ASC Report No. 24/2013

Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, Part I: Weakly-singular integral equation

M. Feischl, T. Führer, M. Karkulik, J.M. Melenk,
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

23/2013 *W. Auzinger, W. Herfort*
Local error structures and order conditions in terms of Lie elements for exponential splitting schemes

22/2013 *Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, and Dirk Praetorius*
Energy norm based error estimators for adaptive BEM for hypersingular integral equations

21/2013 *Michael Karkulik, Carl-Martin Pfeiler, and Dirk Praetorius*
L^2-orthogonal projections onto lowest-order finite elements in R^d are H^1-stable

20/2013 *Markus Faustmann, Jens Markus Melenk, Dirk Praetorius*
$\mathcal{H}$-matrix approximability of the inverses of FEM matrices

19/2013 *Winfried Auzinger, Harald Hofstätter, Othmar Koch, Mechthild Thalhammer*
Defect-based local error estimates for splitting methods, with application to Schrödinger equations
Part III. The nonlinear case

18/2013 *A. Feichtinger, A. Makaruk, E. Weinmüller, A. Friedl, M. Harasek*
Collocation method for the modeling of membrane gas permeation systems

17/2013 *M. Langer, H. Woracek*
Distributional representations of $\mathcal{N}_\kappa^{(\infty)}$-functions

16/2013 *M. Feischl, T. Führer, M. Karkulik, D. Praetorius*
ZZ-type a posteriori error estimators for adaptive boundary element methods on a curve

15/2013 *K.-N. Le, M. Page, D. Praetorius, and T. Tran*
On a decoupled linear FEM integrator for Eddy-Current-LLG

14/2013 *X. Chen, A. Jüngel, and J.-G. Liu*
A note on Aubin-Lions-Dubinskii lemmas

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

QUASI-OPTIMAL CONVERGENCE RATES FOR ADAPTIVE BOUNDARY ELEMENT METHODS WITH DATA APPROXIMATION, PART I: WEAKLY-SINGULAR INTEGRAL EQUATION

M. FEISCHL, T. FÜHRER, M. KARKULIK, J. M. MELENK, AND D. PRAETORIUS

Abstract. We analyze an adaptive boundary element method for Symm's integral equation in 2D and 3D which incorporates the approximation of the Dirichlet data g into the adaptive scheme. We prove quasi-optimal convergence rates for any $H^{1/2}$-stable projection used for data approximation.

1. Introduction & Outline

Data approximation is ubiquitous in numerical algorithms, and reliable, adaptive numerical schemes have to properly account for it. In this direction, the present work proves quasi-optimal convergence rates for an adaptive boundary element method (ABEM) that includes data errors. As a model problem, we study Symm's integral equation

$$V\phi = (1/2 + K)g \quad \text{on } \Gamma := \partial\Omega \tag{1}$$

for given boundary data $g \in H^{1/2}(\Gamma)$ and a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ for $d = 2$, 3. The goal is to prove convergence and quasi-optimality of some standard adaptive algorithm of the type

$$\boxed{\texttt{solve}} \rightarrow \boxed{\texttt{estimate}} \rightarrow \boxed{\texttt{mark}} \rightarrow \boxed{\texttt{refine}}$$

steered by a residual-based error estimator plus data approximation terms. A focus of our analysis will be on the fact that the data g is not given exactly but approximated as part of the algorithm. Our theory covers several of commonly used techniques to approximate g.

In the framework of h-adaptive finite element methods (AFEM) for second order elliptic PDEs, algorithms of this type have been studied in several works, and convergence with quasi-optimal algebraic rates can be proven (see e.g. [BDD04, CKNS08, Dör96, FFP12, Ste07] and the references therein). Naturally, one is interested in the very same questions like convergence of the approximations and convergence rates also for ABEM. The recent works [FKMP13] and [Tso13] lay out the path for proving quasi-optimal convergence rates of ABEM with respect to the error estimator, or even for the energy error [AFF$^+$14]. However, the above mentioned works [AFF$^+$14, FKMP13, Tso13] are restricted to lowest-order discretizations and, more importantly, do not deal explicitly with data approximation. The present paper raises ABEM to the same level of mathematical understanding as AFEM already is. More precisely, the improvements over the state of the art are fourfold and read as follows:

First, in contrast to the FEM, the right-hand sides in BEM typically involve boundary integral operators, which cannot be evaluated exactly in practice. Thus, the analysis of data error is mandatory. To compute the right-hand side term $(1/2 + K)g$ numerically

Date: August 28, 2013.

2000 *Mathematics Subject Classification.* 65N30, 65N15, 65N38.

Key words and phrases. boundary element method, weakly-singular integral equation, a posteriori error estimate, adaptive algorithm, convergence, optimality.

in our model problem (1), we follow the earlier works [AFLG$^+$12, KOP13] and replace the exact data g by an approximate, piecewise polynomial data G_ℓ. Our approach thus decouples the problem of integrating the singular kernel of the integral operator K from integrating the possibly singular data g to compute Kg. On their own, both problems are well understood and can be solved with standard methods. Moreover, in 2D (see [Mai01]) one can even find analytic formulas to compute the term KG_ℓ exactly. For $d = 3$, there also exist black-box quadrature algorithms to compute KG_ℓ (see, e.g., [SS11]).

Second, in contrast to [FKMP13, Tso13], where only lowest-order BEM is considered, the present analysis works for arbitrary, but fixed-order discretizations.

Third, we provide an improved analysis for the optimality of the Dörfler marking. This eliminates the efficiency constant in the estimates (even the weak efficiency used in [FKMP13] is not needed) and in contrast to e.g. [CKNS08, FKMP13, Tso13] no lower error bound of any kind is involved.

Finally, to deal with several non-local data approximation terms, we introduce a modified mesh size function. This may be of independent interest in the context of AFEM and ABEM since it is pointwise equivalent to the usual mesh size function, but contractive not only on the refined elements, but also on an arbitrary but fixed number of element layers around them.

An overall advantage of the presented approach is that the possible implementation only has to deal with operator matrices of the discrete integral operators. This is advantageous in terms of fast boundary element methods as e.g. $\mathcal{H}$-matrices. Consequently, adaptive approximation of the Dirichlet data g seems to be the natural next step to the final goal of a fast, fully discrete, and black-box ABEM algorithm.

Several other works deal with data approximations for adaptive BEM. However, they focus on convergence of the error estimator instead of proving quasi-optimal algebraic convergence rates. In the 2D case, [AFLG$^+$12] proves estimator convergence of ABEM for the Laplace problem with mixed boundary conditions. The algorithm is steered by an $(h - h/2)$-based error estimator and also approximates the given data adaptively by nodal interpolation. The work [KOP13] uses the L^2-projection to prove estimator convergence of ABEM with data approximation in 3D. Both works do not guarantee any convergence rate of the estimator, and convergence of the error can only by proved under the *saturation assumption*, which is widely believed to hold true in practice, but still remains mathematically open for BEM (see [AFF$^+$14] for the proof of a weaker form of this assumption). In contrast to this, the present approach with residual-based error estimator guarantees convergence even with optimal rates. In particular, we prove that the optimal convergence is independent of the chosen data approximation operator.

The remainder of this work is organized as follows: We present the model problem as well as the adaptive algorithm for data approximation by means of the Scott-Zhang projection for $d = 2, 3$ or nodal interpolation for $d = 2$ in Section 2. In Section 3, we develop some crucial tools which are used to prove convergence of the adaptive scheme in Section 4 and quasi-optimality in Section 5. Bootstrapping the foregoing results, we are able to prove quasi-optimal convergence of a slightly modified algorithm in Section 6 for each $H^{1/2}$-stable projection P_ℓ used for data approximation, i.e. $G_\ell = P_\ell g$. Finally, Section 7 presents a numerical experiment which underlines the results of the work.

Throughout the work, the symbol $\lesssim$ abbreviates $\leq$ up to a multiplicative constant, and $\simeq$ means that both estimates $\lesssim$ and $\gtrsim$ hold. Finally, $\#\mathcal{M}$ denotes the cardinality of a finite set $\mathcal{M}$.

2. Model Problem & Adaptive Algorithm

2.1. **Model problem.** We consider Symm's integral equation (1) where $\Gamma := \partial\Omega$ is the boundary of a polygonal resp. polyhedral Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$. For $d = 2$, we ensure $\operatorname{diam}(\Omega) < 1$ by scaling of the domain to guarantee the ellipticity of the simple-layer operator V (see (3) below). With $n(x) \in \mathbb{R}^d$ denoting the exterior normal unit field at $x \in \Gamma$ and the fundamental solution of the Laplacian

$$G(z) := \begin{cases} -\frac{1}{2\pi}\log|z| & d = 2 \\ \frac{1}{4\pi}|z|^{-1} & d = 3 \end{cases} \quad \text{for all } z \in \mathbb{R}^d \setminus \{0\}, \tag{2}$$

the simple-layer operator V and the double-layer operator K formally read

$$(V\phi)(x) := \int_\Gamma G(x-y)\phi(y)\,dy \quad \text{and} \quad (Kg)(x) := \text{p.v.}\int_\Gamma \partial_{n(y)}G(x-y)g(y)\,dy \tag{3}$$

for all $x \in \Gamma$. Here, p.v. $\int_\Gamma$ denotes Cauchy's principal value. Then, (1) is an equivalent formulation of the Dirichlet problem

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{on } \Gamma \end{aligned} \tag{4}$$

in the following sense [McL00]: The normal derivative $\phi = \partial_n u \in H^{-1/2}(\Gamma)$ of the solution $u \in H^1(\Omega)$ of (4) solves (1) (Note that $\partial_n u \in H^{-1/2}(\Gamma)$ is well-defined, since $\Delta u \in L^2(\Gamma)$). Conversely, with $\phi \in H^{-1/2}(\Gamma)$ being a solution of (1), the representation formula

$$u = V\phi - Kg \in H^1(\Omega)$$

gives the solution of (4), where the operators V and K are now evaluated in Ω instead of Γ.

2.2. **Variational form and unique solvability.** The operator $V : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ is an elliptic and symmetric isomorphism (see e.g. the monographs [HW08, McL00, SS11]). It thus provides a scalar product defined by $\langle\!\langle \phi\,,\,\psi \rangle\!\rangle := \langle V\phi\,,\,\psi\rangle_{L^2(\Gamma)}$. This scalar product induces an equivalent norm on $H^{-1/2}(\Gamma)$, which will be denoted by $|\!|\!|\psi|\!|\!| := \langle\!\langle \psi\,,\,\psi \rangle\!\rangle^{1/2}$. For some Γ-dependent constant $C_{\rm norm} > 0$, it thus holds

$$C_{\rm norm}^{-1}|\!|\!|\psi|\!|\!| \le \|\psi\|_{H^{-1/2}(\Gamma)} \le C_{\rm norm}|\!|\!|\psi|\!|\!| \quad \text{for all } \psi \in H^{-1/2}(\Gamma). \tag{5}$$

With this, we may state (1) equivalently as

$$\langle\!\langle \phi\,,\,\psi \rangle\!\rangle = \langle (K + \tfrac{1}{2})g\,,\,\psi\rangle \quad \text{for all } \psi \in H^{-1/2}(\Gamma). \tag{6}$$

The fact that $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle$ is a scalar product allows us to apply the Lax-Milgram lemma and hence guarantees existence and uniqueness of the solution $\phi \in H^{-1/2}(\Gamma)$ of (1). Whereas $g \in H^{1/2}(\Gamma)$ is sufficient to guarantee the solvability of (1), even without data approximation, it is necessary to require $g \in H^1(\Gamma)$ to formulate the weighted residual error estimator η_ℓ. In the present case (see Section 2.5 below), we make also use of the fact that we approximate an $H^1(\Gamma)$ function.

2.3. **Sobolev spaces and mapping properties.** Consider the Hilbert space

$$H^1(\Omega) := \Big\{ v \in L^2(\Omega) \;:\; \nabla v \in L^2(\Omega)^d \Big\},$$

equipped with the norm $\|v\|_{H^1(\Omega)}^2 := \|v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2$. We define the trace $v|_\Gamma$ of a function $v \in H^1(\Omega)$ by continuous extension of the classical trace for smooth functions. This permits the definition of

- the trace space

$$H^{1/2}(\Gamma) := \Big\{ v \in L^2(\Gamma) \,:\, \text{exists } w \in H^1(\Omega) \text{ with } v = w|_\Gamma \Big\}$$

  associated with the norm $\|v\|_{H^{1/2}(\Gamma)} := \inf \Big\{ \|w\|_{H^1(\Omega)} \,:\, v = w|_\Gamma \text{ for } w \in H^1(\Omega) \Big\}$ and
- its dual space

$$H^{-1/2}(\Gamma) := H^{1/2}(\Gamma)^*.$$

We revisit the integral operators from (3) and continuously extend them to the following boundary integral operators:

$$\begin{aligned} V &: H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma), \\ K &: H^{1/2}(\Gamma) \to H^{1/2}(\Gamma). \end{aligned}$$

According to [McL00, Chapter 7] there holds also well-posedness and continuity of

$$\begin{aligned} V &: H^{-1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma), \\ K &: H^{1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma). \end{aligned}$$

for all $|s| \leq 1/2$.

2.4. **Discretization.** Let $\mathcal{T}_\ell$ denote a regular triangulation of Γ into compact and flat boundary simplices $T \in \mathcal{T}_\ell$ with Euclidean diameter $\mathrm{diam}(T)$ and surface area $|T|$. Note that $\mathrm{diam}(T) = |T|$ for $d = 2$. For $d = 3$, we restrict to γ-shape regular meshes, i.e.

$$\gamma^{-1} \,\mathrm{diam}(T) \leq |T|^{1/2} \leq \gamma \,\mathrm{diam}(T) \quad \text{for all } T \in \mathcal{T}_\ell \tag{7}$$

for a fixed constant $\gamma \geq 1$. For $d = 2$, γ-shape regularity is understood in the sense of

$$\max \Big\{ \mathrm{diam}(T)/\mathrm{diam}(T') \,:\, T, T' \in \mathcal{T}_\ell \text{ and } T \cap T' \neq \emptyset \Big\} \leq \gamma. \tag{8}$$

Note that this assumption does not exclude strongly adapted meshes.

