ASC Report No. 34/2010

Convergence and Quasi-Optimality of Adaptive FEM with Inhomogeneous Dirichlet Data

Michael Feischl, Marcus Page, Dirk Praetorius

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-03-2

Most recent ASC Reports

33/2010	*Michael Feischl, Marcus Page, Dirk Praetorius* Convergence of Adaptive FEM for some Elliptic Obstacle Problem with Inhomogeneous Dirichlet Data
32/2010	*Othmar Koch, Friedrich Kupka, Bernhard Löw-Baselli, A. Mayrhofer, F. Zaussinger* SDIRK Methods for the ANTARES Code
31/2010	*Roland Donninger, Birgit Schörkhuber, Peter C. Aichelburg* On Stable Self-similar Blow Up for Equivalent Wave Maps: The Linearized Problem
30/2010	*Matthias Langer, Harald Woracek* The Exponential Type of the Fundamental Solution of an Indefinite Hamiltonian System
29/2010	*Harald Woracek* An Addendum to M.G. Krein's Inverse Spectral Theorem for Strings
28/2010	*Philipp Dörsek, Josef Teichmann* A Semigroup Point of View On Splitting Schemes For Stochastic (Partial) Differential Equations
27/2010	*Ansgar Jüngel* Dissipative quantum fluid models
26/2010	*Sabine Hittmeir, Ansgar Jüngel* Cross Diffusion Preventing Blow Up in the Two-dimensional Keller-Segel Model
25/2010	*Anton Arnold, Irene M. Gamba, Maria Pia Gualdani, Stéphane Mischler, Clément Mouhot, Christof Sparber* The Wigner-Fokker-Planck Equation: Stationary States and Large Time Behavior
24/2010	*Lehel Banjai, Christian Lubich, Jens Markus Melenk* Runge-Kutta Convolution Quadrature for Operators Arising in Wave Propagation

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 978-3-902627-03-2

CONVERGENCE AND QUASI-OPTIMALITY OF ADAPTIVE FEM WITH INHOMOGENEOUS DIRICHLET DATA

M. FEISCHL, M. PAGE, AND D. PRAETORIUS

Abstract. We consider the solution of a second order elliptic PDE in 2D with inhomogeneous Dirichlet data by means of adaptive lowest-order FEM. As is usually done in practice, the given Dirichlet data are discretized by nodal interpolation. As model example serves the Poisson equation with mixed Dirichlet-Neumann boundary conditions. For error estimation, we use an edge-based residual error estimator which replaces the volume residual contributions by edge oscillations. We consider two marking strategies from the literature and prove that either of them is convergent with quasi-optimal convergence behaviour. Numerical experiments conclude the work.

1. Introduction

1.1. Model problem. Recently, there has been a major breakthrough in the thorough mathematical understanding of convergence and quasi-optimality of h-adaptive FEM for second-order elliptic PDEs. However, the focus of the numerical analysis usually lied on model problems with homogeneous Dirichlet conditions, i.e. $\Delta u = f$ in Ω with $u = 0$ on $\Gamma = \partial\Omega$, see e.g. [CKNS, D, KS, MNS, S07]. Instead, our model problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \Gamma_D, \\ \partial_n u &= \phi \quad \text{on } \Gamma_N \end{aligned} \tag{1}$$

considers mixed Dirichlet-Neumann boundary conditions. Here, Ω is a bounded Lipschitz domain in $\mathbb{R}^2$ with polygonal boundary $\Gamma = \partial\Omega$ which is split into two relatively open boundary parts, namely the Dirichlet boundary Γ_D and the Neumann boundary Γ_N, i.e. $\Gamma_D \cap \Gamma_N = \emptyset$ and $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \Gamma$. We stress that the surface measure of the Dirichlet boundary has to be positive $|\Gamma_D| > 0$, whereas Γ_N is allowed to be empty. The given data formally satisfy $f \in \widetilde{H}^{-1}(\Omega)$, $g \in H^{1/2}(\Gamma_D)$, and $\phi \in H^{-1/2}(\Gamma_N)$. As is usually required to derive (localized) a posteriori error estimators, we assume additional regularity of the given data, namely $f \in L^2(\Omega)$, $g \in H^1(\Gamma_D)$, and $\phi \in L^2(\Gamma_N)$.

Whereas the inclusion of inhomogeneous Neumann conditions ϕ into the numerical analysis seems to be obvious, incorporating inhomogeneous Dirichlet g conditions is technically more demanding. This is mainly due to the fractional-order Sobolev space $H^{1/2}(\Gamma_D)$. Since the $H^{1/2}$-norm is non-local, the a posteriori error analysis requires appropriate localization techniques. These have recently been developed in the context of adaptive boundary element methods [AGP, CP1, CP2, EFFP, EFGP, KP].

Date: December 29, 2010.

2000 *Mathematics Subject Classification.* 65N30, 65N50.

Key words and phrases. adaptive finite element methods, convergence analysis, quasi-optimality, inhomogeneous Dirichlet data.

It is well-known that the Poisson problem (1) admits a unique weak solution $u \in H^1(\Omega)$ with $u = g$ on Γ_D in the sense of traces which solves the variational formulation

$$\langle \nabla u\,,\,\nabla v\rangle_\Omega = \langle f\,,\,v\rangle_\Omega + \langle \phi\,,\,v\rangle_{\Gamma_N} \quad \text{for all } v \in H^1_D(\Omega). \tag{2}$$

Here, the test space reads $H^1_D(\Omega) = \big\{v \in H^1(\Omega)\,:\, v = 0 \text{ on } \Gamma_D \text{ in the sense of traces}\big\}$, and $\langle\cdot\,,\,\cdot\rangle$ denotes the respective L^2-scalar products.

1.2. Discretization. For the Galerkin discretization, let $\mathcal{T}_\ell$ be a regular triangulation of Ω into triangles $T \in \mathcal{T}_\ell$. We use lowest-order conforming elements, where the ansatz space reads

$$\mathcal{S}^1(\mathcal{T}_\ell) = \big\{V_\ell \in C(\overline{\Omega})\,:\, V_\ell|_T \text{ is affine for all } T \in \mathcal{T}_\ell\big\}. \tag{3}$$

Since a discrete function $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ cannot satisfy continuous Dirichlet conditions, we have to discretize the given data $g \in H^1(\Gamma_D)$. According to the Sobolev inequality on the 1D manifold Γ_D, the given Dirichlet data are continuous on $\overline{\Gamma}_D$. Therefore, the nodal interpoland g_ℓ of g is well-defined. As is usually done in practice, we approximate $g \approx g_\ell$. Again, it is well-known that there is a unique $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ with $U_\ell = g_\ell$ on Γ_D which solves the Galerkin formulation

$$\langle \nabla U_\ell\,,\,\nabla V_\ell\rangle_\Omega = \langle f\,,\,V_\ell\rangle_\Omega + \langle \phi\,,\,V_\ell\rangle_{\Gamma_N} \quad \text{for all } V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell). \tag{4}$$

Here, the test space is given by $\mathcal{S}^1_D(\mathcal{T}_\ell) = \mathcal{S}^1(\mathcal{T}_\ell) \cap H^1_D(\Omega) = \big\{V_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)\,:\, V_\ell = 0 \text{ on } \Gamma_D\big\}$.

1.3. A posteriori error estimation. An element-based residual error estimator for this discretization reads

$$\rho_\ell^2 = \sum_{T\in\mathcal{T}_\ell} \rho_\ell(T)^2 \tag{5}$$

with corresponding refinement indicators

$$\begin{aligned}\rho_\ell(T)^2 := |T|\,\|f\|^2_{L^2(T)}\\ + |T|^{1/2}\big(\|[\partial_n U_\ell]\|^2_{L^2(\partial T\cap\Omega)} + \|\phi - \partial_n U_\ell\|^2_{L^2(\partial T\cap\Gamma_N)} + \|(g-g_\ell)'\|^2_{L^2(\partial T\cap\Gamma_D)}\big),\end{aligned} \tag{6}$$

where $[\cdot]$ denotes the jump across edges and $(\cdot)'$ denotes the arclength derivative. We prove reliability and efficiency of ρ_ℓ (Proposition 2) and discrete local reliability (Proposition 3). Inspired by [PP], we introduce an edge-based error estimator ϱ_ℓ which reads

$$\varrho_\ell^2 = \sum_{E\in\mathcal{E}_\ell} \varrho_\ell(E)^2. \tag{7}$$

For an edge $E \in \mathcal{E}_\ell$, its local contributions read

$$\varrho_\ell(E)^2 = \begin{cases} |E|\|[\partial_n U_\ell]\|^2_{L^2(E)} + |\omega_{\ell,E}|\|f - f_{\omega_{\ell,E}}\|^2_{\omega_{\ell,E}} & \text{if } E \subset \Omega,\\ |E|\|\phi - \partial_n U_\ell\|^2_{L^2(E)} & \text{if } E \subseteq \Gamma_N,\\ |E|\|(g-g_\ell)'\|^2_{L^2(E)} & \text{if } E \subseteq \Gamma_D.\end{cases} \tag{8}$$

Here, $\omega_{\ell,E} \subset \Omega$ denotes the edge patch, and $f_{\omega_{\ell,E}}$ denotes the corresponding integral mean. The advantage of ϱ_ℓ is that the volume residual terms $|T|^{1/2}\|f\|_{L^2(T)}$ in (6) are replaced by the edge oscillations $|\omega_{\ell,E}|^{1/2}\|f - f_{\omega_{\ell,E}}\|_{\omega_{\ell,E}}$, which are generically of higher order. We prove that ρ_ℓ and ϱ_ℓ are locally equivalent (Lemma 4) and thus obtain reliability and efficiency of ϱ_ℓ (Proposition 5) as well as discrete local reliability (Proposition 6).

1.4. Adaptive algorithm. We use the local contributions of ϱ_ℓ to mark edges for refinement in two realizations (Algorithm 7 and Algorithm 8) of the standard adaptive loop (AFEM)

$$\boxed{\texttt{solve}} \quad \rightarrow \quad \boxed{\texttt{estimate}} \quad \rightarrow \quad \boxed{\texttt{mark}} \quad \rightarrow \quad \boxed{\texttt{refine}} \tag{9}$$

Our adaptive algorithms use variants of the well-studied Dörfler marking [D] to mark certain edges for refinement. Throughout, we use newest vertex bisection, and at least marked edges are bisected. Given some initial mesh $\mathcal{T}_0$, both algorithms generate successively locally refined meshes $\mathcal{T}_\ell$ with corresponding discrete solutions $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ of (4).

1.5. Main results. The first main result (Theorem 13) states that either algorithm leads to contraction

$$\Delta_{\ell+1} \leq \kappa\, \Delta_\ell \quad \text{for all } \ell \in \mathbb{N}_0 \text{ and some constant } 0 < \kappa < 1 \tag{10}$$

for some quasi-error quantity $\Delta_\ell \simeq \varrho_\ell$ which is equivalent to the error estimator. In particular, this proves linear convergence of the adaptively generated solutions $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ to the (unknown) weak solution $u \in H^1(\Omega)$ of (2). The main ingredients of the proof are an equivalent error estimator $\widetilde{\varrho}_\ell \simeq \varrho_\ell$ for which we prove some estimator reduction

$$\widetilde{\varrho}_{\ell+1}^2 \leq q\, \widetilde{\varrho}_\ell^2 + C\, \|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}^2 \quad \text{for all } \ell \in \mathbb{N}_0 \text{ and some } 0 < \kappa < 1 \text{ and } C > 0, \tag{11}$$

see Lemma 11 and a quasi-Galerkin orthogonality in Lemma 12, whereas the general concept follows that of [CKNS].

The second main result is Theorem 17 which states that the outcome of the adaptive algorithms is quasi-optimal in the sense of Stevenson [S07]: Provided the given data $(f, g, \phi) \in L^2(\Omega) \times H^1(\Gamma_D) \times L^2(\Gamma_N)$ and the corresponding weak solution $u \in H^1(\Omega)$ of (2) belong to the approximation class

$$\mathbb{A}_s := \big\{(u, f, g, \phi) \,:\, \|(u, f, g, \phi)\|_{\mathbb{A}_s} := \sup_{N \in \mathbb{N}} \big(N^s \sigma(N, u, f, g, \phi)\big) < \infty\big\} \tag{12}$$

with

$$\begin{aligned} \sigma(N, u, f, g, \phi)^2 := \inf_{\mathcal{T}_* \in \mathbb{T}_N} \inf_{W_* \in \mathcal{S}^1(\mathcal{T}_*)} \big(\|\nabla(u - W_*)\|_{L^2(\Omega)}^2 + \|h_*^{1/2}(g - W_*|_\Gamma)'\|_{L^2(\Gamma_D)}^2\big) \\ + \operatorname{osc}_{\mathcal{T},*}^2 + \operatorname{osc}_{N,*}^2, \end{aligned} \tag{13}$$

the adaptively generated solutions also yield convergence order $\mathcal{O}(N^{-s})$, i.e.

$$\|u - U_\ell\|_{H^1(\Omega)} \lesssim \big(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \|h_\ell^{1/2}(g - g_\ell)'\|_{L^2(\Gamma_D)}^2\big)^{1/2} \lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s}. \tag{14}$$

Here, $\mathbb{T}_N$ denotes the set of all triangulations $\mathcal{T}_*$ which can be obtained by local refinement of the initial mesh $\mathcal{T}_0$ such that $\#\mathcal{T}_* - \#\mathcal{T}_0 \leq N$. Moreover, $\operatorname{osc}_{\mathcal{T},*}$ and $\operatorname{osc}_{N,*}$ denote the data oscillations of the volume data f and the Neumann data ϕ, see Section 3.1. The resolution of the Dirichlet data g is part of the error norm. With $h_*|_E = |E|$ the local edge-length function $h_* \in L^\infty(\Gamma_D)$, the $h_*^{1/2}$-weighted L^2-norm is a realization of the $H^{1/2}(\Gamma_D)$-norm, see [AGP, CP2, EFGP].

The ingredients for the proof are the observation that the proposed marking strategies are optimal (Proposition 14 and Proposition 15) and the Céa-type estimate

$$\begin{aligned} &\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \|h_\ell^{1/2}(g - g_\ell)'\|_{L^2(\Gamma_D)}^2 \\ &\qquad \leq C_{\rm cea} \inf_{W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)} \big(\|\nabla(u - W_\ell)\|_{L^2(\Omega)}^2 + \|h_\ell^{1/2}(g - W_\ell|_\Gamma)'\|_{L^2(\Gamma_D)}^2\big) \end{aligned} \tag{15}$$

for the Galerkin solution $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ in Lemma 16.

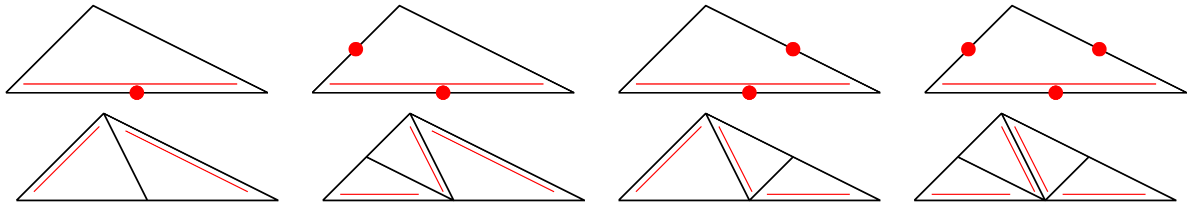

FIGURE 1. For each triangle $T \in \mathcal{T}_\ell$, there is one fixed *reference edge*, indicated by the double line (left, top). Refinement of T is done by bisecting the reference edge, where its midpoint becomes a new node. The reference edges of the son triangles $T' \in \mathcal{T}_{\ell+1}$ are opposite to this newest vertex (left, bottom). To avoid hanging nodes, one proceeds as follows: We assume that certain edges of T, but at least the reference edge, are marked for refinement (top). Using iterated newest vertex bisection, the element is then split into 2, 3, or 4 son triangles (bottom).

1.6. Outline. The remainder of this paper is organized as follows: We first collect some necessary preliminaries on, e.g., newest vertex bisection (Section 2.2) and the Scott-Zhang quasi-interpolation operator (Section 2.3). Section 3 contains the analysis of the a posteriori error estimators ρ_ℓ from (5)–(6) and ϱ_ℓ from (7)–(8). Moreover, we state two possible versions of the adaptive Algorithm in Section 3.4–3.5. The convergence of both algorithms in proven in Section 4, while the quasi-optimality results are found in Section 5. Whereas the major part of the paper is concerned with the 2D model problem, a short Section 6 considers convergence of AFEM for 3D. Finally, some numerical experiments conclude the work.

