CrossMark

Adaptive BEM with inexact PCG solver yields almost optimal computational costs

Thomas Führer[1] · Alexander Haberl[2] · Dirk Praetorius[2] · Stefan Schimanko[2]

Abstract

We consider the preconditioned conjugate gradient method (PCG) with optimal preconditioner in the frame of the boundary element method for elliptic first-kind integral equations. Our adaptive algorithm steers the termination of PCG as well as the local mesh-refinement. Besides convergence with optimal algebraic rates, we also prove almost optimal computational complexity. In particular, we provide an additive Schwarz preconditioner which can be computed in linear complexity and which is optimal in the sense that the condition numbers of the preconditioned systems are uniformly bounded. As model problem serves the 2D or 3D Laplace operator and the associated weakly-singular integral equation with energy space $\widetilde{H}^{-1/2}(\Gamma)$. The main results also hold for the hyper-singular integral equation with energy space $H^{1/2}(\Gamma)$.

Mathematics Subject Classification 65N38 · 65N22 · 65F08 · 65N50 · 41A25 · 65Y20

✉ Stefan Schimanko
  stefan.schimanko@asc.tuwien.ac.at

  Thomas Führer
  tofuhrer@mat.uc.cl

  Alexander Haberl
  alexander.haberl@asc.tuwien.ac.at

  Dirk Praetorius
  dirk.praetorius@asc.tuwien.ac.at

[1] Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Vickuña Mackenna 4860, Santiago, Chile

[2] Institute for Analysis and Scientific Computing, TU Wien, Wiedner Hauptstr. 8–10/E101/4, 1040 Vienna, Austria

🖄 Springer

1 Introduction

1.1 Model problem

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with polyhedral boundary $\partial\Omega$. Let $\Gamma \subseteq \partial\Omega$ be a (relatively) open and connected subset. Given $f : \Gamma \to \mathbb{R}$, we seek the density $\phi^\star : \Gamma \to \mathbb{R}$ of the weakly-singular integral equation

$$(V\phi^\star)(x) := \int_\Gamma G(x - y)\phi^\star(y)\,\mathrm{d}y = f(x) \quad \text{for all } x \in \Gamma, \tag{1}$$

where $G(\cdot)$ denotes the fundamental solution of the Laplace operator, i.e.,

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

Given a triangulation $\mathcal{T}_\bullet$ of Γ, we employ a lowest-order Galerkin boundary element method (BEM) to compute a $\mathcal{T}_\bullet$-piecewise constant function $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$ such that

$$\int_\Gamma (V\phi_\bullet^\star)(x)\,\psi_\bullet(x)\,\mathrm{d}x = \int_\Gamma f(x)\,\psi_\bullet(x)\,\mathrm{d}x \quad \text{for all } \psi_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet). \tag{3}$$

With the numbering $\mathcal{T}_\bullet = \{T_1, \ldots, T_N\}$, consider the standard basis $\{\chi_{\bullet,j} : j = 1, \ldots, N\}$ of $\mathcal{P}^0(\mathcal{T}_\bullet)$ consisting of characteristic functions $\chi_{\bullet,j}$ of $T_j \in \mathcal{T}_\bullet$. We make the ansatz

$$\phi_\bullet^\star = \sum_{k=1}^N x_\bullet^\star[k]\,\chi_{\bullet,k} \quad \text{with coefficient vector } x_\bullet^\star = (x_\bullet^\star[1], \ldots, x_\bullet^\star[N]) \in \mathbb{R}^N. \tag{4}$$

Then, the Galerkin formulation (3) is equivalent to the linear system

$$A_\bullet x_\bullet^\star = b_\bullet \text{ with } A_\bullet[j, k] := \int_{T_j} (V\chi_{\bullet,k})(x)\,\mathrm{d}x, \quad b_\bullet[j] := \int_{T_j} f(x)\,\mathrm{d}x, \tag{5}$$

where the matrix $A_\bullet \in \mathbb{R}^{N \times N}$ is positive definite and symmetric. For a given initial triangulation $\mathcal{T}_0$, we consider an adaptive mesh-refinement strategy of the type

$$\boxed{solve} \longrightarrow \boxed{estimate} \longrightarrow \boxed{mark} \longrightarrow \boxed{refine} \tag{6}$$

which generates a sequence $\mathcal{T}_\ell$ of successively refined triangulations $\mathcal{T}_\ell$ for all $\ell \in \mathbb{N}_0$. We note that the condition number of the Galerkin matrix A_ℓ from (5) depends on the number of elements of $\mathcal{T}_\ell$, as well as the minimal and maximal diameter. Therefore, the step *solve* requires an efficient preconditioner as well as an appropriate iterative solver.

1.2 State of the art

In the last decade, the mathematical understanding of adaptive mesh-refinement has matured. We refer to [8,13,18,21,37,41] for some milestones for adaptive finite element methods for second-order linear elliptic equations, [3,19,20,24,27] for adaptive BEM, and [11] for a general framework of rate-optimality of adaptive mesh-refining algorithms. The interplay between adaptive mesh-refinement, optimal convergence rates, and inexact solvers has been addressed and analyzed for adaptive FEM for linear problems in [4,6,41], for eigenvalue problems in [12], and recently also for strongly monotone nonlinearities in [28]. In particular, all available results for adaptive BEM [3,19,20,24,27] assume that the Galerkin system (5) is solved exactly. Instead, the present work analyzes an adaptive algorithm which steers both, the local mesh-refinement and the iterations of the PCG algorithm.

In principle, it is known [11, Section 7] that convergence and optimal convergence rates are preserved if the linear system is solved inexactly, but with sufficient accuracy. The purpose of this work is to guarantee the latter by incorporating an appropriate stopping criterion for the PCG solver into the adaptive algorithm. Moreover, to prove that the proposed algorithm does not only lead to optimal algebraic convergence rates, but also to (almost) optimal computational costs, we provide an appropriate symmetric and positive definite preconditioner $P_\ell \in \mathbb{R}^{N \times N}$ such that

- first, the matrix-vector products with P_ℓ^{-1} can be computed at linear cost;
- second, the system matrix $P_\ell^{-1/2} A_\ell P_\ell^{-1/2}$ of the preconditioned linear system

$$P_\ell^{-1/2} A_\ell P_\ell^{-1/2} \widetilde{x}_\ell^\star = P_\ell^{-1/2} b_\ell \tag{7}$$

has a uniformly bounded condition number which is independent of $\mathcal{T}_\ell$.

Then, $x_\ell^\star = P_\ell^{-1/2} \widetilde{x}_\ell^\star$ solves the original system (5). To that end, we exploit the multilevel structure of adaptively generated meshes in the framework of adaptive Schwarz methods. For hyper-singular integral equations, such a multilevel additive Schwarz preconditioner has been proposed and analyzed in [22,25] for $d = 2, 3$ and for weakly-singular integral equations in [23] for $d = 2$. In particular, the present work closes this gap by analyzing an optimal additive Schwarz preconditioner for weakly-singular integral equations for $d = 3$. Besides, we refer to the recent work [43] on optimal preconditioning in Hilbert spaces of negative order. We note that the proofs of [22,23] do not transfer to weakly-singular integral equations for $d = 3$. Instead, we build on recent results for finite element discretizations [5,33] which are then transferred to the present BEM setting by use of an abstract concept from [39].

1.3 Outline and main results

Section 2 introduces the functional analytic framework and fixes the necessary notation. Section 3 states our main results. In Sect. 3.1, we define a local multilevel additive Schwarz preconditioner (24) for a sequence of locally refined meshes. Theorem 3 states that the ℓ_2-condition number of the preconditioned systems is uniformly bounded for

all these meshes, i.e., the preconditioner is optimal. In Sect. 3.2, we first state our adaptive algorithm which steers the local mesh-refinement as well as the stopping of the PCG iteration (Algorithm 5). Theorem 8 proves

- that the overall error in the energy norm can be controlled *a posteriori*,
- that the *quasi-error* (which consists of energy norm error plus error estimator) is linearly convergent in each step of the adaptive algorithm (i.e., independent of whether the algorithm decides for local mesh-refinement or for one step of the PCG iteration),
- that the quasi-error even decays with optimal rate (i.e., with each possible algebraic rate) with respect to the degrees of freedom, i.e., Algorithm 5 is *rate optimal* in the sense of, e.g., [11,13,24,41].

Finally, Sect. 3.3 considers the computational costs. Under realistic assumptions on the treatment of the arising discrete integral operators, Corollary 10 states that the quasi-error converges at almost optimal rate (i.e., with rate $s - \varepsilon$ for any $\varepsilon > 0$ if rate $s > 0$ is possible for the exact Galerkin solution) with respect to computational costs, i.e., Algorithm 5 requires *almost optimal computational time*. Section 4 underpins our theoretical findings by some 2D and 3D experiments. The proof of Theorem 3 is given in Sect. 5, the proofs of Theorem 8 and Corollary 10 are given in Sect. 6. The final Sect. 7 shows that our main results also apply to the hyper-singular integral equation.

2 Preliminaries and notation

2.1 Functional analytic setting

We briefly recall the most important facts and refer to [36] for further details and proofs. With the Sobolev space $H^\alpha(\partial\Omega)$ defined as in [36] for $0 \leq \alpha \leq 1$, let $H^\alpha(\Gamma) := \{v|_\Gamma : v \in H^\alpha(\partial\Omega)\}$ be associated with the natural quotient norm. Let $\widetilde{H}^{-\alpha}(\Gamma)$ be the dual space of $H^\alpha(\Gamma)$ with respect to the extended $L^2(\Gamma)$ scalar product $\langle \psi, f \rangle = \int_\Gamma \psi(x) f(x) \, dx$. Then, the single-layer potential V from (1) gives rise to a bounded linear operator $V : \widetilde{H}^{-1/2+s}(\Gamma) \to H^{1/2+s}(\Gamma)$ for all $-1/2 \leq s \leq 1/2$ which is even an isomorphism for $-1/2 < s < 1/2$. For $d = 2$, the latter requires $\mathrm{diam}(\Omega) < 1$ which can always be ensured by scaling of Ω. For $s = 0$, the operator V is even symmetric and elliptic, i.e.,

$$\langle \phi, \psi \rangle := \int_\Gamma (V\phi)(x) \, \psi(x) \, dx \qquad \text{for all } \phi, \psi \in \widetilde{H}^{-1/2}(\Gamma) \tag{8}$$

defines a scalar product and $\|\phi\|^2 := \langle \phi, \phi \rangle$ is an equivalent norm on $\widetilde{H}^{-1/2}(\Gamma)$. For a given right-hand side $f \in H^{1/2}(\Gamma)$, the weakly-singular integral equation (1) can thus equivalently be reformulated as

$$\langle \phi^\star, \psi \rangle = \langle f, \psi \rangle \quad \text{for all } \psi \in \widetilde{H}^{-1/2}(\Gamma). \tag{9}$$

In particular, the Lax–Milgram theorem proves existence and uniqueness of the solution $\phi^\star \in \widetilde{H}^{-1/2}(\Gamma)$ to (9).

2.2 Boundary element method (BEM)

Given a mesh $\mathcal{T}_\bullet$ of Γ, let

$$\mathcal{P}^0(\mathcal{T}_\bullet) := \left\{ \psi_\bullet : \Gamma \to \mathbb{R} \; : \; \forall T \in \mathcal{T}_\bullet \quad \psi_\bullet|_T \text{ is constant} \right\} \tag{10}$$

be the space of $\mathcal{T}_\bullet$-piecewise constant functions. Note that $\mathcal{P}^0(\mathcal{T}_\bullet) \subset L^2(\Gamma) \subset \widetilde{H}^{-1/2}(\Gamma)$. The Galerkin formulation (3) can be reformulated as

$$\langle \phi_\bullet^\star , \; \psi_\bullet \rangle = \langle f , \; \psi_\bullet \rangle \quad \text{for all } \psi_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet). \tag{11}$$

Therefore, the Lax–Milgram theorem proves existence and uniqueness of the discrete solution $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$.

2.3 Mesh-refinement for 2D BEM

For $d = 2$, a mesh $\mathcal{T}_\bullet$ of Γ is a partition into non-degenerate compact line segments. It is called γ-shape regular, if

$$\max \left\{ h_T / h_{T'} \; : \; T, T' \in \mathcal{T}_\bullet \text{ with } T \cap T' \neq \emptyset \right\} \leq \gamma. \tag{12}$$

Here, $h_T := \mathrm{diam}(T) > 0$ denotes the Euclidean diameter of T, i.e., the length of the line segment.

We employ the extended bisection algorithm from [2]. For a mesh $\mathcal{T}_\bullet$ and $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$, let $\mathcal{T}_\circ := \mathrm{refine}(\mathcal{T}_\bullet, \mathcal{M}_\bullet)$ be the coarsest mesh such that all marked elements $T \in \mathcal{M}_\bullet$ have been refined, i.e., $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet \backslash \mathcal{T}_\circ$. We write $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, if there exists $n \in \mathbb{N}_0$, conforming triangulations $\mathcal{T}_0, \dots, \mathcal{T}_n$ and corresponding sets of marked elements $\mathcal{M}_j \subseteq \mathcal{T}_j$ such that

- $\mathcal{T}_\bullet = \mathcal{T}_0$,
- $\mathcal{T}_{j+1} = \mathrm{refine}(\mathcal{T}_j, \mathcal{M}_j)$ for all $j = 0, \dots, n-1$,
- $\mathcal{T}_\circ = \mathcal{T}_n$,

i.e., $\mathcal{T}_\circ$ is obtained from $\mathcal{T}_\bullet$ by finitely many steps of refinement. Note that the bisection algorithm from [2] guarantees, in particular, that all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ are uniformly γ-shape regular, where γ depends only on $\mathcal{T}_\bullet$.

2.4 Mesh-refinement for 3D BEM

For $d = 3$, a mesh $\mathcal{T}_\bullet$ of Γ is a conforming triangulation into non-degenerate compact surface triangles. In particular, we avoid hanging nodes. To ease the presentation, we suppose that the elements $T \in \mathcal{T}_\bullet$ are flat. The triangulation is called γ-shape regular, if
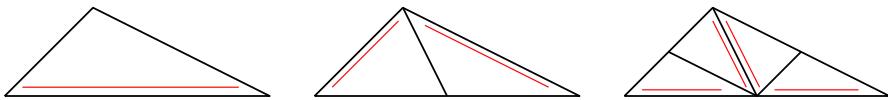
Fig. 1 For newest vertex bisection (NVB) in 2D, each triangle $T \in \mathcal{T}$ has one *reference edge*, indicated by the double line (left). Bisection of T is achieved by halving the reference edge (middle). The reference edges of the sons are always opposite to the new vertex. Recursive application of this refinement rule leads to conforming triangulations

$$\max_{T \in \mathcal{T}_\bullet} \frac{\mathrm{diam}(T)}{h_T} \leq \gamma. \tag{13}$$

Here, $\mathrm{diam}(T)$ denotes the Euclidean diameter of T and $h_T := |T|^{1/2}$ with $|T|$ being the two-dimensional surface measure. Note that γ-shape regularity implies that $h_T \leq \mathrm{diam}(T) \leq \gamma \, h_T$ and hence excludes anisotropic elements.

For 3D BEM, we employ 2D newest vertex bisection (NVB) to refine triangulations locally; see [35,42] for details on the refinement algorithm and Fig. 1 for an illustration. For a mesh $\mathcal{T}_\bullet$ and $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$, we employ the same notation $\mathcal{T}_\circ := \mathrm{refine}(\mathcal{T}_\bullet, \mathcal{M}_\bullet)$ resp. $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ as for $d = 2$.

2.5 A posteriori BEM error control

For $\psi_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet)$ and $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet$, define

$$\eta_\bullet(\mathcal{U}_\bullet, \psi_\bullet)^2 := \sum_{T \in \mathcal{U}_\bullet} \eta_\bullet(T, \psi_\bullet)^2,$$

$$\text{where} \quad \eta_\bullet(T, \psi_\bullet)^2 := h_T \, \|\nabla_\Gamma (f - V\psi_\bullet)\|^2_{L^2(T)}. \tag{14}$$

Here $\nabla_\Gamma(\cdot)$ denotes the arclength derivative for $d = 2$ resp. the surface gradient for $d = 3$. To abbreviate notation, let $\eta_\bullet(\psi_\bullet) := \eta_\bullet(\mathcal{T}_\bullet, \psi_\bullet)$. If $\psi_\bullet = \phi_\bullet^\star$ is the discrete solution to (11), then there holds the reliability estimate (i.e., the global upper bound)

$$\|\!|\!| \phi^\star - \phi_\bullet^\star |\!|\!| \leq C_{\mathrm{rel}} \, \eta_\bullet(\phi_\bullet^\star), \tag{15}$$

where $C_{\mathrm{rel}} > 0$ depends only on Γ and γ-shape regularity of $\mathcal{T}_\bullet$; see [10,17] for $d = 2$ resp. [15] for $d = 3$. Provided that $\phi^\star \in L^2(\Gamma)$, the following weak efficiency

$$\|\!|\!| \phi^\star - \phi_\bullet^\star |\!|\!| + \eta_\bullet(\phi_\bullet^\star) \leq C_{\mathrm{eff}} \, \|h_\bullet^{1/2} (\phi^\star - \phi_\bullet^\star)\|_{L^2(\Gamma)} \tag{16}$$

has recently been proved in [3], where $C_{\mathrm{eff}} > 0$ depends only on Γ and γ-shape regularity of $\mathcal{T}_\bullet$. We note that the weighted L^2-norm on the right-hand side of (16) is only slightly stronger than $\|\!|\!| \cdot |\!|\!| \simeq \|\cdot\|_{\widetilde{H}^{-1/2}(\Gamma)}$, so that one empirically observes $\eta_\bullet(\phi_\bullet^\star) \lesssim \|\!|\!| \phi - \phi_\bullet^\star |\!|\!|$ in practice, cf. [10,15,17]. In certain situations (e.g., weakly-singular integral formulation of the interior 2D Dirichlet problem), one can rigorously prove the latter (strong) efficiency estimate up to higher-order data oscillations; see [2].

2.6 Preconditioned conjugate gradient method (PCG)

Suppose that $P_\bullet, A_\bullet \in \mathbb{R}^{N \times N}$ are symmetric and positive definite matrices. Given $b_\bullet \in \mathbb{R}^N$ and an initial guess $x_{\bullet 0}$, PCG (see [29, Algorithm 11.5.1]) aims to approximate the solution $x_\bullet^\star \in \mathbb{R}^N$ to (5). We note that each step of PCG has the following computational costs:

- $\mathcal{O}(N)$ cost for vector operations (e.g., assignment, addition, scalar product),
- computation of *one* matrix-vector product with $A_\bullet$,
- computation of *one* matrix-vector product with $P_\bullet^{-1}$.

