is obtained by finitely many steps of one-level refinements. In this sense, it also holds $\mathbb{T} = \texttt{refine}(\mathcal{T}_0)$.

We assume that, for any triangulation $\mathcal{T} \in \mathbb{T}$, we can compute a discrete solution $U_\mathcal{T}$ and a corresponding error estimator $\eta_\mathcal{T} = \big(\sum_{T\in\mathcal{T}} \eta_\mathcal{T}(T)^2\big)^{1/2}$. The local contributions $\eta_\mathcal{T}(T)$, called *refinement indicators*, are used to single-out certain elements for refinement. Overall, we consider the following realization of the adaptive loop (1), where we abbreviate $U_\ell := U_{\mathcal{T}_\ell}$ and $\eta_\ell := \eta_{\mathcal{T}_\ell}$.

Algorithm 1 Input: *Initial triangulation $\mathcal{T}_0$ and bulk parameter $0 < \theta < 1$.*
Loop*: For $\ell = 0, 1, 2, \ldots,$ do* (i)–(iv).

(i) *Compute discrete solution U_ℓ for $\mathcal{T}_\ell$.*

(ii) *Compute refinement indicators $\eta_\ell(T)$ for all $T \in \mathcal{T}_\ell$.*

(iii) *Determine set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of minimal cardinality such that*

$$\theta\,\eta_\ell^2 \le \sum_{T\in\mathcal{T}_\ell} \eta_\ell(T)^2. \tag{3}$$

(iv) *Generate new triangulation $\mathcal{T}_{\ell+1} = \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$.*

Output: *Sequence of locally refined triangulations $\mathcal{T}_\ell$ with corresponding discrete solutions U_ℓ and error estimators η_ℓ for all $\ell \in \mathbb{N}_0$.* ■

The mathematical analysis of Algorithm 1 requires some properties on the formal setting. First, we assume that uniform mesh-refinement leads to guaranteed convergence, i.e., there exists some (unknown) limit u and some appropriate norm $\|\cdot\|$ such that

$$\forall \varepsilon > 0 \, \forall \mathcal{T} \in \texttt{refine}(\mathcal{T}_0) \, \exists \mathcal{T}' \in \texttt{refine}(\mathcal{T}) : \quad \|u - U_{\mathcal{T}'}\| \leq \varepsilon. \tag{4}$$

Note that this assumption does *not* guarantee convergence of Algorithm 1 in the sense of $\|u - U_\ell\| \to 0$ as $\ell \to \infty$, but is essential to allow for convergence at all. Moreover, we assume that we can measure the difference $\|U_{\mathcal{T}'} - U_{\mathcal{T}}\|$ of two discrete solutions corresponding to $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$. With these notations, the performance of Algorithm 1 can be characterized with the below axioms which have been identified in [3]: Main Result 1 holds if (A1)–(A3) are satisfied. Main Result 2 holds if (A1)–(A4) are satisfied.

(A1) Stability on non-refined simplices: There exists some constant $C_{\rm stab} > 0$ such that, for all $\mathcal{T} \in \mathbb{T}$ and $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$, it holds

$$\Big| \Big(\sum_{T \in \mathcal{T} \cap \mathcal{T}'} \eta_{\mathcal{T}'}(T)^2 \Big)^{1/2} - \Big(\sum_{T \in \mathcal{T} \cap \mathcal{T}'} \eta_{\mathcal{T}}(T)^2 \Big)^{1/2} \Big| \leq C_{\rm stab} \, \|U_{\mathcal{T}'} - U_{\mathcal{T}}\|.$$

(A2) Reduction on refined simplices: There exist constants $C_{\rm red} > 0$ and $0 < q_{\rm red} < 1$ such that, for all $\mathcal{T} \in \mathbb{T}$ and $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$, it holds

$$\sum_{T \in \mathcal{T}' \backslash \mathcal{T}} \eta_{\mathcal{T}'}(T)^2 \leq q_{\rm red} \sum_{T \in \mathcal{T} \backslash \mathcal{T}'} \eta_{\mathcal{T}}(T)^2 + C_{\rm red} \, \|U_{\mathcal{T}'} - U_{\mathcal{T}}\|^2.$$

(A3) Quasi-orthogonality: With u from (4), there exists some constants $0 \leq \varepsilon_{\rm orth} \ll 1$ and $C_{\rm orth}(\varepsilon_{\rm orth}) > 0$ such that the output of Algorithm 1 satisfies for all $\ell, n \in \mathbb{N}_0$

$$\sum_{k=\ell}^{\ell+n} \big(\|U_{k+1} - U_k\|^2 - \varepsilon_{\rm orth} \|u - U_k\|^2 \big) \leq C_{\rm orth}(\varepsilon_{\rm orth}) \, \eta_\ell^2.$$