2. PRELIMINARIES

2.1. Notation. Throughout, $\mathcal{T}_\ell$ denotes a regular triangulation which is obtained by ℓ steps of (local) newest vertex bisection for a given initial triangulation $\mathcal{T}_0$. By $\mathcal{K}_\ell$, we denote the set of all nodes of $\mathcal{T}_\ell$. By $\mathcal{E}_\ell$, we denote the set of all edges of $\mathcal{T}_\ell$ which is split into the interior edges $\mathcal{E}_\ell^\Omega = \big\{E \in \mathcal{E}_\ell \,:\, E \cap \Omega \neq \emptyset\big\}$ and boundary edges $\mathcal{E}_\ell^\Gamma = \mathcal{E}_\ell \backslash \mathcal{E}_\ell^\Omega$. We assume that the partition of Γ into Dirichlet boundary Γ_D and Neumann boundary Γ_N is resolved, i.e. $\mathcal{E}_\ell^\Gamma$ is split into $\mathcal{E}_\ell^D = \big\{E \in \mathcal{E}_\ell \,:\, E \subseteq \overline{\Gamma}_D\big\}$ and $\mathcal{E}_\ell^N = \big\{E \in \mathcal{E}_\ell \,:\, E \subseteq \overline{\Gamma}_N\big\}$. Note that $\mathcal{E}_\ell^D$ (resp. $\mathcal{E}_\ell^N$) provides a partition of Γ_D (resp. Γ_N).

For a node $z \in \mathcal{K}_\ell$, the corresponding patch is defined by

$$\omega_{\ell,z} = \bigcup \big\{T \in \mathcal{T}_\ell \,:\, z \in \partial T\big\}. \tag{16}$$

For an edge $E \in \mathcal{E}_\ell$, the edge patch is defined by

$$\omega_{\ell,E} = \bigcup \big\{T \in \mathcal{T}_\ell \,:\, E \subset \partial T\big\}. \tag{17}$$

Moreover, for a given node $z \in \mathcal{K}_\ell$,

$$\mathcal{E}_{\ell,z} = \bigcup \big\{E \in \mathcal{E}_\ell \,:\, z \in E\big\} \tag{18}$$

denotes the star of edges originating at z.

2.2. Newest vertex bisection. Throughout, we assume that newest vertex bisection is used for mesh-refinement, see Figure 1. Let $\mathcal{T}_\ell$ be a given mesh and $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ an arbitrary set of marked edges. Then,

$$\mathcal{T}_{\ell+1} = \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell) \tag{19}$$

denotes the coarsest regular triangulation such that all marked edges $E \in \mathcal{M}_\ell$ have been bisected. Moreover, we write

$$\mathcal{T}_* = \texttt{refine}(\mathcal{T}_\ell) \tag{20}$$

if $\mathcal{T}_*$ is a finite refinement of $\mathcal{T}_\ell$, i.e., there are finitely many triangulations $\mathcal{T}_{\ell+1}, \dots, \mathcal{T}_n$ and sets of marked edges $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell, \dots, \mathcal{M}_{n-1} \subseteq \mathcal{T}_{n-1}$ such that $\mathcal{T}_* = \mathcal{T}_n$ and $\mathcal{T}_{j+1} = \texttt{refine}(\mathcal{T}_j, \mathcal{M}_j)$ for all $j = \ell, \dots, n-1$.

We stress that, for a fixed initial mesh $\mathcal{T}_0$, only finitely many shapes of triangles $T \in \mathcal{T}_\ell$ appear. In particular, only finitely many shapes of patches (16)–(17) appear. This observation will be used below. Moreover, newest vertex bisection guarantees that any sequence $\mathcal{T}_\ell$ of generated meshes with $\mathcal{T}_{\ell+1} = \texttt{refine}(\mathcal{T}_\ell)$ is uniformly shape regular in the sense of

$$\sup_{\ell\in\mathbb{N}} \sigma(\mathcal{T}_\ell) < \infty, \quad \text{where} \quad \sigma(\mathcal{T}_\ell) = \max_{T\in\mathcal{T}} \frac{\mathrm{diam}(T)^2}{|T|}. \tag{21}$$

Further details are found in [V, Chapter 5].

2.3. Scott-Zhang quasi-interpolation and discrete lifting operator. Our analysis below makes heavy use of the Scott-Zhang projection $P_\ell : H^1(\Omega) \to \mathcal{S}^1(\mathcal{T}_\ell)$ from [SZ]: For all nodes $z \in \mathcal{K}_\ell$, one chooses an edge $E_z \in \mathcal{E}_\ell$ with $z \in E_z$. For $z \in \Gamma$, this choice is restricted to $E_z \subset \Gamma$. Moreover, for $z \in \overline{\Gamma}_D$, we even enforce $E_z \subset \overline{\Gamma}_D$. For $w \in H^1(\Omega)$, $P_\ell w$ is then defined by

$$(P_\ell w)(z) := \langle \psi_z \,,\, w \rangle_{E_z},$$

for a node $z \in \mathcal{K}_\ell$. Here, $\psi_z \in L^2(E_z)$ denotes the dual basis function defined by $\langle \psi_z \,,\, \varphi_{z'} \rangle_{E_z} = \delta_{zz'}$, and $\varphi_z \in \mathcal{S}^1(\mathcal{T}_\ell)$ denotes the hat function associated with $z \in \mathcal{K}_\ell$. By definition, we then have the following projection properties

- $P_\ell W_\ell = W_\ell$ for all $W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$,
- $(P_\ell w)|_\Gamma = w|_\Gamma$ for all $w \in H^1(\Omega)$ and $W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ with $w|_\Gamma = W_\ell|_\Gamma$,
- $(P_\ell w)|_{\Gamma_D} = w|_{\Gamma_D}$ for all $w \in H^1(\Omega)$ and $W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ with $w|_{\Gamma_D} = W_\ell|_{\Gamma_D}$,

i.e. the projection P_ℓ preserves discrete (Dirichlet) boundary data. Moreover, P_ℓ satisfies the following stability property

$$\|(1 - P_\ell)w\|_{H^1(\Omega)} \le C_{\mathrm{sz}} \, \|\nabla w\|_{L^2(\Omega)} \quad \text{for all } w \in H^1(\Omega) \tag{22}$$

and approximation property

$$\|(1 - P_\ell)w\|_{L^2(\Omega)} \le C_{\mathrm{sz}} \, \|h_\ell \nabla w\|_{L^2(\Omega)} \quad \text{for all } w \in H^1(\Omega) \tag{23}$$

where $C_{\mathrm{sz}} > 0$ depends only on $\sigma(\mathcal{T}_\ell)$. Together with the projection property onto $\mathcal{S}^1(\mathcal{T}_\ell)$, it is an easy consequence of the stability (22) of P_ℓ that

$$\|(1 - P_\ell)w\|_{H^1(\Omega)} = \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|(1 - P_\ell)(w - W_\ell)\|_{H^1(\Omega)} \lesssim \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(w - W_\ell)\|_{L^2(\Omega)} \tag{24}$$

for all $w \in H^1(\Omega)$. In particular, P_ℓ is quasi-optimal in the sense of the Céa lemma with respect to $\|\cdot\|_{H^1(\Omega)}$ and $\|\nabla(\cdot)\|_{L^2(\Omega)}$, i.e.

$$\begin{aligned}\|(1-P_\ell)w\|_{H^1(\Omega)} &\lesssim \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|w-W_\ell\|_{H^1(\Omega)},\\ \|\nabla(1-P_\ell)w\|_{L^2(\Omega)} &\lesssim \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(w-W_\ell)\|_{L^2(\Omega)}.\end{aligned} \tag{25}$$

Moreover, P_ℓ allows to define a discrete lifting operator

$$\mathcal{L}_\ell := P_\ell\mathcal{L} : \mathcal{S}^1(\mathcal{T}_\ell|_\Gamma) \to \mathcal{S}^1(\mathcal{T}_\ell), \quad \text{i.e. } \mathcal{L}_\ell(W_\ell|_\Gamma)|_\Gamma = W_\ell|_\Gamma \quad \text{for all } W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell) \tag{26}$$

whose operator norm is uniformly bounded in terms of $\sigma(\mathcal{T}_\ell)$. Here, $\mathcal{L} \in L(H^{1/2}(\Gamma); H^1(\Omega))$ denotes an arbitrary lifting operator, i.e. $(\mathcal{L}w)|_\Gamma = w$ for all $w \in H^{1/2}(\Gamma)$, see e.g. [McL].

Finally, we put emphasis on the fact that our definition of P_ℓ also provides an operator $P_\ell = P_\ell^\Gamma : L^2(\Gamma) \to \mathcal{S}^1(\mathcal{T}_\ell|_\Gamma)$ which is consistent in the sense that $(P_\ell v)|_\Gamma = P_\ell^\Gamma(v|_\Gamma)$ for all $v \in H^1(\Omega)$. Using the definition of $H^{1/2}(\Gamma)$ as the trace space of $H^1(\Omega)$ and the stability (22), we see

$$\begin{aligned}\|\widehat{g} - P_\ell\widehat{g}\|_{H^{1/2}(\Gamma)} &:= \inf\big\{\|w\|_{H^1(\Omega)} \,:\, w \in H^1(\Omega), w|_\Gamma = \widehat{g} - P_\ell\widehat{g}\big\}\\ &\le \inf\big\{\|w - P_\ell w\|_{H^1(\Omega)} \,:\, w \in H^1(\Omega), w|_\Gamma = \widehat{g}\big\}\\ &\lesssim \inf\big\{\|\nabla w\|_{L^2(\Omega)} \,:\, w \in H^1(\Omega), w|_\Gamma = \widehat{g}\big\}\\ &\le \inf\big\{\|w\|_{H^1(\Omega)} \,:\, w \in H^1(\Omega), w|_\Gamma = \widehat{g}\big\} = \|\widehat{g}\|_{H^{1/2}(\Gamma)}\end{aligned}$$

for all $\widehat{g} \in H^{1/2}(\Gamma)$, i.e. $P_\ell : H^{1/2}(\Gamma) \to \mathcal{S}^1(\mathcal{T}_\ell|_\Gamma)$ is a continuous projection with respect to the $H^{1/2}$-norm. In particular, P_ℓ also provides a continuous projection $P_\ell = P_\ell^D : H^{1/2}(\Gamma_D) \to \mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D})$, since

$$\begin{aligned}\|g - P_\ell g\|_{H^{1/2}(\Gamma_D)} &= \inf\big\{\|\widehat{g} - P_\ell\widehat{g}\|_{H^{1/2}(\Gamma)} \,:\, \widehat{g} \in H^{1/2}(\Gamma), \widehat{g}|_{\Gamma_D} = g\big\}\\ &\lesssim \inf\big\{\|\widehat{g}\|_{H^{1/2}(\Gamma)} \,:\, \widehat{g} \in H^{1/2}(\Gamma), \widehat{g}|_{\Gamma_D} = g\big\} = \|g\|_{H^{1/2}(\Gamma_D)}\end{aligned}$$

for all $g \in H^{1/2}(\Gamma_D)$. As before, this definition is consistent with the previous notation of P_ℓ since $(P_\ell^\Gamma\widehat{g})|_{\Gamma_D} = P_\ell^D(\widehat{g}|_{\Gamma_D})$ for all $\widehat{g} \in H^{1/2}(\Gamma)$.

3. A Posteriori Error Estimation and Adaptive Mesh-Refinement

3.1. Data oscillations. We start with the element data oscillations

$$\mathrm{osc}_{\mathcal{T},\ell}^2 := \sum_{T\in\mathcal{T}_\ell} \mathrm{osc}_{\mathcal{T},\ell}(T)^2, \text{ where } \mathrm{osc}_{\mathcal{T},\ell}(T)^2 := |T|\,\|f - f_T\|_{L^2(T)}^2 \quad \text{for all } T \in \mathcal{T}_\ell \tag{27}$$

and where $f_T := |T|^{-1}\int_T f\,dx \in \mathbb{R}$ denotes the integral mean over an element $T \in \mathcal{T}_\ell$. These arise in the efficiency estimate for residual error estimators.

Our residual error estimator will involve the edge data oscillations

$$\mathrm{osc}_{\mathcal{E},\ell}^2 := \sum_{E\in\mathcal{E}_\ell^\Omega} \mathrm{osc}_{\mathcal{E},\ell}(E)^2, \text{ where } \mathrm{osc}_{\mathcal{E},\ell}(E)^2 := |\omega_{\ell,E}|\,\|f - f_{\omega_{\ell,E}}\|_{L^2(\omega_{\ell,E})}^2 \text{ for all } E \in \mathcal{E}_\ell^\Omega. \tag{28}$$

Here, $\omega_{\ell,E} \subset \Omega$ is the edge patch, and $f_{\omega_{\ell,E}} \in \mathbb{R}$ is the corresponding integral mean of f.

For the analysis, we shall additionally need the node data oscillations

$$\mathrm{osc}_{\mathcal{K},\ell}^2 := \sum_{z\in\mathcal{K}_\ell} \mathrm{osc}_{\mathcal{K},\ell}(z)^2, \text{ where } \mathrm{osc}_{\mathcal{K},\ell}(z)^2 := |\omega_{\ell,z}|\,\|f - f_{\omega_{\ell,z}}\|_{L^2(\omega_{\ell,z})}^2 \text{ for all } z \in \mathcal{K}_\ell. \tag{29}$$

Here, $\omega_{\ell,z} \subset \Omega$ is the node patch, and $f_{\omega_{\ell,z}} \in \mathbb{R}$ is the corresponding integral mean of f.

Moreover, the efficiency needs the Neumann data oscillations

$$\mathrm{osc}_{N,\ell}^2 := \sum_{E\in\mathcal{E}_\ell^N} \mathrm{osc}_{N,\ell}(E)^2, \text{ where } \mathrm{osc}_{N,\ell}(E)^2 := |E|\,\|\phi-\phi_E\|_{L^2(E)}^2 \text{ for all } E\in\mathcal{E}_\ell^N \tag{30}$$

and where $\phi_E := |E|^{-1}\int_E \phi\,dx$ denotes the integral mean over an edge $E\in\mathcal{E}_\ell^N$.

Finally, the approximation of the Dirichlet data $g\approx g_\ell$ is controlled by the Dirichlet data oscillations

$$\mathrm{osc}_{D,\ell} := \sum_{E\in\mathcal{E}_\ell^D} \mathrm{osc}_{D,\ell}(E)^2, \text{ where } \mathrm{osc}_{D,\ell}(E)^2 := |E|\|(g-g_\ell)'\|_{L^2(E)}^2 \text{ for all } E\in\mathcal{E}_\ell^D. \tag{31}$$

Here, $(\cdot)'$ denotes the arclength derivative. The following result is found in [AGP, Lemma 3.3].

Lemma 1. *Let $g\in H^1(\Gamma_D)$ and let g_ℓ denote the nodal interpoland of g_ℓ on $\overline{\Gamma}_D$. Then,*

$$\|g-g_\ell\|_{H^{1/2}(\Gamma_D)} \le C_1\,\mathrm{osc}_{D,\ell}, \tag{32}$$

where the constant $C_1>0$ depends only on the shape regularity constant $\sigma(\mathcal{T}_\ell)$ and Ω. □

Finally, recall that, on the 1D manifold Γ_D, the derivative of the nodal interpoland is the elementwise best approximation of the derivative by piecewise constants, i.e.,

$$\|(g-g_\ell)'\|_{L^2(E)} = \min_{c\in\mathbb{R}} \|g'-c\|_{L^2(E)} \quad \text{for all } E\in\mathcal{E}_\ell^D. \tag{33}$$

According to the elementwise Pythagoras theorem, this implies

$$\|(g-g_\ell)'\|_{L^2(E)}^2 + \|(g_\ell-\widetilde{g}_\ell)'\|_{L^2(E)}^2 = \|(g-\widetilde{g}_\ell)'\|_{L^2(E)}^2 \text{ for all } \widetilde{g}_\ell\in\mathcal{S}^1(\mathcal{T}_\ell|_\Gamma) \tag{34}$$

and all Dirichlet edges $E\in\mathcal{E}_\ell^D$. This observation will be crucial in the analysis below.

To keep the notation simple, we extend the Dirichlet and the Neumann data oscillations from (30)–(31) by zero to all edges $E\in\mathcal{E}_\ell$, e.g. $\mathrm{osc}_{D,\ell}(E)=0$ for $E\in\mathcal{E}_\ell\backslash\mathcal{E}_\ell^D$. Moreover, we will write

$$\mathrm{osc}_{\mathcal{T},\ell}(\omega_{\ell,z})^2 = \sum_{\substack{T\in\mathcal{T}_\ell\\ T\subset\omega_{\ell,z}}} \mathrm{osc}_{\mathcal{T},\ell}(T)^2 \quad \text{resp.} \quad \mathrm{osc}_{N,\ell}(\mathcal{E}_{\ell,z})^2 = \sum_{\substack{E\in\mathcal{E}_\ell^N\\ E\subset\mathcal{E}_{\ell,z}}} \mathrm{osc}_{N,\ell}(E)^2 \tag{35}$$

to abbreviate the notation.