Let $\widetilde{x}_\bullet^\star \in \mathbb{R}^N$ be the solution to (7) and recall that $x_\bullet^\star = P_\bullet^{-1/2} \widetilde{x}_\bullet^\star$. We note that PCG formally applies the conjugate gradient method (CG, see [29, Algorithm 11.3.2]) for the matrix $\widetilde{A}_\bullet := P_\bullet^{-1/2} A_\bullet P_\bullet^{-1/2}$ and the right-hand side $\widetilde{b}_\bullet = P_\bullet^{-1/2} b_\bullet$. The iterates $x_{\bullet k} \in \mathbb{R}^N$ of PCG (applied to $P_\bullet, A_\bullet, b_\bullet$, and the initial guess $x_{\bullet 0}$) and the iterates $\widetilde{x}_{\bullet k}$ of CG (applied to $\widetilde{A}_\bullet, \widetilde{b}_\bullet$, and the initial guess $\widetilde{x}_{\bullet 0} := P_\bullet^{1/2} x_{\bullet 0}$) are formally linked by

$$x_{\bullet k} = P_\bullet^{-1/2} \widetilde{x}_{\bullet k};$$

see [29, Section 11.5]. Moreover, direct computation proves that

$$\|\widetilde{y}_\bullet\|_{\widetilde{A}_\bullet}^2 := \widetilde{y}_\bullet \cdot \widetilde{A}_\bullet \widetilde{y}_\bullet = y_\bullet \cdot A_\bullet y_\bullet =: \|y_\bullet\|_{A_\bullet}^2 \quad \text{for all } \widetilde{y}_\bullet \in \mathbb{R}^N \text{ and } y_\bullet = P_\bullet^{-1/2} \widetilde{y}_\bullet. \tag{17}$$

Consequently, [29, Theorem 11.3.3] for CG (applied to $\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0}$) yields the following lemma for PCG (which follows from the implicit steepest decent approach of CG).

Lemma 1 *Let $A_\bullet, P_\bullet \in \mathbb{R}^{N \times N}$ be symmetric and positive definite, $b_\bullet \in \mathbb{R}^N$, $x_\bullet^\star := A_\bullet^{-1} b_\bullet$, and $x_{\bullet 0} \in \mathbb{R}^N$. Suppose the ℓ_2-condition number estimate*

$$\text{cond}_2(P_\bullet^{-1/2} A_\bullet P_\bullet^{-1/2}) \le C_{\text{pcg}}. \tag{18}$$

Then, the iterates $x_{\bullet k}$ of the PCG algorithm satisfy the contraction property

$$\|x_\bullet^\star - x_{\bullet(k+1)}\|_{A_\bullet} \le q_{\text{pcg}} \|x_\bullet^\star - x_{\bullet k}\|_{A_\bullet} \quad \text{for all } k \in \mathbb{N}_0, \tag{19}$$

where $q_{\text{pcg}} := (1 - 1/C_{\text{pcg}})^{1/2} < 1$. □

If the matrix $A_\bullet \in \mathbb{R}^{N \times N}$ stems from the Galerkin discretization (5) for $\mathcal{T}_\bullet = \{T_1, \ldots, T_N\}$, there is a one-to-one correspondence of vectors $y_\bullet \in \mathbb{R}^N$ and discrete functions $\psi_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet)$ via $\psi_\bullet = \sum_{j=1}^N y_\bullet[j] \chi_{\bullet,j}$. Let $\phi_{\bullet k} \in \mathcal{P}^0(\mathcal{T}_\bullet)$ denote the discrete function corresponding to the PCG iterate $x_{\bullet k} \in \mathbb{R}^N$, while the Galerkin solution $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$ of (11) corresponds to $x_\bullet^\star = A_\bullet^{-1} b_\bullet$. We note the elementary identity

$$\|\phi_\bullet^\star - \phi_{\bullet k}\|^2 = (x_\bullet^\star - x_{\bullet k}) \cdot A_\bullet (x_\bullet^\star - x_{\bullet k}) = \|x_\bullet^\star - x_{\bullet k}\|_{A_\bullet}^2. \tag{20}$$

2.7 Optimal preconditioners

We say that $P_\bullet$ is an *optimal preconditioner*, if $C_{\mathrm{pcg}} \geq 1$ in the ℓ_2-condition number estimate (18) depends only on γ-shape regularity of $\mathcal{T}_\bullet$ and the initial mesh $\mathcal{T}_0$ (and is hence essentially independent of the mesh $\mathcal{T}_\bullet$).

3 Main results

3.1 Optimal additive Schwarz preconditioner

In this work, we consider multilevel additive Schwarz preconditioners that build on the adaptive mesh-hierarchy.

Let $\mathcal{E}_\bullet$ denote the set of all nodes ($d = 2$) resp. edges ($d = 3$) of the mesh $\mathcal{T}_\bullet$ which do not belong to the relative boundary $\partial\Gamma$. Only for $\Gamma = \partial\Omega$, $\mathcal{E}_\bullet$ contains all nodes resp. edges of $\mathcal{T}_\bullet$. For $E \in \mathcal{E}_\bullet$, let $T^\pm \in \mathcal{T}_\bullet$ denote the two unique elements with $T^+ \cap T^- = E$. We define the Haar-type function $\varphi_{\bullet,E} \in \mathcal{P}^0(\mathcal{T}_\bullet)$ (associated to $E \in \mathcal{E}_\bullet$) by

$$\varphi_{\bullet,E}|_T := \begin{cases} \pm\frac{|E|}{|T^\pm|} & \text{for } T \in \{T^+, T^-\}, \\ 0 & \text{else,} \end{cases} \tag{21}$$

where $|E| := 1$ for $d = 2$ and $|E| := \mathrm{diam}(E)$ for $d = 3$. Note that

$$\varphi_{\bullet,E} \in \mathcal{P}^0_*(\mathcal{T}_\bullet) := \left\{ \psi \in \mathcal{P}^0(\mathcal{T}_\bullet) : \int_\Gamma \psi \, \mathrm{d}x = 0 \right\}. \tag{22}$$

For $d = 3$, we additionally suppose that the orientation of each edge E is arbitrary but fixed. We choose $T^+ \in \mathcal{T}_\bullet$ such that ∂T^+ and $E \subset \partial T^+$ have the same orientation.

Given a mesh $\mathcal{T}_0$, suppose that $\mathcal{T}_\ell$ is a sequence of locally refined meshes, i.e., for all $\ell \in \mathbb{N}_0$, there exists a set $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ such that $\mathcal{T}_{\ell+1} = \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$. Then, define

$$\mathcal{E}^\star_\ell := \mathcal{E}_\ell \backslash \mathcal{E}_{\ell-1} \cup \left\{ E \in \mathcal{E}_\ell : \mathrm{supp}(\varphi_{\ell,E}) \subsetneqq \mathrm{supp}(\varphi_{\ell-1,E}) \right\} \quad \text{for all } \ell \geq 1,$$

which consist of new (interior) nodes/edges plus some of their neighbours. We note the following subspace decomposition which is, in general, *not* direct.

Lemma 2 *With $\mathcal{X}_\bullet := \mathcal{P}^0(\mathcal{T}_\bullet)$ and $\mathcal{X}_{\bullet,E} := \mathrm{span}\{\varphi_{\bullet,E}\}$, it holds that*

$$\mathcal{X}_L = \mathcal{X}_0 + \sum_{\ell=1}^L \sum_{E \in \mathcal{E}^\star_\ell} \mathcal{X}_{\ell,E} \quad \text{for all } L \in \mathbb{N}_0. \tag{23}$$

$\square$

Additive Schwarz preconditioners are based on (not necessarily direct) subspace decompositions. Following the standard theory (see, e.g., [44, Chapter 2]), (23) yields

a (local multilevel) preconditioner. To provide its matrix formulation, let $\boldsymbol{I}_{k,\ell} \in \mathbb{R}^{\#\mathcal{T}_\ell \times \#\mathcal{T}_k}$ be the matrix representation of the canonical embedding $\mathcal{P}^0(\mathcal{T}_k) \hookrightarrow \mathcal{P}^0(\mathcal{T}_\ell)$ for $k < \ell$, i.e.,

$$\sum_{i=1}^{\#\mathcal{T}_k} \boldsymbol{x}_k[i]\,\chi_{k,i} = \sum_{i=1}^{\#\mathcal{T}_\ell} \boldsymbol{x}_\ell[i]\,\chi_{\ell,i} \quad \text{for all } \boldsymbol{x}_k \in \mathbb{R}^{\#\mathcal{T}_k} \text{ and } \boldsymbol{x}_\ell := \boldsymbol{I}_{k,\ell}\boldsymbol{x}_k \in \mathbb{R}^{\#\mathcal{T}_\ell}.$$

Let $\boldsymbol{H}_\ell \in \mathbb{R}^{\#\mathcal{T}_\ell \times \#\mathcal{E}_\ell}$ denote the matrix that represents Haar-type functions, i.e.,

$$\varphi_{\ell,E_j} = \sum_{i=1}^{\#\mathcal{T}_\ell} \boldsymbol{H}_\ell[i,j]\chi_{\ell,i} \quad \text{for all } E_j \in \mathcal{E}_\ell.$$

Since only two coefficients per column are non-zero, $\boldsymbol{H}_\ell$ is sparse, while $\boldsymbol{I}_{k,\ell}$ is non-sparse in general. Finally, define the (non-invertible) diagonal matrix $\boldsymbol{D}_\ell \in \mathbb{R}^{\#\mathcal{E}_\ell \times \#\mathcal{E}_\ell}$ by

$$(\boldsymbol{D}_\ell)_{jk} := \begin{cases} \|\,\varphi_{\ell,E_j}\,\|^{-2} & E_j \in \mathcal{E}_\ell^\star \text{ and } j = k, \\ 0 & \text{else.} \end{cases}$$

Then, the matrix representation of the preconditioner associated to (23) reads

$$\boldsymbol{P}_L^{-1} := \boldsymbol{I}_{0,L}\boldsymbol{A}_0^{-1}\boldsymbol{I}_{0,L}^T + \sum_{\ell=1}^{L} \boldsymbol{I}_{\ell,L}\boldsymbol{H}_\ell\boldsymbol{D}_\ell\boldsymbol{H}_\ell^T\boldsymbol{I}_{\ell,L}^T. \tag{24}$$

For $d = 2$, the subsequent Theorem 3 is already proved in [23, Section III.B] for $\Gamma = \partial\Omega$ and in [26, Section 6.3] for $\Gamma \subsetneqq \partial\Omega$. For $d = 3$, we need the following additional assumptions:

- First, suppose that $\Omega \subset \mathbb{R}^3$ is simply connected and $\Gamma = \partial\Omega$.
- Second, let $\widehat{\mathcal{T}_0}$ be a conforming triangulation of Ω into non-degenerate compact simplices such that $\mathcal{T}_0 = \widehat{\mathcal{T}_0}|_\Gamma$ is the induced boundary partition on Γ.

Then, the following theorem is our first main result. The proof is given in Sect. 5.

Theorem 3 *Under the foregoing assumptions, the preconditioner $\boldsymbol{P}_L$ from (24) is optimal, i.e., there holds (18), where $C_{\mathrm{pcg}} \geq 1$ depends only on Ω and $\widehat{\mathcal{T}_0}$, but is independent of $L \in \mathbb{N}$.*

We stress that the matrix in (24) will never be assembled in practice. The PCG algorithm only needs the action of $\boldsymbol{P}_L^{-1}$ on a vector. This can be done recursively by using the embeddings $\boldsymbol{I}_{\ell,\ell+1}$ which are, in fact, sparse. Up to (storing and) inverting $\boldsymbol{A}_0$ on the coarse mesh, the evaluation of $\boldsymbol{P}_L^{-1}\boldsymbol{x}$ can be done in $\mathcal{O}(\#\mathcal{T}_L)$ operations; see, e.g., [22, Section 3.1] for a detailed discussion. If the mesh $\mathcal{T}_L$ is fine compared to the initial mesh $\mathcal{T}_0$ (or if $\boldsymbol{A}_0$ is realized with, e.g., $\mathcal{H}$-matrix techniques), then the computational costs and storage requirements associated with $\boldsymbol{A}_0$ can be neglected.

Remark 4 Our proof for $d = 3$ requires additional assumptions on Ω, $\Gamma = \partial\Omega$, and $\mathcal{T}_0$. As stated above, the case $d = 2$ allows for a different proof (which, however, does not transfer to $d = 3$) and can thus avoid these assumptions; see [23,26]. We believe that Theorem 3 also holds for $d = 3$ and $\Gamma \subsetneqq \partial\Omega$. This is also underpinned by a numerical experiment in Sect. 4.4. The mathematical proof, however, remains open.

3.2 Optimal convergence of adaptive algorithm

We analyze the following adaptive strategy which is driven by the weighted-residual error estimator (14). We note that Algorithm 5 as well as the following results are independent of the precise preconditioning strategy as long as the employed preconditioners are optimal; see Sect. 2.7.

Algorithm 5 *Input: Conforming triangulation $\mathcal{T}_0$ of Γ, adaptivity parameters $0 < \theta \leq 1$ and $\lambda > 0$, and $C_{\mathrm{mark}} > 0$, optimal preconditioning strategy $\boldsymbol{P}_\bullet$ for all $\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}_0)$.*
Loop: *With $k := 0 =: j$ and $\phi_{00} := 0$, iterate the following steps (i)–(vii):*

  (i) *Update counter $(j, k) \mapsto (j, k + 1)$.*
  (ii) *Do one step of the PCG algorithm with the optimal preconditioner $\boldsymbol{P}_j$ to obtain $\phi_{jk} \in \mathcal{P}^0(\mathcal{T}_j)$ from $\phi_{j(k-1)} \in \mathcal{P}^0(\mathcal{T}_j)$.*
  (iii) *Compute the local contributions $\eta_j(T, \phi_{jk})$ of the error estimator for all $T \in \mathcal{T}_j$.*
  (iv) *If $\|\!|\!| \phi_{jk} - \phi_{j(k-1)} \|\!|\!| > \lambda\,\eta_j(\phi_{jk})$, continue with (i).*
  (v) *Otherwise, define $\underline{k}(j) := k$ and determine some set $\mathcal{M}_j \subseteq \mathcal{T}_j$ with up to the multiplicative factor C_{mark} minimal cardinality such that $\theta\,\eta_j(\phi_{jk}) \leq \eta_j(\mathcal{M}_j, \phi_{jk})$.*
  (vi) *Generate $\mathcal{T}_{j+1} := \mathrm{refine}(\mathcal{T}_j, \mathcal{M}_j)$ and define $\phi_{(j+1)0} := \phi_{jk}$.*
  (vii) *Update counter $(j, k) \mapsto (j + 1, 0)$ and continue with (i).*

Output: *Sequences of successively refined triangulations $\mathcal{T}_j$, discrete solutions ϕ_{jk}, and corresponding error estimators $\eta_j(\phi_{jk})$, for all $j \geq 0$ and $k \geq 0$.*   □

Remark 6 The choice $\lambda = 0$ corresponds to the case that (11) is solved exactly, i.e., $\phi_{(j+1)0} = \phi_j^\star$. Then, optimal convergence of Algorithm 5 has already been proved in [2,19,24,27] for weakly-singular integral equations and [20,27] for hyper-singular integral equations. The choice $\theta = 1$ will generically lead to uniform mesh-refinement, where for each mesh all elements $\mathcal{M}_j = \mathcal{T}_j$ are refined in step (vi) of Algorithm 5. Instead, small $0 < \theta \ll 1$, will lead to highly adapted meshes.

Remark 7 Let $\mathcal{Q} := \big\{ (j, k) \in \mathbb{N}_0 \times \mathbb{N}_0 \; : \; \text{index (j, k) is used in Algorithm 5} \big\}$. It holds that $(0, 0) \in \mathcal{Q}$. Moreover, for $j, k \in \mathbb{N}_0$, it holds that

- for $j \geq 1$, $(j, 0) \in \mathcal{Q}$ implies that $(j - 1, 0) \in \mathcal{Q}$,
- for $k \geq 1$, $(j, k) \in \mathcal{Q}$ implies that $(j, k - 1) \in \mathcal{Q}$.

If j is clear from the context, we abbreviate $\underline{k} := \underline{k}(j)$, e.g., $\phi_{j\underline{k}} := \phi_{j\underline{k}(j)}$. In particular, it holds that $\phi_{j\underline{k}} = \phi_{(j+1)0}$. Since PCG (like any Krylov method) provides the exact solution after at most $\#\mathcal{T}_j$ steps, it follows that $1 \leq \underline{k}(j) < \infty$. Finally, we define the ordering

$$(j', k') < (j, k) \quad \overset{\text{def}}{\Longleftrightarrow} \quad \begin{cases} \text{either:} & j' < j \\ \text{or:} & j' = j \text{ and } k' < k \end{cases}$$

for all $(j', k'), (j, k) \in \mathcal{Q}$.

Moreover, let

$$|(j, k)| := \begin{cases} 0, & \text{if } j = 0 = k, \\ \#\{(j', k') \in \mathcal{Q} : (j', k') < (j, k) \text{ and } k' < \underline{k}(j')\}, & \text{if } j > 0 \text{ or } k > 0, \end{cases} \tag{25}$$

be the total number of PCG iterations until the computation of ϕ_{jk}. Note that $j' > j$ and $|(j', k')| = |(j, k)|$ imply that $j' = j + 1$, $k = \underline{k}(j)$, and $k' = 0$ and hence $\phi_{j'k'} = \phi_{jk}$. □

Theorem 8 *The output of Algorithm 5 satisfies the following assertions* (a)–(c). *The constants* $C^{\star}_{\text{rel}}, C^{\star}_{\text{eff}} > 0$ *depend only on* q_{pcg}, Γ, *and the uniform γ-shape regularity of* $\mathcal{T}_j \in \text{refine}(\mathcal{T}_0)$, *whereas* $C_{\text{lin}} > 1$ *and* $0 < q_{\text{lin}} < 1$ *depend additionally only on* θ *and* λ, *and* $C_{\text{opt}} > 0$ *depends additionally only on* s, $\mathcal{T}_0$, *and* $\Lambda_{0\underline{k}}$.

(a) *There exists a constant* $C^{\star}_{\text{rel}} > 0$ *such that*

$$\| \phi^{\star} - \phi_{jk} \| \leq C^{\star}_{\text{rel}} \left(\eta_j(\phi_{jk}) + \| \phi_{jk} - \phi_{j(k-1)} \| \right)$$
$$\text{for all } (j, k) \in \mathcal{Q} \quad \text{with } k \geq 1. \tag{26}$$

There exists a constant $C^{\star}_{\text{eff}} > 0$ *such that, provided that* $\phi^{\star} \in L^2(\Gamma)$, *it holds that*

$$\eta_j(\phi_{jk}) \leq C^{\star}_{\text{eff}} \left(\|h_j^{1/2}(\phi^{\star} - \phi_{jk})\|_{L^2(\Gamma)} + \| \phi_{jk} - \phi_{j(k-1)} \| \right)$$
$$\text{for all } (j, k) \in \mathcal{Q}, k \geq 1. \tag{27}$$

(b) *For arbitrary* $0 < \theta \leq 1$ *and arbitrary* $\lambda > 0$, *there exist constants* $C_{\text{lin}} \geq 1$ *and* $0 < q_{\text{lin}} < 1$ *such that the quasi-error*

$$\Lambda_{jk}^2 := \| \phi^{\star} - \phi_{jk} \|^2 + \eta_j(\phi_{jk})^2 \tag{28}$$

is linearly convergent in the sense of

$$\Lambda_{j'k'} \leq C_{\text{lin}} \, q_{\text{lin}}^{|(j', k')| - |(j, k)|} \, \Lambda_{jk} \quad \text{for all } (j, k), (j', k') \in \mathcal{Q} \text{ with } (j', k') \geq (j, k). \tag{29}$$

(c) *For* $s > 0$, *define the approximation class*

$$\| \phi^{\star} \|_{\mathbb{A}_s} := \sup_{N \in \mathbb{N}_0} \left((N + 1)^s \min_{\substack{\mathcal{T}_{\bullet} \in \text{refine}(\mathcal{T}_0) \\ \#\mathcal{T}_{\bullet} - \#\mathcal{T}_0 \leq N}} \eta_{\bullet}(\phi_{\bullet}^{\star}) \right). \tag{30}$$

Then, for sufficiently small $0 < \theta \ll 1$ and $0 < \lambda \ll 1$, cf. Assumption (68) below, and all $s > 0$, it holds that

$$\|\phi^\star\|_{\mathbb{A}_s} < \infty \Longleftrightarrow \exists\, C_{\mathrm{opt}} > 0 : \sup_{(j,k)\in\mathcal{Q}} \left(\#\mathcal{T}_j - \#\mathcal{T}_0 + 1\right)^s \Lambda_{jk} \le C_{\mathrm{opt}} \|\phi^\star\|_{\mathbb{A}_s} < \infty. \tag{31}$$

Remark 9 By definition, it holds that $\eta_j(\phi_{jk}) \le \Lambda_{jk}$ for all $(j,k) \in \mathcal{Q}$. If $\phi_{jk} \in \{\phi_j^\star, \phi_{j\underline{k}}\}$, then there also holds the converse inequality $\eta_j(\phi_{jk}) \simeq \Lambda_{jk}$. To see this, note that $\phi_{jk} = \phi_j^\star$ and (15) prove that $\Lambda_{jk} \le (1 + C_{\mathrm{rel}})\, \eta_j(\phi_{jk})$. If $\phi_{jk} = \phi_{j\underline{k}}$, then Theorem 8(a) and Step (iv) of Algorithm 5 prove that $\Lambda_{j\underline{k}} \le (1 + C_{\mathrm{rel}}^\star)\, \eta_j(\phi_{j\underline{k}}) + \|\!|\phi_{j\underline{k}} - \phi_{j(\underline{k}-1)}\|\!| \le (1 + C_{\mathrm{rel}}^\star + \lambda)\, \eta_j(\phi_{j\underline{k}})$.