(A4) Discrete reliability: There exist constants $C_{\rm rel}, C_{\rm ref} > 0$ such that for all $\mathcal{T} \in \mathbb{T}$ and all $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$ and some appropriate set $\mathcal{T} \backslash \mathcal{T}' \subseteq \mathcal{R}(\mathcal{T}, \mathcal{T}') \subseteq \mathcal{T}$, it holds

$$\|U_{\mathcal{T}'} - U_{\mathcal{T}}\|^2 \leq C_{\rm rel}^2 \sum_{T \in \mathcal{R}(\mathcal{T}, \mathcal{T}')} \eta_{\mathcal{T}}(T)^2 \quad \text{and} \quad \#\mathcal{R}(\mathcal{T}, \mathcal{T}') \leq C_{\rm ref} \, \#(\mathcal{T} \backslash \mathcal{T}'),$$

where $\#(\cdot)$ denotes the number of simplices.

3 REMARKS ON AXIOMS

● **Discretization.** Let $\mathcal{X}$ denote the space which contains the target u. For each triangulation $\mathcal{T} \in \mathbb{T}$, let $\mathcal{X}_{\mathcal{T}}$ be the discrete space which contains the discrete solution $U_{\mathcal{T}}$. We note that we do neither assume conformity of the discretization, i.e., $\mathcal{X}_{\mathcal{T}} \subseteq \mathcal{X}$, nor nestedness of the discrete spaces, i.e., $\mathcal{X}_\mathcal{T} \subseteq \mathcal{X}_{\mathcal{T}'}$ for $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$. Moreover, it is essentially immaterial whether $\mathcal{X}$ and $\mathcal{X}_\mathcal{T}$ have a vector space structure and if $\|\cdot\|$ is really a norm or metric [3, Section 2].

● **Applicability.** The axioms are independent of the actual formulations which determine the target u as well as the approximations $U_\mathcal{T}$ for $\mathcal{T} \in \mathbb{T}$. In particular, they are independent of the method, e.g., (conforming, nonconforming, mixed) finite elements or boundary elements, as well as of linearity or nonlinearity of the problem. Moreover, all results are independent of any Céa-type quasi-optimality

$$\|u - U_\mathcal{T}\| \le C_{\text{Céa}} \min_{V_\mathcal{T} \in \mathcal{X}_\mathcal{T}} \|u - V_\mathcal{T}\|, \tag{5}$$

where $\mathcal{X}_\mathcal{T}$ is the ansatz set for $U_\mathcal{T}$. Instead, all these properties only enter for the verification of the axioms in particular situations, see Section 5 as well as [3, Section 5 and 9].

● **Existing results.** The set of axioms (A1)–(A4) does neither contain the classical reliability, nor the efficiency estimate: The error estimator is ***reliable*** if there exists a constant $C'_{\text{rel}} > 0$ such that for all $\mathcal{T} \in \mathbb{T}$ holds

$$\|u - U_\mathcal{T}\| \le C'_{\text{rel}}\, \eta_\mathcal{T}. \tag{6}$$

The error estimator is ***efficient*** if there exists a constant $C_{\text{eff}} > 0$ such that for all $\mathcal{T} \in \mathbb{T}$ and some quantity $\text{osc}_\mathcal{T}$ holds

$$C_{\text{eff}}^{-1}\, \eta_\mathcal{T} \le \|u - U_\mathcal{T}\| + \text{osc}_\mathcal{T} \quad \text{and} \quad \text{osc}_\mathcal{T} \le C_{\text{eff}}\, \eta_\mathcal{T}. \tag{7}$$

Instead, we note that, first, discrete reliability (A4) together with (4) implies reliability (6) even with $C'_{\text{rel}} = C_{\text{rel}}$, see [3, Section 3.3]. Second, the mainstream literature on optimality of adaptive algorithms, e.g., [10, 4, 7, 8] relies on (6)–(7) and uses the *total error* $\|u - U_\mathcal{T}\| + \gamma\, \text{osc}_\mathcal{T}$ with some (generically small) constant $\gamma > 0$ to formulate their respective Main Result 1–2. We note that (6)–(7) imply

$$\min\{1, \gamma\}\, C_{\text{eff}}^{-1}\, \eta_\mathcal{T} \le \|u - U_\mathcal{T}\| + \gamma\, \text{osc}_\mathcal{T} \le (C_{\text{rel}} + \gamma C_{\text{eff}})\eta_\mathcal{T}, \tag{8}$$

so that our estimator-based formulation, in fact, generalizes those from the literature.