3.2. Element-based residual error estimator. Our first proposition states reliability and efficiency of the error estimator ρ_ℓ from (5)–(6).

Proposition 2 (reliability and efficiency of ρ_ℓ). *The error estimator ρ_ℓ is reliable*

$$\|u-U_\ell\|_{H^1(\Omega)} \le C_2\,\rho_\ell \tag{36}$$

and efficient

$$C_3^{-1}\,\rho_\ell \le \big(\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \mathrm{osc}_{\mathcal{T},\ell}^2 + \mathrm{osc}_{N,\ell}^2 + \mathrm{osc}_{D,\ell}^2\big)^{1/2}. \tag{37}$$

The constants $C_2, C_3>0$ depend only on the shape regularity constant $\sigma(\mathcal{T}_\ell)$ and on Ω.

Sketch of proof. We consider a continuous auxiliary problem

$$\begin{aligned} -\Delta w &= 0 && \text{in } \Omega,\\ w &= g-g_\ell && \text{on } \Gamma_D,\\ \partial_n w &= 0 && \text{on } \Gamma_N, \end{aligned} \tag{38}$$

with unique solution $w \in H^1(\Omega)$. We then have norm equivalence $\|w\|_{H^1(\Omega)} \simeq \|g - g_\ell\|_{H^{1/2}(\Gamma_D)}$ as well as $u - U_\ell - w \in H^1_D(\Omega)$. From this, we obtain

$$\|u - U_\ell\|^2_{H^1(\Omega)} \lesssim \|\nabla(u - U_\ell - w)\|^2_{L^2(\Omega)} + \|g - g_\ell\|^2_{H^{1/2}(\Gamma_D)}.$$

Whereas the second term is controlled by Lemma 1, the first can be handled as for homogeneous Dirichlet data, i.e. use of the Galerkin orthogonality combined with approximation estimates for a Clément-type quasi-interpolation operator. Details are found e.g. in [BCD]. This proves reliability (36).

By use of bubble functions and local scaling arguments, one obtains the estimates

$$\begin{aligned} |T|\,\|f\|^2_{L^2(T)} &\lesssim \|\nabla(u - U_\ell)\|^2_{L^2(T)} + \mathrm{osc}_{\mathcal{T},\ell}(T)^2 + \mathrm{osc}_{N,\ell}(\partial T \cap \Gamma_N),\\ |T|^{1/2}\,\|[\partial_n U_\ell]\|^2_{L^2(E\cap\Omega)} &\lesssim \|\nabla(u - U_\ell)\|^2_{L^2(\omega_{\ell,E})} + \mathrm{osc}_{\mathcal{T},\ell}(\omega_{\ell,E})^2\\ |T|^{1/2}\,\|\phi - \partial_n U_\ell\|^2_{L^2(E\cap\Gamma_N)} &\lesssim \|\nabla(u - U_\ell)\|^2_{L^2(\omega_{\ell,E})} + \mathrm{osc}_{\mathcal{T},\ell}(\omega_{\ell,E})^2 + \mathrm{osc}_{N,\ell}(E \cap \Gamma_N)^2 \end{aligned}$$

where $\omega_{\ell,E}$ denotes the edge patch of $E \in \mathcal{E}_\ell$. Details are found e.g. in [AO, V]. Summing these estimates over all elements, one obtains the efficiency estimate (37). □

Proposition 3 (discrete local reliability of ρ_ℓ)**.** *Let $\mathcal{T}_* = \texttt{refine}(\mathcal{T}_\ell)$ be an arbitrary refinement of $\mathcal{T}_\ell$ with associated Galerkin solution $U_* \in \mathcal{S}^1(\mathcal{T}_*)$. Let $\mathcal{R}_\ell(\mathcal{T}_*) := \mathcal{T}_\ell \backslash \mathcal{T}_*$ be the set of all elements $T \in \mathcal{T}_\ell$ which are refined to generate $\mathcal{T}_*$. Then, there holds*

$$\|U_* - U_\ell\|_{H^1(\Omega)} \le C_4\, \rho_\ell(\mathcal{R}_\ell(\mathcal{T}_*)) \tag{39}$$

with some constant $C_4 > 0$ which depends only on $\sigma(\mathcal{T}_\ell)$ and Ω.

Proof. We consider a discrete auxiliary problem

$$\langle \nabla W_*\,,\,\nabla V_*\rangle_\Omega = 0 \quad \text{for all } V_* \in \mathcal{S}^1_D(\mathcal{T}_*)$$

with unique solution $W_* \in \mathcal{S}^1(\mathcal{T}_*)$ with $W_*|_{\Gamma_D} = g_* - g_\ell$. To estimate the H^1-norm of W_* in terms of the boundary data, let $\mathcal{L}_* : H^{1/2}(\Gamma) \to \mathcal{S}^1(\mathcal{T}_*)$ denote the discrete lifting operator from (26). Let $\widehat{g}_*, \widehat{g}_\ell \in H^{1/2}(\Gamma)$ be arbitrary extensions of g_* and g_ℓ, respectively. Then, we have $V_* = W_* - \mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell) \in \mathcal{S}^1_D(\mathcal{T}_*)$. According to the triangle inequality and a Poincaré inequality for $V_* \in \mathcal{S}^1_D(\mathcal{T}_*)$, we first observe

$$\begin{aligned} \|W_*\|_{L^2(\Omega)} &\le \|V_*\|_{L^2(\Omega)} + \|\mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell)\|_{L^2(\Omega)}\\ &\lesssim \|\nabla V_*\|_{L^2(\Omega)} + \|\mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell)\|_{L^2(\Omega)}\\ &\lesssim \|\nabla W_*\|_{L^2(\Omega)} + \|\mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell)\|_{H^1(\Omega)}. \end{aligned}$$

Moreover, the variational formulation for $W_* \in \mathcal{S}^1(\mathcal{T}_*)$ yields

$$0 = \langle \nabla W_*\,,\,\nabla V_*\rangle_\Omega = \|\nabla W_*\|^2_{L^2(\Omega)} - \langle \nabla W_*\,,\,\nabla \mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell)\rangle_\Omega,$$

whence by the Cauchy-Schwarz inequality

$$\|\nabla W_*\|_{L^2(\Omega)} \le \|\nabla \mathcal{L}_*(\widehat{g}_* - \widehat{g}_\ell)\|_{L^2(\Omega)} \lesssim \|\widehat{g}_* - \widehat{g}_\ell\|_{H^{1/2}(\Gamma)}.$$

Altogether, this proves $\|W_*\|_{H^1(\Omega)} \lesssim \|\widehat{g}_* - \widehat{g}_\ell\|_{H^{1/2}(\Gamma)}$. Since the extensions $\widehat{g}_*, \widehat{g}_\ell$ were arbitrary and by definition of the $H^{1/2}(\Gamma_D)$-norm, this proves

$$\|W_*\|_{H^1(\Omega)} \lesssim \|g_* - g_\ell\|_{H^{1/2}(\Gamma_D)} \lesssim \|h_\ell^{1/2}(g_* - g_\ell)'\|_{L^2(\Gamma_D)}, \tag{40}$$

where we have finally used that g_ℓ is also the nodal interpoland of g_* so that Lemma 1 applies. For an element $T \in \mathcal{T}_\ell \cap \mathcal{T}_*$ holds $g_*|_{\partial T\cap\Gamma_D} = g_\ell|_{\partial T\cap\Gamma_D}$, and the last term thus satisfies

$$\|h_\ell^{1/2}(g_* - g_\ell)'\|^2_{L^2(\Gamma_D)} \simeq \sum_{T\in\mathcal{T}_\ell} |T|^{1/2}\|(g_* - g_\ell)'\|^2_{L^2(\partial T\cap\Gamma_D)} = \sum_{T\in\mathcal{R}_\ell(\mathcal{T}_*)} |T|^{1/2}\|(g_* - g_\ell)'\|^2_{L^2(\partial T\cap\Gamma_D)}.$$

With the orthogonality relation (34) applied for $g_* \in \mathcal{S}^1(\mathcal{T}_*|_{\Gamma_D})$, we see

$$\|W_*\|^2_{H^1(\Omega)} \lesssim \sum_{T\in\mathcal{R}_\ell(\mathcal{T}_*)} |T|^{1/2}\|(g_* - g_\ell)'\|^2_{L^2(\partial T\cap\Gamma_D)} \le \sum_{T\in\mathcal{R}_\ell(\mathcal{T}_*)} |T|^{1/2}\|(g - g_\ell)'\|^2_{L^2(\partial T\cap\Gamma_D)}.$$

Finally, we observe $U_* - U_\ell - W_* \in \mathcal{S}^1_D(\mathcal{T}_*)$ with

$$\langle\nabla(U_* - U_\ell - W_*)\,,\,\nabla V_\ell\rangle = 0 \quad \text{for all } V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell).$$

Arguing as in [CKNS, Lemma 3.6], we see

$$\begin{aligned}&\|\nabla(U_* - U_\ell - W_*)\|^2_{L^2(\Omega)}\\ &\quad\lesssim \sum_{T\in\mathcal{R}_\ell(\mathcal{T}_*)} \big(|T|\,\|f\|^2_{L^2(T)} + |T|^{1/2}\,\|[\partial_n U_\ell]\|^2_{L^2(\partial T\cap\Omega)} + |T|^{1/2}\,\|\phi - \partial_n U_\ell\|^2_{L^2(\partial T\cap\Gamma_N)}\big)\end{aligned}$$

Finally, we again use the triangle inequality and the Poincaré inequality to see

$$\|U_* - U_\ell\|^2_{H^1(\Omega)} \lesssim \|W_*\|^2_{H^1(\Omega)} + \|\nabla(U_* - U_\ell - W_*)\|^2_{L^2(\Omega)}$$

and thus obtain the discrete local reliability (39). The constant $C_4 > 0$ depends only on $C_1 > 0$ and on local estimates for the Scott-Zhang projection which are controlled by boundedness of $\sigma(\mathcal{T}_\ell)$. □

3.3. Edge-based residual error estimator. In the following, we define an edge-based error estimator ϱ_ℓ which is (locally) equivalent to the element-based error estimator ρ_ℓ from the previous section. The main advantage is that ϱ_ℓ replaces the volume residuals

$$\operatorname{res}_\ell(T) := |T|\,\|f\|_{L^2(T)} \tag{41}$$

by the edge oscillations $\operatorname{osc}_{\mathcal{E},\ell}$. We define the edge jump contributions

$$\eta_\ell(E)^2 := \begin{cases}|E|\,\|[\partial_n U_\ell]\|^2_{L^2(E)} & \text{for } E \in \mathcal{E}^\Omega_\ell,\\ |E|\,\|\phi - \partial_n U_\ell\|^2_{L^2(E)} & \text{for } E \in \mathcal{E}^N_\ell\end{cases} \tag{42}$$

where $[\cdot]$ denotes the jump across an interior edge. Together with the edge oscillations from (28) and the Dirichlet oscillations from (31), our version of the residual error estimator from (7)–(8) reads

$$\varrho_\ell^2 = \sum_{E\in\mathcal{E}_\ell} \varrho_\ell(E)^2 = \sum_{E\in\mathcal{E}^\Omega_\ell\cup\mathcal{E}^N_\ell} \eta_\ell(E)^2 + \sum_{E\in\mathcal{E}^\Omega_\ell} \operatorname{osc}_{\mathcal{E},\ell}(E)^2 + \sum_{E\in\mathcal{E}^D_\ell} \operatorname{osc}_{D,\ell}(E)^2. \tag{43}$$

Note that $\operatorname{osc}_{\mathcal{E},\ell}(\mathcal{E}_{\ell,z})$, $\eta_\ell(\mathcal{E}_{\ell,z})$, and $\operatorname{res}_\ell(\omega_{\ell,E})$ are defined analogously to (35). The following lemma implies local equivalence of the estimators ρ_ℓ and ϱ_ℓ.

Lemma 4. *The following local estimates hold:*

(i) $\operatorname{osc}_{\mathcal{T},\ell}(\omega_{\ell,E}) \le \operatorname{osc}_{\mathcal{E},\ell}(E) \le C_5\operatorname{res}_\ell(\omega_{\ell,E})$ *for all* $E \in \mathcal{E}^\Omega_\ell$.
(ii) $\operatorname{res}_\ell(\omega_{\ell,z}) \le C_6\big(\eta_\ell(\mathcal{E}_{\ell,z}) + \operatorname{osc}_{\mathcal{K},\ell}(z)\big)$ *for all* $z \in \mathcal{K}_\ell$.
(iii) $C_7^{-1}\operatorname{osc}_{\mathcal{E},\ell}(\mathcal{E}_{\ell,z}) \le \operatorname{osc}_{\mathcal{K},\ell}(z) \le C_8\operatorname{osc}_{\mathcal{E},\ell}(\mathcal{E}_{\ell,z})$ *for all* $z \in \mathcal{K}_\ell$.

The constants $C_5, C_6, C_7 > 0$ depend only on the shape regularity constant $\sigma(\mathcal{T}_\ell)$, whereas $C_8 > 0$ depends on the use of newest vertex bisection and the initial mesh $\mathcal{T}_0$.

Sketch of proof. The proof of (i) follows from the fact that taking the integral mean f_ω is the L^2 best approximation by a constant, i.e.

$$\|f - f_\omega\|_{L^2(\omega)} = \min_{c\in\mathbb{R}} \|f - c\|_{L^2(\omega)} \quad \text{for all measurable } \omega \subseteq \Omega,$$

and that the area of neighboring elements can only change up to $\sigma(\mathcal{T}_\ell)$. The estimate (ii) is well-known and found, e.g., in [KS, Section 2.2.4]. The lower estimate in (iii) follows from the same arguments as (i), namely

$$\|f - f_{\omega_{\ell,E}}\|_{L^2(\omega_{\ell,E})} \le \|f - f_{\omega_{\ell,z}}\|_{L^2(\omega_{\ell,E})} \le \|f - f_{\omega_{\ell,z}}\|_{L^2(\omega_{\ell,z})}$$

and the fact that —up to shape regularity— only finitely many edges belong to $\mathcal{E}_{\ell,z}$. For f being a piecewise polynomial, the upper estimate in (iii) follows from a scaling argument since both terms, $\mathrm{osc}_{\mathcal{E},\ell}(\mathcal{E}_{\ell,z}) \simeq \mathrm{osc}_{\mathcal{K},\ell}(z)$ define seminorms on $\mathcal{P}^p(\{T \in \mathcal{T}_\ell \,:\, z \in T\})$ with kernel being the constant functions. Note that the equivalence constants depend on the shape of the node patch $\omega_{\ell,z}$, but newest vertex bisection leads only to finitely many shapes of the patches. For arbitrary $f \in L^2(\Omega)$, we first observe that the $\mathcal{T}_\ell$-piecewise integral mean $f_\ell \in \mathcal{P}^0(\mathcal{T}_\ell)$, defined by $f_\ell|_T = f_T$ for all $T \in \mathcal{T}_\ell$, satisfies $(f_\ell)_{\omega_{\ell,E}} = f_{\omega_{\ell,E}}$ as well as $(f_\ell)_{\omega_{\ell,z}} = f_{\omega_{\ell,z}}$, e.g.

$$(f_\ell)_{\omega_{\ell,z}} = \frac{1}{|\omega_{\ell,z}|}\int_{\omega_{\ell,z}} f_\ell\,dx = \frac{1}{|\omega_{\ell,z}|}\sum_{T\subset\omega_{\ell,z}}\int_T f_\ell\,dx = \frac{1}{|\omega_{\ell,z}|}\sum_{T\subset\omega_{\ell,z}}\int_T f\,dx = f_{\omega_{\ell,z}}.$$

This and the Pythagoras theorem for the integral mean f_ℓ prove

$$\begin{aligned}
\|f - f_{\omega_{\ell,z}}\|^2_{L^2(\omega_{\ell,z})} &= \|f - f_\ell\|^2_{L^2(\omega_{\ell,z})} + \|f_\ell - f_{\omega_{\ell,z}}\|^2_{L^2(\omega_{\ell,z})} \\
&\lesssim \sum_{E\in\mathcal{E}_{\ell,z}} \|f - f_\ell\|^2_{L^2(\omega_{\ell,E})} + \sum_{E\in\mathcal{E}_{\ell,z}} \|f_\ell - f_{\omega_{\ell,z}}\|^2_{L^2(\omega_{\ell,E})} \\
&= \sum_{E\in\mathcal{E}_{\ell,z}} \|f - f_{\omega_{\ell,z}}\|^2_{L^2(\omega_{\ell,E})}.
\end{aligned}$$

Scaling with $|\omega_{\ell,z}| \simeq |\omega_{\ell,E}|$ concludes the proof. □

Proposition 5 (reliability and efficiency of ϱ_ℓ)**.** *The error estimator ϱ_ℓ is reliable*

$$\|u - U_\ell\|_{H^1(\Omega)} \le C_{\rm rel}\,\varrho_\ell \tag{44}$$

and efficient

$$C_{\rm eff}^{-1}\,\varrho_\ell \le \big(\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2_{\mathcal{T},\ell} + \mathrm{osc}^2_{N,\ell} + \mathrm{osc}^2_{D,\ell}\big)^{1/2}. \tag{45}$$

The constants $C_{\rm rel}, C_{\rm eff} > 0$ depend only on Ω, the use of newest vertex bisection, and the initial mesh $\mathcal{T}_0$.