3.3 Almost optimal computational complexity

Suppose that we use $\mathcal{H}^2$-matrices for the efficient treatment of the discrete single-layer integral operator. Recall that the storage requirements (resp. the cost for one matrix-vector multiplication) of an $\mathcal{H}^2$-matrix are of order $\mathcal{O}(Np^2)$, where N is the matrix size and $p \in \mathbb{N}$ is the local block rank. For $\mathcal{H}^2$-matrices (unlike $\mathcal{H}$-matrices), these costs are, in particular, independent of a possibly unbalanced binary tree which underlies the hierarchical data structure [30].

For a mesh $\mathcal{T}_\bullet \in \mathbb{T}$, we employ the local block rank $p = \mathcal{O}(\log(1 + \#\mathcal{T}_\bullet))$ to ensure that the matrix compression is asymptotically exact as $N = \#\mathcal{T}_\bullet \to \infty$, i.e., the error between the exact matrix and the $\mathcal{H}$-matrix decays exponentially fast; see [30]. We stress that we neglect this error in the following and assume that the matrix-vector multiplication (based on the $\mathcal{H}^2$-matrix) yields the exact matrix-vector product.

The computational cost for storing $A_\bullet$ (as well as for one matrix-vector multiplication) is $\mathcal{O}((\#\mathcal{T}_\bullet) \log^2(1 + \#\mathcal{T}_\bullet))$. In an idealized optimal case, the computation of $\phi_\bullet^\star$ is hence (at least) of cost $\mathcal{O}((\#\mathcal{T}_\bullet) \log^2(1 + \#\mathcal{T}_\bullet))$.

We consider the computational costs for one step of Algorithm 5:

- We assume that one step of the PCG algorithm with the employed optimal preconditioner is of cost $\mathcal{O}((\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j))$; cf. the preconditioner from Sect. 3.1.
- We assume that we can compute $\eta_j(\psi_j)$ for any $\psi_j \in \mathcal{P}^0(\mathcal{T}_j)$ (by means of numerical quadrature) with $\mathcal{O}((\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j))$ operations.
- Clearly, the Dörfler marking in Step (v) can be done in $\mathcal{O}((\#\mathcal{T}_j) \log(1 + \#\mathcal{T}_j))$ operations by sorting. Moreover, for $C_{\mathrm{mark}} = 2$, Stevenson [41] proposed a realization of the Dörfler marking based on binning, which can be performed at linear cost $\mathcal{O}(\#\mathcal{T}_j)$.
- Finally, the mesh-refinement in Step (vi) can be done in linear complexity $\mathcal{O}(\#\mathcal{T}_j)$ if the data structure is appropriate.

Overall, one step of Algorithm 5 is thus done in $\mathcal{O}((\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j))$ operations. However, an adaptive step $(j', k') \in \mathcal{Q}$ depends on the full history of previous steps.

- Hence, the cumulative computational complexity for the adaptive step $(j', k') \in \mathcal{Q}$ is of order $\mathcal{O}\big(\sum_{(j,k)\le(j',k')}(\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j)\big)$.

The following corollary proves that Algorithm 5 does not only lead to convergence of the quasi-error Λ_{jk} with optimal rate with respect to the degrees of freedom (see Theorem 8), but also with *almost* optimal rate with respect to the computational costs.

Corollary 10 *For $j \in \mathbb{N}_0$, let $\widehat{\mathcal{T}}_{j+1} = \mathrm{refine}(\widehat{\mathcal{T}}_j, \widehat{\mathcal{M}}_j)$ with arbitrary $\widehat{\mathcal{M}}_j \subseteq \widehat{\mathcal{T}}_j$ and $\widehat{\mathcal{T}}_0 = \mathcal{T}_0$. Let $s > 0$ and suppose that the corresponding error estimator $\widehat{\eta}_j(\widehat{\phi}_j^\star)$ converges at rate s with respect to the single-step computational costs, i.e.,*

$$\sup_{j \in \mathbb{N}_0} \left[(\#\widehat{\mathcal{T}}_j) \, \log^2(1 + \#\widehat{\mathcal{T}}_j) \right]^s \widehat{\eta}_j(\widehat{\phi}_j^\star) < \infty. \tag{32}$$

Suppose that λ and θ satisfy the assumptions of Theorem 8(c). Then, the quasi-errors Λ_{jk} generated by Algorithm 5 converge almost at rate s with respect to the cumulative computational costs, i.e.,

$$\sup_{(j',k') \in \mathcal{Q}} \left[\sum_{(j,k) \le (j',k')} (\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j) \right]^{s-\varepsilon} \Lambda_{j'k'} < \infty \quad \text{for all } \varepsilon > 0. \tag{33}$$

4 Numerical experiments

In this section, we present numerical experiments that underpin our theoretical findings. We use lowest-order BEM for direct and indirect formulations in 2D as well as 3D. For each problem, we compare the performance of Algorithm 5 for

- different values of $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$,
- different values of $\theta \in \{0.2, 0.4, 0.6, 0.8, 1\}$,

where $\theta = 1$ corresponds to uniform mesh-refinement. In particular, we monitor the condition numbers of the arising BEM systems for diagonal preconditioning [7], the proposed additive Schwarz preconditioning from Sect. 3.1, and no preconditioning. The 2D implementation is based on our MATLAB implementation HILBERT [1], while the 3D implementation relies on an extension of the BEM++ library [40].

4.1 Slit problem in 2D

Let $\Gamma := (-1, 1) \times \{0\}$, cf. Fig. 2. We consider

$$V\phi = 1 \quad \text{on } \Gamma. \tag{34}$$

The unique exact solution of (34) reads $\phi^\star(x, 0) := -2x/\sqrt{1 - x^2}$. For uniform mesh-refinement, we thus expect a convergence order of $\mathcal{O}(N^{-1/2})$, while the optimal rate is $\mathcal{O}(N^{-3/2})$ with respect to the number of elements.

In Fig. 3, we compare Algorithm 5 for different values for θ and λ as well as uniform mesh-refinement. Uniform mesh-refinement leads only to the rate $\mathcal{O}(N^{-1/2})$, while
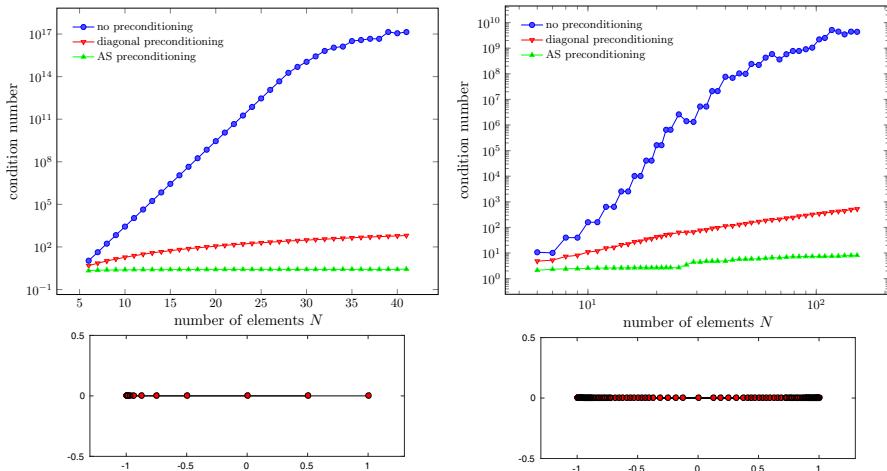
Fig. 2 Example 4.1: condition numbers of the preconditioned and non-preconditioned Galerkin matrix for an artificial refinement towards the left end point (left) and for the matrices arising from Algorithm 5 (right)
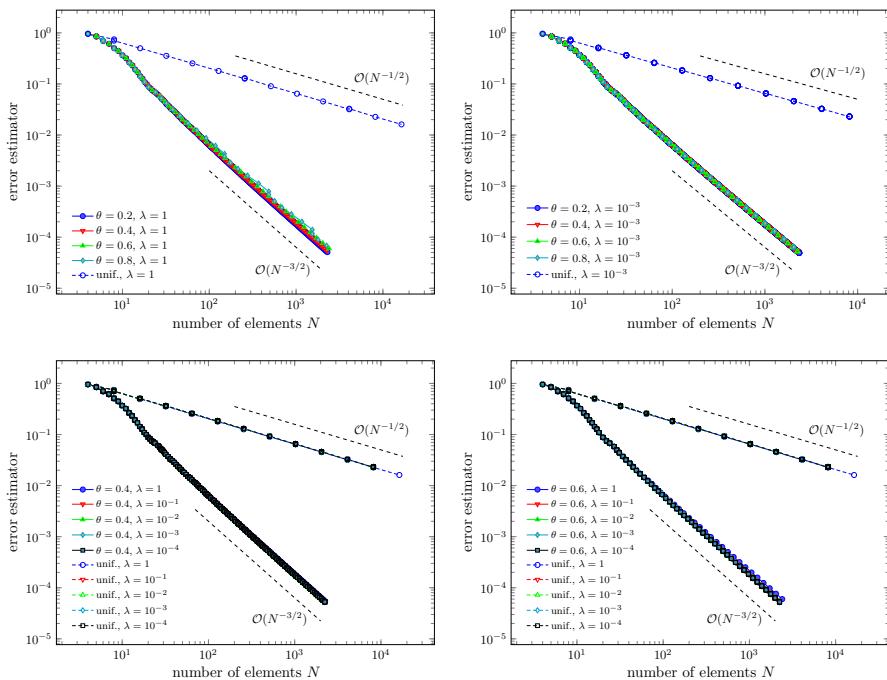

Fig. 3 Example 4.1: estimator convergence for fixed values of λ (left: $\lambda = 1$, right: $\lambda = 10^{-3}$) and $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ (top) and for fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, \ldots, 10^{-4}\}$ (bottom)
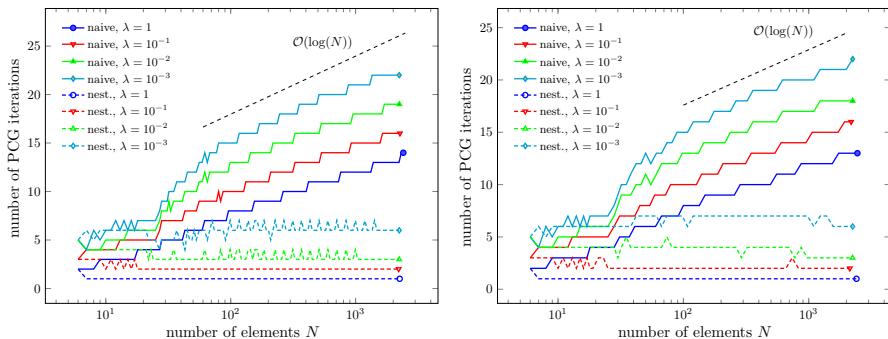
Fig. 4 Example 4.1: number of PCG iterations in Algorithm 5 for nested iteration (dashed lines), i.e., $\phi_{(j+1)0} := \phi_{j\underline{k}}$ in Step (vi), and naive initial guess (solid lines), i.e., $\phi_{(j+1)0} := 0$. We compare fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}\}$

adaptivity, independently of the value of θ and λ, regains the optimal rate $\mathcal{O}(N^{-3/2})$. A naive initial guess in Step (vi) of Algorithm 5 (i.e., if $\phi_{(j+1)0} := 0$) leads to a logarithmical growth of the number of PCG iterations, whereas for nested iteration $\phi_{(j+1)0} := \phi_{j\underline{k}}$ (as formulated in Algorithm 5) the number of PCG iterations stays uniformly bounded, cf. Fig. 4. Finally, Fig. 2 shows the condition numbers for an artificial refinement towards the left end point and for Algorithm 5 with $\lambda = 10^{-3}$ and $\theta = 0.5$.

4.2 Z-shaped domain in 2D

Let $\Gamma := \partial\Omega$ be the boundary of the Z-shaped domain with reentrant corner at the origin $(0, 0)$, cf. Fig. 5. The right-hand side is given by $f = (K + 1/2)g$ with the double-layer operator $K : H^{1/2}(\Gamma) \to H^{1/2}(\Gamma)$. We note that the weakly-singular integral equation (1) is then equivalent to the Dirichlet problem

$$-\Delta u = 0 \quad \text{in } \Omega \qquad \text{subject to} \qquad u = g \quad \text{on } \Gamma. \tag{35}$$

We prescribe the exact solution in 2D polar coordinates as

$$u(x) = r^{4/7}\cos(4\,\boldsymbol{x}i/7) \quad \text{with } x = r(\cos\boldsymbol{x}i, \sin\boldsymbol{x}i). \tag{36}$$

Then, u admits a generic singularity at the reentrant corner. The exact solution $\phi^\star$ of (1) is just the normal derivative of the solution u.

We expect a convergence order of $\mathcal{O}(N^{-4/7})$ for uniform mesh-refinement, and the optimal rate $\mathcal{O}(N^{-3/2})$ for the adaptive strategy, which is seen in Fig. 6 for different values of θ and λ. A naive initial guess in Step (vi) of Algorithm 5 (i.e., if $\phi_{(j+1)0} := 0$) leads to a logarithmical growth of the number of PCG iterations, whereas for nested iteration $\phi_{(j+1)0} := \phi_{j\underline{k}}$ the number of PCG iterations stays uniformly bounded, cf. Fig. 7. Figure 5 shows the condition numbers for an artificial refinement towards the
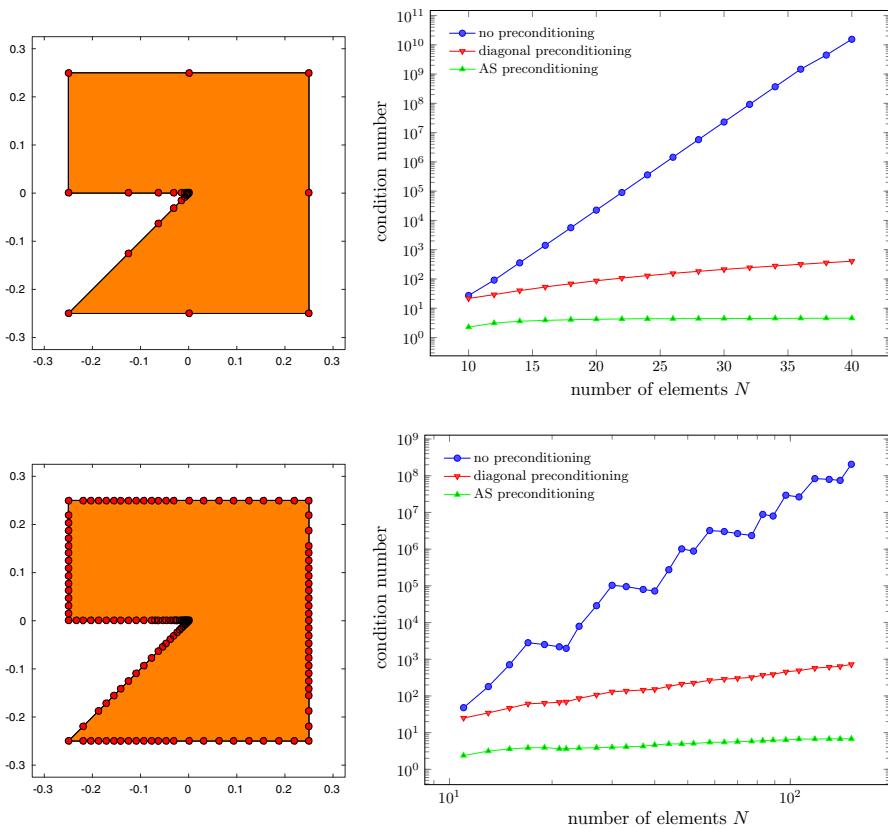
Fig. 5 Example 4.2: condition numbers of the preconditioned and non-preconditioned Galerkin matrix for an artificial refinement towards the reentrant corner (top) and for Algorithm 5 (bottom)

reentrant corner as well as the condition numbers for Algorithm 5 with $\lambda = 10^{-3}$ and $\theta = 0.5$.