Given a mesh $\mathcal{T}_\ell$ and $p \in \mathbb{N} \cup \{0\}$, we define the space

$$\mathcal{P}^p(\mathcal{T}_\ell) := \Big\{ \Psi_\ell \in L^2(\Gamma) \,:\, \Psi_\ell|_T \text{ is a polynomial of degree at most } p \text{ on each } T \in \mathcal{T}_\ell \Big\}$$

of piecewise polynomials of degree p as well as the space

$$\mathcal{S}^{p+1}(\mathcal{T}_\ell) := \Big\{ V_\ell \in C(\Gamma) \,:\, V_\ell|_T \text{ is a polynomial of degree at most } p+1 \text{ on each } T \in \mathcal{T}_\ell \Big\}$$

of continuous spline functions of degree $p+1$. Note that there holds $\mathcal{P}^p(\mathcal{T}_\ell) \subset H^{-1/2}(\Gamma)$ and $\mathcal{S}^{p+1}(\mathcal{T}_\ell) \subset H^1(\Gamma)$.

Moreover, we need generalized patches of subsets $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$. We define the k-patch $\omega_\ell^k(\mathcal{E}_\ell) \subseteq \mathcal{T}_\ell$ inductively by

$$\begin{aligned} k = 1 : &\quad \omega_\ell(\mathcal{E}_\ell) := \omega_\ell^1(\mathcal{E}_\ell) := \Big\{ T \in \mathcal{T}_\ell \,:\, T \cap T' \neq \emptyset \text{ for some } T' \in \mathcal{E}_\ell \Big\}, \\ k > 1 : &\quad \omega_\ell^k(\mathcal{E}_\ell) := \omega_\ell(\omega_\ell^{k-1}(\mathcal{E}_\ell)). \end{aligned}$$

Note that due to γ-shape regularity, there holds $\#\mathcal{E}_\ell \simeq \#\omega_\ell^k(\mathcal{E}_\ell)$, where the hidden constant depends only on $\gamma > 0$ and $k \in \mathbb{N}$, but not on $\mathcal{E}_\ell$ or $\mathcal{T}_\ell$. Finally, for $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$, we write $\bigcup \mathcal{E}_\ell := \bigcup_{T \in \mathcal{E}_\ell} T \subseteq \Gamma$.

2.5. **Data approximation.** To compute the right-hand side $(1/2+K)g$ in the Galerkin scheme, we replace the exact Dirichlet data g by an approximation $G_\ell \in \mathcal{S}^{p+1}(\mathcal{T}_\ell)$. It remains to control the error introduced in this way. To that end, we define the auxiliary solution $\phi_\ell \in L^2(\Gamma) \subset H^{-1/2}(\Gamma)$

$$V\phi_\ell = (1/2+K)G_\ell. \tag{9}$$

The regularity of the solution ϕ_ℓ is a consequence of $G_\ell \in H^1(\Gamma)$ and the mapping properties of $V^{-1} : H^1(\Gamma) \to L^2(\Gamma)$ and $K : H^1(\Gamma) \to H^1(\Gamma)$.

The problem we actually solve on the discrete level reads: Find $\Phi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$ such that

$$\langle\!\langle \Phi_\ell\,,\, \Psi_\ell \rangle\!\rangle = \langle (1/2+K)G_\ell\,,\, \Psi_\ell\rangle_{L^2(\Gamma)} \tag{10}$$

for all $\Psi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$. As for the continuous problem, the Lax-Milgram lemma applies and proves the unique solvability of (10). In Sections 3–5, we analyze data approximation by means of the Scott-Zhang projection $J_\ell :\, L^2(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$, i.e.,

$$G_\ell := J_\ell g.$$

The mapping property $J_\ell : L^2(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$ deserves a comment. The original construction of [SZ90] is defined on $H^1(\Gamma)$ so as to be able to conserve boundary conditions. However, since we are not interested in conservation of boundary data (Γ has no boundary), we may define the Scott-Zhang projection for a mesh $\mathcal{T}_\ell$ in an element-based way on $L^2(\Gamma)$. We briefly sketch the construction. We choose a nodal basis of $\mathcal{S}^{p+1}(\mathcal{T}_\ell)$ with basis functions ξ_z associated with the Lagrangian nodes $z \in \mathcal{N}_\ell$ of $\mathcal{T}_\ell$ (i.e. $\xi_z(z') = \delta_{zz'}$ for all $z, z' \in \mathcal{N}_\ell$ and Kronecker's delta $\delta_{zz'} \in \{0,1\}$). Note that $\mathrm{supp}(\xi_z) \subseteq \omega_\ell(T_z)$ for some arbitrary, but fixed element $T_z \in \mathcal{T}_\ell$ with $z \in T_z$. Let $\zeta_z \in \mathcal{P}^{p+1}(T_z)$ denote the L^2-dual basis function with respect to $\xi_z|_{T_z}$, i.e. $\int_{T_z} \xi_{z'}\zeta_z\,dx = \delta_{zz'}$ for all $z, z' \in \mathcal{N}_\ell$. Then, J_ℓ is defined as

$$J_\ell g := \sum_{z\in\mathcal{N}_\ell} \Big(\int_{T_z} g\zeta_z\,dx\Big)\,\xi_z,$$

Obviously, $g \in L^2(\Gamma)$ is sufficient to define $J_\ell g$, and the results of [SZ90] hold accordingly.

Data approximation via other $H^{1/2}$-stable projections such as the L^2-orthogonal projection (for H^1-stability see [KPP13]) or the nodal interpolation for $d = 2$ is analyzed in Section 6.

2.6. **Mesh refinement.** For local mesh refinement, we use the bisection algorithm from [AFF+14] for $d = 2$ and *newest vertex bisection*, see e.g. [Ver96, Chapter 4], for $d = 3$. For a set of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$, we denote by $\mathcal{T}_\star = \mathtt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ the coarsest refinement (with respect to the above mentioned bisection algorithms), where at least all elements $T \in \mathcal{M}_\ell$ are refined. To ensure uniform γ-shape regularity (7) resp. (8), further refinements are made so that $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_\star$. We write $\mathcal{T}_\star \in \mathtt{refine}(\mathcal{T}_\ell)$ if there exists a finite sequence of meshes $\mathcal{T}_{\ell,0} = \mathcal{T}_\ell, \mathcal{T}_{\ell,1}, \dots, \mathcal{T}_{\ell,N} = \mathcal{T}_\star$ and sets of marked elements $\mathcal{M}_{\ell,j} \subset \mathcal{T}_{\ell,j}$ for $j = 0, \dots, N-1$ such that $\mathcal{T}_{\ell,j+1} := \mathtt{refine}(\mathcal{T}_{\ell,j}, \mathcal{M}_{\ell,j})$ for all $j = 0, \dots, N-1$. The set of all meshes which can be obtained by refinement of the initial mesh $\mathcal{T}_0$ is denoted by

$$\mathbb{T} := \{\mathcal{T} \in \mathtt{refine}(\mathcal{T}_0)\}. \tag{11}$$

We emphasize the following two crucial properties: First, the number of additional refinements which ensure regularity and γ-shape regularity, does not dominate the number of marked elements. More precisely, for $\mathcal{T}_\ell = \texttt{refine}(\mathcal{T}_0)$ with $\mathcal{T}_{j+1} = \texttt{refine}(\mathcal{T}_j, \mathcal{M}_j)$ and $\mathcal{M}_j \subseteq \mathcal{T}_j$ for $j = 0, \dots, \ell-1$, it holds that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \le C_1 \sum_{j=0}^{\ell-1} \#\mathcal{M}_j, \tag{12}$$

where $C_1 > 0$ depends only on $\mathcal{T}_0$. Furthermore, for two meshes $\mathcal{T}_\star, \mathcal{T}_\ell \in \mathbb{T}$, there exists a coarsest common refinement $\mathcal{T}_\star \oplus \mathcal{T}_\ell \in \texttt{refine}(\mathcal{T}_\ell) \cap \texttt{refine}(\mathcal{T}_\star)$ such that

$$\#(\mathcal{T}_\star \oplus \mathcal{T}_\ell) \le \#\mathcal{T}_\star + \#\mathcal{T}_\ell - \#\mathcal{T}_0. \tag{13}$$

Both properties (12) and (13) are proved for $d \ge 2$. The overlay estimate (13) was first proved for $d = 2$ in [Ste07] and then for $d \ge 2$ in [CKNS08]. For newest vertex bisection, the first proofs of (12) go back to [BDD04] for $d = 2$ and later [Ste08] for $d \ge 2$. Their proofs rely on a certain condition on the labelling of the reference edges in the initial mesh $\mathcal{T}_0$. For two-dimensional meshes, i.e. $d = 3$, the recent work [KPP13] proved that this particular condition is not necessary and can be dropped. The bisection algorithm for $d = 2$ is analysed in [AFF$^+$14] for $d = 2$. Moreover, uniform γ-shape regularity (7) resp. (8) holds for all meshes $\mathcal{T}_\ell \in \mathbb{T}$ with a constant $\gamma > 0$ which depends only on the initial mesh $\mathcal{T}_0$.

2.7. **Error estimator.** For error estimation, we employ the weighted residual error estimator η_ℓ which dates back to the seminal works [CS95, CS95, Car97] for 2D was extended to 3D in [CMS01]. Given a mesh $\mathcal{T}_\ell$ as well as a solution Φ_ℓ of (10), the local contributions read

$$\eta_\ell^2(T) := |T|^{1/(d-1)} \|\nabla(V\Phi_\ell - (1/2 + K)G_\ell)\|_{L^2(T)}^2 \quad \text{for all } T \in \mathcal{T}_\ell. \tag{14}$$

Here, ∇ denotes the surface gradient on Γ in the 3D case. For 2D, ∇ reduces to the arc-length derivative along Γ. For any subset $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$, we write

$$\eta_\ell^2(\mathcal{E}_\ell) := \sum_{T \in \mathcal{E}_\ell} \eta_\ell^2(T).$$

The global error estimator then reads

$$\eta_\ell := \eta_\ell(\mathcal{T}_\ell) = \|h_\ell^{1/2} \nabla(V\Phi_\ell - (1/2 + K)G_\ell)\|_{L^2(\Gamma)},$$

where $h_\ell \in L^\infty(\Gamma)$ with $h_\ell|_T := |T|^{1/(d-1)}$ for $d = 2, 3$ denotes the local mesh width function.

To control the error of the data approximation, we introduce the so called data oscillation term

$$\mathrm{osc}_\ell^2(T) := |T|^{1/(d-1)} \|(1 - \Pi_\ell)\nabla g\|_{L^2(T)}^2 \quad \text{for all } T \in \mathcal{T}_\ell, \tag{15}$$

where $\Pi_\ell : L^2(\Gamma) \to \mathcal{P}^p(\mathcal{T}_\ell)$ denotes the L^2-orthogonal projection onto $\mathcal{P}^p(\mathcal{T}_\ell)$. For $p = 0$, this is just the piecewise integral mean. To abbreviate notation, we write $\mathrm{osc}_\ell^2(\mathcal{E}_\ell) = \sum_{T \in \mathcal{E}_\ell} \mathrm{osc}_\ell^2(T)$ for all subsets $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$ and the global oscillation term is defined as $\mathrm{osc}_\ell = \mathrm{osc}_\ell(\mathcal{T}_\ell)$.

The error estimator and the oscillation term are combined to

$$\rho_\ell^2(T) := \eta_\ell^2(T) + \mathrm{osc}_\ell^2(T) \quad \text{for all } T \in \mathcal{T}_\ell.$$

Again, we write $\rho_\ell^2(\mathcal{E}_\ell) := \sum_{T \in \mathcal{E}_\ell} \rho_\ell^2(T)$ for each subset $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$ and $\rho_\ell^2 = \eta_\ell^2 + \mathrm{osc}_\ell^2$.

2.8. **The adaptive algorithm.** Now, we are able to state the adaptive algorithm which will be proved to converge even with optimal rate.

Algorithm 1. ***Input:*** *initial mesh* $\mathcal{T}_0$*, adaptivity parameter* $0 < \theta < 1$*. Set* $\ell := 0$

(i) *Compute approximate data* $G_\ell = J_\ell g$ *by use of the Scott-Zhang projection.*
(ii) *Compute solution* Φ_ℓ *of* (10).
(iii) *Compute error estimator* $\rho_\ell(T)$ *for all* $T \in \mathcal{T}_\ell$.
(iv) *Determine a set of marked elements* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *with minimal cardinality which satisfies combined Dörfler marking*

$$\theta \rho_\ell^2 \le \rho_\ell^2(\mathcal{M}_\ell). \tag{16}$$

(v) *Refine at least the marked elements to obtain* $\mathcal{T}_{\ell+1} = \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$.
(vi) *Increment* $\ell \mapsto \ell + 1$ *and goto* (i).

Output: *sequence of error estimators* $(\rho_\ell)_{\ell\in\mathbb{N}}$ *and sequence of Galerkin solutions* $(\Phi_\ell)_{\ell\in\mathbb{N}}$.

Remark. *To achieve the minimal cardinality of the set* $\mathcal{M}_\ell$ *in step* (iv)*, one usually sorts the quantities* $\rho_\ell^2(T_j)$ *in descending order* $\rho_\ell(T_1) \ge \rho_\ell(T_2) \ge \dots$ *and defines* $\mathcal{M}_\ell := \{T_1, \dots, T_j\}$ *for the minimal number* $j \in \mathbb{N}$ *such that* (16) *is satisfied. In general, the set* $\mathcal{M}_\ell$ *may not be unique.* □

3. Preliminaries

This section states some facts which are used throughout the work. First, we shall need certain inverse estimates.