Proof. With the help of the preceding lemma, we obtain equivalence $\varrho_\ell \simeq \rho_\ell$. Consequently, reliability and efficiency of ϱ_ℓ follow from the respective properties of the element-based estimator ρ_ℓ, see Proposition 2. □

Proposition 6 (discrete local reliability of ϱ_ℓ)**.** *Let $\mathcal{T}_* = \texttt{refine}(\mathcal{T}_\ell)$ be an arbitrary refinement of $\mathcal{T}_\ell$ with associated Galerkin solution $U_* \in \mathcal{S}^1(\mathcal{T}_*)$. Let $\mathcal{R}_\ell(\mathcal{T}_*) := \mathcal{T}_\ell \backslash \mathcal{T}_*$ be the set of all elements $T \in \mathcal{T}_\ell$ which are refined to generate $\mathcal{T}_*$ and*

$$\mathcal{R}_\ell(\mathcal{E}_*) := \big\{ E \in \mathcal{E}_\ell \,:\, \exists T \in \mathcal{R}_\ell(\mathcal{T}_*) \quad E \cap T \neq \emptyset \big\} \tag{46}$$

be the set of all edges which touch a refined element. Then,

$$\#\mathcal{R}_\ell(\mathcal{E}_*) \le C_{\rm ref}\, \#\mathcal{R}_\ell(\mathcal{T}_*) \tag{47}$$

and

$$\|U_* - U_\ell\|_{H^1(\Omega)} \le C_{\rm dlr}\, \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*)) \tag{48}$$

with constants $C_{\rm ref}, C_{\rm dlr} > 0$ which depend only on Ω, the use of newest vertex bisection, and the initial mesh $\mathcal{T}_0$.

Proof. According to shape regularity, the number of elements which share a node $z \in \mathcal{K}_\ell$ is uniformly bounded. Consequently, so is the number of edges which touch an element $T \in \mathcal{R}_\ell(\mathcal{T}_*)$ which will be refined. This proves the estimate $\#\mathcal{R}_\ell(\mathcal{E}_*) \le C_{\rm ref}\, \#\mathcal{R}_\ell(\mathcal{T}_*)$. To prove (48), we use the discrete local reliability of ρ_ℓ from Proposition 3. With the help of Lemma 4, each refinement indicator $\rho_\ell(T)$ for $T \in \mathcal{R}_\ell(\mathcal{T}_*)$ is dominated by finitely many indicators $\varrho_\ell(E)$ for $E \in \mathcal{R}_\ell(\mathcal{E}_*)$, where the number depends only on the shape regularity constant $\sigma(\mathcal{T}_\ell)$. □

3.4. Adaptive algorithm based on Dörfler marking. The first version of the adaptive algorithm has been well-studied in the literature mainly for element-based estimators, cf. e.g. [CKNS].

Algorithm 7. *Let adaptivity parameter $0 < \theta < 1$ and initial triangulation $\mathcal{T}_0$ be given. For each $\ell = 0, 1, 2, \dots$ do:*

- (i) *Compute discrete solution $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$.*
- (ii) *Compute refinement indicators $\varrho_\ell(E)$ for all $E \in \mathcal{E}_\ell$.*
- (iii) *Choose set $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ with minimal cardinality such that*

$$\theta\, \varrho_\ell^2 \le \varrho_\ell(\mathcal{M}_\ell)^2. \tag{49}$$

- (iv) *Generate new mesh $\mathcal{T}_{\ell+1} := \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$.*
- (v) *Update counter $\ell \mapsto \ell + 1$ and go to* (i).

3.5. Adaptive algorithm based on modified Dörfler marking. For (piecewise) smooth data $f \in H^1$ and $g \in H^2$, uniform mesh-refinement guarantees $\mathrm{osc}_{\mathcal{E},\ell} = \mathcal{O}(h^2)$ as well as $\mathrm{osc}_{D,\ell} = \mathcal{O}(h^{3/2})$, whereas the error and hence the error estimator ϱ_ℓ may at most decay as $\mathcal{O}(h)$. Consequently, we may expect that the normal jump terms dominate the error estimator [CV]. This observation led to the following version of the adaptive algorithm which has essentially been proposed in [BMS]. We stress, however, that the algorithm in [BMS, BM] is stated with node oscillations $\mathrm{osc}_{\mathcal{K},\ell}$ instead of edge oscillations $\mathrm{osc}_{\mathcal{E},\ell}$. Moreover, certain details in the proofs of [BM] seem to be dubious.

Algorithm 8. *Let adaptivity parameters $0 < \theta_1, \theta_2 < 1$ and $\vartheta > 0$ and an initial triangulation $\mathcal{T}_0$ be given. For each $\ell = 0, 1, 2, \dots$ do:*

- (i) *Compute discrete solution $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$.*

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

(iii.1) *If* $\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 \le \vartheta\,\eta_\ell^2$, *choose set* $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ *with minimal cardinality such that*

$$\theta_1\,\eta_\ell^2 \le \eta_\ell(\mathcal{M}_\ell)^2. \tag{50}$$

(iii.2) *Otherwise, choose set* $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ *with minimal cardinality such that*

$$\theta_2\,(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2) \le \mathrm{osc}_{\mathcal{E},\ell}(\mathcal{M}_\ell)^2 + \mathrm{osc}_{D,\ell}(\mathcal{M}_\ell)^2. \tag{51}$$

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

(v) *Update counter* $\ell \mapsto \ell+1$ *and go to* (i).

4. Convergence of Adaptive Algorithm

In this section, we prove a contraction property $\Delta_{\ell+1} \le \kappa\,\Delta_\ell$ for some quasi-error quantity $\Delta_\ell \simeq \varrho_\ell$. To that end, we first note that the modified Dörfler marking (50)–(51) implies the Dörfler marking (49).

Lemma 9. *If the set* $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ *satisfies the modified Dörfler marking* (50)–(51) *with parameters* $0 < \theta_1, \theta_2 < 1$ *and* $\vartheta > 0$. *Then,* $\mathcal{M}_\ell$ *satisfies the Dörfler marking* (49) *with parameter* $0 < \theta := \min\{\theta_1/(1+\vartheta)\,,\,\theta_2/(1+\vartheta^{-1})\} < 1$.

Proof. In case of $\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 \le \vartheta\,\eta_\ell^2$, it holds that $\varrho_\ell^2 \le (1+\vartheta)\,\eta_\ell^2$. This implies

$$\frac{\theta_1}{1+\vartheta}\,\varrho_\ell^2 \le \theta_1\,\eta_\ell^2 \le \eta_\ell(\mathcal{M}_\ell)^2 \le \varrho_\ell(\mathcal{M}_\ell)^2.$$

Otherwise, it holds that $\varrho_\ell^2 \le (1+\vartheta^{-1})\,(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2)$ which yields

$$\frac{\theta_2}{1+\vartheta^{-1}}\,\varrho_\ell^2 \le \theta_2\,\big(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2\big) \le \big(\mathrm{osc}_{\mathcal{E},\ell}(\mathcal{M}_\ell)^2 + \mathrm{osc}_{D,\ell}(\mathcal{M}_\ell)^2\big) \le \varrho_\ell(\mathcal{M}_\ell)^2.$$

This concludes the proof. □

Lemma 10 (equivalent error estimator)**.** *Consider the extended error estimator*

$$\widetilde{\varrho}_\ell^2 = \sum_{E\in\mathcal{E}_\ell^\Omega\cup\mathcal{E}_\ell^N} \eta_\ell(E)^2 + \sum_{E\in\mathcal{E}_\ell} \widetilde{\mathrm{osc}}_{\mathcal{E},\ell}(E)^2 + \sum_{E\in\mathcal{E}_\ell^D} \mathrm{osc}_{D,\ell}(E)^2, \tag{52}$$

where the oscillation terms $\widetilde{\mathrm{osc}}_{\mathcal{E},\ell}(E)$ *read*

$$\widetilde{\mathrm{osc}}_{\mathcal{E},\ell}(E)^2 := \begin{cases} \mathrm{osc}_{\mathcal{E},\ell}(E)^2 & \text{for } E \in \mathcal{E}_\ell^\Omega, \\ |T_E|\,\|f\|_{L^2(T_E)}^2 & \text{for } E \in \mathcal{E}_\ell^\Gamma \text{ and } T \in \mathcal{T}_\ell \text{ with } E \subset \partial T_E. \end{cases} \tag{53}$$

Then, there holds equivalence in the following sense

$$C_9^{-1}\,\widetilde{\varrho}_\ell^2 \le \varrho_\ell^2 \le \widetilde{\varrho}_\ell^2 \quad \text{and} \quad \varrho_\ell(E) \le \widetilde{\varrho}_\ell(E) \text{ for all } E \in \mathcal{E}_\ell,$$

where $C_9 \ge 1$ *depends only on* $\sigma(\mathcal{T}_\ell)$. *Particularly, if* $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ *satisfies the Dörfler marking* (49) *with* ϱ_ℓ *and* $\theta > 0$, *then* $\mathcal{M}_\ell$ *satisfies the Dörfler marking with* $\widetilde{\varrho}_\ell$ *for some modified parameter* $0 < \widetilde{\theta} := \theta/C_9 < 1$.

Proof. The estimates $\varrho_\ell(E) \le \widetilde{\varrho}_\ell(E)$ for all $E \in \mathcal{E}_\ell$ are obvious and imply $\varrho_\ell^2 \le \widetilde{\varrho}_\ell^2$. The estimate $C_9^{-1}\,\widetilde{\varrho}_\ell^2 \le \varrho_\ell^2$ follows from Lemma 4 (ii) & (iii). Now, we obtain

$$\widetilde{\theta}\,\widetilde{\varrho}_\ell^2 \le \theta\,\varrho_\ell^2 \le \varrho_\ell(\mathcal{M}_\ell)^2 \le \widetilde{\varrho}_\ell(\mathcal{M}_\ell)^2,$$

i.e. the estimator $\widetilde{\varrho}_\ell$ satisfies the Dörfler marking (49) with $\widetilde{\theta} := \theta/C_9$. □

Lemma 11 (estimator reduction). *Assume that the set $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ of marked edges satisfies the Dörfler marking (49) with ϱ_ℓ and some fixed parameter $0 < \theta < 1$ and that $\mathcal{T}_{\ell+1} = \mathtt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ is obtained by local newest vertex bisection of $\mathcal{T}_\ell$. Then, there holds the estimator reduction estimate*

$$\widetilde{\varrho}_{\ell+1}^2 \le q\,\widetilde{\varrho}_\ell^2 + C_{10}\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}^2 \tag{54}$$

with some contraction constant $q \in (0,1)$ which depends only on $\theta \in (0,1)$. The constant $C_{10} > 0$ additionally depends only on the initial mesh $\mathcal{T}_0$.

Sketch of proof. For the sake of completeness, we include the idea of the proof of (54). To keep the notation simple, we define $\eta_\ell(E) = 0$ for $E \in \mathcal{E}_\ell^D$ and $\mathrm{osc}_{D,\ell}(E) = 0$ for $E \in \mathcal{E}_\ell^\Omega \cup \mathcal{E}_\ell^N$ so that all contributions of $\widetilde{\varrho}_\ell$ are defined on the entire set of edges $\mathcal{E}_\ell$.

First, we employ a triangle inequality and the Young inequality to see

$$\begin{aligned}\widetilde{\varrho}_{\ell+1}^2 \le (1+\delta)\Big(&\sum_{E\in\mathcal{E}_{\ell+1}^\Omega} |E|\|[\partial_n U_\ell]\|_{L^2(E)}^2 + \sum_{E\in\mathcal{E}_{\ell+1}^N} |E|\|\phi - \partial_n U_\ell\|_{L^2(E)}^2\Big)\\ &+ (1+\delta^{-1})\Big(\sum_{E\in\mathcal{E}_{\ell+1}^\Omega} |E|\|[\partial_n(U_{\ell+1} - U_\ell)]\|_{L^2(E)}^2 + \sum_{E\in\mathcal{E}_{\ell+1}^N} |E|\|\partial_n(U_{\ell+1} - U_\ell)\|_{L^2(E)}^2\Big)\\ &+ \widetilde{\mathrm{osc}}_{\mathcal{E},\ell+1}^2 + \mathrm{osc}_{D,\ell+1}^2,\end{aligned}$$

where $\delta > 0$ is arbitrary. Second, a scaling argument proves

$$\sum_{E\in\mathcal{E}_{\ell+1}^\Omega} |E|\|[\partial_n(U_{\ell+1} - U_\ell)]\|_{L^2(E)}^2 + \sum_{E\in\mathcal{E}_{\ell+1}^\Gamma} |E|\|\partial_n(U_{\ell+1} - U_\ell)\|_{L^2(E)}^2 \le C\,\|\nabla(U_{\ell+1} - U_\ell)\|_{L^2(\Omega)}^2,$$

and the constant $C > 0$ depends only on $\sigma(\mathcal{T}_\ell)$. Third, we argue as in [CKNS, Corollary 3.4] to see

$$\sum_{E\in\mathcal{E}_{\ell+1}^\Omega} |E|\|[\partial_n U_\ell]\|_{L^2(E)}^2 + \sum_{E\in\mathcal{E}_{\ell+1}^N} |E|\|\phi - \partial_n U_\ell\|_{L^2(E)}^2 \le \eta_\ell^2 - \frac{1}{2}\,\eta_\ell(\mathcal{M}_\ell)^2.$$

Fourth, it is part of the proof of [AGP, Theorem 4.2] that

$$\mathrm{osc}_{D,\ell+1}^2 \le \mathrm{osc}_{D,\ell}^2 - \frac{1}{2}\,\mathrm{osc}_{D,\ell}(\mathcal{M}_\ell)^2,$$

which essentially follows from the orthogonality relation (34). Fifth, in [PP, Lemma 6] it is proven that

$$\widetilde{\mathrm{osc}}_{\mathcal{E},\ell+1}^2 \le \widetilde{\mathrm{osc}}_{\mathcal{E},\ell}^2 - \frac{1}{4}\,\widetilde{\mathrm{osc}}_{\mathcal{E},\ell}(\mathcal{M}_\ell)^2. \tag{55}$$

Plugging everything together, we see

$$\begin{aligned}\widetilde{\varrho}_{\ell+1}^2 &\le (1+\delta)\big(\widetilde{\varrho}_\ell^2 - \frac{1}{4}\,\widetilde{\varrho}_\ell(\mathcal{M}_\ell)^2\big) + C(1+\delta^{-1})\,\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2\\ &\le (1+\delta)(1-\widetilde{\theta}/4)\widetilde{\varrho}_\ell^2 + C(1+\delta^{-1})\,\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2,\end{aligned}$$

where we have used that Lemma 10 guarantees the Dörfler marking for $\widetilde{\varrho}_\ell$ in the second estimate. Finally, it only remains to choose $\delta > 0$ sufficiently small so that $q := (1+\delta)(1-\widetilde{\theta}/4) < 1$. □

The following lemma states some quasi-Galerkin orthogonality property which allows to overcome the lack of Galerkin orthogonality used in [CKNS].

Lemma 12 (quasi-Galerkin orthogonality)**.** *Let $\mathcal{T}_* = \mathtt{refine}(\mathcal{T}_\ell)$ be an arbitrary refinement of $\mathcal{T}_\ell$ with the associated Galerkin solution $U_* \in \mathcal{S}^1(\mathcal{T}_*)$. Then,*

$$2\,|\langle\nabla(u-U_*)\,,\,\nabla(U_*-U_\ell)\rangle_\Omega| \le \alpha\,\|\nabla(u-U_*)\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2, \tag{56}$$

for all $\alpha > 0$, and consequently

$$\begin{aligned}(1-\alpha)\|\nabla(u-U_*)\|_{L^2(\Omega)}^2 &\le \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 - \|\nabla(U_*-U_\ell)\|_{L^2(\Omega)}^2\\ &\quad + \alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2\end{aligned} \tag{57}$$

as well as

$$\begin{aligned}\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 &\le (1+\alpha)\|\nabla(u-U_*)\|_{L^2(\Omega)}^2 + \|\nabla(U_*-U_\ell)\|_{L^2(\Omega)}^2\\ &\quad + \alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2.\end{aligned} \tag{58}$$

The constant $C_{\rm orth} > 0$ depends only on the shape regularity of $\sigma(\mathcal{T}_\ell)$ and $\sigma(\mathcal{T}_*)$ and on Ω.