4.3 L-shaped domain in 3D

Let $\Gamma := \partial\Omega$ be the boundary of the L-shaped domain $\Omega = (-1, 1)^3 \setminus ([-1, 0] \times [0, 1] \times [-1, 1])$, cf. Fig. 8. The right-hand side is given by $f = (K + 1/2)g$ with the double-layer operator $K \colon H^{1/2}(\Gamma) \to H^{1/2}(\Gamma)$. Again, the weakly-singular integral equation (1) is then equivalent to the Dirichlet problem (35). We prescribe the exact solution in 3D cylindrical coordinates as

$$u(x) = z\, r^{2/3} \cos(2/3\, (xi - \pi/4)) \quad \text{with } x = (r\cos xi, r\sin xi, z). \qquad (37)$$

Note that u admits a singularity along the reentrant edge. The exact solution $\phi^\star$ of (1) is just the normal derivative of the exact solution u.
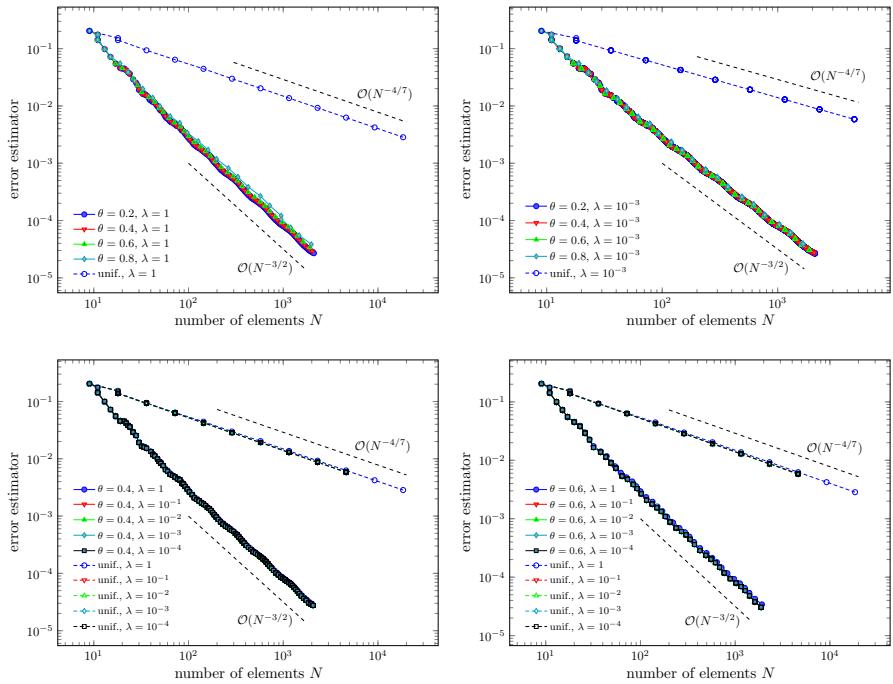
Fig. 6 Example 4.2: estimator convergence for fixed values of λ (left: $\lambda = 1$, right: $\lambda = 10^{-3}$) and $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ (top) and for fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, \ldots, 10^{-4}\}$ (bottom)
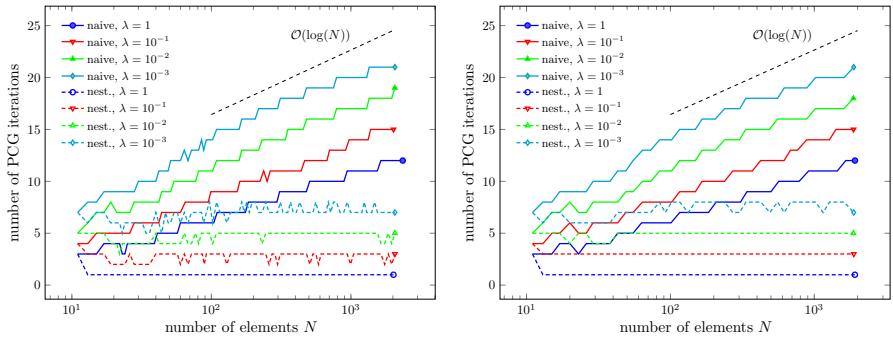
Fig. 7 Example 4.2: number of PCG iterations in Algorithm 5 for nested iteration (dashed lines), i.e., $\phi_{(j+1)0} := \phi_{j\underline{k}}$ in Step (vi), and naive initial guess (solid lines), i.e., $\phi_{(j+1)0} := 0$. We compare fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, 10^{-2}, 10^{-3}\}$

In Fig. 9, we compare Algorithm 5 with different values for θ and λ to uniform mesh-refinement. Un7767iform mesh-refinement leads only to a reduced rate of $\mathcal{O}(N^{-1/2})$, while adaptivity, independently of θ and λ, leads to the improved rate of approximately $\mathcal{O}(N^{-2/3})$. While one would expect $\mathcal{O}(N^{-3/4})$ for smooth exact solutions $\phi^\star$, this would require anisotropic elements along the reentrant edge for the present solution
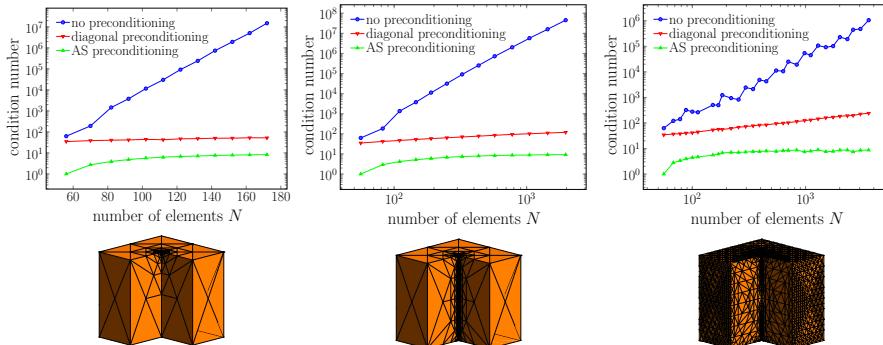
Fig. 8 Example 4.3: condition numbers of the preconditioned and non-preconditioned Galerkin matrix for an artificial refinement towards one reentrant corner (left) or edge (middle), and for Algorithm 5 (right)
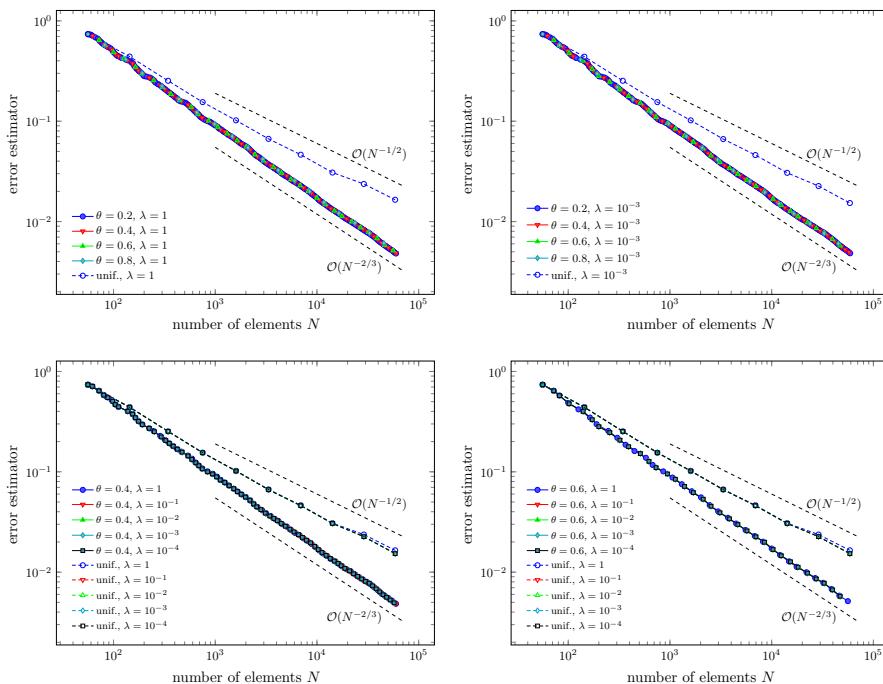

Fig. 9 Example 4.3: estimator convergence for fixed values of λ (left: $\lambda = 1$, right: $\lambda = 10^{-3}$) and $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ (top) and for fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, \ldots, 10^{-4}\}$ (bottom)

$\phi^\star = \partial_n u$. Since NVB guarantees uniform γ-shape regularity of the meshes, the latter is not possible and hence leads to a reduced optimal rate. Finally, Fig. 8 shows the condition numbers for (diagonal or additive Schwarz) preconditioning and no preconditioning for artificial refinements towards one reentrant corner or the reentrant edge as well as the condition numbers of the matrices arising from Algorithm 5 with $\lambda = 10^{-3}$ and $\theta = 0.5$.
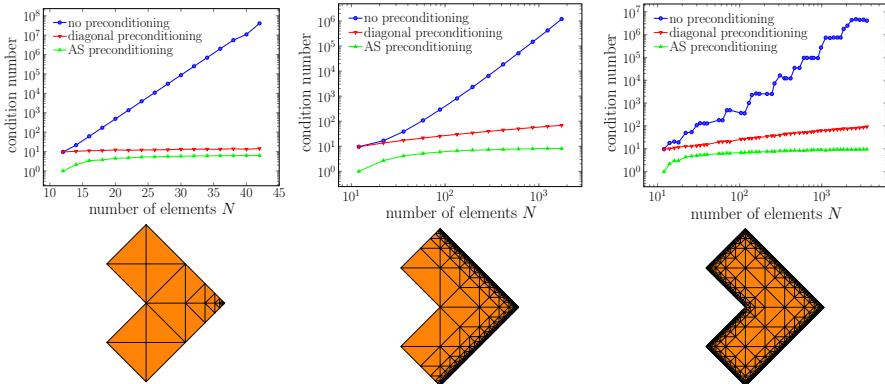
Fig. 10 Example 4.4: condition numbers of the preconditioned and non-preconditioned Galerkin matrix for an artificial refinement towards the right corner (left), the right edges (middle), and for Algorithm 5 (right)
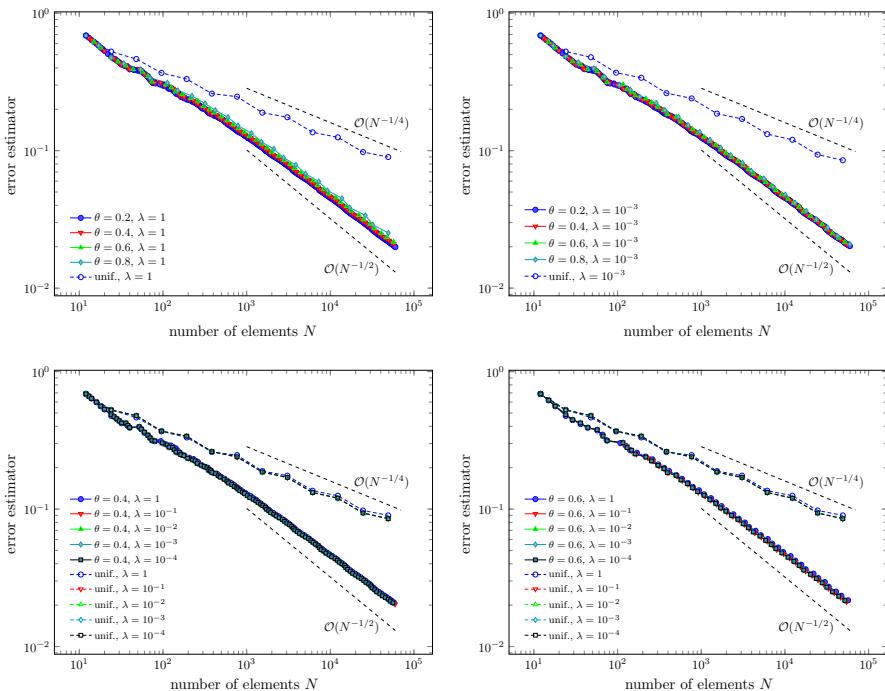

Fig. 11 Example 4.4: estimator convergence for fixed values of λ (left: $\lambda = 1$, right: $\lambda = 10^{-3}$) and $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ (top) and for fixed values of θ (left: $\theta = 0.4$, right: $\theta = 0.6$) and $\lambda \in \{1, 10^{-1}, \ldots, 10^{-4}\}$ (bottom)

4.4 Screen problem in 3D

Let $\Gamma := ((-1, 1)^2 \backslash [0, 1]^2) \times \{0\}$, rotated by $3\pi/4$, cf. Figure 10. We consider the weakly-singular integral equation $V\phi^\star = 1$ on Γ. The exact solution $\phi^\star \in \widetilde{H}^{-1/2}(\Gamma)$ is unknown.
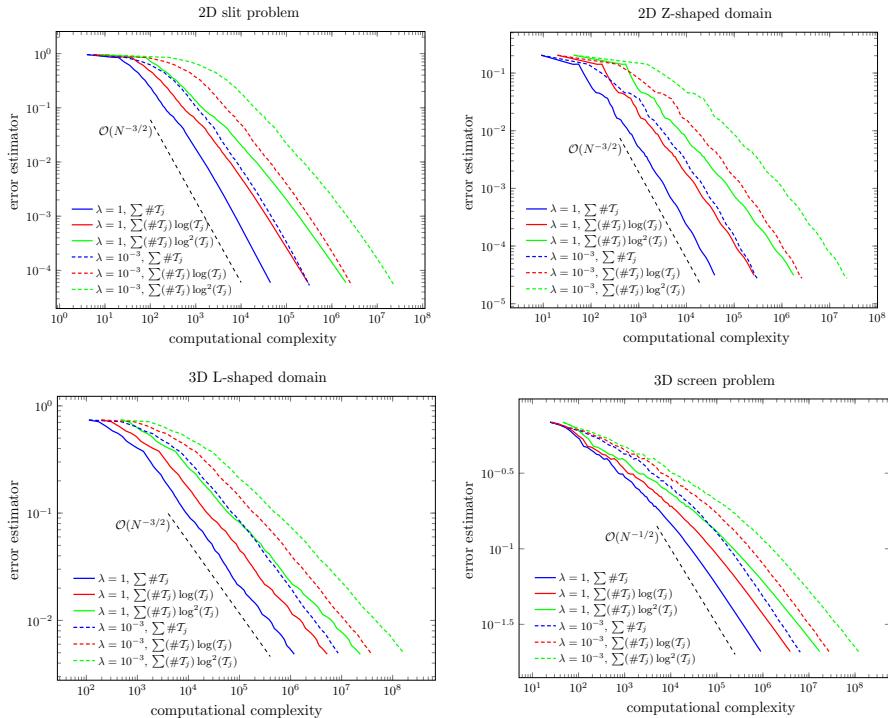
Fig. 12 To underline the quasi-optimal computational complexity of Algorithm 5, we plot the error estimator $\eta_j(\phi_{j\underline{k}})$ in the different experiments over the cumulative quantities $\sum_{(j,k)\leq(j',k')} \#\mathcal{T}_j$, $\sum_{(j,k)\leq(j',k')} (\#\mathcal{T}_j) \log(\#\mathcal{T}_j)$ and $\sum_{(j,k)\leq(j',k')} (\#\mathcal{T}_j) \log^2(\#\mathcal{T}_j)$ for $\theta = 0.4$ and $\lambda \in \{1, 10^{-3}\}$

For the numerical solution of the Galerkin system, we employ PCG with the additive Schwarz preconditioner from Sect. 3.1. We note that Theorem 3 does not cover this setting. In particular, we note that the proposed additive Schwarz preconditioner from Sect. 3.1 appears to be optimal (cf. Fig. 10), while the mathematical optimality proof still remains open for screens.

In Fig. 11, we compare Algorithm 5 with different values for θ and λ to uniform mesh-refinement. We see that uniform mesh-refinement leads only to a reduced rate of $\mathcal{O}(N^{-1/4})$, while adaptivity, independently of θ and λ, leads to the improved rate of approximately $\mathcal{O}(N^{-1/2})$.

4.5 Computational complexity

With Fig. 12, we aim to underpin the almost optimal computational complexity of Algorithm 5 (see Corollary 10). To this end, we plot the error estimator $\eta_j(\phi_{jk})$ over the cumulative sums

$$\sum_{(j,k)\leq(j',k')} \mathrm{cost}(\mathcal{T}_j) \quad \text{with} \quad \mathrm{cost}(\mathcal{T}_j) \in \left\{ \#\mathcal{T}_j \,, \, (\#\mathcal{T}_j)\log(\#\mathcal{T}_j) \,, \, (\#\mathcal{T}_j)\log^2(\#\mathcal{T}_j) \right\}$$

for $\theta = 0.4$ and $\lambda \in \{1, 10^{-3}\}$. The negative impact of the logarithmic terms on the (preasymptotic) convergence rate is clearly visible.

5 Proof of Theorem 3 (optimal multilevel preconditioner)

For $d = 2$, we refer to [23,26] and thus focus only on $d = 3$ and $\Gamma = \partial\Omega$. Due to our additional assumption, $\mathcal{T}_0 = \widehat{\mathcal{T}}_0|_\Gamma$ is the restriction of a conforming simplicial triangulation $\widehat{\mathcal{T}}_0$ of Ω to the boundary Γ. Moreover, 2D NVB refinement of $\mathcal{T}_0$ (on the boundary Γ) is a special case of 3D NVB refinement of $\widehat{\mathcal{T}}_0$ (in the volume Ω) plus restriction to the boundary; see, e.g., [42]. Hence, each mesh $\mathcal{T}_\bullet \in \mathbb{T} = \mathrm{refine}(\mathcal{T}_0)$ is the restriction of a conforming NVB refinement $\widehat{\mathcal{T}}_\bullet \in \widehat{\mathbb{T}} := \mathrm{refine}(\widehat{\mathcal{T}}_0)$, i.e., $\mathcal{T}_\bullet = \widehat{\mathcal{T}}_\bullet|_\Gamma$. Throughout, let $\widehat{\mathcal{T}}_\bullet \in \widehat{\mathbb{T}}$ be the coarsest extension of $\mathcal{T}_\bullet \in \mathbb{T}$. Recall that NVB is a binary refinement rule. Therefore, $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ also implies that $\widehat{\mathcal{T}}_\circ \in \mathrm{refine}(\widehat{\mathcal{T}}_\bullet)$. Finally, we note that all triangulations $\widehat{\mathcal{T}}_\bullet \in \widehat{\mathbb{T}}$ are uniformly γ-shape regular, i.e.,

$$\max_{\widehat{T} \in \widehat{\mathcal{T}}_\bullet} \frac{\mathrm{diam}(\widehat{T})}{|\widehat{T}|^{1/3}} \le \gamma < \infty.$$

where γ depends only on $\widehat{\mathcal{T}}_0$.

Our argument adapts ideas from [32], where a subspace decomposition for the lowest-order Nédélec space $\mathcal{ND}^1(\widehat{\mathcal{T}}_\bullet)$ (see, e.g., [34]) in $\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)$ implies a decomposition of the corresponding discrete trace space. While the original idea dates back to [39], a nice summary of the argument is found in [32, Section 2].

Remark 11 (i) Our proof is based on the construction of an extension operator from $\mathcal{P}_*^0(\mathcal{T}_\bullet)$ to $\mathcal{ND}^1(\widehat{\mathcal{T}}_\bullet)$, see Lemma 13 below. It is not clear if such an operator can be constructed for the case $\Gamma \subsetneqq \partial\Omega$.

(ii) In [31], a subspace decomposition of the lowest-order Raviart–Thomas space $\mathcal{RT}^0(\widehat{\mathcal{T}}_\bullet)$ (see, e.g., [45]) in $\boldsymbol{H}(\mathbf{div}\,;\,\Omega)$ implies a decomposition of the corresponding normal trace space $\mathcal{P}^0(\mathcal{T}_\bullet)$. Due to different scaling properties of the Raviart–Thomas basis functions [in the $\boldsymbol{H}(\mathbf{div}\,;\,\Omega)$ norm] and their normal trace (in the $H^{-1/2}(\Gamma)$ norm), this argument does not apply in our case.