Proposition 2. *Let* $\mathcal{T}_\ell \in \mathbb{T}$*. Then, there exists a constant* $C_{\rm inv} > 0$ *such that for all* $\psi \in L^2(\Gamma)$ *and* $v \in H^1(\Gamma)$*, it holds*

$$C_{\rm inv}^{-1} \| h_\ell^{1/2} \nabla V \psi \|_{L^2(\Gamma)} \le |||\psi||| + \| h_\ell^{1/2} \psi \|_{L^2(\Gamma)}, \tag{17}$$

$$C_{\rm inv}^{-1} \| h_\ell^{1/2} \nabla (1/2 + K) v \|_{L^2(\Gamma)} \le \| v \|_{H^{1/2}(\Gamma)} + \| h_\ell^{1/2} \nabla v \|_{L^2(\Gamma)}. \tag{18}$$

Particularly, for all $\Psi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$ *and* $W_\ell \in \mathcal{S}^{p+1}(\mathcal{T}_\ell)$*, it holds that*

$$\| h_\ell^{1/2} \Psi_\ell \|_{L^2(\Gamma)} + \| h_\ell^{1/2} \nabla V \Psi_\ell \|_{L^2(\Gamma)} \le C_{\rm inv} |||\Psi_\ell|||, \tag{19}$$

$$\| h_\ell^{1/2} \nabla W_\ell \|_{L^2(\Gamma)} + \| h_\ell^{1/2} \nabla (1/2 + K) W_\ell \|_{L^2(\Gamma)} \le C_{\rm inv} \| W_\ell \|_{H^{1/2}(\Gamma)}. \tag{20}$$

The constant $C_{\rm inv} > 0$ *depends only on* $\mathcal{T}_0$ *and* $p \ge 0$.

Proof. The estimates (17) and (18) are proved in [AFF+12, Theorem 1] resp. [Kar, Section 4.2]. The estimates (19) and (20) follow directly by employing the inverse estimates from [GHS05, Theorem 3.6] for $|||\cdot||| \simeq \|\cdot\|_{H^{-1/2}(\Gamma)} \gtrsim \| h_\ell^{1/2} (\cdot) \|_{L^2(\Gamma)}$, and from [CP07, Corollary 3.2] for $\|\cdot\|_{H^{1/2}(\Gamma)} \gtrsim \| h_\ell^{1/2} \nabla (\cdot) \|_{L^2(\Gamma)}$. □

The following lemma states some properties of the Scott-Zhang projection.

Lemma 3. *Let* $\mathcal{T}_\ell \in \mathbb{T}$ *and* $g \in H^1(\Gamma)$*. For* Γ *denoting a* $(d-1)$*-dimensional manifold, the Scott-Zhang projection* $J_\ell : L^2(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$ *satisfies*

$$\| h_\ell^{1/2} \nabla (1 - J_\ell) g \|_{L^2(\Gamma)} + \| (1 - J_\ell) g \|_{H^{1/2}(\Gamma)} \le C_{\rm osc} {\rm osc}_\ell. \tag{21}$$

Furthermore, for all refinements $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\ell)$*, there holds the discrete local upper bound*

$$\| (J_\star - J_\ell) g \|_{H^{1/2}(\Gamma)} \le C_{\rm osc} \, {\rm osc}_\ell(\omega_\ell^5(\mathcal{T}_\ell \setminus \mathcal{T}_\star)) \tag{22}$$

for all $\ell \in \mathbb{N}$*. The constant* $C_{\rm osc} > 0$ *depends only on* $\mathcal{T}_0$ *and* $p \ge 0$.

Proof. The estimate (21) is a combination of [KOP13, Theorem 3] and [AFK+13, Proposition 8], and the discrete upper bound (22) is proved in [AFK+13, Proposition 21]. □

Now, we take a closer look at the error estimator ρ_ℓ.

Proposition 4. *The error estimator ρ_ℓ satisfies*

(E1) *Stability on non-refined elements: There exists a constant $C_{\rm stab} > 0$ such that*

$$C_{\rm stab}^{-1}\Big|\rho_\star(\mathcal{T}_\star \cap \mathcal{T}_\ell) - \rho_\ell(\mathcal{T}_\star \cap \mathcal{T}_\ell)\Big| \le |||\Phi_\star - \Phi_\ell||| + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}$$

for all meshes $\mathcal{T}_\ell$, $\mathcal{T}_\star \in \mathbb{T}$.

(E2) *Reduction property on refined elements: There exist constants $0 < q_{\rm red} < 1$ and $C_{\rm red} > 0$ such that*

$$\rho_\star^2(\mathcal{T}_\star \setminus \mathcal{T}_\ell) \le q_{\rm red}\rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star) + C_{\rm red}\Big(|||\Phi_\star - \Phi_\ell|||^2 + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2\Big)$$

for all meshes $\mathcal{T}_\star \in$ `refine`$(\mathcal{T}_\ell)$ with $\mathcal{T}_\ell \in \mathbb{T}$.

(E3) *Reliability: There exists $C_{\rm rel} > 0$ such that*

$$|||\phi - \Phi_\ell||| \le C_{\rm rel}\rho_\ell$$

for all meshes $\mathcal{T}_\ell \in \mathbb{T}$.

(E4) *Discrete local reliability: There exists a constant $C_{\rm dlr} > 0$ such that*

$$|||\Phi_\star - \Phi_\ell||| \le C_{\rm dlr}\rho_\ell(\mathcal{R}_\ell), \tag{23}$$

for all meshes $\mathcal{T}_\ell \in \mathbb{T}$ and refinements $\mathcal{T}_\star \in$ `refine`$(\mathcal{T}_\ell)$ with corresponding Galerkin solution $\Phi_\star$. Here

$$\mathcal{R}_\ell := \omega_\ell^5(\mathcal{T}_\ell \setminus \mathcal{T}_\star)$$

denotes an extended set of refined elements.

Proof. We first consider (E1). The fact that $h_\ell|_{\cup(\mathcal{T}_\ell\cap\mathcal{T}_\star)} = h_\star|_{\cup(\mathcal{T}_\ell\cap\mathcal{T}_\star)}$ shows $\mathrm{osc}_\ell(\mathcal{T}_\ell\cap\mathcal{T}_\star) = \mathrm{osc}_\star(\mathcal{T}_\ell \cap \mathcal{T}_\star)$. Together with the triangle inequality, this yields

$$\begin{aligned}\Big|\rho_\star(\mathcal{T}_\star \cap \mathcal{T}_\ell) - \rho_\ell(\mathcal{T}_\star \cap \mathcal{T}_\ell)\Big| &\le \|h_\star^{1/2}\nabla V(\Phi_\star - \Phi_\ell)\|_{L^2(\Gamma)} + \|h_\star^{1/2}\nabla(1/2 + K)(G_\star - G_\ell)\|_{L^2(\Gamma)}\\ &\lesssim |||\Phi_\star - \Phi_\ell||| + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)},\end{aligned}$$

where we have used the inverse inequalities from Proposition 2 to get the final estimate.

To see (E2), we use Young's inequality with arbitrary $\delta > 0$ and estimate

$$\begin{aligned}\rho_\star^2(\mathcal{T}_\star \setminus \mathcal{T}_\ell) \le (1+\delta)\Big(&\|h_\star^{1/2}\nabla\Big(V\Phi_\ell - (1/2-K)G_\ell\Big)\|_{L^2(\cup\mathcal{T}_\ell\setminus\mathcal{T}_\star)}^2 + \|h_\star^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\cup\mathcal{T}_\ell\setminus\mathcal{T}_\star)}^2\Big)\\ &+ (1+\delta^{-1})C_{\rm inv}\Big(|||\Phi_\star - \Phi_\ell|||^2 + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2\Big),\end{aligned}$$

where we again applied the inverse estimates from Proposition 2. Exploiting $h_\star|_{\cup\mathcal{T}_\star\setminus\mathcal{T}_\ell} \le 2^{-1/(d-1)}h_\ell|_{\cup\mathcal{T}_\ell\setminus\mathcal{T}_\star}$ for $d = 2, 3$, we conclude the proof with $q_{\rm red} = (1+\delta)2^{-1/(d-1)}$ for sufficiently small $\delta > 0$.

Reliability of η_ℓ with unperturbed right-hand side is proved in [CS95, Theorem 2] for the 2D case and in [CMS01, Corollary 4.3] for the 3D case, i.e.

$$|||\phi_\ell - \Phi_\ell||| \lesssim \eta_\ell \quad \text{for all } \mathcal{T}_\ell \in \mathbb{T}. \tag{24}$$

To obtain (E3), we incorporate the data oscillations via

$$|||\phi - \Phi_\ell||| \le |||\phi - \phi_\ell||| + |||\phi_\ell - \Phi_\ell||| \lesssim \|g - G_\ell\|_{H^{1/2}(\Gamma)} + \eta_\ell,$$

where we used the definition of ϕ_ℓ in (9) as well as the stability of $V^{-1} : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ and $K : H^{1/2}(\Gamma) \to H^{1/2}(\Gamma)$. Now, estimate (21) implies

$$|||\phi - \Phi_\ell||| \lesssim \rho_\ell.$$

Finally, discrete local reliability (E4) of η_ℓ is proved in [FKMP13, Proposition 4.3] for $p = 0$. The proof holds verbatim for fixed $p \geq 0$. Therefore, we get

$$|||\Phi_\star^\ell - \Phi_\ell||| \lesssim \eta_\ell(\omega_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_\star)) \leq \eta_\ell(\mathcal{R}_\ell), \tag{25}$$

where $\Phi_\star^\ell \in \mathcal{P}^p(\mathcal{T}_\star)$ is the solution of

$$\langle\!\langle \Phi_\star^\ell \,,\, \Psi_\star \rangle\!\rangle = \langle (1/2 + K) G_\ell \,,\, \Psi_\star \rangle_{L^2(\Gamma)} \quad \text{for all } \Psi_\star \in \mathcal{P}^p(\mathcal{T}_\star).$$

We employ the definition of $\Phi_\star^\ell$ and the stability of Galerkin schemes to see

$$|||\Phi_\star^\ell - \Phi_\star||| \lesssim \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)} \lesssim \mathrm{osc}_\ell(\mathcal{R}_\ell),$$

where the last estimate follows from (22).

Finally, we combine the last estimate with (25) and prove

$$|||\Phi_\star - \Phi_\ell||| \leq |||\Phi_\star - \Phi_\star^\ell||| + |||\Phi_\star^\ell - \Phi_\ell||| \lesssim \rho_\ell(\mathcal{R}_\ell).$$

This concludes the proof. $\square$

Corollary 5. *The error estimator is quasi-monotone, i.e. for $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\ell)$ an arbitrary refinement of $\mathcal{T}_\ell \in \mathbb{T}$, there holds*

$$\rho_\star \leq C_{\mathrm{mon}} \rho_\ell$$

with some constant $C_{\mathrm{mon}} > 0$ which depends only on $\mathcal{T}_0$ and $p \geq 0$.

Proof. The triangle inequality and $h_\star \leq h_\ell$ yield

$$\begin{aligned} \rho_\star &\leq \rho_\ell + \|h_\star^{1/2} \nabla V(\Phi_\star - \Phi_\ell)\|_{L^2(\Gamma)} + \|h_\star^{1/2} \nabla (1/2 + K)(G_\star - G_\ell)\|_{L^2(\Gamma)} \\ &\lesssim \rho_\ell + |||\Phi_\star - \Phi_\ell||| + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}, \end{aligned}$$

where we applied the inverse estimates (19) and (20) from Proposition 2 to obtain the last estimate. Now, with discrete local reliability (E4) and (22), we get

$$\rho_\star \lesssim \rho_\ell + \rho_\ell(\mathcal{R}_\ell) + \mathrm{osc}_\ell(\mathcal{R}_\ell) \lesssim \rho_\ell.$$

This concludes the proof. $\square$

4. Convergence of Algorithm 1

This section analyzes the convergence of Algorithm 1. Although the results of this section may be interesting on their own (since convergence of the adaptive algorithm is not clear a priori), they also provide the crucial foundation for the optimality proof of Section 5.

4.1. **Estimator reduction.** The following estimator reduction result is a very general concept and applies to numerous situations in the context of *a posteriori* error estimation in BEM and FEM, see e.g. [AFLP12]

Proposition 6. *Let $\mathcal{T}_\star := \texttt{refine}(\mathcal{T}_\ell)$ denote an arbitrary refinement of $\mathcal{T}_\ell$ such that the set of refined elements $\mathcal{T}_\ell \backslash \mathcal{T}_\star$ satisfies combined Dörfler marking* (16) *for some $0 < \theta < 1$. Then, there exist constants $0 < q_{\mathrm{est}} < 1$ and $C_{\mathrm{est}} > 0$ such that ρ_ℓ satisfies the perturbed contraction estimate*

$$\rho_\star^2 \leq q_{\mathrm{est}} \rho_\ell^2 + C_{\mathrm{est}} \Big(|||\Phi_\star - \Phi_\ell|||^2 + \|G_\star - G_\ell\|^2_{H^{1/2}(\Gamma)} \Big). \tag{26}$$

The constants $q_{\mathrm{est}}, C_{\mathrm{est}} > 0$ depend only on θ, $\mathcal{T}_0$, and $p \geq 0$.

Proof. Recall Young's inequality $(a+b)^2 \le (1+\delta)a^2 + (1+\delta^{-1})b^2$ for all $\delta > 0$ and $a, b \in \mathbb{R}$. We exploit stability (E1) and reduction (E2) to see

$$\begin{aligned}\rho_\star^2 &= \rho_\star^2(\mathcal{T}_\ell \cap \mathcal{T}_\star) + \rho_\star^2(\mathcal{T}_\star \setminus \mathcal{T}_\ell)\\ &\le (1+\delta)\rho_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_\star) + q_{\mathrm{red}}\rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star)\\ &\qquad + \Big((1+\delta^{-1})2C_{\mathrm{stab}}^2 + C_{\mathrm{red}}\Big)\Big(|||\Phi_\star - \Phi_\ell|||^2 + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2\Big).\end{aligned}$$

Now, Dörfler marking (16) for $\mathcal{T}_\ell \setminus \mathcal{T}_\star$ implies

$$\begin{aligned}(1+\delta)\rho_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_\star) + q_{\mathrm{red}}\rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star) &\le (1+\delta)\rho_\ell^2 - (1+\delta-q_{\mathrm{red}})\rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star)\\ &\le \Big(1+\delta-\theta(1+\delta-q_{\mathrm{red}})\Big)\rho_\ell^2.\end{aligned}$$

For sufficiently small $\delta > 0$, the combination of the last two estimates proves (26) with $q_{\mathrm{est}} = 1+\delta-\theta(1+\delta-q_{\mathrm{red}})$ and $C_{\mathrm{est}} = (1+\delta^{-1})2C_{\mathrm{stab}}^2 + C_{\mathrm{red}}$. □

4.2. **Contraction of quasi-error.** In this section, we make explicit use of the fact that $G_\ell = J_\ell g$ is obtained via the Scott-Zhang projection. As a theoretical tool, we introduce an equivalent mesh width function that takes care of the fact that in many of the local estimates below the patches of the elements come into play.