Proof. We recall the Galerkin orthogonality

$$\langle\nabla(u-U_*)\,,\,\nabla V_*\rangle_\Omega = 0 \quad \text{for all } V_* \in \mathcal{S}^1_D(\mathcal{T}_*).$$

Now, let $U_*^\ell \in \mathcal{S}^1(\mathcal{T}_*)$ be the unique Galerkin solution solution of (4) with $U_*^\ell|_{\Gamma_D} = g_\ell$. We use the Galerkin orthogonality with $V_* = U_*^\ell - U_\ell \in \mathcal{S}^1_D(\mathcal{T}_*)$. This and the Young inequality allow to estimate the L^2-scalar product by

$$\begin{aligned}2\,|\langle\nabla(u-U_*)\,,\,\nabla(U_*-U_\ell)\rangle_\Omega| &= 2\,|\langle\nabla(u-U_*)\,,\,\nabla(U_*-U_*^\ell)\rangle_\Omega|\\ &\le \alpha\,\|\nabla(u-U_*)\|_{L^2(\Omega)}^2 + \alpha^{-1}\,\|\nabla(U_*-U_*^\ell)\|_{L^2(\Omega)}^2\end{aligned}$$

for all $\alpha > 0$. To estimate the second contribution on the right-hand side, we proceed as in the proof of Proposition 3 and choose arbitrary extensions $\widehat{g}_*, \widehat{g}_\ell \in H^{1/2}(\Gamma)$ of the nodal interpolands g_*, g_ℓ from Γ_D to Γ. Then, we use the test function $V_* = (U_*-U_*^\ell)-\mathcal{L}_*(\widehat{g}_*-\widehat{g}_\ell) \in \mathcal{S}^1_D(\mathcal{T}_*)$ and the Galerkin orthogonalities for $U_*, U_*^\ell \in \mathcal{S}^1(\mathcal{T}_*)$ to see

$$0 = \langle\nabla(u-U_*^\ell)\,,\,\nabla V_*\rangle_\Omega - \langle\nabla(u-U_*)\,,\,\nabla V_*\rangle_\Omega = \langle\nabla(U_*-U_*^\ell)\,,\,\nabla V_*\rangle_\Omega.$$

Arguing as above, we obtain

$$\|\nabla(U_*-U_*^\ell)\|_{L^2(\Omega)} \lesssim \|g_*-g_\ell\|_{H^{1/2}(\Gamma_D)} \lesssim \|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}. \tag{59}$$

This concludes the proof of (56).

To verify (57)–(58), we use the identity

$$\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 = \|\nabla\big((u-U_*)+(U_*-U_\ell)\big)\|_{L^2(\Omega)}^2$$
$$= \|\nabla(u-U_*)\|_{L^2(\Omega)}^2 + 2\,\langle\nabla(u-U_*)\,,\,\nabla(U_*-U_\ell)\rangle_\Omega + \|\nabla(U_*-U_\ell)\|_{L^2(\Omega)}^2.$$

Rearranging the terms accordingly and use of the quasi-Galerkin orthogonality (56) to estimate the scalar product, concludes the proof. □

Theorem 13 (contraction of quasi-error). *For both adaptive algorithms stated in Algorithm 7 and Algorithm 8 above, there are constants $\gamma,\lambda>0$ and $0<\kappa<1$ such that the combined error quantity*

(60) $$\Delta_\ell := \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \lambda\,\mathrm{osc}_{D,\ell}^2 + \gamma\,\widetilde\varrho_\ell^2 \ge 0$$

satisfies a contraction property

(61) $$\Delta_{\ell+1} \le \kappa\,\Delta_\ell \quad \textit{for all } \ell\in\mathbb{N}_0.$$

In particular, this implies $\lim\limits_{\ell\to\infty}\varrho_\ell = 0 = \lim\limits_{\ell\to\infty}\|u-U_\ell\|_{H^1(\Omega)}$.

Proof. According to Lemma 9, we may restrict our attention to Algorithm 7 with given parameter $0<\theta<1$ since we will not use the minimality of $\mathcal{M}_\ell$ throughout the proof. Using the quasi-Galerkin orthogonality (57) with $\mathcal{T}_* = \mathcal{T}_{\ell+1}$, we see

$$(1-\alpha)\,\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 \le \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 - \|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2$$
$$+ \alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_{\ell+1}-g_\ell)'\|_{L^2(\Gamma_D)}^2.$$

The orthogonality relation (34) applied for $g_{\ell+1}\in\mathcal{S}^1(\mathcal{T}_{\ell+1}|_{\Gamma_D})$ yields

$$\mathrm{osc}_{D,\ell+1}^2 + \|h_\ell^{1/2}(g_{\ell+1}-g_\ell)'\|_{L^2(\Gamma_D)}^2 \le \|h_\ell^{1/2}(g-g_\ell)'\|_{L^2(\Gamma_D)}^2 = \mathrm{osc}_{D,\ell}^2.$$

Together with the aforegoing estimate, we obtain

$$(1-\alpha)\,\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell+1}^2$$
$$\le \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell}^2 - \|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2.$$

Now, we add the error estimator and use the estimator reduction (54) to see, for $\beta>0$,

$$(1-\alpha)\,\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell+1}^2 + \beta\,\widetilde\varrho_{\ell+1}^2$$
$$\le \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell}^2 + \beta\,q\,\widetilde\varrho_\ell^2 + (\beta C_{10}-1)\,\|\nabla(U_{\ell+1}-U_\ell)\|_{L^2(\Omega)}^2.$$

We choose $\beta>0$ sufficiently small to guarantee $\beta C_{10}-1\le 0$. Then, we use the reliability (44) of $\varrho_\ell\le\widetilde\varrho_\ell$ in the form

$$C_{\rm rel}^{-1}\,\|\nabla(u-U_\ell)\|_{L^2(\Omega)} \le C_{\rm rel}^{-1}\,\|u-U_\ell\|_{H^1(\Omega)} \le \widetilde\varrho_\ell$$

to see, for $\varepsilon>0$,

$$(1-\alpha)\,\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell+1}^2 + \beta\,\widetilde\varrho_{\ell+1}^2$$
$$\le (1-\varepsilon\beta C_{\rm rel}^{-2})\,\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell}^2 + \beta(q+\varepsilon)\,\widetilde\varrho_\ell^2.$$

Moreover, since $\mathrm{osc}_{D,\ell}$ is a contribution of $\widetilde\varrho_\ell$, we have $\mathrm{osc}_{D,\ell}\le\widetilde\varrho_\ell$, whence, for $\delta>0$,

$$(1-\alpha)\,\|\nabla(u-U_{\ell+1})\|_{L^2(\Omega)}^2 + \alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell+1}^2 + \beta\,\widetilde\varrho_{\ell+1}^2$$
$$\le (1-\varepsilon\beta C_{\rm rel}^{-2})\,\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2 + (1-\delta\beta)\,\alpha^{-1}C_{\rm orth}\,\mathrm{osc}_{D,\ell}^2 + \beta(q+\varepsilon+\delta\,\alpha^{-1}C_{\rm orth})\,\widetilde\varrho_\ell^2.$$

For $0 < \alpha < 1$, we may now rearrange this estimate to end up with

$$\begin{aligned} \|\nabla(u - U_{\ell+1})\|_{L^2(\Omega)}^2 + \frac{C_{\rm orth}}{\alpha(1-\alpha)} \operatorname{osc}_{D,\ell+1}^2 + \frac{\beta}{1-\alpha} \widetilde{\varrho}_{\ell+1}^2 \\ \le \frac{1-\varepsilon\beta C_{\rm rel}^{-2}}{1-\alpha} \|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + (1-\delta\beta) \frac{C_{\rm orth}}{\alpha(1-\alpha)} \operatorname{osc}_{D,\ell}^2 \\ + (q + \varepsilon + \delta\, \alpha^{-1} C_{\rm orth}) \frac{\beta}{1-\alpha} \widetilde{\varrho}_\ell^2. \end{aligned}$$

It remains to choose the free constants $0 < \alpha, \delta, \varepsilon < 1$, whereas $\beta > 0$ has already been fixed:

- First, choose $0 < \varepsilon < C_{\rm rel}^2/\beta$ sufficiently small to guarantee $0 < q + \varepsilon < 1$.
- Second, choose $0 < \alpha < 1$ sufficiently small such that $0 < (1 - \varepsilon\beta C_{\rm rel}^{-2})/(1-\alpha) < 1$.
- Third, choose $\delta > 0$ sufficiently small with $0 < q + \varepsilon + \delta\, \alpha^{-1} C_{\rm orth} < 1$.

With $\gamma := \beta/(1-\alpha)$, $\lambda := \alpha^{-1}\, C_{\rm orth}/(1-\alpha)$, and $0 < \kappa < 1$ the maximal contraction constant of the three contributions, we conclude the proof of (61). $\square$

5. Quasi-Optimality of Adaptive Algorithm

5.1. Optimality of marking strategy. In the following, we first observe that the Dörfler marking (49) resp. the modified Dörfler marking (50)–(51) is not only sufficient but in some sense also necessary to obtain contraction of the error.

Proposition 14 (optimality of Dörfler marking). *Let $\alpha > 0$ and assume that the adaptivity parameter $0 < \theta < 1$ is sufficiently small, more precisely*

$$q_\star := \frac{1 - \theta(C_{\rm dlr}^2 + 1 + \alpha^{-1} C_{\rm orth}) C_{\rm eff}^2}{1+\alpha} > 0. \tag{62}$$

Let $0 < q \le q_\star$ and $\mathcal{T}_* = \texttt{refine}(\mathcal{T}_\ell)$ and assume that

$$\begin{aligned} \big(\|\nabla(u - U_*)\|_{L^2(\Omega)}^2 + \operatorname{osc}_{\mathcal{E},*}^2 + \operatorname{osc}_{D,*}^2 + \operatorname{osc}_{N,*}^2\big) \\ \le q\, \big(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \operatorname{osc}_{\mathcal{E},\ell}^2 + \operatorname{osc}_{D,\ell}^2 + \operatorname{osc}_{N,\ell}^2\big). \end{aligned} \tag{63}$$

Then, there holds the Dörfler marking for the set $\mathcal{R}_\ell(\mathcal{E}_*) \subseteq \mathcal{E}_\ell$ defined in (46), i.e.

$$\theta\, \varrho_\ell^2 \le \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2. \tag{64}$$

Proof. We start with the elementary observation that $q \le q_\star$ is equivalent to

$$\theta \le \frac{1 - q(1+\alpha)}{(C_{\rm dlr}^2 + 1 + \alpha^{-1} C_{\rm orth}) C_{\rm eff}^2}.$$

Using the discrete local reliability (48) and the quasi-Galerkin orthogonality (58), we see

$$
\begin{aligned}
C_{\rm dlr}^2\varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2 &\ge \|\nabla(U_*-U_\ell)\|_{L^2(\Omega)}^2\\
&\ge \|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2-(1+\alpha)\,\|\nabla(u-U_*)\|_{L^2(\Omega)}^2-\alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2\\
&= \big(\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2+{\rm osc}_{\mathcal{E},\ell}^2+{\rm osc}_{D,\ell}^2+{\rm osc}_{N,\ell}^2\big)\\
&\quad-(1+\alpha)\big(\|\nabla(u-U_*)\|_{L^2(\Omega)}^2+{\rm osc}_{\mathcal{E},*}^2+{\rm osc}_{D,*}^2+{\rm osc}_{N,*}^2\big)\\
&\quad-{\rm osc}_{\mathcal{E},\ell}^2-{\rm osc}_{D,\ell}^2-{\rm osc}_{N,\ell}^2+(1+\alpha)({\rm osc}_{\mathcal{E},*}^2+{\rm osc}_{D,*}^2+{\rm osc}_{N,*}^2)\\
&\quad-\alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2\\
&\ge \big(1-q(1+\alpha)\big)\big(\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2+{\rm osc}_{\mathcal{E},\ell}^2+{\rm osc}_{D,\ell}^2+{\rm osc}_{N,\ell}^2\big)\\
&\quad-{\rm osc}_{\mathcal{E},\ell}^2-{\rm osc}_{D,\ell}^2-{\rm osc}_{N,\ell}^2+(1+\alpha)({\rm osc}_{\mathcal{E},*}^2+{\rm osc}_{D,*}^2+{\rm osc}_{N,*}^2)\\
&\quad-\alpha^{-1}C_{\rm orth}\,\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2,
\end{aligned}
$$

where we have finally used Assumption (63). As in the proof of Proposition 3, we have

$$\|h_\ell^{1/2}(g_*-g_\ell)'\|_{L^2(\Gamma_D)}^2 \le {\rm osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 \le \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2.$$

Moreover, the identities ${\rm osc}_{D,\ell}(E) = {\rm osc}_{D,*}(E)$, ${\rm osc}_{\mathcal{E},\ell}(E) = {\rm osc}_{\mathcal{E},*}(E)$ and ${\rm osc}_{N,\ell}(E) = {\rm osc}_{N,*}(E)$ for $E\in\mathcal{E}_\ell\backslash\mathcal{R}_\ell(\mathcal{E}_*)$ prove

$$\tag{65} {\rm osc}_{D,\ell}^2-{\rm osc}_{D,*}^2 \le {\rm osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2,$$

$$\tag{66} {\rm osc}_{\mathcal{E},\ell}^2-{\rm osc}_{\mathcal{E},*}^2 \le {\rm osc}_{\mathcal{E},\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2,$$

$$\tag{67} {\rm osc}_{N,\ell}^2-{\rm osc}_{N,*}^2 \le {\rm osc}_{N,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2.$$

Note that (66) led to the definition of $\mathcal{R}_\ell(\mathcal{E}_*)$ given above. Together with the efficiency (45) and ${\rm osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2+{\rm osc}_{\mathcal{E},\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2+{\rm osc}_{N,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 \le \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2$, we may now conclude

$$\big(C_{\rm dlr}^2+1+\alpha^{-1}C_{\rm orth}\big)\,\varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2 \ge \big(1-q(1+\alpha)\big)\,C_{\rm eff}^{-2}\,\varrho_\ell^2.$$

This is equivalent to $\theta\,\varrho_\ell^2 \le \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2$ and led to the definition of $q_\star$. □

Proposition 15 (optimality of modified Dörfler marking)**.** *Let $\alpha>0$ and $0<\theta_2<1$ and assume that the adaptivity parameters $0<\theta_1,\vartheta<1$ are sufficiently small, more precisely*

$$\tag{68} q_\star := \max\Big\{\frac{1-C_{\rm eff}^2\big(\theta_1(1+C_{\rm dlr}^2)+\vartheta(1+C_{\rm dlr}^2+\alpha^{-1}C_{\rm orth})\big)}{1+\alpha}\,,\,\frac{1-\theta_2}{(1+\vartheta^{-1})(C_{\rm rel}^2+1)}\Big\} > 0.$$

Let $0<q\le q_\star$ and $\mathcal{T}_*=\texttt{refine}(\mathcal{T}_\ell)$ and assume that

$$\tag{69} \begin{aligned}&\big(\|\nabla(u-U_*)\|_{L^2(\Omega)}^2+{\rm osc}_{\mathcal{E},*}^2+{\rm osc}_{D,*}^2+{\rm osc}_{N,*}^2\big)\\ &\qquad\le q\,\big(\|\nabla(u-U_\ell)\|_{L^2(\Omega)}^2+{\rm osc}_{\mathcal{E},\ell}^2+{\rm osc}_{D,\ell}^2+{\rm osc}_{N,\ell}^2\big).\end{aligned}$$

Then, there holds the modified Dörfler marking for the set $\mathcal{R}_\ell(\mathcal{E}_*)\subseteq\mathcal{E}_\ell$, i.e. there holds either

$$\tag{70} \theta_1\,\eta_\ell^2 \le \eta_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2$$

in case of ${\rm osc}_{\mathcal{E},\ell}^2+{\rm osc}_{D,\ell}^2\le\vartheta\,\eta_\ell^2$ or

$$\tag{71} \theta_2\,\big({\rm osc}_{\mathcal{E},\ell}^2+{\rm osc}_{D,\ell}^2\big) \le {\rm osc}_{\mathcal{E},\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2+{\rm osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2$$

otherwise.