5.1 Discrete spaces and extensions

Let $\widehat{\mathcal{E}}_\bullet$ (resp. $\widehat{\mathcal{N}}_\bullet$) denote the set of all edges (resp. all nodes) of $\widehat{\mathcal{T}}_\bullet \in \widehat{\mathbb{T}}$. For each node $\boldsymbol{x} \in \widehat{\mathcal{N}}_\bullet$, let $\eta_{\bullet,\boldsymbol{x}} \in \mathcal{S}^1(\widehat{\mathcal{T}}_\bullet)$ be the corresponding hat function, i.e., $\eta_{\bullet,\boldsymbol{x}}$ is $\widehat{\mathcal{T}}_\bullet$-piecewise affine and globally continuous with $\eta_{\bullet,\boldsymbol{x}}(\boldsymbol{y}) = \delta_{\boldsymbol{xy}}$ for all $\boldsymbol{x}, \boldsymbol{y} \in \widehat{\mathcal{N}}_\bullet$. For $E \in \widehat{\mathcal{E}}_\bullet$, let $\boldsymbol{u}_{\bullet,E} \in \mathcal{ND}^1(\widehat{\mathcal{T}}_\bullet)$ denote the corresponding Nédélec basis function, i.e., for $K \in \widehat{\mathcal{T}}_\bullet$ with $E = \mathrm{conv}\{\boldsymbol{x}, \boldsymbol{y}\} \subset \partial K$, it holds that

$$\boldsymbol{u}_{\bullet,E}|_K = C(\eta_{\bullet,\boldsymbol{x}}\nabla\eta_{\bullet,\boldsymbol{y}} - \eta_{\bullet,\boldsymbol{y}}\nabla\eta_{\bullet,\boldsymbol{x}}), \tag{38}$$

where $C > 0$ is chosen such that $\int_{E'} \boldsymbol{u}_{\bullet,E}\,ds = |E|\,\delta_{EE'}$ for all $E, E' \in \widehat{\mathcal{E}}_\bullet$. Scaling arguments yield the next lemma. The proof follows the lines of [32, Lemma 5.7].

Lemma 12 *For $E \in \mathcal{E}_\bullet$, recall the Haar function $\varphi_{\bullet,E} \in \mathcal{P}^0(\mathcal{T}_\bullet)$ from* (21). *Let $\boldsymbol{u}_{\bullet,E} \in \mathcal{ND}^1(\widehat{\mathcal{T}_\bullet})$ denote the corresponding Nédélec basis function; see* (38). *Then,*

$$\varphi_{\bullet,E} = \operatorname{\mathbf{curl}} \boldsymbol{u}_{\bullet,E} \cdot \boldsymbol{n}|_\Gamma \text{ and}$$

$$C^{-1}\|\varphi_{\bullet,E}\|_{H^{-1/2}(\Gamma)} \leq \|\boldsymbol{u}_{\bullet,E}\|_{H(\mathbf{curl}\,;\Omega)} \leq C\,\|\varphi_{\bullet,E}\|_{H^{-1/2}(\Gamma)}, \tag{39}$$

where $C > 0$ depends only on Ω and the γ-shape regularity of $\widehat{\mathcal{T}_\bullet}$.

Proof A straightforward calculation shows that $\varphi_{\bullet,E} = \operatorname{\mathbf{curl}} \boldsymbol{u}_{\bullet,E} \cdot \boldsymbol{n}|_\Gamma$. Then, continuity of the normal trace operator yields that

$$\|\varphi_{\bullet,E}\|_{H^{-1/2}(\Gamma)} \lesssim \|\operatorname{\mathbf{curl}} \boldsymbol{u}_{\bullet,E}\|_{H(\mathbf{div}\,;\Omega)} = \|\operatorname{\mathbf{curl}} \boldsymbol{u}_{\bullet,E}\|_{L^2(\Omega)} \leq \|\boldsymbol{u}_{\bullet,E}\|_{H(\mathbf{curl}\,;\Omega)}.$$

Furthermore, scaling arguments prove that

$$\|\boldsymbol{u}_{\bullet,E}\|_{H(\mathbf{curl}\,;\Omega)} \simeq \|\operatorname{\mathbf{curl}} \boldsymbol{u}_{\bullet,E}\|_{L^2(\Omega)} \simeq |E|^{1/2}$$
$$\simeq |E|^{1/2}\|\varphi_{\bullet,E}\|_{L^2(\Gamma)} \lesssim \|\varphi_{\bullet,E}\|_{H^{-1/2}(\Gamma)},$$

where we have finally applied an inverse estimate. This concludes the proof. □

The following lemma holds for (simply) connected Lipschitz domains Ω and follows essentially from [5]. Recall $\mathcal{P}_*^0(\mathcal{T}_\bullet)$ from (22).

Lemma 13 *There exists a linear operator $\boldsymbol{E}_\bullet : \mathcal{P}_*^0(\mathcal{T}_\bullet) \to \mathcal{ND}^1(\widehat{\mathcal{T}_\bullet})$ such that*

$$\operatorname{\mathbf{curl}}(\boldsymbol{E}_\bullet\psi_\bullet) \cdot \boldsymbol{n}|_\Gamma = \psi_\bullet \text{ and } \|\boldsymbol{E}_\bullet\psi_\bullet\|_{H(\mathbf{curl}\,;\Omega)}$$
$$\leq C\,\|\psi_\bullet\|_{H^{-1/2}(\Gamma)} \text{ for all } \psi_\bullet \in \mathcal{P}_*^0(\mathcal{T}_\bullet). \tag{40}$$

The constant $C > 0$ depends only on Ω and γ-shape regularity of $\widehat{\mathcal{T}_\bullet}$.

Proof Let $\psi_\bullet \in \mathcal{P}_*^0(\mathcal{T}_\bullet)$. First, [5, Theorem 2.1] provides $\boldsymbol{\sigma}_\bullet \in \mathcal{RT}^0(\widehat{\mathcal{T}_\bullet})$ with

$$\boldsymbol{\sigma}_\bullet \cdot \boldsymbol{n}|_\Gamma = \psi_\bullet, \quad \operatorname{\mathbf{div}} \boldsymbol{\sigma}_\bullet = 0, \quad \text{and} \quad \|\boldsymbol{\sigma}_\bullet\|_{H(\mathbf{div}\,;\Omega)} \lesssim \|\psi_\bullet\|_{H^{-1/2}(\partial\Omega)}.$$

Then, [5, Lemma 4.3] provides $\boldsymbol{E}_\bullet\psi_\bullet := \boldsymbol{v}_\bullet \in \mathcal{ND}^1(\widehat{\mathcal{T}_\bullet})$ such that

$$\operatorname{\mathbf{curl}} \boldsymbol{v}_\bullet = \boldsymbol{\sigma}_\bullet \quad \text{and} \quad \|\boldsymbol{v}_\bullet\|_{H(\mathbf{curl}\,;\Omega)} \lesssim \|\boldsymbol{\sigma}_\bullet\|_{H(\mathbf{div}\,;\Omega)}.$$

Combining these results, we conclude the proof. □

5.2 Abstract additive Schwarz preconditioners

Let $\mathcal{X}$ denote some finite dimensional Hilbert space with norm $\|\cdot\|_\mathcal{X}$ and subspace decomposition

$$\mathcal{X} = \sum_{i \in \mathcal{I}} \mathcal{X}_i,$$

where $\mathcal{I}$ is a finite index set. The additive Schwarz operator is given by $\mathcal{S} = \sum_{i \in \mathcal{I}} \mathcal{S}_i$, where $\mathcal{S}_i$ is the $\mathcal{X}$-orthogonal projection onto $\mathcal{X}_i$, i.e.,

$$\langle \mathcal{S}_i x \,, \, x_i \rangle_{\mathcal{X}} = \langle x \,, \, x_i \rangle_{\mathcal{X}} \quad \text{for all } x_i \in \mathcal{X}_i \text{ and all } x \in \mathcal{X},$$

where $\langle \cdot \,, \, \cdot \rangle_{\mathcal{X}}$ denotes the scalar product on $\mathcal{X}$. Then, the operator $\mathcal{S}$ is positive definite and symmetric (with respect to $\langle \cdot \,, \, \cdot \rangle_{\mathcal{X}}$). Define the multilevel norm

$$\interleave x \interleave_{\mathcal{X}}^2 := \inf \left\{ \sum_{i \in \mathcal{I}} \|x_i\|_{\mathcal{X}}^2 : x = \sum_{i \in \mathcal{I}} x_i \ \text{ with } x_i \in \mathcal{X}_i \text{ for all } i \in \mathcal{I} \right\}. \qquad (41)$$

It is proved, e.g., in [38, Theorem 16] that $\langle \mathcal{S}^{-1} x \,, \, x \rangle_{\mathcal{X}} = \interleave x \interleave_{\mathcal{X}}^2$. Let $C > 0$. If

$$C^{-1} \|x\|_{\mathcal{X}} \leq \interleave x \interleave_{\mathcal{X}} \leq C \|x\|_{\mathcal{X}} \quad \text{for all } x \in \mathcal{X},$$

then the extreme eigenvalues of $\mathcal{S}^{-1}$ (and hence those of $\mathcal{S}$) are bounded (from above and below). In particular, the additive Schwarz operator $\mathcal{S}$ is optimal in the sense that its condition number (ratio of largest and smallest eigenvalues) depends only on $C > 0$.

Let S denote the matrix representation of $\mathcal{S}$. Then, the norm equivalence from above and the latter observations imply that the condition number of S is bounded. The abstract theory on additive Schwarz operators given in [44, Chapter 2] shows that S has the form $S = P^{-1} A$, where A is the Galerkin matrix of $\langle \cdot \,, \, \cdot \rangle_{\mathcal{X}}$. Therefore, boundedness of the condition number of S implies *optimality* of the preconditioner P^{-1}.

We shortly discuss the matrix representation (24) of the additive Schwarz preconditioner P^{-1}. Following [44, Chapter 2], let A_i denote the Galerkin matrix of $\langle \cdot \,, \, \cdot \rangle_{\mathcal{X}}$ restricted to $\mathcal{X}_i$, and let I_i denote the matrix that realizes the embedding from $\mathcal{X}_i \to \mathcal{X}$. We consider the matrix representation of $\mathcal{S}_i : \mathcal{X} \to \mathcal{X}_i \subset \mathcal{X}$. Let $x \in \mathcal{X}$ with coordinate vector x, and let $x_i \in \mathcal{X}_i$ be arbitrary with coordinate vector x_i. The defining relation

$$\langle \mathcal{S}_i x \,, \, x_i \rangle_{\mathcal{X}} = \langle x \,, \, x_i \rangle_{\mathcal{X}} \quad \text{for all } x_i \in \mathcal{X}_i$$

of $\mathcal{S}_i$ then reads in matrix-vector form (with S_i being the matrix representation of $\mathcal{S}_i$) as

$$x_i \cdot (A_i S_i x) = (I_i x_i) \cdot (A x) \quad \text{for all coefficient vectors } x_i,$$

or equivalently

$$A_i S_i x = I_i^T A x.$$

Since A_i is invertible, we have that

$$S_i = A_i^{-1} I_i^T A.$$

Note that the range of the operator $\mathcal{S}_i$ is $\mathcal{X}_i$ and correspondingly for the matrix representation S_i. We therefore apply the embedding I_i and obtain the representation

$$S = P^{-1} A, \quad \text{where} \quad P^{-1} = \sum_{i \in \mathcal{I}} I_i A_i^{-1} I_i^T.$$

To finally prove (24), note that for one-dimensional subspaces $\mathcal{X}_i$, A_i reduces to the diagonal entry of the matrix A. Overall, we thus derive the matrix representation (24).

5.3 Subspace decomposition of $\mathcal{ND}^1(\widehat{\mathcal{T}}_\bullet)$ in $H(\mathbf{curl}\,;\,\Omega)$

The following result is taken from [33, Theorem 4.1]; see also the references therein. In particular, we note that their proof requires the assumption that Ω is simply connected.

Proposition 14 *Let* $\mathcal{Y}_\bullet := \mathcal{ND}^1(\widehat{\mathcal{T}}_\bullet)$, $\mathcal{Y}_{\bullet,E} := \mathrm{span}\{u_{\bullet,E}\}$, $\mathcal{Y}_{\bullet,x} := \mathrm{span}\{\nabla \eta_{\bullet,x}\}$, *and*

$$\widehat{\mathcal{E}}_\ell^\star := (\widehat{\mathcal{E}}_\ell \backslash \widehat{\mathcal{E}}_{\ell-1}) \cup \big\{ E \in \widehat{\mathcal{E}}_\ell \,:\, \mathrm{supp}\, u_{\ell,E} \subsetneqq \mathrm{supp}\, u_{\ell-1,E} \big\},$$
$$\widehat{\mathcal{N}}_\ell^\star := (\widehat{\mathcal{N}}_\ell \backslash \widehat{\mathcal{N}}_{\ell-1}) \cup \big\{ x \in \widehat{\mathcal{N}}_\ell \,:\, \mathrm{supp}\, \eta_{\ell,x} \subsetneqq \mathrm{supp}\, \eta_{\ell-1,x} \big\}.$$

Then, it holds that

$$\mathcal{Y}_L = \mathcal{Y}_0 + \sum_{\ell=1}^{L} \left(\sum_{E \in \widehat{\mathcal{E}}_\ell^\star} \mathcal{Y}_{\ell,E} + \sum_{x \in \widehat{\mathcal{N}}_\ell^\star} \mathcal{Y}_{\ell,x} \right). \tag{42}$$

Moreover, it holds that

$$C^{-1} \|v\|_{H(\mathbf{curl}\,;\,\Omega)} \leq \|\!|\, v \,|\!\|_{\mathcal{Y}_L} \leq C \, \|v\|_{H(\mathbf{curl}\,;\,\Omega)} \quad \text{for all } v \in \mathcal{Y}_L, \tag{43}$$

where $C > 0$ *depends only on* Ω *and* $\widehat{\mathcal{T}}_0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.4 Subspace decomposition of $\mathcal{P}^0(\mathcal{T}_\bullet)$ in $H^{-1/2}(\Gamma)$

It remains to prove the following proposition to conclude the proof of Theorem 3.

Proposition 15 *The multilevel norm* $\|\!| \cdot |\!\|_{\mathcal{X}_L}$ *associated with the decompomposition* (23) *satisfies the equivalence*

$$C^{-1} \|\psi\|_{H^{-1/2}(\Gamma)} \leq \|\!|\, \psi \,|\!\|_{\mathcal{X}_L} \leq C \, \|\psi\|_{H^{-1/2}(\Gamma)} \quad \text{for all } \psi \in \mathcal{P}^0(\mathcal{T}_L), \tag{44}$$

where $C > 0$ *depends only on* Ω *and* $\widehat{\mathcal{T}}_0$.

Proof of lower estimate in (44). Let $\psi \in \mathcal{P}^0(\mathcal{T}_L)$ with arbitrary decomposition

$$\psi = \psi_0 + \psi_*, \quad \psi_* = \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}_\ell^\star} \psi_{\ell,E} \quad \text{with} \quad \psi_0 \in \mathcal{X}_0 \text{ and } \psi_{\ell,E} \in \mathcal{X}_{\ell,E}. \quad (45)$$

Note that $\mathcal{X}_{\ell,E} \subset \mathcal{P}_*^0(\mathcal{T}_\ell)$. Recall the extension operator $\boldsymbol{E}_\ell$ from Lemma 13. Define $\boldsymbol{v}_* := \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}_\ell^\star} \boldsymbol{E}_\ell \psi_{\ell,E} \in \mathcal{Y}_L$. Then, $\mathbf{curl}\, \boldsymbol{v}_* \cdot \boldsymbol{n}|_\Gamma = \psi_*$ and hence

$$\|\psi_*\|_{H^{-1/2}(\Gamma)} \lesssim \|\mathbf{curl}\, \boldsymbol{v}_*\|_{H(\mathbf{div}\,;\,\Omega)} = \|\mathbf{curl}\, \boldsymbol{v}_*\|_{L^2(\Omega)} \overset{(43)}{\lesssim} \|\!|\, \boldsymbol{v}_* \,|\!\|_{\mathcal{Y}_L}$$

from the continuity of the trace operator in $H(\mathbf{div}\,;\,\Omega)$. Moreover, the triangle inequality, the definition of the multilevel norm $\|\!|\,\cdot\,|\!\|_{\mathcal{Y}_L}$, and Lemma 13 show that

$$\begin{aligned}
\|\psi\|_{H^{-1/2}(\Gamma)}^2 \\
&\lesssim \|\psi_0\|_{H^{-1/2}(\Gamma)}^2 + \|\psi_*\|_{H^{-1/2}(\Gamma)}^2 \lesssim \|\psi_0\|_{H^{-1/2}(\Gamma)}^2 + \|\!|\, \boldsymbol{v}_* \,|\!\|_{\mathcal{Y}_L}^2 \\
&\overset{(41)}{\le} \|\psi_0\|_{H^{-1/2}(\Gamma)}^2 + \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}_\ell^\star} \|\boldsymbol{E}_\ell \psi_{\ell,E}\|_{H(\mathbf{curl}\,;\,\Omega)}^2 \overset{(40)}{\lesssim} \|\psi_0\|_{H^{-1/2}(\Gamma)}^2 \\
&\quad + \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}_\ell^\star} \|\psi_{\ell,E}\|_{H^{-1/2}(\Gamma)}^2 .
\end{aligned}$$

Taking the infimum over all possible decompositions (45), we derive the lower estimate in (44) by definition (41) of the multilevel norm. □

Proof of upper estimate in (44). Let $\psi \in \mathcal{P}^0(\mathcal{T}_L)$. Define $\psi_{00} := \langle \psi\,,\,1 \rangle_\Gamma / |\Gamma|$ and $\psi_* := \psi - \psi_{00} \in \mathcal{P}_*^0(\mathcal{T}_L)$. Note that

$$\|\psi_*\|_{H^{-1/2}(\Gamma)} \le \|\psi\|_{H^{-1/2}(\Gamma)} + \|\psi_{00}\|_{H^{-1/2}(\Gamma)} \le \big(1 + \|1/|\Gamma|\|_{H^{1/2}(\Gamma)}\big) \|\psi\|_{H^{-1/2}(\Gamma)}. \quad (46)$$

With Lemma 13, choose $\boldsymbol{v} = \boldsymbol{E}_L \psi_* \in \mathcal{Y}_L = \mathcal{N}\mathcal{D}^1(\widehat{\mathcal{T}}_L)$. Note that $\psi_* = \mathbf{curl}\, \boldsymbol{v} \cdot \boldsymbol{n}|_\Gamma$ and $\|\boldsymbol{v}\|_{H(\mathbf{curl}\,;\,\Omega)} \lesssim \|\psi_*\|_{H^{-1/2}(\Gamma)}$. The upper bound in Proposition 14 further provides $\boldsymbol{v}_0 \in \mathcal{Y}_0$, $\boldsymbol{v}_{\ell,E} \in \mathcal{Y}_{\ell,E}$, and $\boldsymbol{v}_{\ell,x} \in \mathcal{Y}_{\ell,x}$ such that

$$\boldsymbol{v} = \boldsymbol{v}_0 + \sum_{\ell=1}^{L} \left(\sum_{E \in \widehat{\mathcal{E}}_\ell^\star} \boldsymbol{v}_{\ell,E} + \sum_{x \in \widehat{\mathcal{N}}_\ell^\star} \boldsymbol{v}_{\ell,x} \right)$$

as well as

$$\|\boldsymbol{v}_0\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)} + \sum_{\ell=1}^{L} \left(\sum_{E \in \widehat{\mathcal{E}}^{\star}_{\ell}} \|\boldsymbol{v}_{\ell,E}\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)} \right.$$

$$\left. + \sum_{\boldsymbol{x} \in \widehat{\mathcal{N}}^{\star}_{\ell}} \|\boldsymbol{v}_{\ell,\boldsymbol{x}}\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)} \right) \overset{(43)}{\lesssim} \|\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)}. \tag{47}$$