● **Estimator reduction.** Stability (A1) and reduction (A2) imply the existence of $C_{\text{est}} > 0$ and $0 < q_{\text{est}} < 1$ such that for all $\mathcal{T} \in \mathbb{T}$ and all $\mathcal{T}' \in \texttt{refine}(\mathcal{T})$ it holds

$$\theta\, \eta_\mathcal{T}^2 \le \sum_{T \in \mathcal{T} \setminus \mathcal{T}'} \eta_\mathcal{T}(T)^2 \quad \Longrightarrow \quad \eta_{\mathcal{T}'}^2 \le q_{\text{est}}\, \eta_\mathcal{T}^2 + C_{\text{est}}\, \|U_{\mathcal{T}'} - U_\mathcal{T}\|^2, \tag{9}$$

see [3, Section 4.3], where q_{est} and C_{est} depend also on the bulk parameter $0 < \theta < 1$.

● **Quasi-orthogonality (A3) and *R*-linear convergence.** Suppose that the error estimator satisfies the estimator reduction (9) and reliability (6). (For instance, assume that stability (A1), reduction (A2), and discrete reliability (A4) are satisfied.) Then, quasi-orthogonality (A3) is satisfied if and only if Algorithm 1 is R-linearly convergent (2). In this case, (A3) holds even with $\varepsilon_{\mathrm{orth}} = 0$, see [3, Section 4.4]. The subsequent Section 4 provides mathematical frameworks, where the quasi-orthogonality (A3) is problem inherently satisfied.

- **Discrete reliability (A4) and optimality of marking strategy.** Suppose stability (A1) and reduction (A2). Then, discrete reliability (A4) proves that the marking criterion (3) is not only sufficient for linear convergence, but also necessary: For $0 < \theta < \theta_\star := 1/(1 + C_{\mathrm{stab}}^2 C_{\mathrm{rel}}^2)$, there exists $0 < \kappa < 1$ such that for all $\mathcal{T} \in \mathbb{T}$ and all $\mathcal{T}' \in \mathtt{refine}(\mathcal{T})$, it holds

$$\eta_{\mathcal{T}'} \le \kappa\, \eta_{\mathcal{T}} \quad \Longrightarrow \quad \theta\, \eta_{\mathcal{T}}^2 \le \sum_{T \in \mathcal{R}(\mathcal{T}, \mathcal{T}')} \eta_{\mathcal{T}}(T)^2. \tag{10}$$

This means that if an adaptive algorithm yields R-linear convergence (2) of the estimator sequence $(\eta_\ell)_{\ell \in \mathbb{N}_0}$, there exists some $n \in \mathbb{N}$ such that the marking strategy (3) is at least satisfied every n steps of the adaptive loop (1).

4 VALIDITY OF QUASI-ORTHOGONALITY AXIOM (A3)

In many important situations of mixed FEM and conforming FEM, the abstract frame is the following: On the continuous level, the target belongs to a real Hilbert space $\mathcal{X} \ni u$ with norm $\|\cdot\|$ and solves the variational formulation

$$a(u, v) = f(v) \quad \text{for all } v \in \mathcal{X} \tag{11}$$

with a continuous bilinear form $a(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and a continuous linear functional $f : \mathcal{X} \to \mathbb{R}$. On the discrete level, each triangulation $\mathcal{T} \in \mathbb{T}$ induces a conforming subspace $\mathcal{X}_{\mathcal{T}} \subseteq \mathcal{X}$ with $U_{\mathcal{T}} \in \mathcal{X}_{\mathcal{T}}$ which satisfies nestedness $\mathcal{X}_{\mathcal{T}} \subseteq \mathcal{X}_{\mathcal{T}'}$ for all $\mathcal{T}' \in \mathtt{refine}(\mathcal{T})$. Moreover, the discrete solution solves the variational formulation

$$a(U_{\mathcal{T}}, V_{\mathcal{T}}) = f(V_{\mathcal{T}}) \quad \text{for all } V_{\mathcal{T}} \in \mathcal{X}_{\mathcal{T}}. \tag{12}$$

In this situation, the quasi-orthogonality (A3) is problem inherent (see Corollary 4 below). The following result from [6, Theorem 1] covers conforming discretizations as, e.g., FEM, BEM, and the FEM-BEM coupling, in the frame of the Lax-Milgram lemma.