Lemma 7. *Let $k \in \mathbb{N}$ be arbitrary and let $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ denote the sequence of meshes generated by Algorithm 1. Then, there exists a modified mesh width function $\widetilde{h}_\ell$ such that*

$$\widetilde{h}_\ell \le h_\ell \le C_2\widetilde{h}_\ell \quad \textit{for all } \ell \in \mathbb{N}, \tag{27}$$

which is monotone and additionally provides a contraction on the k-patch of each refined element, i.e. for all $\ell \ge 1$ it holds

$$\widetilde{h}_\ell \le \widetilde{h}_{\ell-1} \qquad \textit{pointwise almost everywhere in } \Omega, \tag{28a}$$

$$\widetilde{h}_\ell|_T \le q\widetilde{h}_{\ell-1}|_T \quad \textit{for all } T \in \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}). \tag{28b}$$

The constants $0 < q < 1$ and $C_2 > 0$ depend only on $\mathcal{T}_0$ and on $k \in \mathbb{N}$.

Proof. First, we observe that due to γ-shape regularity, the number of elements in the k-patch is bounded, i.e.

$$\#\omega_\ell^k(T) \le C_3 \quad \text{for all } T \in \mathcal{T}_\ell. \tag{29}$$

The constant $C_3 > 0$ depends only on the γ-shape regularity and on $k \in \mathbb{N}$.

Recall the level function $\mathrm{level}(\cdot) : \bigcup_{\ell\in\mathbb{N}} \mathcal{T}_\ell \to \mathbb{N}$, which counts the number of bisections needed to generate an element $T \in \mathcal{T}_\ell$ from its ancestor $T \subset T_0 \in \mathcal{T}_0$. By definition, there holds $\mathrm{level}(T_0) = 0$ for all $T_0 \in \mathcal{T}_0$ and for $T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}$, with father $T \subset T' \in \mathcal{T}_{\ell-1}$ there holds $\mathrm{level}(T) > \mathrm{level}(T')$. According to [KPP13, Lemma 18] the level difference of two neighbouring elements $T, T' \in \mathcal{T}_\ell$ is less than or equal to two for $d = 3$ and bounded for $d = 2$ (see [AFF+14, Section 3]). Hence, the level difference of two elements $T, T' \in \mathcal{T}_\ell$ which lie within one k-patch is also bounded, i.e.

$$|\mathrm{level}(T) - \mathrm{level}(T')| \le C_4 \quad \text{for } T' \in \omega_\ell^k(T), \tag{30}$$

for some constant $C_4 > 0$. We define a modified level-function inductively as follows:

$$\widetilde{\mathrm{level}}_0(T) := 0 \quad \text{for all } T \in \mathcal{T}_0$$

as well as for all $\ell > 0$ and all $T \in \mathcal{T}_\ell$

$$\widetilde{\text{level}}_\ell(T) := \begin{cases} \text{level}(T) & T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1} \\ \widetilde{\text{level}}_{\ell-1}(T) + 1/(2C_4C_3+1) & T \in \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}) \setminus (\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}) \\ \widetilde{\text{level}}_{\ell-1}(T) & \text{else.} \end{cases}$$

Note that for $T \in \mathcal{T}_\ell$, the modified level $\widetilde{\text{level}}_\ell(T)$ depends on the chosen sequence of meshes $\mathcal{T}_0, \ldots, \mathcal{T}_\ell$, in contrast to level$(T)$.

Below, we define the modified mesh width function $\widetilde{h}_\ell$ via the modified level function $\widetilde{\text{level}}_\ell(\cdot)$. To obtain the equivalence (27), we first show that the level functions are equivalent, i.e.

$$\text{level}(T) \le \widetilde{\text{level}}_\ell(T) \le \text{level}(T) + 2C_3C_4/(2C_3C_4+1) \quad \text{for all } \ell \in \mathbb{N} \text{ and all } T \in \mathcal{T}_\ell. \tag{31}$$

The lower bound follows from the definition of $\widetilde{\text{level}}_\ell(\cdot)$ by induction on ℓ. For the upper bound we argue by contradiction and assume the existence of an element $T \in \mathcal{T}_\ell$ with $\widetilde{\text{level}}_\ell(T) > \text{level}(T) + 2C_3C_4/(2C_3C_4+1)$. With the convention $\mathcal{T}_{-1} := \emptyset$, let $\ell_0 \le \ell$ be such that $T \in \mathcal{T}_{\ell_0} \setminus \mathcal{T}_{\ell_0-1}$. By definition of the modified level function, this implies $\widetilde{\text{level}}_{\ell_0}(T) = \text{level}(T)$ (obviously, this also holds for $\ell_0 = 0$). Again, by definition of the modified level function, we know that the case

$$T \in \omega_{\ell_j}^k(\mathcal{T}_{\ell_j} \setminus \mathcal{T}_{\ell_j-1}) \setminus (\mathcal{T}_{\ell_j} \setminus \mathcal{T}_{\ell_j-1}) \quad \text{for } \ell_0 < \ell_j \le \ell$$

must have occurred at least $2C_3C_4+1$ times, since otherwise $\widetilde{\text{level}}_\ell(T) \le \text{level}(T) + 2C_3C_4/(2C_3C_4+1)$. Put differently,

$$\omega_{\ell_j}^k(T) \cap (\mathcal{T}_{\ell_j} \setminus \mathcal{T}_{\ell_j-1}) \neq \emptyset \quad \text{for all these (at least } 2C_3C_4+1\text{) many indices } \ell_0 < \ell_j \le \ell.$$

Since $\omega_{\ell_0}^k(T)$ contains at most C_3 elements, there is at least one element $T' \in \omega_{\ell_0}^k(T)$ with

$$T' \cap (\mathcal{T}_{\ell_j} \setminus \mathcal{T}_{\ell_j-1}) \neq \emptyset \quad \text{for at least } N_{\max} \ge \frac{2C_3C_4+1}{C_3} \text{ many indices } \ell_0 < \ell_j \le \ell.$$

Hence, there exists an element $T'' \in \omega_\ell^k(T)$ with $\text{level}(T'') \ge N_{\max} + \text{level}(T')$. Combining this with (30), we obtain

$$\begin{aligned} C_4 < C_4 + 1/C_3 \le N_{\max} - C_4 &\le |\text{level}(T'') - \text{level}(T')| - |\text{level}(T') - \text{level}(T)| \\ &\le |\text{level}(T'') - \text{level}(T)| \le C_4. \end{aligned}$$

This contradiction proves (31). Finally, with $T_0(T) \in \mathcal{T}_0$ denoting the unique father of $T \subseteq T_0(T)$ in the initial mesh $\mathcal{T}_0$, we may define

$$\widetilde{h}_\ell|_T := \Big(2^{-\widetilde{\text{level}}_\ell(T)}|T_0(T)|\Big)^{1/(d-1)} \quad \text{for all } T \in \mathcal{T}_\ell \text{ and all } \ell \in \mathbb{N}.$$

The property (27) follows immediately from the equivalence (31) as well as the fact that newest vertex bisection for $d \ge 3$ and simple bisection for $d = 2$ guarantee

$$|T|^{1/(d-1)} = \Big(2^{-\text{level}(T)}|T_0(T)|\Big)^{1/(d-1)}.$$

For (28b), we consider an element $T \in \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1})$ with father $T \subseteq T' \in \mathcal{T}_{\ell-1}$. For $T \subsetneqq T'$, i.e. $T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}$, there holds

$$\widetilde{\text{level}}_\ell(T) \ge \text{level}(T) \ge \text{level}(T') + 1 \ge \widetilde{\text{level}}_{\ell-1}(T') + 1/(2C_3C_4+1), \tag{32}$$

and therefore

$$\widetilde{h}_\ell|_T \le \widetilde{h}_{\ell-1}|_{T'} \Big(2^{-1/(2C_3C_4+1)}\Big)^{1/(d-1)}, \tag{33}$$

i.e. contraction in (28b) for $T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}$ with $q = \Big(2^{-1/(2C_3C_4+1)}\Big)^{1/(d-1)}$. For $T = T'$, i.e. $T \in \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}) \setminus (\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1})$, we see by definition of the modified level function that there also holds

$$\widetilde{\text{level}}_\ell(T) \ge \widetilde{\text{level}}_{\ell-1}(T') + 1/(2C_3C_4+1)$$

which implies (33) and hence (28b) for $T \in \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1}) \setminus (\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1})$. For $T \notin \omega_\ell^k(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell-1})$, there holds $\widetilde{\text{level}}_{\ell-1}(T) = \widetilde{\text{level}}_\ell(T)$ and hence $\widetilde{h}_{\ell-1}|_T = \widetilde{h}_\ell|_T$. This yields (28a) and concludes the proof. □

The next lemma provides a decisive improvement of our analysis compared to [AFK+13]. Instead of treating all data approximation methods with the techniques of Section 6 below, we use the locality of the Scott-Zhang projection together with the augmented contraction area of the modified mesh width function $\widetilde{h}_\ell$, to prove a certain orthogonality relation of the approximate Dirichlet data. This allows us to use the standard Dörfler marking (16) in the adaptive algorithm (instead of the separate Dörfler marking (62) as in Section 6) and is exploited in the proof of the contraction result of Theorem 9.

Lemma 8. *Let $\mathcal{T}_\star \in$ `refine`$(\mathcal{T}_\ell)$ denote a refinement of $\mathcal{T}_\ell$. Then, there holds*

$$C_5^{-1}\|(J_\star - J_\ell)g\|^2_{H^{1/2}(\Gamma)} \le \|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|^2_{L^2(\Gamma)} - \|\widetilde{h}_\star^{1/2}(1-\Pi_\star)\nabla g\|^2_{L^2(\Gamma)} \tag{34}$$

for all $\ell \in \mathbb{N}$. Here, $\widetilde{h}_\ell$ denotes the modified mesh width function from Lemma 7 for $k = 5$. The constant $C_5 > 0$ depends only on $\mathcal{T}_0$ and $p \ge 0$.

Proof. We employ Lemma 7 for $k = 5$ and obtain in combination with (22)

$$\|(J_\star - J_\ell)g\|_{H^{1/2}(\Gamma)} \lesssim \|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\omega_\ell^5(\mathcal{T}_\ell\setminus\mathcal{T}_\star))} \quad \text{for all } \ell \in \mathbb{N}. \tag{35}$$

With monotonicity and the contraction property (28) of $\widetilde{h}_\ell$, we see

$$\widetilde{h}_\ell - \widetilde{h}_\star \ge (1-q)\widetilde{h}_\ell\, \chi_{\omega_\ell^5(\mathcal{T}_\ell\setminus\mathcal{T}_\star)} \quad \text{for all } \ell \in \mathbb{N},$$

where $\chi_{\omega_\ell^5(\mathcal{T}_\ell\setminus\mathcal{T}_\star)}$ denotes the characteristic function with respect to the set $\omega_\ell^5(\mathcal{T}_\ell \setminus \mathcal{T}_\star)$. Together with (35), we therefore conclude

$$\begin{aligned}
\|(J_\star - J_\ell)g\|^2_{H^{1/2}(\Gamma)} &\lesssim \int_\Gamma (\widetilde{h}_\ell - \widetilde{h}_\star)\Big|(1-\Pi_\ell)\nabla g\Big|^2\,dx \\
&= \|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|^2_{L^2(\Gamma)} - \|\widetilde{h}_\star^{1/2}(1-\Pi_\ell)\nabla g\|^2_{L^2(\Gamma)} \\
&\le \|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|^2_{L^2(\Gamma)} - \|\widetilde{h}_\star^{1/2}(1-\Pi_\star)\nabla g\|^2_{L^2(\Gamma)}
\end{aligned}$$

according to the elementwise best approximation property of $\Pi_\star$. □

Theorem 9. *Let $\mathcal{T}_\star =$ `refine`$(\mathcal{T}_\ell)$ denote a refinement of $\mathcal{T}_\ell \in \mathbb{T}$ such that $\mathcal{T}_\ell \setminus \mathcal{T}_\star$ satisfies the combined Dörfler marking* (16) *(e.g. in Algorithm 1 with $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$). Then, there exist constants $0 < \alpha < 1$, $\beta > 0$ and $0 < \kappa < 1$ such that the quasi-error*

$$\Delta_\ell := |\!|\!|\phi_\ell - \Phi_\ell|\!|\!|^2 + \alpha\rho_\ell^2 + \beta\|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|^2_{L^2(\Gamma)} \tag{36}$$

satisfies the contraction property

$$\Delta_\star \le \kappa\Delta_\ell.$$

Moreover, it holds $\alpha\rho_\ell^2 \le \Delta_\ell \le (C_7^2 + \alpha + \beta)\rho_\ell^2$ for all $\ell \in \mathbb{N}$. The constants $\alpha, \beta, \kappa > 0$ depend only on the use of newest vertex bisection, $\mathcal{T}_0$, q_{est}, and $p \ge 0$.

Proof. First, we observe that the data oscillations with the modified mesh width function $\widetilde{h}_\ell$ are still dominated by the error estimator, i.e.

$$\|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\Gamma)} \le \mathrm{osc}_\ell \le \rho_\ell \tag{37}$$

for all $\ell \in \mathbb{N}$.

Second, we recall that stability of the problem allows us to control the influence of the data approximation in the sense of

$$|||\phi_\star - \phi_\ell|||^2 \le C_6 \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2. \tag{38}$$

The constant $C_6 > 0$ depends only on the norms of V^{-1} and K. Furthermore, one has reliability

$$|||\phi_\ell - \Phi_\ell||| \le C_7 \eta_\ell \le C_7 \rho_\ell \tag{39}$$

with $C_7 > 0$ independent of $\ell \in \mathbb{N}$, cf. (24).