Proof. We first assume $\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 \le \vartheta\,\eta_\ell^2$. Arguing as in the proof of Proposition 14, we see

$$\begin{aligned}
C_{\mathrm{dlr}}^2 \varrho_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2 &\ge \|\nabla(U_* - U_\ell)\|_{L^2(\Omega)}^2 \\
&\ge \big(1 - q(1+\alpha)\big)\big(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 + \mathrm{osc}_{N,\ell}^2\big) \\
&\quad - \mathrm{osc}_{\mathcal{E},\ell}^2 - \mathrm{osc}_{D,\ell}^2 - \mathrm{osc}_{N,\ell}^2 + (1+\alpha)(\mathrm{osc}_{\mathcal{E},*}^2 + \mathrm{osc}_{D,*}^2 + \mathrm{osc}_{N,*}^2) \\
&\quad - \alpha^{-1} C_{\mathrm{orth}}\,\|h_\ell^{1/2}(g_* - g_\ell)'\|_{L^2(\Gamma_D)}^2 \\
&\ge \big(1 - q(1+\alpha)\big)\,C_{\mathrm{eff}}^{-2}\,\eta_\ell^2 - \mathrm{osc}_{\mathcal{E},\ell}^2 - \mathrm{osc}_{D,\ell}^2 - \mathrm{osc}_{N,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 \\
&\quad - \alpha^{-1} C_{\mathrm{orth}}\,\|h_\ell^{1/2}(g_* - g_\ell)'\|_{L^2(\Gamma_D)}^2,
\end{aligned}$$

where we have used (67). Next, we recall the edge-wise definition $\varrho_\ell^2 = \eta_\ell^2 + \mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2$ and collect all oscillation terms on the right-hand side. Together with $\|h_\ell^{1/2}(g_* - g_\ell)'\|_{L^2(\Gamma_D)} \le \mathrm{osc}_{D,\ell}$ and $\mathrm{osc}_{N,\ell}(E) \le \eta_\ell(E)$, this leads to

$$\begin{aligned}
C_{\mathrm{dlr}}^2\,\eta_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2 &\ge \big(1 - q(1+\alpha)\big)\,C_{\mathrm{eff}}^{-2}\,\eta_\ell^2 - (1 + C_{\mathrm{dlr}}^2)\mathrm{osc}_{\mathcal{E},\ell}^2 \\
&\quad - (1 + C_{\mathrm{dlr}}^2 + \alpha^{-1} C_{\mathrm{orth}})\mathrm{osc}_{D,\ell}^2 - \mathrm{osc}_{N,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 \\
&\ge \big[\big(1 - q(1+\alpha)\big)\,C_{\mathrm{eff}}^{-2} - \vartheta(1 + C_{\mathrm{dlr}}^2 + \alpha^{-1} C_{\mathrm{orth}})\big]\,\eta_\ell^2 - \eta_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2.
\end{aligned}$$

We then conclude

$$\eta_\ell(\mathcal{R}_\ell(\mathcal{E}_*))^2 \ge \frac{1 - q(1+\alpha) - \vartheta C_{\mathrm{eff}}^2(1 + C_{\mathrm{dlr}}^2 + \alpha^{-1} C_{\mathrm{orth}})}{(1 + C_{\mathrm{dlr}}^2) C_{\mathrm{eff}}^2}\,\eta_\ell^2 \ge \theta_1\,\eta_\ell^2,$$

which follows from our assumption on $0 < q \le q_\star < 1$ and the definition of $q_\star$ in (68). This concludes the proof of (70).

Second, we assume $\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 > \vartheta\,\eta_\ell^2$. Recall the estimates (65)–(67). Then, reliability (44) of $\varrho_\ell^2 = \eta_\ell^2 + \mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2$ and $\mathrm{osc}_{N,\ell} \le \eta_\ell$ yield

$$\begin{aligned}
\big(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2\big) &- \big(\mathrm{osc}_{\mathcal{E},\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 + \mathrm{osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2\big) \\
&\le \mathrm{osc}_{\mathcal{E},*}^2 + \mathrm{osc}_{D,*}^2 \\
&\le q\,\big(\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 + \mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2 + \mathrm{osc}_{N,\ell}^2\big) \\
&\le q\,\big((C_{\mathrm{rel}}^2 + 1)\,(\eta_\ell^2 + \mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2)\big) \\
&< q\,(1 + \vartheta^{-1})(C_{\mathrm{rel}}^2 + 1)\,(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2).
\end{aligned}$$

Rearranging the terms, we obtain

$$\begin{aligned}
\theta_2\,\big(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2\big) &\le \big[1 - q\,(1 + \vartheta^{-1})(C_{\mathrm{rel}}^2 + 1)\big]\,\big(\mathrm{osc}_{\mathcal{E},\ell}^2 + \mathrm{osc}_{D,\ell}^2\big) \\
&\le \mathrm{osc}_{\mathcal{E},\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2 + \mathrm{osc}_{D,\ell}(\mathcal{R}_\ell(\mathcal{E}_*))^2,
\end{aligned}$$

where the first estimate follows from $0 < q \le q_\star < 1$ and the definition of $q_\star$ in (68). □

5.2. Optimality of newest vertex bisection. The quasi-optimality analysis for adaptive FEM involves two properties of the mesh-refinement which are, so far, only mathematically guaranteed for newest vertex bisection.

First, assume that the reference edges of the initial mesh $\mathcal{T}_0$ are chosen such that an interior edge $E = T_+ \cap T_- \in \mathcal{E}_0^\Omega$ is either the reference edge of both elements $T_+, T_- \in \mathcal{T}_0$ or of none.

It has originally been proven in [BDD] and lateron improved in [S08, KS] that the sequence of meshes defined inductively by $\mathcal{T}_{\ell+1} := \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ with arbitrary $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ satisfies

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \le C_{\rm nvb} \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \quad \text{for all } \ell \in \mathbb{N} \tag{72}$$

with some constant $C_{\rm nvb} > 0$ which depends only on $\mathcal{T}_0$. This proves that the closure step in newest vertex bisection which avoids hanging nodes and leads to possible bisections of edges $E \in \mathcal{E}_\ell \backslash \mathcal{M}_\ell$ may not lead to a arbitrary many refinements. It has already been observed in [BDD] that such a labeling of the reference edges of $\mathcal{T}_0$ is always possible although no algorithm is known yet which provides this labeling in linear complexity.

Second, for two meshes $\mathcal{T}' = \texttt{refine}(\mathcal{T}_0)$ and $\mathcal{T}'' = \texttt{refine}(\mathcal{T}_0)$ obtained by newest vertex bisection of the initial mesh $\mathcal{T}_0$, there is a unique coarsest common refinement $\mathcal{T}' \oplus \mathcal{T}'' = \texttt{refine}(\mathcal{T}_0)$ which is a refinement of both $\mathcal{T}'$ and $\mathcal{T}''$. It is shown in [S07, CKNS] that $\mathcal{T}' \oplus \mathcal{T}''$ is, in fact, the overlay of these meshes. Moreover, it holds that

$$\#(\mathcal{T}' \oplus \mathcal{T}'') \le \#\mathcal{T}' + \#\mathcal{T}'' - \#\mathcal{T}_0. \tag{73}$$

5.3. Definition of approximation class. To state the optimality result, we have to introduce the appropriate approximation class. Let

$$\mathbb{T} := \big\{\mathcal{T} \,:\, \mathcal{T} = \texttt{refine}(\mathcal{T}_0)\big\} \tag{74}$$

be the set of all triangulations which can be obtained from $\mathcal{T}_0$ by newest vertex bisection. Moreover, let

$$\mathbb{T}_N := \big\{\mathcal{T} \in \mathbb{T} \,:\, \#\mathcal{T} - \#\mathcal{T}_0 \le N\big\} \tag{75}$$

be the set of triangulations which have at most $N \in \mathbb{N}$ elements more than the initial mesh $\mathcal{T}_0$. For $s > 0$, the approximation class $\mathbb{A}_s$ has already been defined in (12)–(13). The first step is to prove that, up to constants, nodal interpolation of the boundary data yields the best possible approximation of the exact solution.

Lemma 16. *The Galerkin solution $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ of (4) satisfies*

$$\begin{aligned} \|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 &+ \|h_\ell^{1/2}(g - g_\ell)'\|_{L^2(\Gamma_D)}^2 \\ &\le C_{\rm cea} \inf_{W_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)} \big(\|\nabla(u - W_\ell)\|_{L^2(\Omega)}^2 + \|h_\ell^{1/2}(g - W_\ell|_\Gamma)'\|_{L^2(\Gamma_D)}^2\big), \end{aligned} \tag{76}$$

where $C_{\rm cea} > 0$ depends only on Γ and $\sigma(\mathcal{T}_\ell)$.

Proof. Let $\widehat{g}, \widehat{g}_\ell \in H^{1/2}(\Gamma)$ denote arbitrary extensions of $g = u|_{\Gamma_D}$ resp. g_ℓ. Note that $(\mathcal{L}_\ell P_\ell \widehat{g})|_{\Gamma_D} = (P_\ell u)|_{\Gamma_D}$ as well as $(\mathcal{L}_\ell P_\ell \widehat{g}_\ell)|_{\Gamma_D} = g_\ell$, where $\mathcal{L}_\ell$ denotes the discrete lifting operator from (26). For $V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell)$, we thus have $U_\ell - (V_\ell + \mathcal{L}_\ell P_\ell \widehat{g}_\ell) \in \mathcal{S}^1_D(\mathcal{T}_\ell)$, whence

$$\|\nabla(u - U_\ell)\|_{L^2(\Omega)}^2 = \langle \nabla(u - U_\ell)\,,\, \nabla(u - (V_\ell + \mathcal{L}_\ell P_\ell \widehat{g}_\ell))\rangle_\Omega$$

according to the Galerkin orthogonality. Therefore, the Cauchy-Schwarz inequality provides the Céa-type quasi-optimality

$$\|\nabla(u - U_\ell)\|_{L^2(\Omega)} \le \min_{V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell)} \|\nabla(u - (V_\ell + \mathcal{L}_\ell P_\ell \widehat{g}_\ell))\|_{L^2(\Omega)}.$$

We now plug-in $V_\ell = P_\ell u - \mathcal{L}_\ell P_\ell \widehat{g} \in \mathcal{S}^1_D(\mathcal{T}_\ell)$ to see

$$\begin{aligned}\|\nabla(u-U_\ell)\|_{L^2(\Omega)} &\le \|\nabla(u - P_\ell u + \mathcal{L}_\ell P_\ell(\widehat{g}-\widehat{g}_\ell))\|_{L^2(\Omega)}\\ &\lesssim \|\nabla(u-P_\ell u)\|_{L^2(\Omega)} + \|\widehat{g}-\widehat{g}_\ell\|_{H^{1/2}(\Gamma)}.\end{aligned}$$

Since the extensions $\widehat{g}, \widehat{g}_\ell$ of g, g_ℓ were arbitrary, we obtain

$$\begin{aligned}\|\nabla(u-U_\ell)\|_{L^2(\Omega)} &\lesssim \|\nabla(u-P_\ell u)\|_{L^2(\Omega)} + \|g-g_\ell\|_{H^{1/2}(\Gamma_D)}\\ &\lesssim \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(u-W_\ell)\|_{L^2(\Omega)} + \|h_\ell^{1/2}(g-g_\ell)'\|_{L^2(\Gamma_D)},\end{aligned}$$

where we have used the quasi-optimality of the Scott-Zhang projection, see Section 2.3, and Lemma 1. Finally, we use the orthogonality relation (34) to see

$$\begin{aligned}\|\nabla(u-U_\ell)\|^2_{L^2(\Omega)} + \|h_\ell^{1/2}(g-g_\ell)'\|^2_{L^2(\Gamma_D)} &\lesssim \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \|\nabla(u-W_\ell)\|^2_{L^2(\Omega)} + \|h_\ell^{1/2}(g-g_\ell)'\|^2_{L^2(\Gamma_D)}\\ &\le \min_{W_\ell\in\mathcal{S}^1(\mathcal{T}_\ell)} \big(\|\nabla(u-W_\ell)\|^2_{L^2(\Omega)} + \|h_\ell^{1/2}(g-W_\ell|_\Gamma)'\|^2_{L^2(\Gamma_D)}\big).\end{aligned}$$

This concludes the proof. □

5.4. Quasi-optimality result. Finally, we may formally state the optimality result (14) described in the introduction.

Theorem 17. *Suppose that the adaptivity parameter* $0 < \theta < 1$ *in Algorithm 7 satisfies* (62) *and that the adaptivity parameters* $0 < \theta_1, \theta_2, \vartheta < 1$ *satisfy* (68) *so that either marking strategy is optimal in the sense of Proposition 14 resp. Proposition 15. Suppose that the initial mesh* $\mathcal{T}_0$ *satisfies the assumption of Section 5.2 so that* (72)–(73) *are guaranteed. Let* $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ *denote the sequence of discrete solutions generated by either Algorithm 7 or Algorithm 8. If the given data and the corresponding weak solution of* (2) *satisfy* $(u, f, g, \phi) \in \mathbb{A}_s$*, there holds*

$$\|u-U_\ell\|_{H^1(\Omega)} \le C_{\rm opt}(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s}, \tag{77}$$

i.e. each possible convergence rate $s > 0$ *is asymptotically achieved by AFEM. The constant* $C_{\rm opt} > 0$ *depends only on* $\|(u,f,g,\phi)\|_{\mathbb{A}_s}$*, the initial mesh* $\mathcal{T}_0$*, and the adaptivity parameters.*

Proof. Since the proof follows essentially the lines of [S07, CKNS], we leave the elaborate details to the reader. For any $\varepsilon > 0$, the definition of the approximation class $\mathbb{A}_s$ guarantees some triangulation $\mathcal{T}_\varepsilon \in \mathbb{T}$ such that

$$\inf_{W_\varepsilon\in\mathcal{S}^1(\mathcal{T}_\varepsilon)} \big(\|\nabla(u-W_\varepsilon)\|^2_{L^2(\Omega)} + \|h_\varepsilon^{1/2}(g-W_\varepsilon|_\Gamma)'\|^2_{L^2(\Gamma_D)} + \mathrm{osc}^2_{\mathcal{T},\varepsilon} + \mathrm{osc}^2_{N,\varepsilon}\big)^{1/2} \le \varepsilon$$

and

$$\#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s},$$

where the constant depends only on $\|(u,f,g,\phi)\|_{\mathbb{A}_s}$. We now consider the overlay $\mathcal{T}_* := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$. With the help of Lemma 16 as well as the elementary estimates $\mathrm{osc}_{\mathcal{T},*} \le \mathrm{osc}_{\mathcal{T},\varepsilon}$ and $\mathrm{osc}_{N,*} \le \mathrm{osc}_{N,\varepsilon}$, we observe

$$\Lambda_* := \big(\|\nabla(u-U_*)\|^2_{L^2(\Omega)} + \mathrm{osc}^2_{D,*} + \mathrm{osc}^2_{\mathcal{T},*} + \mathrm{osc}^2_{N,*}\big)^{1/2} \lesssim \varepsilon,$$

since $\mathcal{S}^1(\mathcal{T}_\varepsilon) \subseteq \mathcal{S}^1(\mathcal{T}_*)$. Moreover, the overlay estimate (73) predicts

$$\#\mathcal{R}_\ell(\mathcal{T}_*) \le \#\mathcal{T}_* - \#\mathcal{T}_\ell \le \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s}.$$

Note that Lemma 4 together with reliability and efficiency of ϱ_* yield

$$\Lambda_* \simeq \big(\|\nabla(u - U_*)\|^2_{L^2(\Omega)} + \mathrm{osc}^2_{\mathcal{E},*} + \mathrm{osc}^2_{D,*} + \mathrm{osc}^2_{N,*}\big)^{1/2},$$

where $\mathrm{osc}_{\mathcal{T},*}$ is replaced by $\mathrm{osc}_{\mathcal{E},*}$. Choosing $\varepsilon = \lambda\big(\|\nabla(u-U_\ell)\|^2_{L^2(\Omega)}+\mathrm{osc}^2_{D,\ell}+\mathrm{osc}^2_{\mathcal{E},\ell}+\mathrm{osc}^2_{N,\ell}\big)^{1/2}$ with $\lambda > 0$ sufficiently small, we enforce the reduction (63) resp. (69) and derive that $\mathcal{R}_\ell(\mathcal{E}_*) \subseteq \mathcal{E}_\ell$ satisfy the respective marking criterion, cf. Proposition 14 resp. Proposition 15. Minimality of $\mathcal{M}_\ell$ thus gives

$$\#\mathcal{M}_\ell \le \#\mathcal{R}_\ell(\mathcal{E}_*) \lesssim \#\mathcal{R}_\ell(\mathcal{T}_*) \lesssim \varepsilon^{-1/s} \simeq \big(\|\nabla(u - U_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2_{\mathcal{E},\ell} + \mathrm{osc}^2_{D,\ell} + \mathrm{osc}^2_{N,\ell}\big)^{-1/(2s)}.$$