Theorem 2 *Suppose that $a(\cdot,\cdot)$ is elliptic, i.e., there exists $C_{\mathrm{ell}} > 0$ such that*

$$C_{\mathrm{ell}}\, \|v\|^2 \le a(v, v) \quad \textit{for all } v \in \mathcal{X}. \tag{13}$$

Then, the discrete solutions $U_\ell \in \mathcal{X}_\ell$ generated by Algorithm 1 satisfy the generalized Pythagoras theorem

$$\sum_{k=\ell}^{\infty} \|U_{k+1} - U_k\|^2 \le C_{\mathrm{pyth}}\, \|u - U_\ell\|^2 \quad \textit{for all } \ell \in \mathbb{N}_0. \tag{14}$$

The constant $C_{\mathrm{pyth}} > 0$ depends only on $a(\cdot,\cdot)$ and the sequence $(\mathcal{X}_\ell)_{\ell \in \mathbb{N}_0}$ of nested discrete spaces. ■

Adopting the analysis of [6], the work [9] proves the following similar result [9, Theorem 2.1] which covers certain mixed FEM discretizations like the Taylor-Hood element for the stationary Stokes system. While Theorem 2 covers elliptic, but possibly non-symmetric bilinear forms, the following Theorem 3 assumes symmetry of $a(\cdot,\cdot)$.

Theorem 3 *Let $A \in L(\mathcal{X},\mathcal{X}^*)$ denote the linear operator induced by $Aw := a(w,\cdot)$ for all $w \in \mathcal{X}$. For $\mathcal{T} \in \mathbb{T}$, let $J_\mathcal{T} : \mathcal{X}_\mathcal{T} \to \mathcal{X}$ be the natural injection with adjoint $J_\mathcal{T}^*$. Suppose that $a(\cdot,\cdot)$ is symmetric, i.e., $a(v,w) = a(w,v)$ for all $v,w \in \mathcal{X}$, and satisfies some uniform LBB-condition, i.e., A as well as $A_\mathcal{T} := J_\mathcal{T}^* A J_\mathcal{T} \in L(\mathcal{X}_\mathcal{T},\mathcal{X}_\mathcal{T}^*)$ are isomorphisms with uniformly bounded operators norms*

$$\|A\|,\ \|A^{-1}\|,\ \|A_\mathcal{T}\|,\ \|A_\mathcal{T}^{-1}\| \le M \quad \text{for all } \mathcal{T} \in \mathbb{T}. \tag{15}$$

Then, the discrete solutions $U_\ell \in \mathcal{X}_\ell$ generated by Algorithm 1 satisfy the generalized Pythagoras theorem

$$\sum_{k=\ell}^{\infty} \|U_{k+1} - U_k\|^2 \le C_{\mathrm{pyth}}\, \|u - U_\ell\|^2 \quad \text{for all } \ell \in \mathbb{N}_0. \tag{16}$$

The constant $C_{\mathrm{pyth}} > 0$ depends only on M and the sequence $(\mathcal{X}_\ell)_{\ell\in\mathbb{N}_0}$ of nested discrete spaces. ■

The following corollary is an immediate consequence and proves that the (on a first glance critical) quasi-orthogonality (A3) is automatically satisfied.

Corollary 4 *Suppose that the assumptions of either Theorem 2 or Theorem 3 are satisfied. Then, discrete reliability* (A4) *(or even plain reliability* (6)*) implies the quasi-orthogonality* (A3) *with $\varepsilon_{\mathrm{orth}} = 0$ and $C_{\mathrm{orth}} = C_{\mathrm{pyth}} C_{\mathrm{rel}}$.* ■

One may expect that Theorem 2 and Theorem 3 (and consequently Corollary 4) still allow for structural improvements, e.g., avoidance of the Hilbert space structure, avoidance of the symmetry of $a(\cdot,\cdot)$ in Theorem 3, as well as weaker dependencies of C_{pyth}, namely independence of the sequence $(\mathcal{X}_\ell)_{\ell\in\mathbb{N}_0}$. We illustrate this for the most simple model situation, where $a(\cdot,\cdot)$ is a scalar product on $\mathcal{X}$ with $\|v\|^2 = a(\cdot,\cdot)$. Then, nestedness $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ and the discrete formulation (12) yield the Pythagoras theorem

$$\|u - U_{k+1}\|^2 + \|U_{k+1} - U_k\|^2 = \|u - U_k\|^2. \tag{17}$$

Hence, the telescoping series proves

$$\sum_{k=\ell}^{\ell+n} \|U_{k+1} - U_k\|^2 = \|u - U_\ell\|^2 - \|u - U_{\ell+n+1}\|^2 \le \|u - U_\ell\|^2, \tag{18}$$

i.e., $C_{\mathrm{pyth}} = 1$ depends only on $a(\cdot,\cdot)$, but not on $(\mathcal{X}_\ell)_{\ell\in\mathbb{N}_0}$.