Third, by definition of $\phi_\star$, see (9), there holds orthogonality

$$\langle\!\langle \phi_\star - \Phi_\star \,,\, \Phi_\star - \Phi_\ell \rangle\!\rangle = 0.$$

With (38), we infer

$$\begin{aligned} |||\phi_\star - \Phi_\star|||^2 + |||\Phi_\star - \Phi_\ell|||^2 &= |||\phi_\star - \Phi_\ell|||^2 \\ &\le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + (1+\delta^{-1})|||\phi_\star - \phi_\ell|||^2 \\ &\le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + (1+\delta^{-1})C_6\|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2 \end{aligned} \tag{40}$$

for all $\delta > 0$. Next, we use the estimator reduction (26) to obtain

$$\begin{aligned} \Delta_\star &\le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + (1+\delta^{-1})C_6\|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2 - |||\Phi_\star - \Phi_\ell|||^2 \\ &\qquad + \alpha q_{\rm est}\rho_\ell^2 + \alpha C_{\rm est}\Big(|||\Phi_\star - \Phi_\ell|||^2 + \|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2\Big) + \beta\|\widetilde{h}_\star^{1/2}(1-\Pi_\star)\nabla g\|_{L^2(\Gamma)}^2 \\ &\le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + \Big((1+\delta^{-1})C_6 + \alpha C_{\rm est}\Big)\|G_\star - G_\ell\|_{H^{1/2}(\Gamma)}^2 \\ &\qquad + \alpha q_{\rm est}\rho_\ell^2 + (\alpha C_{\rm est} - 1)|||\Phi_\star - \Phi_\ell|||^2 + \beta\|\widetilde{h}_\star^{1/2}(1-\Pi_\star)\nabla g\|_{L^2(\Gamma)}^2. \end{aligned}$$

With Lemma 8, this implies for $\beta = \Big((1+\delta^{-1})C_6 + \alpha C_{\rm est}\Big)C_5$

$$\begin{aligned} \Delta_\star &\le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + (\alpha C_{\rm est} - 1)|||\Phi_\star - \Phi_\ell|||^2 \\ &\qquad + \alpha q_{\rm est}\rho_\ell^2 + \beta\|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\Gamma)}^2. \end{aligned}$$

Now, choose $\alpha C_{\rm est} < 1$ to simplify the estimate above to

$$\Delta_\star \le (1+\delta)|||\phi_\ell - \Phi_\ell|||^2 + \alpha q_{\rm est}\rho_\ell^2 + \beta\|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\Gamma)}^2.$$

We introduce some parameter $\varepsilon > 0$ and use the bounds (37) and (39) to obtain

$$\begin{aligned} \Delta_\star &\le (1+\delta - C_7^{-1}\varepsilon)|||\phi_\ell - \Phi_\ell|||^2 \\ &\qquad + (\alpha q_{\rm est} + 2\varepsilon)\rho_\ell^2 + (\beta - \varepsilon)\|\widetilde{h}_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\Gamma)}^2 \\ &\le \kappa\Delta_\ell \end{aligned}$$

with

$$\kappa := \max\{1+\delta - C_7^{-1}\varepsilon, (\alpha q_{\rm est} + 2\varepsilon)/\alpha, (\beta-\varepsilon)/\beta\}.$$

To ensure $0 < \kappa < 1$, choose $\varepsilon > 0$ such that $\alpha q_{\rm est} + 2\varepsilon < \alpha$ and fix $\delta = C_7^{-1}\varepsilon/2$. The equivalence $\rho_\ell^2 \simeq \Delta_\ell$ follows immediately from (37) and (39). This proves the assertion. □

5. Quasi-Optimality of Algorithm 1

The quasi-optimality proof roughly consists of two parts: First, we prove in Proposition 10 that Dörfler marking (16) is not only sufficient for the contraction property stated in Theorem 9, but in some sense even necessary. Second, Theorem 11 combines the foregoing results with an estimate on the cardinality of the set of marked elements $\mathcal{M}_\ell$ and derives the optimality result.

5.1. **Optimality of marking criterion.** The following proposition can be seen as the converse of Theorem 9.

Proposition 10. *Let $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\ell)$ denote a refinement of $\mathcal{T}_\ell$. Let $\Phi_\ell, \Phi_\star$ denote the solutions of* (10) *corresponding to $G_\ell = J_\ell g, G_\star = J_\star g$. Let $0 < \kappa_\star < 1$ and suppose that the corresponding error estimators satisfy*

$$\rho_\star^2 \leq \kappa_\star \rho_\ell^2. \tag{41}$$

Then, there exists $0 < \theta_\star < 1$ such that ρ_ℓ satisfies the combined Dörfler marking (16)

$$\theta \rho_\ell^2 \leq \rho_\ell^2(\mathcal{R}_\ell) \tag{42}$$

for all $0 < \theta \leq \theta_\star$ with the set $\mathcal{R}_\ell$ from (E4).

Proof. First, we employ stability (E1) as well as the discrete local reliability of the estimator (E4) and of the Scott-Zhang projection (22) to obtain, for arbitrary $\delta > 0$,

$$\begin{aligned}
\rho_\ell^2 &= \rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star) + \rho_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_\star) \\
&\leq \rho_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\star) + (1+\delta)\rho_\star^2 + (1+\delta^{-1}) 2 C_{\rm stab}^2 \Big(|||\Phi_\star - \Phi_\ell|||^2 + \|(J_\star - J_\ell) g\|_{H^{1/2}(\Gamma)}^2 \Big) \\
&\leq (1 + (1+\delta^{-1}) 2 C_{\rm stab}^2 (C_{\rm dlr}^2 + C_{\rm osc}^2)) \rho_\ell^2(\mathcal{R}_\ell) + (1+\delta)\kappa_\star \rho_\ell^2.
\end{aligned}$$

We choose $\delta > 0$ such that $(1+\delta)\kappa_\star < 1$. Rearranging the terms, we thus see

$$\theta_\star \rho_\ell^2 \leq \rho_\ell^2(\mathcal{R}_\ell)$$

with

$$\theta_\star := (1 - (1+\delta)\kappa_\star) / (1 + (1+\delta^{-1}) 2 C_{\rm stab}^2 (C_{\rm dlr}^2 + C_{\rm osc}^2)) \in (0,1).$$

This concludes the proof of (42). □

Remark. *Analyzing the proof of Theorem 11, we see that we may choose $\kappa_\star$ arbitrarily small. This implies that for any $0 < \theta < 1/(1 + 2C_{\rm stab}^2(C_{\rm dlr}^2 + C_{\rm osc}^2)) := \widetilde{\theta}_\star$ one may choose $\delta > 0$ sufficiently large and $\kappa_\star > 0$ sufficiently small such that $\theta < \theta_\star$.* □

5.2. **Quasi-optimal convergence rates.** To conclude the quasi-optimality proof in this section, we introduce the set of all meshes which have at most N elements more than the initial mesh $\mathcal{T}_0$, i.e.

$$\mathbb{T}_N := \Big\{ \mathcal{T}_\star \in \mathbb{T} \,:\, \#\mathcal{T}_\star - \#\mathcal{T}_0 \leq N \Big\}$$

as well as the approximation class $\mathbb{A}_s^{\rho,J}$ characterized by

$$(\phi, g) \in \mathbb{A}_s^{\rho,J} \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad \|(\phi,g)\|_{\mathbb{A}_s^{\rho,J}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star \in \mathbb{T}_N} \rho_\star N^s < \infty. \tag{43}$$

Here, ϕ and g are supposed to be the solution and data of our model problem (1).

Note that the method of data approximation $G_\ell \approx g$ influences the error estimator ρ_ℓ. Hence, the definition of $\mathbb{A}_s^{\rho,J}$ depends on the method of data approximation as well (which is hence indicated by the superscript ρ, J). Now, the quasi-optimality is formulated in the following theorem, which states that each possible algebraic convergence rate for the error estimator will in fact be achieved by the ABEM algorithm.

Theorem 11. *For sufficiently small parameter* $0 < \theta < 1$*, Algorithm 1 is optimal in the sense of*

$$(\phi, g) \in \mathbb{A}_s^{\rho,J} \quad \Longleftrightarrow \quad \rho_\ell \leq C_{\mathrm{opt}}(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \textit{for all } \ell \in \mathbb{N}. \tag{44}$$

The constant $C_{\mathrm{opt}} > 0$ *depends only on* $\|(\phi, g)\|_{\mathbb{A}_s^{\rho,J}}$*,* $0 < \theta < 1$*, and on the constant* $0 < \kappa < 1$ *from Theorem 9.*

Proof. First, Theorem 9 states that Δ_ℓ is a contractive sequence, i.e. with $0 < \kappa < 1$ given there

$$\Delta_{\ell+1} \leq \kappa \Delta_\ell \quad \text{for all } \ell \in \mathbb{N}. \tag{45}$$

Now, let $\lambda > 0$ be a free parameter which is fixed later on. According to the definition of the approximation class $\mathbb{A}_s^{\rho,J}$ in (43), we find for sufficiently small $\varepsilon^2 := \lambda \rho_\ell^2 > 0$ some triangulation $\mathcal{T}_\varepsilon \in \mathbb{T}$ such that

$$\#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s} \quad \text{and} \quad \rho_\varepsilon \leq \varepsilon, \tag{46}$$

where the hidden constant depends only on $\|(\phi, g)\|_{\mathbb{A}_s^{\rho,J}}$. We now consider the coarsest common refinement $\mathcal{T}_\star := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$ and first note that

$$\#\mathcal{T}_\star - \#\mathcal{T}_\ell \leq (\#\mathcal{T}_\varepsilon + \#\mathcal{T}_\ell - \#\mathcal{T}_0) - \#\mathcal{T}_\ell = \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \tag{47}$$

according to (13). Due to $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\varepsilon)$, we may apply Corollary 5 to get

$$\rho_\star^2 \lesssim \rho_\varepsilon^2 \leq \lambda \rho_\ell^2. \tag{48}$$

Choosing $\lambda > 0$ sufficiently small but fixed from now on, we enforce $\rho_\star^2 \leq \kappa_\star \rho_\ell^2$ for some $\kappa_\star \in (0, 1)$ and $\varepsilon \simeq \rho_\ell$. Next, we employ Proposition 10 to obtain that $\mathcal{R}_\ell := \omega_\ell^5(\mathcal{T}_\ell \backslash \mathcal{T}_\star) \subseteq \mathcal{T}_\ell$ satisfies the combined Dörfler marking (16). Recall that $\mathcal{M}_\ell$ is chosen in Step (iv) of Algorithm 1 to be a set with minimal cardinality. Together with the fact that each refinement splits the element into at least two sons, we get

$$\#\mathcal{M}_\ell \leq \#\mathcal{R}_\ell \simeq \#(\mathcal{T}_\ell \setminus \mathcal{T}_\star) \leq \#\mathcal{T}_\star - \#\mathcal{T}_\ell \leq \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s} \simeq \rho_\ell^{-1/s}. \tag{49}$$

Now, with optimality of the mesh closure (12) and contraction (45), we conclude

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \lesssim \sum_{j=0}^{\ell-1} \#\mathcal{M}_\ell \lesssim \sum_{j=0}^{\ell-1} \rho_j^{-1/s} \simeq \sum_{j=0}^{\ell-1} \Delta_j^{-1/(2s)} \leq \Delta_\ell^{-1/(2s)} \sum_{j=0}^{\ell-1} \kappa^{1/(2s)} \lesssim \rho_\ell^{-1/s}$$

by convergence of the geometric series. Finally, this yields

$$\rho_\ell \lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s}$$

for all $\ell \in \mathbb{N}$ and concludes the proof. ☐

Remark. *Reliability* (E3) *of the error estimator shows the equivalence*

$$\varrho_\ell^2 \simeq \rho_\ell^2 + |||\phi - \Phi_\ell|||^2 \quad \textit{for all } \ell \in \mathbb{N}.$$

Therefore, the approximation class $\mathbb{A}_s^{\rho,J}$ *can be equivalently defined in terms of the total error, as is done in* [CKNS08, FKMP13]. ☐

5.3. **Characterization of approximation class.** We aim to decouple the influence of the data approximation from the problem (1) in the approximation class $\mathbb{A}_s^{\rho,J}$. To that end, we introduce

$$\phi \in \mathbb{A}_s^\mu \quad \overset{\text{def}}{\Longleftrightarrow} \quad \|\phi\|_{\mathbb{A}_s^\mu} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star\in\mathbb{T}_N} \mu_\star N^s < \infty, \tag{50}$$

where μ_ℓ denotes the unperturbed estimator studied in [FKMP13, AFF$^+$14] for adaptive BEM, i.e.

$$\mu_\ell := \|h_\ell^{1/2}\nabla(V\Phi_\ell^\mu - (1/2+K)g)\|_{L^2(\Gamma)}$$

with $\Phi_\ell^\mu \in \mathcal{P}^p(\mathcal{T}_\ell)$ being the solution of the problem with exact right-hand side

$$\langle\!\langle \Phi_\ell^\mu\,,\,\Psi_\ell\rangle\!\rangle = \langle (1/2+K)g\,,\,\Psi_\ell\rangle_{L^2(\Gamma)} \tag{51}$$

for all $\Psi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$. Note that the definition of $\mathbb{A}_s^\mu$ does not incorporate the data approximation $G_\ell \approx g$. Finally, we introduce the approximation class

$$g \in \mathbb{A}_s^{\rm osc} \quad \overset{\text{def}}{\Longleftrightarrow} \quad \|g\|_{\mathbb{A}_s^{\rm osc}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star\in\mathbb{T}_N} {\rm osc}_\star N^s < \infty. \tag{52}$$

The relations between the approximation classes introduced are discussed in the following theorem.

Theorem 12. *There holds the implications*

$$\phi \in \mathbb{A}_{s_1}^\mu,\ g \in \mathbb{A}_{s_2}^{\rm osc} \quad \Longrightarrow \quad (\phi,g) \in \mathbb{A}_{\min\{s_1,s_2\}}^{\rho,J}, \tag{53}$$

$$(\phi,g) \in \mathbb{A}_s^{\rho,J} \quad \Longrightarrow \quad \phi \in \mathbb{A}_s^\mu,\ g \in \mathbb{A}_s^{\rm osc} \tag{54}$$

for all $s_1,\ s_2 > 0$.