We next note that

$$\varrho_\ell^2 \simeq \|\nabla(u - U_\ell)\|^2_{L^2(\Omega)} + \mathrm{osc}^2_{\mathcal{E},\ell} + \mathrm{osc}^2_{D,\ell} + \mathrm{osc}^2_{N,\ell} \simeq \Delta_\ell$$

according to reliability and efficiency of ϱ_ℓ and the definition of the contraction quantity Δ_ℓ in Theorem 13. Combining the last two lines, we see

$$\#\mathcal{M}_\ell \lesssim \Delta_\ell^{-1/(2s)} \simeq \varrho_\ell^{-1/s} \quad \text{for all } \ell \in \mathbb{N}_0.$$

By use of the closure estimate (72) of newest vertex bisection, we obtain

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \lesssim \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \lesssim \sum_{j=0}^{\ell-1} \Delta_j^{-1/(2s)}.$$

Note that the contraction property (61) of Δ_j implies $\Delta_\ell \le \kappa^{\ell-j}\Delta_j$, whence $\Delta_j^{-1/(2s)} \le \kappa^{(\ell-j)/(2s)}\Delta_\ell^{-1/(2s)}$. According to $0 < \kappa < 1$ and the geometric series, this gives

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \lesssim \Delta_\ell^{-1/2s} \sum_{j=0}^{\ell-1} \kappa^{(\ell-j)/(2s)} \lesssim \Delta_\ell^{-1/2s} \simeq \varrho_\ell^{-1/s}.$$

Altogether, we may therefore conclude $\|u - U_\ell\|_{H^1(\Omega)} \lesssim \varrho_\ell \lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-s}$. □

Remark. *All convergence and optimality results in this paper are stated for the edge-based error estimator ϱ_ℓ. Nevertheless, it is only a notational modification to see that also the element-based error estimator ρ_ℓ from* (5)–(6) *leads to quasi-optimally convergent versions of AFEM. To that end, Algorithm 7 and Algorithm 8 are slightly modified, and one seeks minimial sets of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ instead. For each marked element $T \in \mathcal{M}_\ell$, we mark its reference edge. The convergence result in Theorem 13 and the optimality result in Theorem 17 hold accordingly.* □

6. Some Remarks on the 3D Case

So far, we have only considered a 2D model problem (1). In 3D, one additional difficulty is that the regularity assumption $g \in H^1(\Gamma_D)$ is not sufficient to guarantee continuity of g. Therefore, one must not use nodal interpolation to discretize $g \approx g_\ell$ and to define the Dirichlet data oscillations $\mathrm{osc}_{D,\ell}$.

If we do not use nodal interpolation to approximate $g \approx g_\ell$, the estimator reduction estimate (54) becomes

$$\varrho_{\ell+1}^2 \le q\,\varrho_\ell^2 + C_{10}\|U_{\ell+1} - U_\ell\|^2_{H^1(\Omega)}, \tag{78}$$

where $C_{10} > 0$ additionally depends on Ω. The reason for this is that the analysis provides an additional term $\|g_{\ell+1} - g_\ell\|^2_{H^{1/2}(\Gamma_D)}$ on the right-hand side of (54) since we loose the orthogonality relation (34) which is used in the form

$$\begin{aligned}\|h_{\ell+1}^{1/2}(g - g_{\ell+1})'\|^2_{L^2(\Gamma_D)} &\le \|h_{\ell+1}^{1/2}(g - g_{\ell+1})'\|^2_{L^2(\Gamma_D)} + \|h_{\ell+1}^{1/2}(g_{\ell+1} - g_\ell)'\|^2_{L^2(\Gamma_D)}\\ &= \|h_{\ell+1}^{1/2}(g - g_\ell)'\|^2_{L^2(\Gamma_D)}.\end{aligned}$$

Instead, an inverse estimate and the Rellich compactness theorem yield

$$\begin{aligned}\|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)} + \|h_\ell^{1/2}(g_{\ell+1} - g_\ell)'\|^2_{L^2(\Gamma_D)} &\lesssim \|\nabla(U_{\ell+1} - U_\ell)\|^2_{L^2(\Omega)} + \|g_{\ell+1} - g_\ell\|^2_{H^{1/2}(\Gamma_D)}\\ &\simeq \|U_{\ell+1} - U_\ell\|^2_{H^1(\Omega)}\end{aligned}$$

which proves (78). Note that this estimate holds for *any* discretization of $g \approx g_\ell \in \mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D})$ and even in 3D, where the arclength derivative $(\cdot)'$ is replaced by the surface gradient $\nabla_\Gamma(\cdot)$; we refer to [GHS] for the inverse estimate.

A possible choice for g_ℓ is $g_\ell = \Pi_\ell g$, where $\Pi_\ell : L^2(\Gamma_D) \to \mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D})$ is the L^2-orthogonal projection [BCD]. Note that newest vertex bisection ensures that Π_ℓ is a stable projection with respect to the $H^1(\Gamma_D)$-norm. In [KP], we prove the approximation estimate

$$\|g - g_\ell\|_{H^{1/2}(\Gamma_D)} \lesssim \|h_\ell^{1/2}\nabla_\Gamma(g - g_\ell)\|_{L^2(\Gamma_D)} =: \mathrm{osc}_{D,\ell} \tag{79}$$

with ∇_Γ the surface gradient. Moreover, we show that the a priori limit $g_\infty := \lim_\ell g_\ell$ exists strongly in $H^\alpha(\Gamma_D)$ for $0 \le \alpha < 1$ and even weakly in $H^1(\Gamma_D)$ provided that the discrete spaces $\mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D})$ are nested, i.e. $\mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D}) \subseteq \mathcal{S}^1(\mathcal{T}_{\ell+1}|_{\Gamma_D})$ for all $\ell \in \mathbb{N}_0$. Note, however, that this is always the case for adaptive mesh-refining algorithms. In particular, we have

$$\mathcal{S}^1(\mathcal{T}_\ell) \subseteq \mathcal{S}^1(\mathcal{T}_{\ell+1}) \quad \text{for all } \ell \in \mathbb{N}_0. \tag{80}$$

In the following, we even aim to prove that nestedness (80) implies the existence of the a priori limit $\lim_\ell U_\ell$ in $H^1(\Omega)$. To that end, we need the following lemma.

Lemma 18 (a priori convergence of Scott-Zhang projection)**.** *We recall the Scott-Zhang projection P_ℓ and make the additional assumption that the edges E_z are chosen appropriately, i.e. for $\omega_{\ell,z} \subset \bigcup(\mathcal{T}_\ell \cap \mathcal{T}_{\ell+1})$ we ensure that the edge E_z is chosen for both operators P_ℓ and $P_{\ell+1}$. Then, the Scott-Zhang interpolands $v_\ell := P_\ell v \in \mathcal{S}^1(\mathcal{T}_\ell)$ of arbitrary $v \in H^1(\Omega)$ converge to some a priori limit in $H^1(\Omega)$, i.e. there holds*

$$\|P_\infty v - P_\ell v\|_{H^1(\Omega)} \xrightarrow{\ell\to\infty} 0 \tag{81}$$

for a certain element $P_\infty v \in \mathcal{S}^1(\mathcal{T}_\infty) := \overline{\bigcup_{\ell\in N}\mathcal{S}^1(\mathcal{T}_\ell)}$.

Proof. We follow the ideas from [MSV] and define the following subsets of Ω:

$$\begin{aligned}\Omega_\ell^0 &:= \bigcup\{T \in \mathcal{T}_\ell : \omega_\ell(T) \subset \bigcup\big(\textstyle\bigcup_{i=0}^\infty \bigcap_{j=i}^\infty \mathcal{T}_j\}\big),\\ \Omega_\ell &:= \bigcup\{T \in \mathcal{T}_\ell : \text{There exists } k \ge 0 \text{ s.t. } \omega_\ell(T) \text{ is at least uniformly refined in } \mathcal{T}_{\ell+k}\},\\ \Omega_\ell^* &:= \Omega \setminus (\Omega_\ell \cup \Omega_\ell^0),\end{aligned}$$

where $\omega_\ell(\omega) := \bigcup\{T \in \mathcal{T}_\ell : T \cap \omega \ne \emptyset\}$ for all measurable $\omega \subset \Omega$. According to [MSV, Corollary 4.1], it holds that

$$\lim_{\ell\to\infty} \|\chi_{\Omega_\ell} h_\ell\|_{L^\infty(\Omega)} = 0. \tag{82}$$

Let $\varepsilon > 0$ be arbitrary. Since the space $H^2(\Omega)$ is dense in $H^1(\Omega)$, we find $v_\varepsilon \in H^2(\Omega)$ such that $\|v - v_\varepsilon\|_{H^1(\Omega)} \le \varepsilon$. Due to local approximation and stability properties of P_ℓ, we obtain

$$\|(1-P_\ell)v\|_{H^1(\Omega_\ell)} \lesssim \|(1-P_\ell)v_\varepsilon\|_{H^1(\Omega_\ell)} + \varepsilon \le \|h_\ell\, D^2 v_\varepsilon\|_{L^2(\omega_\ell(\Omega_\ell))} + \varepsilon,$$

cf. [SZ]. By use of (82), we may choose $\ell_0 \in \mathbb{N}$ sufficiently large to guarantee $\|h_\ell\, D^2 v_\varepsilon\|_{L^2(\omega_\ell(\Omega_\ell))} \le \|h_\ell\|_{L^\infty(\omega_\ell(\Omega_\ell))} \|D^2 v_\varepsilon\|_{L^2(\Omega)} \le \varepsilon$ for all $\ell \ge \ell_0$. Then, there holds

$$\|(1-P_\ell)v\|_{H^1(\Omega_\ell)} \lesssim \varepsilon \quad \text{for all } \ell \ge \ell_0. \tag{83}$$

There holds $\lim_{\ell\to\infty} |\Omega_\ell^*| = 0$, cf. [MSV, Proposition 4.2], and this provides the existence of $\ell_1 \in \mathbb{N}$ such that

$$\|v\|_{H^1(\omega_\ell(\Omega_\ell^*))} \le \varepsilon \quad \text{for all } \ell \ge \ell_1 \tag{84}$$

due to the non-concentration of Lebesgue functions. With these preparations, we finally aim at proving that $P_\ell v$ is a Cauchy sequence in $H^1(\Omega)$. Therefore, let $\ell \ge \max\{\ell_0, \ell_1\}$ and $k \ge 0$ be arbitrary. First, we use that for any $T \in \mathcal{T}_\ell$, $(P_\ell v)|_T$ depends only on $v|_{\omega_\ell(T)}$. Then, by definition of Ω_ℓ^0 and our assumption on the definition of P_ℓ and $P_{\ell+k}$ on $\mathcal{T}_\ell \cap \mathcal{T}_{\ell+k}$, we obtain

$$\|P_\ell v - P_{\ell+k} v\|_{H^1(\Omega_\ell^0)} = 0. \tag{85}$$

Second, due to the local stability of P_ℓ and (84), there holds

$$\begin{aligned} \|P_\ell v - P_{\ell+k} v\|_{H^1(\Omega_\ell^*)} &\le \|P_\ell v\|_{H^1(\Omega_\ell^*)} + \|P_{\ell+k} v\|_{H^1(\Omega_\ell^*)} \\ &\lesssim \|v\|_{H^1(\omega_\ell(\Omega_\ell^*))} + \|v\|_{H^1(\omega_{\ell+k}(\Omega_\ell^*))} \\ &\le 2\|v\|_{H^1(\omega_\ell(\Omega_\ell^*))} \le 2\varepsilon. \end{aligned} \tag{86}$$

Third, we proceed by exploiting (83). We have

$$\|P_\ell v - P_{\ell+k} v\|_{H^1(\Omega_\ell)} \le \|P_\ell v - v\|_{H^1(\Omega_\ell)} + \|v - P_{\ell+k} v\|_{H^1(\Omega_\ell)} \lesssim \varepsilon. \tag{87}$$

Combining the estimates from (85)–(87), we conclude $\|P_\ell v - P_{\ell+k} v\|_{H^1(\Omega)} \lesssim \varepsilon$, i.e. $(P_\ell v)$ is a Cauchy sequence in $H^1(\Omega)$ and hence convergent. □

Now, we are able to prove a priori convergence of U_ℓ towards some a priori limit u_∞.

Proposition 19 (a priori convergence of U_ℓ)**.** *Suppose that the discrete spaces satisfy nestedness* (80) *and that $U_\ell \in \mathcal{S}^1(\mathcal{T}_\ell)$ solves* (4) *with $g_\ell = \Pi_\ell g$ and $\Pi_\ell : L^2(\Gamma_D) \to \mathcal{S}^1(\mathcal{T}_\ell|_{\Gamma_D})$ the L^2-projection. Then, the a priori limit $u_\infty := \lim_{\ell\to\infty} U_\ell \in H^1(\Omega)$ exists.*

Proof. For $g_\ell \in H^{1/2}(\Gamma)$, we consider the continuous auxiliary problem

$$\begin{aligned} -\Delta w_\ell &= 0 \quad \text{in } \Omega, \\ w_\ell &= g_\ell \quad \text{on } \Gamma_D, \\ \partial_n w_\ell &= 0 \quad \text{on } \Gamma_N. \end{aligned}$$

Let $w_\ell \in H^1(\Omega)$ be the unique (weak) solution and note that the trace $\widehat{g}_\ell := w_\ell|_\Gamma \in H^{1/2}(\Gamma)$ provides an extension of g_ℓ with

$$\|\widehat{g}_\ell\|_{H^{1/2}(\Gamma)} \le \|w_\ell\|_{H^1(\Omega)} \lesssim \|g_\ell\|_{H^{1/2}(\Gamma_D)} \le \|\widehat{g}_\ell\|_{H^{1/2}(\Gamma)}.$$

For arbitrary $k, \ell \in \mathbb{N}$, the same type of arguments proves

$$\|\widehat{g}_\ell - \widehat{g}_k\|_{H^{1/2}(\Gamma)} \simeq \|g_\ell - g_k\|_{H^{1/2}(\Gamma_D)}.$$

Since (g_ℓ) is a Cauchy sequence in $H^{1/2}(\Gamma_D)$, cf. [KP], we obtain that $(\widehat{g}_\ell)$ is a Cauchy sequence in $H^{1/2}(\Gamma)$, whence convergent with limit $\widehat{g}_\infty \in H^{1/2}(\Gamma)$.

Second, note that $(\mathcal{L}_\ell\widehat{g}_\ell)|_{\Gamma_D} = g_\ell$, where $\mathcal{L}_\ell = P_\ell\mathcal{L}$ denotes the discrete lifting from (26). Therefore, $\widetilde{U}_\ell := U_\ell - \mathcal{L}_\ell\widehat{g}_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell)$ is the unique solution of the variational form

$$\langle\nabla\widetilde{U}_\ell\,,\,\nabla V_\ell\rangle_\Omega = \langle\nabla u\,,\,\nabla V_\ell\rangle_\Omega - \langle\nabla\mathcal{L}_\ell\widehat{g}_\ell\,,\,\nabla V_\ell\rangle_\Omega \quad \text{for all } V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell). \tag{88}$$

Third, Lemma 18 implies

$$\begin{aligned}\|\mathcal{L}_\ell\widehat{g}_\ell - P_\infty\mathcal{L}\widehat{g}_\infty\|_{H^1(\Omega)} &\le \|P_\ell(\mathcal{L}\widehat{g}_\ell - \mathcal{L}\widehat{g}_\infty)\|_{H^1(\Omega)} + \|P_\ell\mathcal{L}\widehat{g}_\infty - P_\infty\mathcal{L}\widehat{g}_\infty\|_{H^1(\Omega)}\\ &\lesssim \|\widehat{g}_\ell - \widehat{g}_\infty\|_{H^{1/2}(\Gamma)} + \|P_\ell\mathcal{L}\widehat{g}_\infty - P_\infty\mathcal{L}\widehat{g}_\infty\|_{H^1(\Omega)} \xrightarrow{\ell\to\infty} 0.\end{aligned}$$

Fourth, let $\widetilde{U}_{\ell,\infty} \in \mathcal{S}^1_D(\mathcal{T}_\ell)$ be the unique solution of the discrete auxiliary problem

$$\langle\nabla\widetilde{U}_{\ell,\infty}\,,\,\nabla V_\ell\rangle_\Omega = \langle\nabla u\,,\,\nabla V_\ell\rangle_\Omega - \langle\nabla P_\infty\mathcal{L}\widehat{g}_\infty\,,\,\nabla V_\ell\rangle_\Omega \quad \text{for all } V_\ell \in \mathcal{S}^1_D(\mathcal{T}_\ell). \tag{89}$$

Due to the nestedness of the ansatz spaces $\mathcal{S}^1_D(\mathcal{T}_\ell)$, we derive a priori convergence $\widetilde{U}_{\ell,\infty} \xrightarrow{\ell\to\infty} \widetilde{u}_\infty \in H^1(\Omega)$, where $\widetilde{u}_\infty$ denotes the Galerkin solution with respect to the closure of $\bigcup_{\ell=0}^\infty \mathcal{S}^1_D(\mathcal{T}_\ell)$ in $H^1_0(\Omega)$, see e.g. [BV, Lemma 6.1]. With the stability of (88) and (89), we obtain

$$\|\nabla(\widetilde{U}_{\ell,\infty} - \widetilde{U}_\ell)\|_{L^2(\Omega)} \lesssim \|\mathcal{L}_\ell\widehat{g}_\ell - P_\infty\mathcal{L}\widehat{g}_\infty\|_{H^1(\Omega)} \xrightarrow{\ell\to\infty} 0,$$

and therefore $\widetilde{U}_\ell \xrightarrow{\ell\to\infty} \widetilde{u}_\infty$ in $H^1(\Omega)$. Finally, we conclude

$$U_\ell = \widetilde{U}_\ell + \mathcal{L}_\ell\widehat{g}_\ell \xrightarrow{\ell\to\infty} \widetilde{u}_\infty + P_\infty\mathcal{L}\widehat{g}_\infty =: u_\infty \in H^1(\Omega),$$

which concludes the proof. □

With Proposition 19, elementary calculus and the estimator reduction (78) prove estimator convergence $\lim_\ell \varrho_\ell = 0$, cf. [AFP] for the concept of estimator reduction. According to reliability of ϱ_ℓ, this yields convergence of AFEM in the sense that the a priori limit satisfies $u = \lim_\ell U_\ell$. Note, however, that this convergence result is much weaker than the contraction result of Theorem 13. By now, it is unclear how to prove a contraction result if the additional orthogonality relation (34) fails to hold.