Remark 5 *Suppose reliability* (6) *of the error estimator. Under the assumptions of either Theorem 2 or Theorem 3, the t stability of the Galerkin projections reveals for all $\mathcal{T}' \in$* `refine`$(\mathcal{T}_\ell)$ *and hence $\mathcal{X}_\ell \subseteq \mathcal{X}_{\mathcal{T}'}$*

$$C_{\mathrm{Gal}}^{-2} \, \|U_{\mathcal{T}'} - U_\ell\|^2 \leq \|u - U_\ell\|^2 \leq (C'_{\mathrm{rel}})^2 \, \eta_\ell^2 \leq (C'_{\mathrm{rel}})^2 \theta^{-1} \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2.$$

Consequently, this proves the discrete reliability (A4) *with $C_{\mathrm{rel}} = C_{\mathrm{Gal}} C'_{\mathrm{rel}} \theta^{-1/2}$ at least along the sequence of meshes $\mathcal{T}_\ell$ generated by Algorithm 1. Hence,* (A4) *appears to be not only sufficient, but also necessary.* ■

5 EXEMPLARY VERIFICATION OF AXIOMS

In the following, we verify the axioms (A1)–(A4) for certain examples. Throughout, $\Omega \subset \mathbb{R}^d$ for $d \geq 2$ is a bounded Lipschitz domain with polyhedral boundary $\partial\Omega$. We refer to [3, Section 5 and 9], where those axioms are proved for further examples with nonconforming and mixed FEM as well as conforming FEM for nonlinear problems.

5.1 Conforming FEM for Poisson model problem

For given $f \in L^2(\Omega)$, we consider

$$-\Delta u = f \text{ in } \Omega \quad \text{with homogeneous Dirichlet boundary conditions} \quad u = 0 \text{ on } \partial\Omega. \tag{19}$$

The corresponding bilinear form $a(u, v) := \int_\Omega \nabla u \cdot \nabla v \, dx$ is an equivalent scalar product on $\mathcal{X} = H_0^1(\Omega)$ with norm $\|u\| = \|\nabla u\|_{L^2(\Omega)}$. Given $\mathcal{T} \in \mathbb{T}$, let $\mathcal{P}^p(\mathcal{T})$ be the space of all $\mathcal{T}$-piecewise polynomials of degree p and $\mathcal{X}_\mathcal{T} := \mathcal{P}^p(\mathcal{T}) \cap H_0^1(\Omega)$. Overall, the problem fits in the abstract frame of Section 4, and the Lax-Milgram lemma proves existence and uniqueness of the solutions $u \in \mathcal{X}$ of (11) and $U_\mathcal{T} \in \mathcal{X}_\mathcal{T}$ of (12).

As implicitly proved in [4, Corollary 3.4], the standard residual error estimator with

$$\eta_\mathcal{T}(T)^2 := |T|^{2/d} \|f + \Delta U_\mathcal{T}\|_{L^2(T)}^2 + |T|^{1/d} \|[\nabla U_\mathcal{T} \cdot n]\|_{L^2(\partial T \cap \Omega)}^2 \quad \text{for all } T \in \mathcal{T} \tag{20}$$

satisfies stability (A1) and reduction (A2). We note that the proof of (A1) essentially follows from scaling arguments and inverse-type estimates, and the proof of (A2) additionally employs that the local mesh-size is uniformly shrunken $|T'| \leq |T|/2$ for successors $T' \in \mathcal{T}' \in$ `refine`$(\mathcal{T})$ of $T \in \mathcal{T} \backslash \mathcal{T}'$ with $T' \subsetneqq T$. The proof of (A4) is essentially the same as for the classical reliability, but relies on the choice of the Scott-Zhang projection and the construction of clever test functions [4, Lemma 3.6]. Finally, the quasi-orthogonality (A3) follows either from Corollary 4 or the elementary calculation (17)–(18). Overall, this verifies the axioms (A1)–(A4) and yields the following

Consequence. Algorithm 1 for conforming FEM for the Poisson model problem converges optimally in the sense of Main Result 1–2. ■

5.2 Conforming FEM for linear second-order elliptic operator

We consider

$$\mathcal{L}u := -\mathrm{div}(\boldsymbol{A}\nabla u) + \boldsymbol{b}\cdot\nabla u + cu = f \text{ in } \Omega \quad \text{with} \quad u = 0 \text{ on } \partial\Omega. \tag{21}$$

We suppose a symmetric diffusion matrix $\boldsymbol{A} \in W_1^\infty(\Omega)$ with $\boldsymbol{A}(x) \in \mathbb{R}^{d\times d}_{sym}$, convection $\boldsymbol{b} \in L^\infty(\Omega)$ with $\boldsymbol{b}(x) \in \mathbb{R}^d$, and reaction $c \in L^\infty(\Omega)$ with $c(x) \in \mathbb{R}$. Here, $L^\infty(\Omega)$ is the space of essentially bounded functions, and $W_1^\infty(\Omega) := \{a \in L^\infty(\Omega) : \nabla a \in L^\infty(\Omega)$ in the weak sense$\}$ coincides with the space of Lipschitz continuous functions.