Proof. First, we see by use of (17) and (18)

$$\begin{aligned}\rho_\ell^2 &\lesssim \mu_\ell^2 + \|h_\ell^{1/2}\nabla(1/2+K)(g-G_\ell)\|_{L^2(\Gamma)}^2 + \|h_\ell^{1/2}\nabla V(\Phi_\ell^\mu - \Phi_\ell)\|_{L^2(\Gamma)}^2 + {\rm osc}_\ell^2\\ &\lesssim \mu_\ell^2 + \|g-G_\ell\|_{H^{1/2}(\Gamma)}^2 + \|h_\ell^{1/2}\nabla(g-G_\ell)\|_{L^2(\Gamma)}^2 + |\!|\!|\Phi_\ell^\mu - \Phi_\ell|\!|\!|^2 + {\rm osc}_\ell^2\\ &\lesssim \mu_\ell^2 + {\rm osc}_\ell^2\end{aligned}$$

for all $\mathcal{T}_\ell \in \mathbb{T}$. Here, we used the stability of Galerkin schemes as well as (21) to obtain the last estimate. Analogously, one proves the converse estimate to obtain

$$\rho_\ell^2 \simeq \mu_\ell^2 + {\rm osc}_\ell^2. \tag{55}$$

Now, assume $(\phi,g)\in\mathbb{A}_s^{\rho,J}$. For all $N\in\mathbb{N}$, we obtain a mesh $\mathcal{T}_\star \in \mathbb{T}_N$ with

$$N^s{\rm osc}_\star + N^s\mu_\star \lesssim N^s\rho_\star \lesssim \|(\phi,g)\|_{\mathbb{A}_s^{\rho,J}},$$

where we used (55). This implies immediately

$$\|\phi\|_{\mathbb{A}_s^\mu} + \|g\|_{\mathbb{A}_s^{\rm osc}} \lesssim \|(\phi,g)\|_{\mathbb{A}_s^{\rho,J}} < \infty,$$

which proves $\phi\in\mathbb{A}_s^\mu$ and $g\in\mathbb{A}_s^{\rm osc}$ and therefore (54). To see (53), assume $\phi\in\mathbb{A}_{s_1}^\mu$ and $g\in\mathbb{A}_{s_2}^{\rm osc}$ and define $s:=\min\{s_1,s_2\}$.

For all $N\in\mathbb{N}$, the definition of $\mathbb{A}_{s_1}^\mu$ and $\mathbb{A}_{s_2}^{\rm osc}$ yields meshes $\mathcal{T}_{\star_\mu}$ and $\mathcal{T}_{\star_{\rm osc}}$ with $\mathcal{T}_{\star_\mu}, \mathcal{T}_{\star_{\rm osc}} \in \mathbb{T}_{N/2}$ and

$$\mu_{\star_\mu}(N/2)^{s_1} \lesssim \|\phi\|_{\mathbb{A}_{s_1}^\mu} \quad\text{ and }\quad {\rm osc}_{\star_{\rm osc}}(N/2)^{s_2} \lesssim \|g\|_{\mathbb{A}_{s_2}^{\rm osc}}.$$

Now, we consider the coarsest common refinement $\mathcal{T}_\star := \mathcal{T}_{\star_\mu} \oplus \mathcal{T}_{\star_{\rm osc}}$. Note that there holds $\#\mathcal{T}_\star - \#\mathcal{T}_0 \le \#\mathcal{T}_{\star_\mu} + \#\mathcal{T}_{\star_{\rm osc}} - 2\#\mathcal{T}_0 \le N$ due to (13). Moreover, we see by use of quasi-monotonicity of the error estimator and the fact that $\mathrm{osc}_\star \le \mathrm{osc}_{\star_{\rm osc}}$

$$\rho_\star \lesssim \mu_\star + \mathrm{osc}_\star \lesssim \mu_{\star_\mu} + \mathrm{osc}_{\star_{\rm osc}}.$$

Altogether, we see

$$\rho_\star N^s \lesssim \mu_{\star_\mu} N^{s_1} + \mathrm{osc}_{\star_{\rm osc}} N^{s_2} \lesssim \|\phi\|_{\mathbb{A}_s^\mu} + \|g\|_{\mathbb{A}_s^{\rm osc}} < \infty.$$

The hidden constant depends only $C_{\rm mon} > 0$ and the constants in Proposition 2. This proves $(\phi, g) \in \mathbb{A}_s^{\rho,J}$ and hence (53). □

Remark. *For $d = 2$ and $g \in H^{2+\varepsilon}(\Gamma)$ for some $\varepsilon > 0$, the error estimator is even efficient up to terms of higher order, i.e.*

$$\mu_\ell^2 \lesssim |||\phi - \Phi_\ell|||^2 + \mathrm{hot}_\ell^2,$$

and consequently

$$\rho_\ell^2 \lesssim |||\phi - \Phi_\ell|||^2 + \mathrm{osc}_\ell^2 + \mathrm{hot}_\ell^2.$$

The higher-order term satisfies $\mathrm{hot}_\ell \simeq h^{3/2+\varepsilon}$ for some $\varepsilon > 0$ on quasi-uniform meshes $\mathcal{T}_\ell$ with mesh width $h = h_\ell > 0$ [AFF$^+$14, Theorem 4]. *With* [AFF$^+$14, Proposition 15] *and analogous arguments as in the proof above, one can characterize the approximation class $\mathbb{A}_s^\mu$ for all $0 < s \le 3/2$ as*

$$\phi \in \mathbb{A}_s^\mu \quad \Longleftrightarrow \quad \phi \in \mathbb{A}_s \text{ and } g \in \mathbb{A}_s^{\rm osc},$$

where

$$\phi \in \mathbb{A}_s \quad \overset{\rm def}{\Longleftrightarrow} \quad \|\phi\|_{\mathbb{A}_s} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star \in \mathbb{T}_N} \min_{\Psi_\star \in \mathcal{P}^0(\mathcal{T}_\star)} |||\phi - \Psi_\star||| N^s < \infty.$$

□

6. Other Methods of Data Approximation

This section analyzes other methods for approximating the Dirichlet data g. We distinguish two cases for G_ℓ:

- For $d = 2$ and $p = 0$, i.e. Γ being a 1D manifold, we have $g \in H^1(\Gamma) \subset C(\Gamma)$. Therefore, it is admissible to use the nodal interpolation operator $I_\ell : C(\Gamma) \to \mathcal{S}^1(\mathcal{T}_\ell)$ for data approximation.
- Each projection $P_\ell : H^{1/2}(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$ with ℓ-independent stability constant

$$C_P := \sup_{\ell\in\mathbb{N}} \sup_{v\in H^{1/2}(\Gamma)\setminus\{0\}} \frac{\|P_\ell v\|_{H^{1/2}(\Gamma)}}{\|v\|_{H^{1/2}(\Gamma)}} < \infty \tag{56}$$

  is a valid choice.

The heart of the matter in the following section is to compensate the loss of orthogonality (34) in the data approximation term. To this end, we will use a modified marking strategy from [Ste07], known as separate Dörfler marking, which forces the oscillation term to contract if it is big compared to the estimator.

To clarify which method of data approximation is used, we write e.g. $\Phi_{\ell,P}$ for the solution of (10) and $\eta_{\ell,P}$ for the error estimator if the projection P_ℓ which is used for data approximation is not the Scott-Zhang projection J_ℓ.

6.1. **Data approximation by nodal interpolation $G_\ell = I_\ell g$ for 2D BEM.** The next lemma states some important properties of I_ℓ. Note that throughout this section, we assume $d = 2$ and $p = 0$. This type of data approximation was already considered in [AFLG+12] for the symmetric BEM formulation of the 2D Laplace problem with inhomogeneous mixed boundary conditions. However, in [AFLG+12], only convergence of an $(h-h/2)$-based error estimator to zero is proved. Similar techniques as employed here, are also found in [FPP13], where the lowest-order AFEM of the 2D Laplace problem with inhomogeneous mixed boundary conditions is considered.

Lemma 13. *Let $\mathcal{T}_\ell \in \mathbb{T}$. For a 1D manifold Γ, the nodal interpoland satisfies*

$$C_{\rm osc}^{-1}\|(1-I_\ell)g\|_{H^{1/2}(\Gamma)} \le \|h_\ell^{1/2}\nabla(1-I_\ell)g\|_{L^2(\Gamma)} = {\rm osc}_\ell. \tag{57}$$

Moreover, it satisfies the discrete upper bound

$$\|(I_\star - I_\ell)g\|_{H^{1/2}(\Gamma)} \le C_{\rm osc}\|h_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\mathcal{T}_\ell\setminus\mathcal{T}_\star)}. \tag{58}$$

The constant $C_{\rm osc} > 0$ depends only on $\mathcal{T}_0$.

Proof. Estimate (57) follows from bootstrapping the estimate

$$\|(1-I_\ell)v\|_{H^{1/2}(\Gamma)} \lesssim \|h_\ell^{1/2}\nabla v\|_{L^2(\Gamma)} \quad \text{for all } v \in H^1(\Gamma)$$

proved in [Car97, Theorem 1], see e.g. [EFGP13, Lemma 2.2], and by use of the well-known identity $\nabla I_\ell v = \Pi_\ell \nabla v \in L^2(\Gamma)$ for all $v \in H^1(\Gamma)$ valid in 1D. Estimate (58) was first observed in [FPP13, Proof of Proposition 3] and follows by similar techniques and $I_\star - I_\ell = (1-I_\ell)I_\star$. □

By comparison with Lemma 3, we see that the nodal interpolation operator I_ℓ has the same properties as (and even stronger than) the Scott-Zhang projection J_ℓ. This implies that all the results of the previous sections hold accordingly. In particular, the convergence result of Theorem 9 remains valid. Moreover, we obtain the optimality result of Theorem 11, if we replace the approximation class $\mathbb{A}_s^{\rho,J}$ with

$$(\phi, g) \in \mathbb{A}_s^{\rho,I} \quad \overset{\text{def.}}{\Longleftrightarrow} \quad \|(\phi,g)\|_{\mathbb{A}_s^{\rho,I}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star\in\mathbb{T}_N} \rho_{\star,P} N^s < \infty, \tag{59}$$

where the error estimator now reads

$$\rho_{\ell,I}^2 := \|h_\ell^{1/2}\nabla(V\Phi_{\ell,I} - (1/2+K)I_\ell g)\|_{L^2(\Gamma)}^2 + {\rm osc}_\ell^2. \tag{60}$$

6.2. **Data approximation by an $H^{1/2}$-stable projection $G_\ell = P_\ell g$ for 2D and 3D BEM.** General projections $P_\ell : H^{1/2}(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$ usually lack the discrete upper bound (22). To overcome this difficulty, we use a slightly modified marking strategy known as separate Dörfler marking [Ste07].

This variant was also used in [AFK+13] to prove quasi-optimal convergence rates of AFEM for the Laplace problem with inhomogeneous Dirichlet data. Unlike [AFK+13], we stress that our analysis of Section 4 above even covers the standard Dörfler marking (16) if one uses the Scott-Zhang projection $P_\ell = J_\ell$ for data approximation.

In [KOP13], the case $G_\ell = \Pi_\ell g$ with $\Pi_\ell : L^2(\Gamma) \to \mathcal{S}^{p+1}(\mathcal{T}_\ell)$ denoting the L^2-orthogonal projection is considered. According to [KPP13], newest vertex bisection guarantees that Π_ℓ is $H^1(\Gamma)$ stable and hence, by interpolation, also $H^{1/2}(\Gamma)$ stable. In contrast to the present work, [KOP13] uses an $(h-h/2)$-based error estimator to steer the adaptive algorithm and proves only convergence of the estimator to zero without guaranteeing any convergence rate.

With the error estimator

$$\rho_{\ell,P}^2 = \eta_{\ell,P}^2 + \mathrm{osc}_\ell^2 = \|h_\ell^{1/2}\nabla(V\Phi_{\ell,P} - (1/2+K)P_\ell g)\|_{L^2(\Gamma)}^2 + \mathrm{osc}_\ell^2 \tag{61}$$

the adaptive algorithm now reads as follows:

Algorithm 14. ***Input:*** *initial mesh $\mathcal{T}_0$, adaptivity parameters $0 < \theta_1, \theta_2, \vartheta < 1$. Set $\ell := 0$*

(i) *Compute approximate data $G_\ell = P_\ell g$.*
(ii) *Compute solution $\Phi_{\ell,P}$ of* (10)*.*
(iii) *Compute error estimator $\eta_{\ell,P}(T)$ and the data oscillations $\mathrm{osc}_\ell(T)$ for all $T \in \mathcal{T}_\ell$.*
(iv) *Determine a set of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ with minimal cardinality which satisfies separate Dörfler marking:*
   - *In case of $\mathrm{osc}_\ell^2 \le \vartheta\eta_{\ell,P}^2$, find $\mathcal{M}_\ell$ such that*
$$\theta_1\eta_{\ell,P}^2 \le \eta_{\ell,P}^2(\mathcal{M}_\ell). \tag{62a}$$
   - *In case of $\mathrm{osc}_\ell^2 > \vartheta\eta_{\ell,P}^2$, find $\mathcal{M}_\ell$ such that*
$$\theta_2\mathrm{osc}_\ell^2 \le \mathrm{osc}_\ell^2(\mathcal{M}_\ell). \tag{62b}$$

(v) *Refine at least the marked elements to obtain $\mathcal{T}_{\ell+1} = \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$.*
(vi) *Increment $\ell \mapsto \ell+1$ and goto* (i)*.*

Output: *sequence of error estimators $(\rho_{\ell,P})_{\ell\in\mathbb{N}}$ and sequence of Galerkin solutions $(\Phi_{\ell,P})_{\ell\in\mathbb{N}}$.*

The next lemma shows the equivalence of the solutions and error estimators for different approximations of the Dirichlet data g up to the Dirichlet data oscillations osc_ℓ.