Observe that $\mathbf{curl}\,\boldsymbol{v}_{\ell,\boldsymbol{x}} = 0$, since $\boldsymbol{v}_{\ell,\boldsymbol{x}} \in \mathcal{Y}_{\ell,\boldsymbol{x}} = \mathrm{span}\{\nabla \eta_{\ell,\boldsymbol{x}}\}$. Thus, we see that

$$\psi = \psi_{00} + \psi_{*} = \psi_{00} + \mathbf{curl}\,\boldsymbol{v} \cdot \boldsymbol{n}|_{\Gamma}$$

$$= \psi_{00} + \mathbf{curl}\,\boldsymbol{v}_0 \cdot \boldsymbol{n}|_{\Gamma} + \sum_{\ell=1}^{L} \sum_{E \in \widehat{\mathcal{E}}^{\star}_{\ell}} \mathbf{curl}\,\boldsymbol{v}_{\ell,E} \cdot \boldsymbol{n}|_{\Gamma}.$$

Due to the restriction $(\cdot)|_{\Gamma}$ to the boundary, the latter sum reduces to a sum over all $E \in \mathcal{E}_{\ell}$ (instead of all $E \in \widehat{\mathcal{E}}^{\star}_{\ell}$). Note that $\psi_{*0} := \mathbf{curl}\,\boldsymbol{v}_0 \cdot \boldsymbol{n}|_{\Gamma} \in \mathcal{X}_0 = \mathcal{P}^0(\mathcal{T}_0)$ and hence $\psi_{00} + \psi_{*0} \in \mathcal{X}_0$. Note that

$$\|\psi_{00} + \psi_{*0}\|_{H^{-1/2}(\Gamma)} \leq \|\psi_{00}\|_{H^{-1/2}(\Gamma)} + \|\mathbf{curl}\,\boldsymbol{v}_0 \cdot \boldsymbol{n}\|_{H^{-1/2}(\Gamma)}$$

$$\lesssim \|\psi\|_{H^{-1/2}(\Gamma)} + \|\mathbf{curl}\,\boldsymbol{v}_0\|_{\boldsymbol{H}(\mathbf{div}\,;\,\Omega)} = \|\psi\|_{H^{-1/2}(\Gamma)} + \|\mathbf{curl}\,\boldsymbol{v}_0\|_{L^2(\Omega)}. \tag{48}$$

Due to Lemma 12 and $\boldsymbol{v}_{\ell,E} \in \mathcal{Y}_{\ell,E} = \mathrm{span}\{\boldsymbol{u}_{\ell,E}\}$, it holds that $\psi_{\ell,E} := \mathbf{curl}\,\boldsymbol{v}_{\ell,E} \cdot \boldsymbol{n}|_{\Gamma} \in \mathcal{X}_{\ell,E} = \mathrm{span}\{\varphi_{\ell,E}\}$ with $\|\psi_{\ell,E}\|_{H^{-1/2}(\Gamma)} \simeq \|\boldsymbol{v}_{\ell,E}\|_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)}$. We hence see that

$$\psi = (\psi_{00} + \psi_{*0}) + \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}^{\star}_{\ell}} \psi_{\ell,E}$$

with

$$\|\!|\psi\|\!|^2_{\mathcal{P}^0(\mathcal{T}_L)} \overset{(41)}{\leq} \|\psi_{00} + \psi_{*0}\|^2_{H^{-1/2}(\Gamma)} + \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}^{\star}_{\ell}} \|\psi_{\ell,E}\|^2_{H^{-1/2}(\Gamma)}$$

$$\overset{(48)}{\lesssim} \|\psi\|^2_{H^{-1/2}(\Gamma)} + \|\boldsymbol{v}_0\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)} + \sum_{\ell=1}^{L} \sum_{E \in \mathcal{E}^{\star}_{\ell}} \|\boldsymbol{v}_{\ell,E}\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)}$$

$$\overset{(47)}{\lesssim} \|\psi\|^2_{H^{-1/2}(\Gamma)} + \|\boldsymbol{v}\|^2_{\boldsymbol{H}(\mathbf{curl}\,;\,\Omega)}$$

$$\overset{(40)}{\lesssim} \|\psi\|^2_{H^{-1/2}(\Gamma)} + \|\psi_{*}\|^2_{H^{-1/2}(\Gamma)} \overset{(46)}{\lesssim} \|\psi\|^2_{H^{-1/2}(\Gamma)}.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

6 Proof of Theorem 8 (rate optimality of adaptive algorithm)

In the spirit of [11], we give an abstract analysis, where the precise problem and discretization (i.e., Galerkin BEM with piecewise constants for the weakly-singular integral equation for the 2D and 3D Laplacian) enter only through certain properties of the error estimator. These properties are explicitly stated in Sect. 6.1, before Sect. 6.2 provides general PCG estimates. The remaining sections (Sects. 6.3–6.6) then only exploit these abstract frameworks to prove Theorem 8 and Corollary 10.

6.1 Axioms of adaptivity

In this section, we recall some structural properties of the residual error estimator (14) which have been identified in [11] to be important and sufficient for the numerical analysis of Algorithm 5. For the proof, we refer to [19,24]. We only note that (A4) already implies (A3) with $C_{\mathrm{rel}} \leq C_{\mathrm{drl}}$ in general; see [11, Section 3.3].

For ease of notation, let $\mathcal{T}_0$ be the fixed initial mesh of Algorithm 5. Let $\mathbb{T} := \mathrm{refine}(\mathcal{T}_0)$ be the set of all possible meshes that can be obtained by successively refining $\mathcal{T}_0$.

Proposition 16 *There exist constants $C_{\mathrm{stb}}, C_{\mathrm{red}}, C_{\mathrm{rel}} > 0$ and $0 < q_{\mathrm{red}} < 1$ which depend only on Γ and the γ-shape regularity, such that the following properties (A1)–(A4) hold:*

(A1) *Stability on non-refined element domains For each mesh $\mathcal{T}_\bullet \in \mathbb{T}$, all refinements $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, arbitrary discrete functions $v_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet)$ and $v_\circ \in \mathcal{P}^0(\mathcal{T}_\circ)$, and an arbitrary set $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet \cap \mathcal{T}_\circ$ of non-refined elements, it holds that*

$$|\eta_\circ(\mathcal{U}_\bullet, v_\circ) - \eta_\bullet(\mathcal{U}_\bullet, v_\bullet)| \leq C_{\mathrm{stb}} \,\|\!|\, v_\circ - v_\bullet \,|\!\|.$$

(A2) *Reduction on refined element domains For each mesh $\mathcal{T}_\bullet \in \mathbb{T}$, all refinements $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, and arbitrary $v_\bullet \in \mathcal{P}^0(\mathcal{T}_\bullet)$ and $v_\circ \in \mathcal{P}^0(\mathcal{T}_\circ)$, it holds that*

$$\eta_\circ(\mathcal{T}_\circ \backslash \mathcal{T}_\bullet, v_\circ)^2 \leq q_{\mathrm{red}} \,\eta_\bullet(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ, v_\bullet)^2 + C_{\mathrm{red}} \,\|\!|\, v_\circ - v_\bullet \,|\!\|^2.$$

(A3) *Reliability For each mesh $\mathcal{T}_\bullet \in \mathbb{T}$, the error of the exact discrete solution $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$ of (11) is controlled by*

$$\|\!|\, \phi^\star - \phi_\bullet^\star \,|\!\| \leq C_{\mathrm{rel}} \,\eta_\bullet(\phi_\bullet^\star).$$

(A4) *Discrete reliability For each mesh $\mathcal{T}_\bullet \in \mathbb{T}$ and all refinements $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, there exists a set $\mathcal{R}_{\bullet,\circ} \subseteq \mathcal{T}_\bullet$ with $\mathcal{T}_\bullet \backslash \mathcal{T}_\circ \subseteq \mathcal{R}_{\bullet,\circ}$ as well as $\#\mathcal{R}_{\bullet,\circ} \leq C_{\mathrm{drl}} \#(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ)$ such that the difference of $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$ and $\phi_\circ^\star \in \mathcal{P}^0(\mathcal{T}_\circ)$ is controlled by*

$$\|\!|\, \phi_\circ^\star - \phi_\bullet^\star \,|\!\| \leq C_{\mathrm{drl}} \,\eta_\bullet(\mathcal{R}_{\bullet,\circ}, \phi_\bullet^\star).$$

□

6.2 Energy estimates for the PCG solver

This section collects some auxiliary results which rely on the use of PCG and, in particular, PCG with an optimal preconditioner. We first note the following Pythagoras identity.

Lemma 17 *Let $A_\bullet$, $P_\bullet \in \mathbb{R}^{N \times N}$ be symmetric and positive definite, $b_\bullet \in \mathbb{R}^N$, $x_\bullet^\star :=$ $A_\bullet^{-1} b_\bullet$, $x_{\bullet 0} \in \mathbb{R}^N$ and $x_{\bullet k}$ the iterates of the PCG algorithm.*
There holds the Pythagoras identity

$$\| \phi_\bullet^\star - \phi_{\bullet(k+1)} \|^2 + \| \phi_{\bullet(k+1)} - \phi_{\bullet k} \|^2 = \| \phi_\bullet^\star - \phi_{\bullet k} \|^2 \quad \text{for all } k \in \mathbb{N}_0. \quad (49)$$

Proof According to the definition of PCG (and CG), it holds that

$$\|\widetilde{x}_\bullet^\star - \widetilde{x}_{\bullet k}\|_{\widetilde{A}_\bullet} = \min_{\widetilde{y}_\bullet \in \mathcal{K}_k(\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0})} \|\widetilde{x}_\bullet^\star - \widetilde{y}_\bullet\|_{\widetilde{A}_\bullet},$$

where $\mathcal{K}_k(\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0}) := \operatorname{span}\{\widetilde{r}_{\bullet 0}, \widetilde{A}_\bullet \widetilde{r}_{\bullet 0}, \ldots, \widetilde{A}_\bullet^{k-1} \widetilde{r}_{\bullet 0}\}$ with $\widetilde{r}_{\bullet 0} := \widetilde{b}_\bullet - \widetilde{A}_\bullet \widetilde{x}_{\bullet 0}$. According to Linear Algebra, $\widetilde{x}_{\bullet k}$ is the orthogonal projection of $\widetilde{x}_\bullet^\star$ in $\mathcal{K}_k(\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0})$ with respect to the matrix norm $\| \cdot \|_{\widetilde{A}_\bullet}$. From nestedness $\mathcal{K}_k(\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0}) \subseteq \mathcal{K}_{k+1}(\widetilde{A}_\bullet, \widetilde{b}_\bullet, \widetilde{x}_{\bullet 0})$, it thus follows that

$$\|\widetilde{x}_\bullet^\star - \widetilde{x}_{\bullet k}\|_{\widetilde{A}_\bullet}^2 = \|\widetilde{x}_\bullet^\star - \widetilde{x}_{\bullet(k+1)}\|_{\widetilde{A}_\bullet}^2 + \|\widetilde{x}_{\bullet(k+1)} - \widetilde{x}_{\bullet k}\|_{\widetilde{A}_\bullet}^2.$$

Together with (17) and (20), this proves (49). □

The following lemma collects some estimates which follow from the contraction property (19) of PCG.

Lemma 18 *Algorithm 5 guarantees the following estimates for all $(j, k) \in \mathcal{Q}$ with $k \geq 1$:*

(i) $\| \phi_j^\star - \phi_{jk} \| \leq q_{\mathrm{pcg}} \| \phi_j^\star - \phi_{j(k-1)} \|$
(ii) $\| \phi_{jk} - \phi_{j(k-1)} \| \leq (1 + q_{\mathrm{pcg}}) \| \phi_j^\star - \phi_{j(k-1)} \|$
(iii) $\| \phi_j^\star - \phi_{j(k-1)} \| \leq (1 - q_{\mathrm{pcg}})^{-1} \| \phi_{jk} - \phi_{j(k-1)} \|$
(iv) $\| \phi_j^\star - \phi_{jk} \| \leq q_{\mathrm{pcg}}(1 - q_{\mathrm{pcg}})^{-1} \| \phi_{jk} - \phi_{j(k-1)} \|$

Proof According to (20), estimate (19) proves (i). The estimates (ii)–(iv) follow from (i) and the triangle inequality. □

6.3 Proof of Theorem 8(a)

With reliability (A3) and stability (A1), we see that

$$\||\, \phi^\star - \phi_{jk} \,\|| \leq \||\, \phi^\star - \phi_j^\star \,\|| + \||\, \phi_j^\star - \phi_{jk} \,\||$$

$$\overset{\text{(A3)}}{\lesssim} \eta_j(\phi_j^\star) + \||\, \phi_j^\star - \phi_{jk} \,\||$$

$$\overset{\text{(A1)}}{\lesssim} \eta_j(\phi_{jk}) + \||\, \phi_j^\star - \phi_{jk} \,\|| \quad \text{for all } (j,k) \in \mathcal{Q}.$$

With Lemma 18(iv), we hence prove the reliability estimate (26).

According to [3], it holds that

$$\eta_j(\phi_{jk}) \lesssim \| h_j^{1/2}(\phi^\star - \phi_{jk}) \|_{L^2(\Gamma)} + \||\, \phi^\star - \phi_{jk} \,\||$$

$$\leq \| h_j^{1/2}(\phi^\star - \phi_{jk}) \|_{L^2(\Gamma)} + \||\, \phi^\star - \phi_j^\star \,\|| + \||\, \phi_j^\star - \phi_{jk} \,\||$$

Let $\mathbb{G}_j : \widetilde{H}^{-1/2}(\Gamma) \to \mathcal{P}^0(\mathcal{T}_j)$ be the Galerkin projection. Let $\boldsymbol{Pi}_j : L^2(\Gamma) \to \mathcal{P}^0(\mathcal{T}_j)$ be the L^2-orthogonal projection. With the Céa lemma and a duality argument (see, e.g., [16, Theorem 4.1]), we see that

$$\||\, (1 - \mathbb{G}_j)\psi \,\|| \leq \||\, (1 - \boldsymbol{Pi}_j)\psi \,\|| \lesssim \| h_j^{1/2}\psi \|_{L^2(\Gamma)} \quad \text{for all } \psi \in L^2(\Gamma).$$

Hence, for $\psi = \phi^\star - \phi_{jk}$, it follows that

$$\||\, \phi^\star - \phi_j^\star \,\|| = \||\, (1 - \mathbb{G}_j)\phi^\star \,\|| = \||\, (1 - \mathbb{G}_j)(\phi^\star - \phi_{jk}) \,\|| \lesssim \| h_j^{1/2}(\phi^\star - \phi_{jk}) \|_{L^2(\Gamma)}.$$

Combining the latter estimates, we see that

$$\eta_j(\phi_{jk}) \lesssim \| h_j^{1/2}(\phi^\star - \phi_{jk}) \|_{L^2(\Gamma)} + \||\, \phi_j^\star - \phi_{jk} \,\||.$$

With Lemma 18(iv), we hence prove the efficiency estimate (27). □

6.4 Proof of Theorem 8(b)

The following lemma is the heart of the proof of Theorem 8(b).

Lemma 19 *Consider Algorithm 5 for arbitrary parameters $0 < \theta \leq 1$ and $\lambda > 0$. There exist constants $0 < \mu, q_{\text{ctr}} < 1$ such that*

$$\Delta_{jk} := \mu\, \eta_j(\phi_{jk})^2 + \||\, \phi^\star - \phi_{jk} \,\||^2 \quad \text{for } (j,k) \in \mathcal{Q}$$

satisfies, for all $j \in \mathbb{N}_0$, that

$$\Delta_{j(k+1)} \leqq q_{\text{ctr}}\, \Delta_{jk} \quad \text{for all } 0 \leq k < k+1 < \underline{k} \tag{50}$$

as well as

$$\Delta_{(j+1)0} \leqq q_{\text{ctr}} \, \Delta_{j(\underline{k}-1)}.$$ (51)

Moreover, for all $(j', k'), (j, k) \in \mathcal{Q}$, *it holds that*

$$\Delta_{j'k'} \leqq q_{\text{ctr}}^{|(j',k')|-|(j,k)|} \, \Delta_{jk} \quad \text{provided that}$$
$$(j', k') > (j, k), k' < \underline{k}(j'), \text{ and } k < \underline{k}(j).$$ (52)

The constants $0 < \mu, q_{\text{ctr}} < 1$ *depend only on* λ, θ, q_{pcg}, *and the constants in* (A1)–(A3).

Proof The proof is split into five steps.

Step 1 We fix some constants, which are needed below. We note that all these constants depend on $0 < \theta \leq 1$ and $\lambda > 0$, but do not require any additional constraint. First, define

$$0 < q_{\text{est}} := 1 - (1 - q_{\text{red}})\theta^2 < 1.$$ (53)

Second, choose $\gamma > 0$ such that

$$(1 + \gamma) \, q_{\text{est}} \overset{(53)}{<} 1.$$ (54)

Third, choose $\mu > 0$ such that

$$\mu \, (1 + \gamma^{-1}) \, q_{\text{est}} \, C_{\text{stb}}^2 (1 + q_{\text{pcg}})^2 \; < \; \frac{1 - q_{\text{pcg}}^2}{2} \quad \text{and} \quad \mu \, \lambda^{-2} \; \leq \; \frac{1}{2}.$$ (55)

Fourth, choose $\varepsilon > 0$ such that

$$\varepsilon \, (1 - q_{\text{pcg}})^{-2} + 2 \, \varepsilon \, C_{\text{rel}}^2 \, C_{\text{stb}}^2 \, (1 - q_{\text{pcg}})^{-2} \; \leq \; \frac{1}{2} \quad \text{and} \quad 2 \, \varepsilon \, C_{\text{rel}}^2 \; \leq \; (1 - \varepsilon)\mu.$$ (56)

Fifth, choose $\kappa > 0$ such that

$$2 \, \kappa \, C_{\text{rel}}^2 \overset{(54)}{<} \left(1 - (1 + \gamma) \, q_{\text{est}} \right) \mu \quad \text{and} \quad 2 \, \kappa \, C_{\text{rel}}^2 C_{\text{stb}}^2 \; < \; \frac{1 - q_{\text{pcg}}^2}{2}.$$ (57)

With (55)–(57), we finally define

$$0 < q_{\text{ctr}} := \max \left\{ 1 - \varepsilon, \; \left(\mu \, (1 + \gamma) \, q_{\text{est}} + 2 \, \kappa \, C_{\text{rel}}^2 \right) \mu^{-1}, \; 1 - \kappa, \right.$$
$$\left. \left(\mu \, (1 + \gamma^{-1}) \, q_{\text{est}} \, C_{\text{stb}}^2 \, (1 + q_{\text{pcg}})^2 + q_{\text{pcg}}^2 + 2 \, \kappa \, C_{\text{rel}}^2 C_{\text{stb}}^2 \right) \right\} < 1.$$ (58)

Step 2 Due to reliability (A3), stability (A1), and Lemma 18(iii), it follows that

$$
\begin{aligned}
\| \phi^\star - \phi_j^\star \|^2 &= (1 - \varepsilon) \| \phi^\star - \phi_j^\star \|^2 + \varepsilon \| \phi^\star - \phi_j^\star \|^2 \\
&\overset{(A3)}{\leq} (1 - \varepsilon) \| \phi^\star - \phi_j^\star \|^2 + \varepsilon \, C_{\mathrm{rel}}^2 \, \eta_j(\phi_j^\star)^2 \\
&\overset{(A1)}{\leq} (1 - \varepsilon) \| \phi^\star - \phi_j^\star \|^2 + 2 \varepsilon \, C_{\mathrm{rel}}^2 \left(\eta_j(\phi_{jk})^2 + C_{\mathrm{stb}}^2 \| \phi_j^\star - \phi_{jk} \|^2 \right) \\
&\overset{18(\mathrm{iii})}{\leq} (1 - \varepsilon) \| \phi^\star - \phi_j^\star \|^2 + 2 \varepsilon \, C_{\mathrm{rel}}^2 \, \eta_j(\phi_{jk})^2 \\
&\quad + 2 \varepsilon \, C_{\mathrm{rel}}^2 \, C_{\mathrm{stb}}^2 \, (1 - q_{\mathrm{pcg}})^{-2} \| \phi_{j(k+1)} - \phi_{jk} \|^2 .
\end{aligned}
$$