Note that $\mathcal{L}$ is non-symmetric as

$$\mathcal{L} \neq \mathcal{L}^T = -\mathrm{div}\boldsymbol{A}\nabla u - \boldsymbol{b}\cdot\nabla u + (c - \mathrm{div}\boldsymbol{b})u.$$

We let $f \in L^2(\Omega)$ and assume that the induced bilinear form

$$a(u,v) := \int_\Omega \boldsymbol{A}\nabla u\cdot\nabla v + \boldsymbol{b}\cdot\nabla uv + cuv\,dx \quad \text{for } u,v \in H_0^1(\Omega)$$

is continuous and elliptic on $\mathcal{X} := H_0^1(\Omega)$. With $\mathcal{X}_\mathcal{T} := \mathcal{P}^p(\mathcal{T}) \cap H_0^1(\Omega)$, the problem thus fits in the abstract frame of Section 4, and the Lax-Milgram lemma proves existence and uniqueness of the solutions $u \in \mathcal{X}$ of (11) and $U_\mathcal{T} \in \mathcal{X}_\mathcal{T}$ of (12).

For this problem, the refinement indicators of the residual error estimator read

$$\eta_\mathcal{T}(T)^2 := |T|^{2/d}\|f - \mathcal{L}|_T U_\mathcal{T}\|^2_{L^2(T)} + |T|^{1/d}\|[\boldsymbol{A}\nabla U_\mathcal{T}\cdot n]\|^2_{L^2(\partial T\cap\Omega)} \quad \text{for all } T \in \mathcal{T}, \tag{22}$$

where $|T|$ denotes the volume of the simplex T. Arguing along the lines of [4], it is shown in [5] that the residual error estimator satisfies stability (A1), reduction (A2), and discrete reliability (A4) with $\mathcal{R}(\mathcal{T},\mathcal{T}') = \mathcal{T}\setminus\mathcal{T}'$. The quasi-orthogonality (A3) — and consequently the validity of Main Result 1–2 — was open until [5], and appears there first in the literature. The alternative approach of Theorem 2, influenced by the abstract developments of [3], has only recently been discovered [6]. Overall, this verifies the axioms (A1)–(A4) and yields the following

Consequence. Algorithm 1 for conforming FEM for some general linear second-order elliptic operator converges optimally in the sense of Main Result 1–2. ■

5.3 Conforming BEM for weakly-singular integral equation

Let $\Gamma \subseteq \partial\Omega$ be a relatively open subset. For given $f \in H^{1/2}(\Gamma) := \{\phi|_\Gamma : \phi \in H^1(\Omega)\}$, we consider the weakly-singular first-kind integral equation

$$\mathcal{V}u(x) := \int_\Gamma G(x-y)u(y)\,d\Gamma(y) = f(x) \quad \text{for all } x \in \Gamma, \tag{23a}$$

with the fundamental solution of the Laplacian

$$G(z) := -\frac{1}{2\pi}\log|z| \text{ for } d=2 \quad \text{resp.} \quad G(z) := \frac{1}{4\pi}\frac{1}{|z|} \text{ for } d=3. \tag{23b}$$

The sought solution satisfies $u \in \widetilde{H}^{-1/2}(\Gamma)$ which is the dual space of $H^{1/2}(\Gamma)$ with respect to the extended $L^2(\Gamma)$-scalar product. We note that $\mathcal{V} \in L(\widetilde{H}^{-1/2+s}(\Gamma); H^{1/2+s}(\Gamma))$ is a linear, continuous, and symmetric operator for all $-1/2 < s < 1/2$. For 2D, we assume $\mathrm{diam}(\Omega) < 1$ which can always be achieved by scaling. Then, $\mathcal{V}$ is also $\widetilde{H}^{-1/2}$-elliptic and hence defines an equivalent scalar product on $\mathcal{X} := \widetilde{H}^{-1/2}(\Gamma)$,

$$a(u,v) := \int_\Gamma (\mathcal{V}u)v\,d\Gamma \quad \text{for } u,v \in \widetilde{H}^{-1/2}(\Gamma) \tag{24}$$

with induced norm $\|v\|^2 := a(v,v)$. For $\mathcal{T} \in \mathbb{T}$ being a surface triangulation of Γ, let $\mathcal{X}_\mathcal{T} := \mathcal{P}^p(\mathcal{T})$. Then, the problem fits in the frame of Section 4, and the Lax-Milgram lemma proves existence and uniqueness of the solutions $u \in \mathcal{X}$ of (11) and $U_\mathcal{T} \in \mathcal{X}_\mathcal{T}$ of (12).