Lemma 15. *Let Φ_ℓ and $\Phi_{\ell,P}$ denote solutions of* (10) *corresponding to the different approximations $G_\ell = J_\ell g$ and $G_\ell = P_\ell g$ of the Dirichlet data g. Then, it holds for any subset $\mathcal{E}_\ell \subseteq \mathcal{T}_\ell$*

$$|||\Phi_\ell - \Phi_{\ell,P}||| \le C_8\mathrm{osc}_\ell \quad \text{for all } \ell \in \mathbb{N} \tag{63}$$

as well as

$$|\eta_\ell(\mathcal{E}_\ell) - \eta_{\ell,P}(\mathcal{E}_\ell)| \le C_9\mathrm{osc}_\ell \quad \text{for all } \ell \in \mathbb{N}, \tag{64}$$

where $\eta_\ell, \eta_{\ell,P}$ denote the corresponding error estimators from (14) *resp.* (61)*. The constants $C_8 > 0$ and $C_9 > 0$ depend only on $\mathcal{T}_0$ and $p \ge 0$.*

Proof. We start with (63). By use of stability of $K : H^{1/2}(\Gamma) \to H^{1/2}(\Gamma)$, there holds

$$|||\Phi_\ell - \Phi_{\ell,P}||| \simeq \|(J_\ell - P_\ell)g\|_{H^{1/2}(\Gamma)} \le \|g - J_\ell g\|_{H^{1/2}(\Gamma)} + \|g - P_\ell g\|_{H^{1/2}(\Gamma)}.$$

Now, we conclude with the $H^{1/2}$-stability of P_ℓ

$$\|g - P_\ell g\|_{H^{1/2}(\Gamma)} = \|(1-P_\ell)(1-J_\ell)g\|_{H^{1/2}(\Gamma)} \lesssim \|(1-J_\ell)g\|_{H^{1/2}(\Gamma)} \lesssim \mathrm{osc}_\ell^2,$$

by use of (21). The combination of the last two inequalities shows (63).

It remains to prove (64). To that end, we employ the triangle inequality as well as the inverse estimates (19)–(20) from Proposition 2

$$\begin{aligned}|\eta_\ell(\mathcal{E}_\ell) - \eta_{\ell,P}(\mathcal{E}_\ell)| &\le \|h_\ell^{1/2}\nabla V(\Phi_\ell - \Phi_{\ell,P})\|_{L^2(\Gamma)} + \|h_\ell^{1/2}\nabla(1/2+K)(J_\ell - P_\ell)g\|_{L^2(\Gamma)}\\ &\lesssim |||\Phi_\ell - \Phi_{\ell,P}||| + \|(J_\ell - P_\ell)g\|_{H^{1/2}(\Gamma)}.\end{aligned}$$

Arguing as before to see $\|(J_\ell - P_\ell)g\|_{H^{1/2}(\Gamma)} \lesssim \mathrm{osc}_\ell$, we conclude the proof. □

Now, we show that separate Dörfler marking (62) for $\rho_{\ell,P}^2 = \eta_{\ell,P}^2 + \mathrm{osc}_\ell^2$ implies combined Dörfler marking (16) for $\rho_\ell^2 = \eta_\ell^2 + \mathrm{osc}_\ell^2$.

Lemma 16. *Let Φ_ℓ and $\Phi_{\ell,P}$ denote solutions of* (10) *corresponding to different approximations $G_\ell = J_\ell g$ and $G_\ell = P_\ell g$ of the Dirichlet data g. Let the error estimator $\eta_{\ell,P}$ together with* osc_ℓ *satisfy the separate Dörfler marking* (62) *for arbitrary $0 < \theta_1, \theta_2 < 1$, sufficiently small $0 < \vartheta < 1$, and a set of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$. Then, $\rho_\ell^2 := \eta_\ell^2 + \mathrm{osc}_\ell^2$ satisfies the combined Dörfler marking* (16)

$$\theta \rho_\ell^2 \le \rho_\ell^2(\mathcal{M}_\ell) \tag{65}$$

for some $0 < \theta < 1$.

Proof. First, assume $\mathrm{osc}_\ell^2 \le \vartheta \eta_{\ell,P}^2$. Then, it holds with Lemma 15

$$\eta_{\ell,P}^2(\mathcal{M}_\ell) \le 2\eta_\ell^2(\mathcal{M}_\ell) + 2C_9^2 \mathrm{osc}_\ell^2 \le 2\eta_\ell^2(\mathcal{M}_\ell) + 2C_9^2 \vartheta \eta_{\ell,P}^2.$$

Moving the last term to the left-hand side, we see

$$(1 - 2C_9^2\vartheta)\eta_{\ell,P}^2(\mathcal{M}_\ell) \le 2\eta_\ell^2(\mathcal{M}_\ell) \le 2\rho_\ell^2(\mathcal{M}_\ell).$$

Together with Lemma 15 and (62a), this yields

$$\begin{aligned}\theta_1\rho_\ell^2 \le 2\theta_1\eta_{\ell,P}^2 + \theta_1(2C_9^2+1)\mathrm{osc}_\ell^2 \le \theta_1(2 + (2C_9^2+1)\vartheta)\eta_{\ell,P}^2 \le (2 + (2C_9^2+1)\vartheta)\eta_{\ell,P}^2(\mathcal{M}_\ell) \\ \le 2(2 + (2C_9^2+1)\vartheta)/(1 - 2C_9^2\vartheta)\rho_\ell^2(\mathcal{M}_\ell).\end{aligned} \tag{66}$$

Second, assume $\mathrm{osc}_\ell^2 > \vartheta\eta_{\ell,P}^2$. Then, again with Lemma 15

$$\begin{aligned}\theta_2\rho_\ell^2 &\le 2\theta_2\eta_{\ell,P}^2 + \theta_2(2C_9^2+1)\mathrm{osc}_\ell^2 \le \theta_2(2\vartheta^{-1} + 2C_9^2 + 1)\mathrm{osc}_\ell^2 \\ &\le (2\vartheta^{-1} + 2C_9^2+1)\mathrm{osc}_\ell^2(\mathcal{M}_\ell) \le (2\vartheta^{-1} + 2C_9^2 + 1)\rho_\ell^2(\mathcal{M}_\ell).\end{aligned} \tag{67}$$

Hence, ρ_ℓ satisfies combined Dörfler marking (16) with

$$\theta := \min\{\theta_1(1 - 2C_9^2\vartheta)/(4 + 2(2C_9^2+1)\vartheta), \theta_2/(2\vartheta^{-1} + 2C_9^2 + 1)\}. \tag{68}$$

This concludes the proof. □

Remark. *We established convergence of Algorithm 14 at least for sufficiently small $0 < \vartheta < 1$. The previous lemma shows that ρ_ℓ satisfies combined Dörfler marking* (16) *in each step of the adaptive loop. Therefore, Theorem 9 is applicable and implies $\rho_\ell^2 \lesssim \Delta_\ell \to 0$ as $\ell \to \infty$. The equivalence $\rho_{\ell,P} \simeq \rho_\ell$ which follows immediately from* (64) *proves* $\lim_{\ell\to\infty} \rho_{\ell,P} = 0$. □

Remark. *Arguing as in the proof of Lemma 16, we see that separate Dörfler marking* (62) *for $\rho_{\ell,P}$ also implies combined Dörfler marking* (16) *for $\rho_{\ell,P}$ without any assumption on $0 < \theta_1, \theta_2, \vartheta < 1$. Then, the estimator reduction* (26) *of Proposition 6 also holds in this case. In* [KOP13]*, it is proved that this implies convergence of the adaptive algorithm.* □

Lemma 17. *The error estimator $\rho_{\ell,P}$ satisfies* (E1)–(E3) *and there holds quasi-monotonicity, i.e. for $\mathcal{T}_\star \in \mathtt{refine}(\mathcal{T}_\ell)$ being a refinement of $\mathcal{T}_\ell \in \mathbb{T}$, we have*

$$\rho_{\star,P} \le C_{10}\rho_{\ell,P} \tag{69}$$

for a constant $C_{10} > 0$ which depends only on $\mathcal{T}_0$.

Proof. (E1)–(E3) follow verbatim as in the proof of Proposition 4. For quasi-monotonicity, we apply the equivalence (64) to obtain $\rho_{\star,P} \simeq \rho_\star \lesssim \rho_\ell \simeq \rho_{\ell,P}$. This concludes the proof. □

Proposition 18. *Let $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\ell)$ be a refinement of $\mathcal{T}_\ell \in \mathbb{T}$. Let $\Phi_{\ell,P}$, $\Phi_{\star,P}$ denote the corresponding solutions of* (10). *Suppose that the error estimator satisfies*

$$\rho_{\star,P} \le \kappa_\star \rho_{\ell,P} \tag{70}$$

for some $0 < \kappa_\star < 1$ sufficiently small. Then, the set $\mathcal{R}_\ell := \omega_\ell^5(\mathcal{T}_\ell \setminus \mathcal{T}_\star)$ from (E4) *satisfies separate Dörfler marking* (62) *for sufficiently small $0 < \theta_1, \vartheta < 1$ and arbitrary $0 < \theta_2 < 1$.*

Proof. First, we prove by use of (64)

$$\rho_\star \le \sqrt{2}(C_9+1)\rho_{\star,P} \le \sqrt{2}(C_9+1)\kappa_\star\rho_{\ell,P} \le 2(C_9+1)^2\kappa_\star\rho_\ell, \tag{71}$$

Therefore and with $2(C_9+1)^2\kappa_\star < 1$, we may apply Proposition 10 to conclude

$$\theta_\star \rho_\ell^2 \le \rho_\ell^2(\mathcal{R}_\ell). \tag{72}$$

Now, we distinguish two cases. First, assume $\mathrm{osc}_\ell^2 \le \vartheta\eta_{\ell,P}^2$. Then, it holds with (64)

$$\eta_{\ell,P}^2 \le 2\eta_\ell^2 + 2C_9^2\mathrm{osc}_\ell^2 \le 2\eta_\ell^2 + 2C_9^2\vartheta\eta_{\ell,P}^2.$$

Now, we rearrange the terms in the above equation and employ (72) to get

$$(1-2C_9^2\vartheta)\theta_\star\eta_{\ell,P}^2 \le 2\eta_\ell^2(\mathcal{R}_\ell) + 2\mathrm{osc}_\ell^2(\mathcal{R}_\ell) \le 4\eta_{\ell,P}^2(\mathcal{R}_\ell) + (2C_9^2+2)\vartheta\eta_{\ell,P}^2.$$

For $\vartheta > 0$ sufficiently small, this shows

$$\Big((1-2C_9^2\vartheta)\theta_\star - (2+2C_9^2)\vartheta\Big)/4\,\eta_{\ell,P}^2 \le \eta_{\ell,P}^2(\mathcal{R}_\ell),$$

i.e. (62a) for all $\theta_1 \le ((1-2C_9^2\vartheta)\theta_\star - (2+2C_9^2)\vartheta)/4$.

Second, let $\mathrm{osc}_\ell^2 > \vartheta\eta_{\ell,P}^2$. Here, we use the local definition of osc_ℓ to obtain

$$\begin{aligned}\mathrm{osc}_\ell^2(\mathcal{T}_\ell \cap \mathcal{T}_\star) &= \|h_\ell^{1/2}(1-\Pi_\ell)\nabla g\|_{L^2(\cup\mathcal{T}_\ell\cap\mathcal{T}_\star)} = \mathrm{osc}_\star^2(\mathcal{T}_\ell \cap \mathcal{T}_\star)\\ &\le \rho_\star^2\\ &\le 2(C_9+1)^2\kappa_\star\rho_\ell^2\\ &\le 2(C_9+1)^2\kappa_\star(2\eta_{\ell,P}^2 + (1+2C_9^2)\mathrm{osc}_\ell^2)\\ &\le 2(C_9+1)^2\kappa_\star(2\vartheta^{-1}+1+2C_9^2)\mathrm{osc}_\ell^2.\end{aligned}$$

Now, we conclude

$$\mathrm{osc}_\ell^2 = \mathrm{osc}_\ell^2(\mathcal{T}_\ell\cap\mathcal{T}_\star) + \mathrm{osc}_\ell^2(\mathcal{T}_\ell\setminus\mathcal{T}_\star) \le 2(C_9+1)^2\kappa_\star(2\vartheta^{-1}+1+2C_9^2)\mathrm{osc}_\ell^2 + \mathrm{osc}_\ell^2(\mathcal{R}_\ell),$$

which shows (62b) for all $\theta_2 \le 1 - 2(C_9+1)^2\kappa_\star(2\vartheta^{-1}+1+2C_9^2)$. Moreover, for all $0<\theta_2<1$ we may choose $\kappa_\star$ such that (62b) holds true. This concludes the proof. $\square$

Now, we have collected all ingredients to prove an optimality result similar to Theorem 11. To that end, we define the approximation class

$$(\phi,g) \in \mathbb{A}_s^{\rho,P} \quad \overset{\text{def}}{\Longleftrightarrow} \quad \|(\phi,g)\|_{\mathbb{A}_s^{\rho,P}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}_\star\in\mathbb{T}_N} \rho_{\ell,P}N^s < \infty. \tag{73}$$

Theorem 19. *For sufficiently small parameters $0<\theta_1,\vartheta<1$ and arbitrary $0<\theta_2<1$, Algorithm 14 is optimal in the sense of*

$$(\phi,g)\in\mathbb{A}_s^{\rho,P} \quad \Longleftrightarrow \quad \rho_{\ell,P} \le C_{11}(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s} \quad \textit{for all } \ell\in\mathbb{N}. \tag{74}$$

The constant $C_{11}>0$ depends only on $\|(\phi,g)\|_{\mathbb{A}_s^{\rho,P}}$, $0<\theta_1,\theta_2,\vartheta<1$, and $0<\kappa<1$.

Proof. Analogously to the proof of Theorem 11, the definition of the approximation class provides a mesh $\mathcal{T}_\star \in \texttt{refine}(\mathcal{T}_\ell)$ such that

$$\rho_{\star,P} \le \kappa_\star \rho_{\ell,P} \quad \text{and} \quad \#(\mathcal{T}_\ell \setminus \mathcal{T}_\star) \lesssim \rho_{\ell,P}^{-1/s},$$

where the hidden constant in the second estimate depends on $0 < \kappa_\star < 1$ which can be chosen arbitrarily small.