Step 3. We consider the case $k + 1 < \underline{k}(j)$. Step (iv) of Algorithm 5 yields that

$$
\eta_j(\phi_{j(k+1)})^2 < \lambda^{-2} \| \phi_{j(k+1)} - \phi_{jk} \|^2 . \tag{59}
$$

Moreover, the Pythagoras identity (49) implies that

$$
\begin{aligned}
\| \phi_j^\star - \phi_{j(k+1)} \|^2 &= (1 - \varepsilon) \| \phi_j^\star - \phi_{jk} \|^2 \\
&\quad + \varepsilon \| \phi_j^\star - \phi_{jk} \|^2 - \| \phi_{j(k+1)} - \phi_{jk} \|^2 .
\end{aligned} \tag{60}
$$

Further, we note the Pythagoras identity

$$
\| \phi^\star - \phi_j^\star \|^2 + \| \phi_j^\star - \psi_j \|^2 = \| \phi^\star - \psi_j \|^2 \quad \text{for all } \psi_j \in \mathcal{P}^0(\mathcal{T}_j). \tag{61}
$$

Combining (59)–(61) and applying Lemma 18(iii), we see that

$$
\begin{aligned}
\Delta_{j(k+1)} &= \mu \, \eta_j(\phi_{j(k+1)})^2 + \| \phi_j^\star - \phi_{j(k+1)} \|^2 + \| \phi^\star - \phi_j^\star \|^2 \\
&< (1 - \varepsilon) \| \phi_j^\star - \phi_{jk} \|^2 + \varepsilon \| \phi_j^\star - \phi_{jk} \|^2 + (\mu \, \lambda^{-2} - 1) \| \phi_{j(k+1)} \\
&\quad - \phi_{jk} \|^2 + \| \phi^\star - \phi_j^\star \|^2 \\
&\overset{18(\mathrm{iii})}{\leq} (1 - \varepsilon) \| \phi_j^\star - \phi_{jk} \|^2 + \left(\varepsilon \, (1 - q_{\mathrm{pcg}})^{-2} + \mu \, \lambda^{-2} - 1 \right) \| \phi_{j(k+1)} \\
&\quad - \phi_{jk} \|^2 + \| \phi^\star - \phi_j^\star \|^2 .
\end{aligned}
$$

Step 2 yields that

$$
\begin{aligned}
&\leq (1 - \varepsilon) \left(\| \phi_j^\star - \phi_{jk} \|^2 + \| \phi^\star - \phi_j^\star \|^2 \right) + 2 \varepsilon \, C_{\mathrm{rel}}^2 \, \eta_j(\phi_{jk})^2 \\
&\quad + \left(\varepsilon \, (1 - q_{\mathrm{pcg}})^{-2} + \mu \, \lambda^{-2} - 1 + 2 \varepsilon \, C_{\mathrm{rel}}^2 \, C_{\mathrm{stb}}^2 \, (1 - q_{\mathrm{pcg}})^{-2} \right) \| \phi_{j(k+1)} - \phi_{jk} \|^2 .
\end{aligned}
$$

Using (55)–(56) and (61), we thus see that

$$
\Delta_{j(k+1)} \leq (1 - \varepsilon) \left(\mu \, \eta_j(\phi_{jk})^2 + \| \phi^\star - \phi_{jk} \|^2 \right) \overset{(58)}{\leq} q_{\mathrm{ctr}} \, \Delta_{jk} \quad \text{if } k + 1 < \underline{k}(j).
$$

This concludes the proof of (50).

Step 4 We use the definition $\phi_{(j+1)0} := \phi_{j\underline{k}}$ from Step (vi) of Algorithm 5 to see that

$$
\begin{aligned}
\Delta_{(j+1)0} &= \mu\,\eta_{j+1}(\phi_{(j+1)0})^2 + \|\,\phi^\star - \phi_{(j+1)0}\,\|^2 \\
&= \mu\,\eta_{j+1}(\phi_{j\underline{k}})^2 + \|\,\phi^\star - \phi_{j\underline{k}}\,\|^2.
\end{aligned}
\tag{62}
$$

For the first summand of (62), we use stability (A1) and reduction (A2). Together with the Dörfler marking strategy in Step (v) of Algorithm 5 and $\mathcal{M}_j \subseteq \mathcal{T}_j \backslash \mathcal{T}_{j+1}$, we see that

$$
\begin{aligned}
\eta_{j+1}(\phi_{j\underline{k}})^2 &= \eta_{j+1}(\mathcal{T}_{j+1} \backslash \mathcal{T}_j, \phi_{j\underline{k}})^2 + \eta_{j+1}(\mathcal{T}_{j+1} \cap \mathcal{T}_j, \phi_{j\underline{k}})^2 \\
&\leq q_{\mathrm{red}}\,\eta_j(\mathcal{T}_j \backslash \mathcal{T}_{j+1}, \phi_{j\underline{k}})^2 + \eta_j(\mathcal{T}_{j+1} \cap \mathcal{T}_j, \phi_{j\underline{k}})^2 \\
&= \eta_j(\phi_{j\underline{k}})^2 - (1 - q_{\mathrm{red}})\,\eta_j(\mathcal{T}_j \backslash \mathcal{T}_{j+1}, \phi_{j\underline{k}})^2 \\
&\leq \eta_j(\phi_{j\underline{k}})^2 - (1 - q_{\mathrm{red}})\theta^2\,\eta_j(\phi_{j\underline{k}})^2 \overset{(53)}{=} q_{\mathrm{est}}\,\eta_j(\phi_{j\underline{k}})^2.
\end{aligned}
\tag{63}
$$

With this and stability (A1), the Young inequality and Lemma 18(ii) yield that

$$
\begin{aligned}
\eta_{j+1}(\phi_{j\underline{k}})^2 &\overset{(63)}{\leq} q_{\mathrm{est}}\,\eta_j(\phi_{j\underline{k}})^2 \overset{(A1)}{\leq} (1 + \gamma)\,q_{\mathrm{est}}\,\eta_j(\phi_{j(\underline{k}-1)})^2 \\
&\quad + (1 + \gamma^{-1})\,q_{\mathrm{est}}\,C_{\mathrm{stb}}^2\|\,\phi_{j\underline{k}} - \phi_{j(\underline{k}-1)}\,\|^2 \\
&\overset{(ii)}{\leq} (1 + \gamma)\,q_{\mathrm{est}}\,\eta_j(\phi_{j(\underline{k}-1)})^2 \\
&\quad + (1 + \gamma^{-1})\,q_{\mathrm{est}}\,C_{\mathrm{stb}}^2\,(1 + q_{\mathrm{pcg}})^2\|\,\phi_j^\star - \phi_{j(\underline{k}-1)}\,\|^2.
\end{aligned}
\tag{64}
$$

For the second summand of (62), we apply the Pythagoras identity (61) together with Lemma 18(i) and obtain that

$$
\begin{aligned}
\|\,\phi^\star - \phi_{j\underline{k}}\,\|^2 &\overset{(61)}{=} \|\,\phi^\star - \phi_j^\star\,\|^2 + \|\,\phi_j^\star - \phi_{j\underline{k}}\,\|^2 \\
&\overset{18(i)}{\leq} \|\,\phi^\star - \phi_j^\star\,\|^2 + q_{\mathrm{pcg}}^2\,\|\,\phi_j^\star - \phi_{j(\underline{k}-1)}\,\|^2.
\end{aligned}
\tag{65}
$$

Combining (62)–(65), we end up with

$$
\begin{aligned}
\Delta_{(j+1)0} &\leq \mu\,(1 + \gamma)\,q_{\mathrm{est}}\,\eta_j(\phi_{j(\underline{k}-1)})^2 \\
&\quad + \big(\mu\,(1 + \gamma^{-1})\,q_{\mathrm{est}}\,C_{\mathrm{stb}}^2\,(1 + q_{\mathrm{pcg}})^2 + q_{\mathrm{pcg}}^2\big)\|\,\phi_j^\star - \phi_{j(\underline{k}-1)}\,\|^2 \\
&\quad + \|\,\phi^\star - \phi_j^\star\,\|^2.
\end{aligned}
$$

Using the same arguments as in Step 2, we get that

$$
\begin{aligned}
\Delta_{(j+1)0} &\leq \mu\,(1 + \gamma)\,q_{\mathrm{est}}\,\eta_j(\phi_{j(\underline{k}-1)})^2 \\
&\quad + \big(\mu\,(1 + \gamma^{-1})\,q_{\mathrm{est}}\,C_{\mathrm{stb}}^2\,(1 + q_{\mathrm{pcg}})^2 + q_{\mathrm{pcg}}^2\big)\|\,\phi_j^\star - \phi_{j(\underline{k}-1)}\,\|^2 \\
&\quad + (1 - \kappa)\|\,\phi^\star - \phi_j^\star\,\|^2 + 2\,\kappa\,C_{\mathrm{rel}}^2\,\eta_j(\phi_{j(\underline{k}-1)})^2 \\
&\quad + 2\,\kappa\,C_{\mathrm{rel}}^2\,C_{\mathrm{stb}}^2\,\|\,\phi_j^\star - \phi_{j(\underline{k}-1)}\,\|^2
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\mu\,(1+\gamma)\,q_{\text{est}} + 2\,\kappa\,C_{\text{rel}}^2\right) \eta_j(\phi_{j(\underline{k}-1)})^2 + (1-\kappa)\|\|\phi^\star - \phi_j^\star\|\|^2 \\
&\quad + \left(\mu\,(1+\gamma^{-1})\,q_{\text{est}}\,C_{\text{stb}}^2\,(1+q_{\text{pcg}})^2 + q_{\text{pcg}}^2 \right. \\
&\quad \left. + 2\,\kappa\,C_{\text{rel}}^2 C_{\text{stb}}^2\right) \|\|\phi_j^\star - \phi_{j(\underline{k}-1)}\|\|^2 \\
&\stackrel{(58)}{\leq} q_{\text{ctr}}\,\mu\,\eta_j(\phi_{j(\underline{k}-1)})^2 + q_{\text{ctr}}\,\|\|\phi^\star - \phi_j^\star\|\|^2 + q_{\text{ctr}}\,\|\|\phi_j^\star - \phi_{j(\underline{k}-1)}\|\|^2 \\
&\stackrel{(61)}{=} q_{\text{ctr}}\,\Delta_{j(\underline{k}-1)}.
\end{aligned}
$$

This concludes the proof of (51).

$\quad$ *Step 5* $\;$ Inequality (52) follows by induction. This concludes the proof. $\qquad\qquad$ □

Proof of Theorem 8(b) $\;$ The proof is split into three steps.

$\quad$ *Step 1* $\;$ Let $j \in \mathbb{N}$. Recall the Pythagoras identity (61). We use stability (A1) and Step (iv) of Algorithm 5 to see that

$$
\begin{aligned}
\Delta_{j(\underline{k}-1)} &\stackrel{(61)}{=} \mu\,\eta_j(\phi_{j(\underline{k}-1)})^2 + \|\|\phi_j^\star - \phi_{j(\underline{k}-1)}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \\
&\stackrel{(A1)}{\lesssim} \eta_j(\phi_{j\underline{k}})^2 + \|\|\phi_{j\underline{k}} - \phi_{j(\underline{k}-1)}\|\|^2 + \|\|\phi_j^\star - \phi_{j\underline{k}}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \\
&\stackrel{5(iv)}{\lesssim} \eta_j(\phi_{j\underline{k}})^2 + \|\|\phi_j^\star - \phi_{j\underline{k}}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \stackrel{(61)}{\simeq} \Delta_{j\underline{k}}.
\end{aligned}
$$

With the Pythagoras identity (49), we may argue similarly to obtain that

$$
\begin{aligned}
\Delta_{j\underline{k}} &\stackrel{(61)}{=} \mu\,\eta_j(\phi_{j\underline{k}})^2 + \|\|\phi_j^\star - \phi_{j\underline{k}}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \\
&\stackrel{(A1)}{\lesssim} \eta_j(\phi_{j(\underline{k}-1)})^2 + \|\|\phi_{j\underline{k}} - \phi_{j(\underline{k}-1)}\|\|^2 + \|\|\phi_j^\star - \phi_{j\underline{k}}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \\
&\stackrel{(49)}{=} \eta_j(\phi_{j(\underline{k}-1)})^2 + \|\|\phi_j^\star - \phi_{j(\underline{k}-1)}\|\|^2 + \|\|\phi^\star - \phi_j^\star\|\|^2 \stackrel{(61)}{\simeq} \Delta_{j(\underline{k}-1)}.
\end{aligned}
$$

Hence, it follows that $\Delta_{j\underline{k}} \simeq \Delta_{j(\underline{k}-1)}$.

$\quad$ *Step 2* $\;$ For $0 \leq j \leq j'$, define $\widehat{k}(j) := \widehat{k} \in \mathbb{N}_0$ by

$$
\widehat{k} := \begin{cases} \underline{k}(j) & \text{if } j < j', \\ k' & \text{if } j = j'. \end{cases}
$$

From Step 1, Lemma 19, and the geometric series (for the sum over k), it follows that

$$
\begin{aligned}
\sum_{j=0}^{j'} \sum_{k=0}^{\widehat{k}(j)} \Delta_{jk}^{-1} &\lesssim \Delta_{j'k'}^{-1} + \sum_{j=0}^{j'} \sum_{k=0}^{\widehat{k}(j)-1} \Delta_{jk}^{-1} \\
&\stackrel{(52)}{\leq} \Delta_{j'k'}^{-1} + \sum_{j=0}^{j'} \sum_{k=0}^{\widehat{k}(j)-1} q_{\text{ctr}}^{|(j,\widehat{k}-1)| - |(j,k)|} \Delta_{j(\widehat{k}-1)}^{-1} \\
&\lesssim \Delta_{j'k'}^{-1} + \sum_{j=0}^{j'} \Delta_{j(\widehat{k}-1)}^{-1}.
\end{aligned}
$$

For $k' < \underline{k}(j')$, inequality (52) and the geometric series (for the sum over j) yield that

$$\sum_{j=0}^{j'} \Delta_{j(\widehat{k}-1)}^{-1} \overset{(52)}{\lesssim} \sum_{j=0}^{j'} q_{\mathrm{ctr}}^{|(j',k')|-|(j,\widehat{k}-1)|} \Delta_{j'k'}^{-1} \lesssim \Delta_{j'k'}^{-1}.$$

For $k' = \underline{k}(j')$, inequality (52), the geometric series, and Step 1 yield that

$$\sum_{j=0}^{j'} \Delta_{j(\widehat{k}-1)}^{-1} = \Delta_{j'(\underline{k}-1)}^{-1} + \sum_{j=0}^{j'-1} \Delta_{j(\underline{k}-1)}^{-1} \overset{(52)}{\lesssim} \left(1 + \sum_{j=0}^{j'-1} q_{\mathrm{ctr}}^{|(j',\underline{k}-1)|-|(j,\underline{k}-1)|} \right) \Delta_{j'(\underline{k}-1)}^{-1}$$

$$\lesssim \Delta_{j'(\underline{k}-1)}^{-1} \simeq \Delta_{j'\underline{k}}^{-1} = \Delta_{j'k'}^{-1}.$$

Overall, it follows that

$$\sum_{j=0}^{j'} \sum_{k=0}^{\widehat{k}(j)} \Delta_{jk}^{-1} \lesssim \Delta_{j'k'}^{-1} \quad \text{for all } (j',k') \in \mathcal{Q}. \tag{66}$$

Step 3 According to the proof of [11, Lemma 4.9], estimate (66) guarantees (and is even equivalent to) the existence of $0 < q_{\mathrm{lin}} < 1$ such that

$$\Delta_{j'k'}^{1/2} \lesssim q_{\mathrm{lin}}^{|(j',k')|-|(j,k)|} \Delta_{jk}^{1/2} \quad \text{for all } (j,k), (j',k') \in \mathcal{Q} \text{ with } (j',k') \geq (j,k). \tag{67}$$

Clearly, it holds that $\Lambda_{jk} \simeq \Delta_{jk}^{1/2}$ for all $(j,k) \in \mathcal{Q}$. This and (67) conclude the proof. □

6.5 Proof of Theorem 8(c)

The proof of optimal convergence rates requires the following additional properties of the mesh-refinement strategy. For 3D BEM (and 2D NVB from Sect. 2.4) these properties are verified in [8,41,42], and any assumption on $\mathcal{T}_0$ is removed in [35]. For 2D BEM (and the extended 1D bisection from Sect. 2.3), these properties are verified in [2].

(R1) *Splitting property* Each refined element is split in at least 2 and at most in $C_{\mathrm{son}} \geq 2$ many sons, i.e., for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$, the refined mesh $\mathcal{T}_\circ = \mathrm{refine}(\mathcal{T}_\bullet, \mathcal{M}_\bullet)$ satisfies that

$$\#(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ) + \#\mathcal{T}_\bullet \leq \#\mathcal{T}_\circ \leq C_{\mathrm{son}} \#(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ) + \#(\mathcal{T}_\bullet \cap \mathcal{T}_\circ).$$

(R2) *Overlay estimate* For all meshes $\mathcal{T} \in \mathbb{T}$ and $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T})$ there exists a common refinement $\mathcal{T}_\bullet \oplus \mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet) \cap \mathrm{refine}(\mathcal{T}_\circ) \subseteq \mathrm{refine}(\mathcal{T})$ with

$$\#(\mathcal{T}_\bullet \oplus \mathcal{T}_\circ) \leq \#\mathcal{T}_\bullet + \#\mathcal{T}_\circ - \#\mathcal{T}.$$

(R3) *Mesh-closure estimate* There exists $C_{\mathrm{mesh}} > 0$ such that the sequence $\mathcal{T}_j$ with corresponding $\mathcal{M}_j \subseteq \mathcal{T}_j$, which is generated by Algorithm 5, satisfies that

$$\#\mathcal{T}_j - \#\mathcal{T}_0 \leq C_{\mathrm{mesh}} \sum_{\ell=0}^{j-1} \#\mathcal{M}_\ell.$$

Recall the constants $C_{\mathrm{stb}} > 0$ from (A1) and $C_{\mathrm{drl}} > 0$ from (A4). Suppose that $0 < \theta \leq 1$ and $\lambda > 0$ are sufficiently small such that

$$0 < \theta'' := \frac{\theta + \lambda/\lambda_{\mathrm{opt}}}{1 - \lambda/\lambda_{\mathrm{opt}}} < \theta_{\mathrm{opt}} := \left(1 + C_{\mathrm{stb}}^2 C_{\mathrm{drl}}^2\right)^{-1/2},$$

$$\text{where } \lambda_{\mathrm{opt}} := \left(C_{\mathrm{stb}} \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}}\right)^{-1}. \tag{68}$$

In particular, it holds that $0 < \theta < \theta_{\mathrm{opt}}$ and $0 < \lambda < \lambda_{\mathrm{opt}}$. We need the following comparison lemma which is found in [11, Lemma 4.14].