Under additional regularity of the data $f \in H^1(\Gamma)$, we consider the weighted-residual error estimator, e.g., [7] with local contributions

$$\eta_\mathcal{T}(T)^2 := |T|^{1/(d-1)}\,\|\nabla_\Gamma(f - \mathcal{V}U_\mathcal{T})\|^2_{L^2(T)} \quad \text{for all } T \in \mathcal{T}. \tag{25}$$

Here, $\nabla_\Gamma(\cdot)$ denotes the surface gradient (resp. the arclength derivative for $d = 2$).

Stability (A1) and reduction (A2) are implicitly proved in [7, Proposition 4.2] and rely on new inverse-type estimates which have been independently first proved in [7, 8]. While [7] only treats the lowest-order case $p = 0$, but polyhedral boundaries, [8] treats general $p \geq 0$, but is restricted to smooth $C^{1,1}$ boundaries. General $p \geq 0$ and polyhedral boundaries are analyzed in [1]. Discrete reliability (A4) is proved in [7, Proposition 5.3], where $\mathcal{R}(\mathcal{T},\mathcal{T}')$ consists of the refined elements $\mathcal{T}\backslash\mathcal{T}'$ plus all their neighbors. As for the Poisson model problem, the quasi-orthogonality (A3) follows either from Corollary 4 or the elementary calculation (17)–(18). Overall, this verifies the axioms (A1)–(A4) and yields the following

Consequence. Algorithm 1 for conforming BEM for the weakly-singular integral equation (23) converges optimally in the sense of Main Result 1–2. ■

6 MATHEMATICAL STATEMENT OF MAIN RESULTS

In the following, we give the formal statements of our main results from the introduction.

Main Result 1 *Suppose stability* (A1)*, reduction* (A2)*, and quasi-orthogonality* (A3)*. Then, for all* $0 < \theta \leq 1$*, reliability* (6) *resp. discrete reliability* (A4) *imply R-linear* *convergence of the estimator in the sense that there exists* $0 < q_{\rm conv} < 1$ *and* $C_{\rm conv} > 0$ *such that*

$$\eta_{\ell+n} \le C_{\rm conv} q_{\rm conv}^n \eta_\ell \quad \textit{for all } \ell, n \in \mathbb{N}_0.$$

The constants $0 < q_{\rm conv} < 1$ *and* $C_{\rm conv} > 0$ *depend only on* $C_{\rm stab}$*,* $q_{\rm red}$*,* $C_{\rm red}$*,* $C_{\rm orth}(\varepsilon_{\rm orth})$*,* $C_{\rm rel}$*, and* θ*.* ■

The best possible algebraic convergence rate $0 < s < \infty$ obtained by any local mesh refinement is characterized in terms of

$$\|u\|_{\rm s} := \sup_{N\in\mathbb{N}_0} \min_{\substack{\mathcal{T}\in\mathbb{T}\\ \#\mathcal{T}-\#\mathcal{T}_0\le N}} (N+1)^s \|u - U_\mathcal{T}\| < \infty.$$

The statement $\|u\|_{\rm s} < \infty$ means that $\|u - U_\mathcal{T}\| = \mathcal{O}(N^{-s})$ for the optimal triangulations $\mathcal{T} \in \mathbb{T}$, independently of the error estimator. Since the adaptive algorithm is steered by the error estimator $\eta_\mathcal{T}$, it appears natural to consider the best algebraic convergence rate $\mathcal{O}(N^{-s})$ in terms of $\eta_\mathcal{T}$, characterized by

$$\|\eta\|_{\rm s} := \sup_{N\in\mathbb{N}_0} \min_{\substack{\mathcal{T}\in\mathbb{T}\\ \#\mathcal{T}-\#\mathcal{T}_0\le N}} (N+1)^s \eta_\mathcal{T} < \infty.$$

This implies the convergence rate $\eta_\mathcal{T} = \mathcal{O}(N^{-s})$ for the optimal triangulations $\mathcal{T} \in \mathbb{T}$.