Next, we apply Proposition 18 to obtain that $\mathcal{R}_\ell := \omega_\ell^5(\mathcal{T}_\ell \setminus \mathcal{T}_\star)$ satisfies separate Dörfler marking (62). However, $\mathcal{M}_\ell$ was chosen in Step (iv) of Algorithm 14 to be a set with minimal cardinality. Therefore, it holds

$$\#\mathcal{M}_\ell \le \mathcal{R}_\ell \simeq \#(\mathcal{T}_\ell \setminus \mathcal{T}_\star) \lesssim \rho_{\ell,P}^{-1/s} \lesssim \rho_\ell^{-1/s} \simeq \Delta_\ell^{-1/(2s)}.$$

Lemma 16 shows that ρ_ℓ and $\mathcal{M}_\ell$ satisfy combined Dörfler marking (16). Therefore, Theorem 9 shows

$$\Delta_{\ell+1} \le \kappa \Delta_\ell \quad \text{for all } \ell \in \mathbb{N}.$$

Now, the remainder of the proof follows analogously to the proof of Theorem 11. □

6.3. **Characterization of approximation class.** Finally, we prove that all the approximation classes introduced in this work coincide. In particular, the optimal rate for the error estimator considered, is, in fact, independent of how the Dirichlet data are discretized. Each of the discretizations proposed, will lead to the same convergence rate of ABEM.

Proposition 20. *There holds for all $s > 0$*

$$\mathbb{A}_s^{\rho,J} = \mathbb{A}_s^{\rho,P} = \Big\{(\phi, g) \,:\, \phi \in \mathbb{A}_s^\mu \text{ and } g \in \mathbb{A}_s^{\mathrm{osc}}\Big\} \tag{75}$$

and in case of $d = 2$ and $p = 0$, we even have

$$\mathbb{A}_s^{\rho,J} = \mathbb{A}_s^{\rho,P} = \mathbb{A}_s^{\rho,I} = \Big\{(\phi, g) \,:\, \phi \in \mathbb{A}_s^\mu \text{ and } g \in \mathbb{A}_s^{\mathrm{osc}}\Big\}. \tag{76}$$

Proof. We introduce the error estimators ρ_ℓ, $\rho_{\ell,I}$, and $\rho_{\ell,P}$ corresponding to the different methods of data approximation. Lemma 15 shows $\rho_\ell \simeq \rho_{\ell,P}$ for all $\mathcal{T}_\ell \in \mathbb{T}$. A similar argument also proves $\rho_{\ell,I} \simeq \rho_\ell \simeq \rho_{\ell,P}$ for all $\mathcal{T}_\ell \in \mathbb{T}$ in case of $d = 2$ and $p = 0$. This yields immediately

$$\|(\phi, g)\|_{\mathbb{A}_s^{\rho,J}} \simeq \|(\phi, g)\|_{\mathbb{A}_s^{\rho,P}} \quad \text{and in case of } d = 2 \quad \|(\phi, g)\|_{\mathbb{A}_s^{\rho,I}} \simeq \|(\phi, g)\|_{\mathbb{A}_s^{\rho,J}}.$$

Together with Theorem 12, we prove the assertion. □

7. Numerical experiment

To underline the results of the previous sections, we present a 2D example for $p = 0$ and $p = 1$.

7.1. **Details on the implementation.** The results are visualized with the help of the following quantities:

- Instead of the energy norm error $|||\phi - \Phi_\ell|||$ which can hardly be computed analytically, we plot the following reliable error bound:

$$|||\phi - \Phi_\ell||| \lesssim \mathrm{err}_\ell + \mathrm{osc}_\ell \quad \text{with } \mathrm{err}_\ell := \|h_\ell^{1/2}(\phi - \Phi_\ell)\|_{L^2(\Gamma)}.$$

- We plot the error indicator η_ℓ. The functions $(V\Phi_\ell)(x)$ and $(KG_\ell)(x)$ are computed analytically [Mai01].

To compare the adaptive approach of Algorithm 1 with uniform mesh refinement, we consider the computational times:

- The time t_{unif} to compute the solution $\Phi^{(\ell)}$ of the uniform approach is the time needed to perform ℓ uniform refinements of the initial mesh $\mathcal{T}_0$, plus the time needed to build and solve the linear system corresponding to $\mathcal{T}^{(\ell)}$. Obviously, the second contribution is vastly dominant.
- The time t_{adap} to compute the solution Φ_ℓ of the adaptive approach is the time elapsed in all previous steps, plus the time to build and solve the system corresponding to the mesh $\mathcal{T}_\ell$, to compute the error estimator, and to mark and refine the mesh.

Although this definition seems to favour the uniform approach, we think that it provides a fair comparison between those strategies. All computations were performed on a 64-bit Linux work station with 32 GB of RAM in `Matlab` (Release R2010a). The implementation relies on the open-source `Matlab` BEM library `HILBERT`, see

`http://www.asc.tuwien.ac.at/abem/hilbert/`

Throughout, all the occurring linear systems were solved directly with the MATLAB backslash operator. In all experiments, the adaptivity parameter in Algorithm 1 is chosen

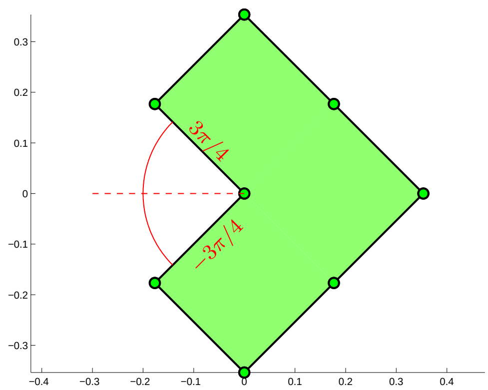

FIGURE 1. L-shaped domain Ω with initial partition of the boundary $\mathcal{T}_0$.

as

$$\theta = 1/2.$$

7.2. **Experiment on L-shaped domain with singular solution and singular data.** Here, Γ is the boundary of the L-shaped domain Ω in Figure 1. We prescribe the solution u of

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{on } \Gamma, \end{aligned} \tag{77}$$

as $u(x,y) := v_{2/3}(x,y) + v_{7/8}(x - z_1, y - z_2)$, where $v_\delta(x,y) := r^\delta \cos(\delta\alpha)$ and $z = (z_1, z_2)$ is the uppermost corner of the L-shaped domain in Figure 1. The solution ϕ has a corner singularity at the re-entrant corner and in addition a singularity resulting from the singular data g. Note that $v_\delta \in H^{1+\delta-\varepsilon}(\Omega)$ for all $\varepsilon > 0$. Figure 2 compares the uniform approach for $p = 0$ with the adaptive approaches for $p = 0$ and $p = 1$. We use the Scott-Zhang projection for data approximation, i.e. $G_\ell = J_\ell g$ for all $\ell \in \mathbb{N}$. We apply Algorithm 1 and confirm the theoretical results from Theorem 11. We see that the uniform approach only leads to a suboptimal convergence rate $N^{-2/3}$, whereas the adaptive approach reaches the optimal order of convergence $N^{-3/2}$ for the lowest-order method $p = 0$ and $N^{-5/2}$ for $p = 1$. The comparison with respect to computational time is also significant. Whereas the uniform approach is slightly faster in the first few steps (due to the overhead which is naturally caused by the adaptive algorithm), we see that on the long run, the adaptive method with $p = 0$ and particularly with $p = 1$ is much superior.

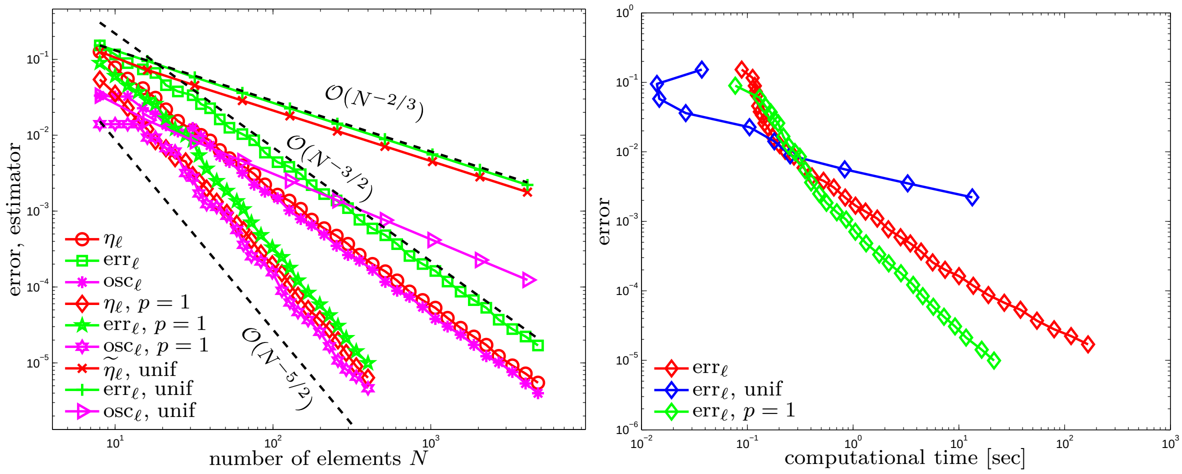

FIGURE 2. Experiment on L-shaped domain with singular solution and singular data. We compare adaptive and uniform mesh refinement in terms of the quantities err_ℓ, η_ℓ, and osc_ℓ plotted over the number of elements $N = \#\mathcal{T}_\ell$ (left). Additionally, err_ℓ is plotted versus the computational time in seconds (right).

Acknowledgements: The authors MF, TF, and DP acknowledge support through the Austrian Science Fund (FWF) under grant P21732 *Adaptive Boundary Element Method*. MK acknowledges support by CONICYT project Anillo ACT1118 (ANANUM).

References

[AFF+12] Markus Aurada, Michael Feischl, Thomas Führer, J. Markus Melenk, and Dirk Praetorius. 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.

[AFF+14] 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. *J. Comput. Appl. Math.*, 255:481–501, 2014.

[AFK+13] Markus Aurada, Michael Feischl, Josef Kemetmüller, Marcus Page, and Dirk Praetorius. Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$. *M2AN, Math. Model. Numer. Anal.*, 47:1207–1235, 2013.

[AFLG+12] M. Aurada, S. Ferraz-Leite, P. Goldenits, M. Karkulik, M. Mayr, and D. Praetorius. Convergence of adaptive BEM for some mixed boundary value problem. *Appl. Numer. Math.*, 62(4):226–245, 2012.

[AFLP12] Markus Aurada, Samuel Ferraz-Leite, and Dirk Praetorius. Estimator reduction and convergence of adaptive BEM. *Appl. Numer. Math.*, 62(6):787–801, 2012.

[BDD04] Peter Binev, Wolfgang Dahmen, and Ronald DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[Car97] Carsten Carstensen. An a posteriori error estimate for a first-kind integral equation. *Math. Comp.*, 66(217):139–155, 1997.

[CKNS08] 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(5):2524–2550, 2008.

[CMS01] 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(2):197–213, 2001.

[CP07] Carsten Carstensen and Dirk Praetorius. Averaging techniques for the a posteriori BEM error control for a hypersingular integral equation in two dimensions. *SIAM J. Sci. Comput.*, 29(2):782–810 (electronic), 2007.

[CS95] Carsten Carstensen and Ernst P. Stephan. A posteriori error estimates for boundary element methods. *Math. Comp.*, 64(210):483–500, 1995.

[Dör96] Willy Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[EFGP13] Christoph Erath, Stephan Funken, Petra Goldenits, and Dirk Praetorius. Simple error estimators for the Galerkin BEM for some hypersingular integral equation in 2D. *Appl. Anal.*, 92:1194–1216, 2013.

[FFP12] Michael Feischl, Thomas Führer, and Dirk Praetorius. Adaptive FEM with optimal convergence rates for a certain class of non-symmetric and possibly non-linear problems. 2012.

[FKMP13] Michael Feischl, Michael Karkulik, J. Markus Melenk, and Dirk Praetorius. Quasi-optimal convergence rate for an adaptive boundary element method. *SIAM J. Numer. Anal.*, 51:1327–1348, 2013.

[FPP13] Michael Feischl, Marcus Page, and Dirk Praetorius. Convergence of adaptive FEM for elliptic obstacle problems. *accepted for publication in Int. J. Numer. Anal. Model.*, 2013.

[GHS05] I. G. Graham, W. Hackbusch, and S. A. Sauter. Finite elements on degenerate meshes: inverse-type inequalities and applications. *IMA J. Numer. Anal.*, 25(2):379–407, 2005.

[HW08] George C. Hsiao and Wolfgang L. Wendland. *Boundary integral equations*, volume 164 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2008.

[Kar] Michael Karkulik. *Zur Konvergenz und Quasioptimalität adaptiver Randelementmethoden.* PhD thesis. (in German).

[KOP13] Michael Karkulik, Günther Of, and Dirk Praetorius. Convergence of adaptive 3D BEM for weakly singular integral equations based on isotropic mesh-refinement. *Numer. Methods Partial Differential Equations*, 2013. in print.

[KPP13] Michael Karkulik, David Pavlicek, and Dirk Praetorius. On 2D newest vertex bisection: Optimality of mesh-closure and H^1-stability of L_2-projection. *accepted for publication in Constr. Approx.*, 2013.

[Mai01] Matthias Maischak. The analytical computation of the Galerkin elements for the Laplace, Lamé and Helmholtz equation in 2D-BEM. *Preprint, Institute for Applied Mathematics, University of Hanover*, 2001.

[McL00] William McLean. *Strongly elliptic systems and boundary integral equations.* Cambridge University Press, Cambridge, 2000.

[SS11] Stefan A. Sauter and Christoph Schwab. *Boundary element methods*, volume 39 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2011. Translated and expanded from the 2004 German original.

[Ste07] Rob Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.

[Ste08] Rob Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.

[SZ90] L. Ridgway Scott and Shangyou Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[Tso13] Gantumur Tsogtgerel. Adaptive boundary element methods with convergence rates. *Numerische Mathematik*, 124(3):471–516, 2013.

[Ver96] Rüdiger Verfürth. A review of a posteriori error estimation and adaptive mesh-refinement techniques, 1996.

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria

E-mail address: {Thomas.Fuehrer,Melenk,Dirk.Praetorius}@tuwien.ac.at

E-mail address: Michael.Feischl@tuwien.ac.at (corresponding author)

Departamento de Matemáticas, Oficina 139, Avenida Vicuña Mackenna 4860, Santiago, Chile
E-mail address: mkarkulik@mat.puc.cl