7. Numerical Experiment

On the Z-shaped domain $(-1,1)^2\backslash\text{conv}\{(0,0),(-1,-1),(0,-1)\}$, we consider the mixed boundary value problem (1), where the partition of the boundary $\Gamma = \partial\Omega$ into Dirichlet boundary Γ_D and Neumann boundary Γ_N as well as the initial mesh are shown in Figure 4. We prescribe the exact solution $u(x)$ in polar coordinates by

$$u(x) = r^{4/7}\cos(4\varphi/7) \quad \text{for } x = r\,(\cos\varphi, \sin\varphi). \tag{90}$$

Then, $f = -\Delta u \equiv 0$, and the solution u as well as its Dirichlet data $g = u|_{\Gamma_D}$ admit a generic singularity at the reentrant corner $r = 0$.

Figure 2 shows a comparison between uniform and adaptive mesh refinement. For Algorithm 8 based on the modified Dörfler marking, we use $\theta := \vartheta = \theta_1 = \theta_2$. For Algorithm 7 and Algorithm 8, we then vary the adaptivity parameter θ between 0.2 and 0.8. We observe that both adaptive algorithms lead to the optimal convergence rate $\mathcal{O}(N^{-1/2})$ for all choices of θ, whereas uniform refinement leads only to suboptimal convergence behaviour of approximately $\mathcal{O}(N^{-2/7})$.

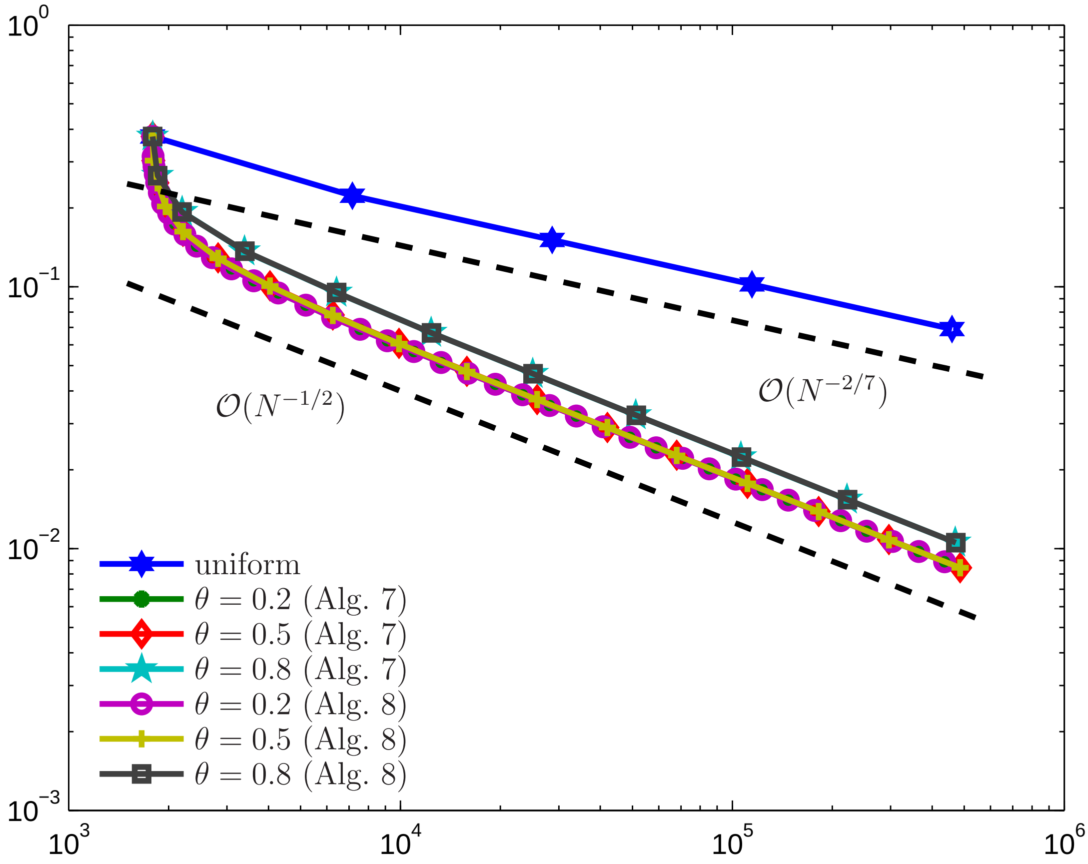

FIGURE 2. Numerical results for ϱ_ℓ for uniform and adaptive mesh-refinement with Algorithm 7 resp. Algorithm 8 and $\theta \in \{0.2, 0.5, 0.8\}$, plotted over the number of elements $N = \#\mathcal{T}_\ell$.

Note that due to $f \equiv 0$, we have $\mathrm{osc}_{\mathcal{E},\ell} \equiv 0$ in this example. In Figure 3, we compare the jump terms

$$\eta_{\Omega,\ell}^2 := \sum_{E \in \mathcal{E}_\ell^\Omega} |E| \| [\partial_n U_\ell] \|_{L^2(E)}^2,$$

the Dirichlet data oscillations $\mathrm{osc}_{D,\ell}$, and the Neumann jump terms

$$\eta_{N,\ell}^2 := \sum_{E \in \mathcal{E}_\ell^N} |E| \| \phi - \partial_n U_\ell \|_{L^2(E)}^2$$

for uniform and adaptive refinement. Due to the corner singularity at $r = 0$, uniform refinement leads to a suboptimal convergence behaviour for $\eta_{\Omega,\ell}$ and even for $\mathrm{osc}_{D,\ell}$ and $\eta_{N,\ell}$, i.e. all contributions of $\varrho_\ell^2 = \eta_{\Omega,\ell}^2 + \eta_{N,\ell}^2 + \mathrm{osc}_{D,\ell}$ show the same poor convergence rate of approximately $\mathcal{O}(N^{-2/7})$. For adaptive mesh-refinement, we observe that the optimal order of convergence is retained, namely $\varrho_\ell \simeq \eta_\ell = \mathcal{O}(N^{-1/2})$ and $\mathrm{osc}_{D,\ell} = \mathcal{O}(N^{-3/4})$. Moreover, we even observe optimal convergence behaviour $\mathrm{osc}_{D,\ell} \simeq \eta_{N,\ell} = \mathcal{O}(N^{-3/4})$ for the boundary contributions of ϱ_ℓ.

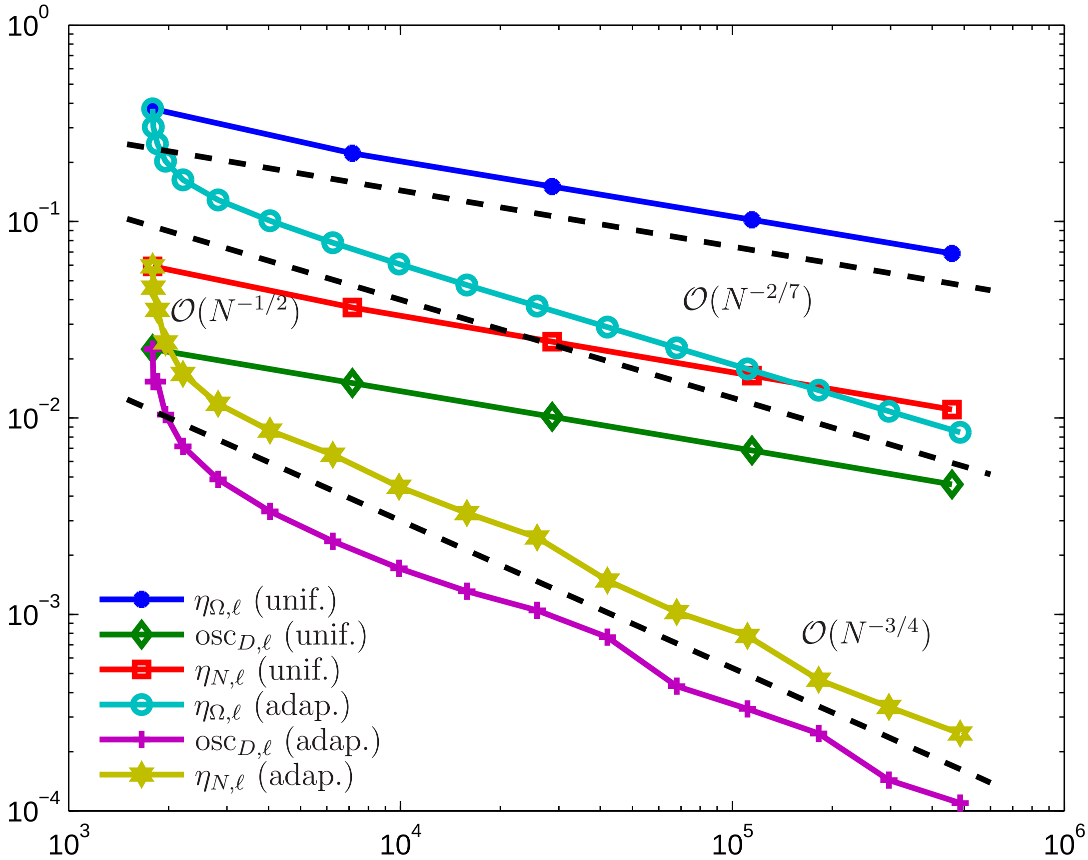

FIGURE 3. Numerical results for $\eta_{\Omega,\ell}$, $\mathrm{osc}_{D,\ell}$, and $\eta_{N,\ell}$ for uniform and adaptive mesh-refinement with Algorithm 7 and $\theta = 0.5$, plotted over the number of elements $N = \#\mathcal{T}_\ell$. Adaptive refinement leads to optimal convergence rates.

Finally, in Figure 4, the initital mesh $\mathcal{T}_0$ and the adaptively generated mesh $\mathcal{T}_9$ with $N =$ 9.074 Elements are visualized. As expected, adaptive refinement is essentially concentrated around the reentrant corner $r = 0$.

REFERENCES

[AFP] M. AURADA, S. FERRAZ-LEITE, D. PRAETORIUS: *Estimator reduction and convergence of adaptive BEM*, ASC Report **09/2010**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2010.

[AGP] M. AURADA, P. GOLDENITS, D. PRAETORIUS: *Convergence of data perturbed adaptive boundary element methods*, ASC Report **40/2009**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2009.

[AO] M. AINSWORTH, T. ODEN: *A posteriori error estimation in finite element analysis*, Wiley-Interscience, New-York, 2000.

[BCD] S. BARTELS, C. CARSTENSEN, G. DOLZMANN: *Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis*, Numer. Math. **99** (2004), 1–24.

[BDD] P. BINEV, W. DAHMEN, R. DEVORE: *Adaptive finite element methods with convergence rates*, Numer. Math. **97** (2004), 219–268.

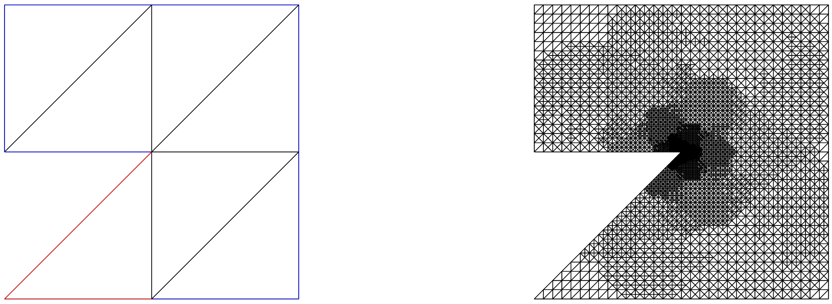

FIGURE 4. Z-shaped domain with initial mesh $\mathcal{T}_0$ and adaptively generated mesh $\mathcal{T}_9$ with $N = 9.074$ for $\theta = 0.5$ in Algorithm 7. The Dirichlet boundary Γ_D is marked red, whereas the blue parts denote the Neumann boundary $\Gamma\backslash\Gamma_D$.

[BM] R. Becker, S. Mao: *Convergence and quasi-optimal complexity of a simple adaptive finite element method*, M2AN Math. Model. Numer. Anal. **43** (2009), 1203–1219.

[BMS] R. Becker, S. Mao, Z. Shi: *A convergent adaptive finite element method with optimal complexity*, Electron. Trans. Numer. Anal. **30** (2008), 291–304.

[BV] I. Babuška, M. Vogelius: *Feedback and adaptive finite element solution of one-dimensional boundary value prbolems*, Numer. Math. **44** (1984), 75–102.

[CP1] C. Carstensen, D. Praetorius: *Averaging techniques for the effective numerical solution of Symms integral equation of the first kind*, SIAM J. Sci. Comp. **27** (2006), 1226–1260.

[CP2] C. Carstensen, D. Praetorius: *Averaging techniques for the a posteriori BEM error control for a hypersingular integral equation in two dimensions*, SIAM J. Sci. Comput. **29** (2007), 782–810.

[CV] C. Carstensen, R. Verfürth: *Edge residuals dominate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal. **36** (1999), 1571–1587.

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

[D] W. Doerfler: *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal. **33** (1996), 1106–1124.

[EFFP] C. Erath, S. Ferraz-Leite, S. Funken, D. Praetorius: *Energy norm based a posteriori error estimation for boundary element methods in two dimensions*, Appl. Numer. Math. **59** (2009), 2713–2734.

[EFGP] C. Erath, S. Funken, P. Goldenits, D. Praetorius: *Simple error estimators for the Galerkin BEM for some hypersingular integral equation in 2D*, ASC Report **20/2009**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2009.

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

[KS] C. Kreuzer, K. Siebert: *Decay rates of adaptive finite elements with Dörfler marking*, Numer. Math., published online first (2010).

[KP] M. Karkulik, D. Praetorius: *Convergence of adaptive 3D BEM for some weakly-singular integral equation*, work in progress (2010).

[McL] W. McLean: *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge 2000.

[MNS] P. Morin, R. Nochetto, K. Siebert: *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal. **18** (2000), 466–488.

[MSV] P. Morin, K. Siebert, A. Veeser: *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Appl. Sci. **18** (2008), 707–737.

[PP] M. Page, D. Praetorius: *Convergence of adaptive FEM for some elliptic obstacle problem*, ASC Report **05/2010**, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2010.

[SZ] L. Scott, S. Zhang: *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), 483–493.

[S07] R. Stevenson: *Optimality of standard adaptive finite element method*, Found. Comput. Math. (2007), 245–269.

[S08] R. Stevenson: *The completion of locally refined simplicial partitions created by bisection*, Math. Comp. **77** (2008), 227–241.

[V] R. Verfürth: *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, 1996.

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

E-mail address: `Michael.Feischl@tuwien.ac.at`

E-mail address: `Marcus.Page@tuwien.ac.at`

E-mail address: `Dirk.Praetorius@tuwien.ac.at` (corresponding author)