Main Result 2 *Suppose stability* (A1)*, reduction* (A2)*, quasi-orthogonality* (A3)*, and discrete reliability* (A4)*. Then, for all* $0 < \theta < 1/(1 + C_{\rm stab}^2 C_{\rm rel}^2)$*, the error estimator converges with the optimal rate in the sense that there exists* $c_{\rm opt}$*,* $C_{\rm opt} > 0$ *such that*

$$c_{\rm opt}\|\eta\|_{\rm s} \le \sup_{\ell\in\mathbb{N}_0} \frac{\eta_\ell}{(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}} \le C_{\rm opt}\|\eta\|_{\rm s}.$$

The constant $c_{\rm opt}$ *depends only on NVB, whereas the constant* $C_{\rm opt}$ *depends additionally on* θ*,* s*,* $C_{\rm stab}$*,* $q_{\rm red}$*,* $C_{\rm red}$*,* $C_{\rm orth}(\varepsilon_{\rm orth})$*,* $C_{\rm rel}$*, and* $C_{\rm ref}$*.* ■

The detailed proofs of Main Result 1–2 are given in [3, Section 4, Theorem 4.1]. If the error estimator is also efficient (7), we can improve the above result. With

$$\|{\rm osc}\|_{\rm s} := \sup_{N\in\mathbb{N}_0} \min_{\substack{\mathcal{T}\in\mathbb{T}\\ \#\mathcal{T}-\#\mathcal{T}_0\le N}} (N+1)^s {\rm osc}_\mathcal{T} < \infty,$$

there holds the following corollary [3, Theorem 4.5].

Corollary 6 *Suppose the claims of Main Result 2 and let* $\eta_\mathcal{T}$ *additionally satisfy efficiency* (7)*. Then, the error converges with optimal rate in the sense that there exists* $c'_{\rm opt}$*,* $C'_{\rm opt} > 0$ *such that*

$$c'_{\rm opt}\|u\|_{\rm s} \le \sup_{\ell\in\mathbb{N}_0} \frac{\|u - U_\ell\|}{(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}} + \|{\rm osc}\|_{\rm s} \le C'_{\rm opt}\big(\|u\|_{\rm s} + \|{\rm osc}\|_{\rm s}\big).$$

The constants $c'_{\rm opt}$*,* $C'_{\rm opt} > 0$ *depend only on* $c_{\rm opt}$*,* $C_{\rm opt} > 0$ *and on* $C_{\rm eff}$*.* ■

REFERENCES

[1] Aurada, M., Feischl, M., Führer, T., Karkulik, M., Melenk, J.M., and Praetorius, D., Inverse estimates for elliptic boundary integral operators and their application to the adaptive coupling of FEM and BEM. *ASC Report* **07/2012**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2012.

[2] Aurada, M., Feischl, M., Kemetmüller, J., Page, M., and Praetorius, D., Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$. *M2AN Math. Model. Numer. Anal.* (2013) **47**:1207–1235.

[3] Carstensen, C., Feischl, M., Page, M., and Praetorius, D., Axioms of adaptivity. *Comput. Math. Appl.* (2014) **67**:1195–1253.
Download (open access): `http://dx.doi.org/10.1016/j.camwa.2013.12.003`

[4] Cascon, J.M., Kreuzer, C., Nochetto, R.H., and Siebert, K., Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* (2008) **46**:2524–2550.

[5] Feischl, M., Führer, T., and Praetorius, D., Adaptive FEM with optimal convergence rates for a certain class of non-symmetric and possibly non-linear problems. *SIAM J. Numer. Anal.* (2014) **52**:601–625.

[6] Feischl, M., Führer, T., and Praetorius, D., Adaptive FEM, BEM, and FEM-BEM coupling with optimal rates for strongly non-symmetric problems. *ASC Report* **39/2013**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2013.

[7] Feischl, M., Karkulik, M., Melenk, J.M., and Praetorius, D., Quasi-optimal convergence rate for an adaptive boundary element method. *SIAM J. Numer. Anal.* (2013) **51**:1327–1348.

[8] Gantumur, T., Adaptive boundary element methods with convergence rates. *Numer. Math.* (2013) **124**:471–516.

[9] Gantumur, T., Optimal convergence rates for an adaptive mixed finite element method for the Stokes problem. Preprint (2014), `http://arxiv.org/abs/1403.0895`.

[10] Stevenson, R., Optimality of a standard adaptive finite element method. *Found. Comput. Math.* (2007) **7**:245–269.

[11] Stevenson, R., The completion of locally refined simplicial partitions created by bisection. *Math. Comp.* (2008) **77**:227–241.