Lemma 20 *Suppose (R2), (A1), (A2), and (A4). Recall the assumption (68). There exist constants $C_1, C_2 > 0$ such that for all $s > 0$ with $\|\phi^\star\|_{\mathbb{A}_s} < \infty$ and all $j \in \mathbb{N}_0$, there exists $\mathcal{R}_j \subseteq \mathcal{T}_j$ which satisfies*

$$\#\mathcal{R}_j \leq C_1 C_2^{-1/s} \|\phi^\star\|_{\mathbb{A}_s}^{1/s} \eta_j(\phi_j^\star)^{-1/s}, \tag{69}$$

as well as the Dörfler marking criterion

$$\theta'' \eta_j(\phi_j^\star) \leq \eta_j(\mathcal{R}_j, \phi_j^\star). \tag{70}$$

The constants C_1, C_2 depend only on the constants of (A1), (A2), and (A4). □

Another lemma, which we need for the proof of Theorem 8(c), shows that the iterates $\phi_{\bullet k}$ of Algorithm 5 are close to the exact Galerkin approximation $\phi_\bullet^\star \in \mathcal{P}^0(\mathcal{T}_\bullet)$.

Lemma 21 *Let $0 < \lambda < \lambda_{\mathrm{opt}}$. For all $j \in \mathbb{N}_0$, it holds that*

$$\|\!|\phi_j^\star - \phi_{j\underline{k}}|\!| \leq \lambda \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}} \, \min\left\{\eta_j(\phi_{j\underline{k}}), \, \frac{1}{1 - \lambda/\lambda_{\mathrm{opt}}} \eta_j(\phi_j^\star)\right\}. \tag{71}$$

Moreover, there holds equivalence

$$(1 - \lambda/\lambda_{\mathrm{opt}}) \, \eta_j(\phi_{j\underline{k}}) \leq \eta_j(\phi_j^\star) \leq (1 + \lambda/\lambda_{\mathrm{opt}}) \, \eta_j(\phi_{j\underline{k}}). \tag{72}$$

Proof Stability (A1) yields that $|\eta_j(\phi_j^\star) - \eta_j(\phi_{j\underline{k}})| \leq C_{\mathrm{stb}} \|\!|\phi_j^\star - \phi_{j\underline{k}}|\!|$. Therefore, Lemma 18(iv) and the assumption on the PCG iterate in Step (iv) of Algorithm 5 imply that

$$\| \phi_j^\star - \phi_{j\underline{k}} \| \overset{18\text{(iv)}}{\leq} \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}} \| \phi_{j\underline{k}} - \phi_{j(\underline{k}-1)} \|$$

$$\leq \lambda \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}} \eta_j(\phi_{j\underline{k}})$$

$$\overset{(A1)}{\leq} \lambda \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}} \left(\eta_j(\phi_j^\star) + C_{\text{stb}} \| \phi_j^\star - \phi_{j\underline{k}} \| \right).$$

Since $0 < \lambda < \lambda_{\text{opt}}$ and hence $\lambda C_{\text{stb}} \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}} = \lambda/\lambda_{\text{opt}} < 1$, this yields that

$$\| \phi_j^\star - \phi_{j\underline{k}} \| \leq \frac{\lambda \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}}}{1 - \lambda C_{\text{stb}} \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}}} \eta_j(\phi_j^\star) = \lambda \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}} \frac{1}{1 - \lambda/\lambda_{\text{opt}}} \eta_j(\phi_j^\star).$$

Altogether, this proves (71). Moreover, with stability (A1), we see that

$$\eta_j(\phi_j^\star) \overset{(A1)}{\leq} \eta_j(\phi_{j\underline{k}}) + C_{\text{stb}} \| \phi_j^\star - \phi_{j\underline{k}} \| \overset{(71)}{\leq} (1 + \lambda/\lambda_{\text{opt}}) \, \eta_j(\phi_{j\underline{k}})$$

as well as

$$\eta_j(\phi_{j\underline{k}}) \overset{(A1)}{\leq} \eta_j(\phi_j^\star) + C_{\text{stb}} \| \phi_j^\star - \phi_{j\underline{k}} \| \overset{(71)}{\leq} \left(1 + \frac{\lambda/\lambda_{\text{opt}}}{1 - \lambda/\lambda_{\text{opt}}} \right) \eta_j(\phi_j^\star)$$

$$= \frac{1}{1 - \lambda/\lambda_{\text{opt}}} \eta_j(\phi_j^\star).$$

This concludes the proof. □

Finally, we need the following lemma which immediately shows "$\Longleftarrow$" in (31).

Lemma 22 *Suppose (R1). For $j \in \mathbb{N}_0$, let $\widehat{\mathcal{T}}_{j+1} = \text{refine}(\widehat{\mathcal{T}}_j, \widehat{\mathcal{M}}_j)$ with arbitrary, but non-empty $\widehat{\mathcal{M}}_j \subseteq \widehat{\mathcal{T}}_j$ and $\widehat{\mathcal{T}}_0 = \mathcal{T}_0$. Let $\widehat{\mathcal{Q}} \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ be an index set and $\widehat{\phi}_{jk} \in \mathcal{P}^0(\widehat{\mathcal{T}}_j)$ for all $(j, k) \in \widehat{\mathcal{Q}}$. Let $s > 0$ and suppose that the corresponding quasi-errors $\widehat{\Lambda}_{jk}^2 := \| \phi^\star - \widehat{\phi}_{jk} \|^2 + \widehat{\eta}_j(\widehat{\phi}_j^\star)^2$ satisfy that*

$$\sup_{(j,k) \in \widehat{\mathcal{Q}}} \left(\#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 + 1 \right)^s \widehat{\Lambda}_{jk} < \infty. \tag{73}$$

Then, it follows that $\| \phi^\star \|_{\mathbb{A}_s} < \infty$.

Proof Due to the Pythagoras identity (61) and stability (A1), it holds that

$$\widehat{\Lambda}_{jk}^2 = \| \phi^\star - \widehat{\phi}_{jk} \|^2 + \widehat{\eta}_j(\widehat{\phi}_{jk})^2$$

$$\overset{(61)}{=} \| \phi^\star - \widehat{\phi}_j^\star \|^2 + \| \widehat{\phi}_j^\star - \widehat{\phi}_{jk} \|^2 + \widehat{\eta}_j(\widehat{\phi}_{jk})^2 \overset{(A1)}{\gtrsim} \widehat{\eta}_j(\widehat{\phi}_j^\star)^2. \tag{74}$$

Additionally, [9, Lemma 22] shows that

$$\#\mathcal{T}_\bullet - \#\mathcal{T}_0 + 1 \le \#\mathcal{T}_\bullet \le \#\mathcal{T}_0 \left(\#\mathcal{T}_\bullet - \#\mathcal{T}_0 + 1 \right) \quad \text{for all } \mathcal{T}_\bullet \in \mathbb{T}. \tag{75}$$

Given $N \in \mathbb{N}_0$, there exists an index $j \in \mathbb{N}_0$ such that

$$\#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 \le N < N + 1 \le \#\widehat{\mathcal{T}}_{j+1} - \#\mathcal{T}_0 + 1 \overset{(75)}{\le} \#\widehat{\mathcal{T}}_{j+1} \overset{(R1)}{\lesssim} \#\widehat{\mathcal{T}}_j \overset{(75)}{\lesssim} \#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 + 1. \tag{76}$$

With (74)–(76), it follows that

$$(N + 1)^s \min_{\substack{\mathcal{T}_\bullet \in \text{refine}(\mathcal{T}_0) \\ \#\mathcal{T}_\bullet - \#\mathcal{T}_0 \le N}} \eta_\bullet(\phi_\bullet^\star) \overset{(76)}{\lesssim} \left(\#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 + 1 \right)^s \widehat{\eta}_j(\widehat{\phi}_j^\star)$$

$$\overset{(74)}{\lesssim} \sup_{(j,k) \in \widehat{\mathcal{Q}}} \left(\#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 + 1 \right)^s \widehat{\Lambda}_{jk} \overset{(73)}{<} \infty.$$

Since the upper bound is finite and independent of N, this implies that $\|\phi^\star\|_{\mathbb{A}_s} < \infty$. $\qquad\square$

Proof of Theorem 8(c) With Lemma 22, it only remains to prove the implication "$\Longrightarrow$" in (31). The proof is split into three steps, where we may suppose that $\|\phi^\star\|_{\mathbb{A}_s} < \infty$.

Step 1 By Assumption (68), Lemma 20 provides a set $\mathcal{R}_j \subseteq \mathcal{T}_j$ with (69)–(70). Due to stability (A1) and $\lambda_{\text{opt}}^{-1} = C_{\text{stb}} \frac{q_{\text{pcg}}}{1 - q_{\text{pcg}}}$, it holds that

$$\eta_j(\mathcal{R}_j, \phi_j^\star) \overset{(A1)}{\le} \eta_j(\mathcal{R}_j, \phi_{j\underline{k}}) + C_{\text{stb}} \|\!|\!| \phi_j^\star - \phi_{j\underline{k}} |\!|\!| \overset{(71)}{\le} \eta_j(\mathcal{R}_j, \phi_{j\underline{k}}) + \lambda/\lambda_{\text{opt}} \, \eta_j(\phi_{j\underline{k}}).$$

Together with $\theta'' \eta_j(\phi_j^\star) \le \eta_\ell(\mathcal{R}_\ell, \phi_j^\star)$, this proves that

$$(1 - \lambda/\lambda_{\text{opt}})\theta'' \, \eta_j(\phi_{j\underline{k}}) \overset{(72)}{\le} \theta'' \, \eta_j(\phi_j^\star) \le \eta_j(\mathcal{R}_j, \phi_j^\star)$$

$$\le \eta_j(\mathcal{R}_j, \phi_{j\underline{k}}) + \lambda/\lambda_{\text{opt}} \, \eta_j(\phi_{j\underline{k}})$$

and results in

$$\theta \, \eta_j(\phi_{j\underline{k}}) \overset{(68)}{=} \left((1 - \lambda/\lambda_{\text{opt}})\theta'' - \lambda/\lambda_{\text{opt}} \right) \eta_j(\phi_{j\underline{k}}) \le \eta_j(\mathcal{R}_j, \phi_{j\underline{k}}). \tag{77}$$

Hence, $\mathcal{R}_j$ satisfies the Dörfler marking for $\phi_{j\underline{k}}$ with parameter θ. By choice of $\mathcal{M}_j$ in Step (v) of Algorithm 5, we thus infer that

$$\#\mathcal{M}_j \overset{(77)}{\lesssim} \#\mathcal{R}_j \overset{(69)}{\lesssim} \eta_j(\phi_j^\star)^{-1/s} \overset{(72)}{\simeq} \eta_j(\phi_{j\underline{k}})^{-1/s} \quad \text{for all } j \in \mathbb{N}_0.$$

The mesh-closure estimate (R3) guarantees that

$$\#\mathcal{T}_j - \#\mathcal{T}_0 + 1 \overset{(R3)}{\lesssim} \sum_{\ell=0}^{j-1} \#\mathcal{M}_\ell \lesssim \sum_{\ell=0}^{j-1} \eta_\ell(\phi_{\ell\underline{k}})^{-1/s} \quad \text{for all } j > 0. \tag{78}$$

Step 2 For $j = 0$ it holds that $1 \lesssim \Lambda_{0\underline{k}}^{-1/s}$. For $j > 0$, we proceed as follows: Remark 9 yields that $\eta_\ell(\phi_{\ell\underline{k}}) \simeq \Lambda_{\ell\underline{k}}$. Theorem 8(b) and the geometric series prove that

$$\sum_{\ell=0}^{j-1} \eta_\ell(\phi_{\ell\underline{k}})^{-1/s} \simeq \sum_{\ell=0}^{j-1} \Lambda_{\ell\underline{k}}^{-1/s} \overset{(29)}{\lesssim} \sum_{\ell=0}^{j-1} (q_{\text{lin}}^{1/s})^{|(j,\underline{k})-(\ell,\underline{k})|} \Lambda_{j\underline{k}}^{-1/s} \lesssim \Lambda_{j\underline{k}}^{-1/s}.$$

Combining this with (78) and including the estimate for $j = 0$, we derive that

$$\#\mathcal{T}_j - \#\mathcal{T}_0 + 1 \lesssim \Lambda_{j\underline{k}}^{-1/s} \quad \text{for all } j \in \mathbb{N}_0. \tag{79}$$

Step 3 Arguing as in (76) and employing Theorem 8(b), we see that

$$\#\mathcal{T}_j - \#\mathcal{T}_0 + 1 \overset{(76)}{\simeq} \#\mathcal{T}_{j-1} - \#\mathcal{T}_0 + 1 \overset{(79)}{\lesssim} \Lambda_{(j-1)\underline{k}}^{-1/s}$$
$$\overset{(29)}{\lesssim} \Lambda_{jk}^{-1/s} \quad \text{for all } (j,k) \in \mathcal{Q} \text{ with } j > 0.$$

Since $\underline{k}(0) \leq \#\mathcal{T}_0 < \infty$, we hence conclude that $\sup_{(j,k)\in\mathcal{Q}} (\#\mathcal{T}_j - \#\mathcal{T}_0 + 1)^s \Lambda_{jk} < \infty$. $\qquad\square$

6.6 Proof of Corollary 10

For all $\delta > 0$, it holds that

$$\#\mathcal{T}_\bullet - \#\mathcal{T}_0 + 1 \overset{(75)}{\simeq} \#\mathcal{T}_\bullet \leq (\#\mathcal{T}_\bullet) \log^2(1 + \#\mathcal{T}_\bullet) \lesssim (\#\mathcal{T}_\bullet)^{1+\delta} \quad \text{for all } \mathcal{T}_\bullet \in \mathbb{T},$$

where the hidden constant depends only on δ. From (32), it thus follows that

$$\sup_{j\in\mathbb{N}_0} \left[\#\widehat{\mathcal{T}}_j - \#\mathcal{T}_0 + 1\right]^s \widehat{\eta}_j(\widehat{\phi}_j^\star) \lesssim \sup_{j\in\mathbb{N}_0} \left[(\#\widehat{\mathcal{T}}_j) \log^2(1 + \#\widehat{\mathcal{T}}_j)\right]^s \widehat{\eta}_j(\widehat{\phi}_j^\star) < \infty.$$

From Lemma 22, we derive that $\|\phi^\star\|_{\mathbb{A}_s} < \infty$. Hence, Theorem 8(c) yields that

$$\sup_{(j,k)\in\mathcal{Q}} \left[\#\mathcal{T}_j\right]^s \Lambda_{jk} \simeq \sup_{(j,k)\in\mathcal{Q}} \left[\#\mathcal{T}_j - \#\mathcal{T}_0 + 1\right]^s \Lambda_{jk} < \infty. \tag{80}$$

Let $0 < \varepsilon < s$ and choose $\delta > 0$ such that

$$0 < s - \varepsilon = \frac{s}{1 + \delta} =: t.$$

This leads to

$$(\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j) \lesssim (\#\mathcal{T}_j)^{1+\delta} \overset{(80)}{\lesssim} \Lambda_{jk}^{-(1+\delta)/s}$$
$$= \Lambda_{jk}^{-1/t} \quad \text{for all } (j, k) \in \mathcal{Q}.$$

From Theorem 8(b) and the geometric series, it follows that

$$\sum_{(j,k) \leq (j',k')} \Lambda_{jk}^{-1/t} \overset{(29)}{\lesssim} \sum_{(j,k) \leq (j',k')} (q_{\text{lin}}^{1/t})^{|(j',k')| - |(j,k)|} \Lambda_{j'k'}^{-1/t}$$
$$\lesssim \Lambda_{j'k'}^{-1/t} \quad \text{for all } (j', k') \in \mathcal{Q}.$$

Combining the last two estimates, we see that

$$\left[\sum_{(j,k) \leq (j',k')} (\#\mathcal{T}_j) \log^2(1 + \#\mathcal{T}_j)) \right]^{s-\varepsilon} \lesssim \Lambda_{j'k'}^{-(s-\varepsilon)/t} = \Lambda_{j'k'}^{-1} \text{ for all } (j', k') \in \mathcal{Q}.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7 Hyper-singular integral equation

We only sketch the setting and refer to [36] for further details and proofs. Given $f : \Gamma \to \mathbb{R}$, the hyper-singular integral equation seeks $u^\star : \Gamma \to \mathbb{R}$ such that

$$(W u^\star)(x) := -\partial_{\boldsymbol{n}(x)} \int_\Gamma \partial_{\boldsymbol{n}(y)} G(x - y) u^\star(y) \, \mathrm{d}y = f(x) \quad \text{for all } x \in \Gamma, \quad (81)$$

where $\partial_{\boldsymbol{n}}$ denotes the normal derivative with the outer unit normal vector $\boldsymbol{n}(\cdot)$ on $\Gamma \subseteq \partial\Omega$. For $0 \leq \alpha \leq 1$, define $\widetilde{H}^\alpha(\Gamma) := \{v \in H^\alpha(\Gamma) : \text{supp}(v) \subseteq \overline{\Gamma}\}$ and let $H^{-\alpha}(\Gamma)$ be its dual space with respect to $\langle \cdot, \cdot \rangle$. Note that $\widetilde{H}^{\pm\alpha}(\Gamma) = H^{\pm\alpha}(\Gamma)$ for $\Gamma = \partial\Omega$. The hyper-singular integral operator $W : \widetilde{H}^{1/2+s}(\Gamma) \to H^{-1/2+s}(\Gamma)$ is a bounded linear operator for all $-1/2 \leq s \leq 1/2$ which is even an isomorphism for $-1/2 < s < 1/2$. For $s = 0$, the operator W is symmetric and (since Γ is connected) positive semi-definite with kernel being the constant functions. For $\Gamma \subsetneqq \partial\Omega$, the operator $W : \widetilde{H}^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ is hence an elliptic isomorphism. Moreover, for $\Gamma = \partial\Omega$ and $H_*^{\pm 1/2}(\Gamma) := \{\psi \in H^{\pm 1/2}(\Gamma) : \langle \psi, 1 \rangle = 0\}$, $W : H_*^{1/2}(\Gamma) \to H_*^{-1/2}(\Gamma)$ is an elliptic isomorphism. Therefore,

$$\langle u\,,\,v\rangle := \begin{cases} \langle Wu\,,\,v\rangle, & \text{if } \Gamma \subsetneqq \partial\Omega, \\ \langle Wu\,,\,v\rangle + \langle u\,,\,v\rangle, & \text{if } \Gamma = \partial\Omega \end{cases}$$

defines a scalar product on $\widetilde{H}^{1/2}(\Gamma)$, and the induced norm $\|\!|\,u\,|\!\| := \langle u\,,\,u\rangle^{1/2}$ is an equivalent norm on $\widetilde{H}^{1/2}(\Gamma)$. Let $f \in H^{-1/2}(\Gamma)$. If $\Gamma \subsetneqq \partial\Omega$, suppose additionally that $f \in H_*^{-1/2}(\partial\Omega)$. Then, (81) admits a unique solution $u^\star \in \widetilde{H}^{1/2}(\Gamma)$ resp. $u^\star \in H_*^{1/2}(\partial\Omega)$, which is also the unique solution $u^\star \in \widetilde{H}^{1/2}(\Gamma)$ of the variational formulation

$$\langle u^\star\,,\,v\rangle = \langle f\,,\,v\rangle \quad \text{for all } v \in \widetilde{H}^{1/2}(\Gamma).$$

Given a mesh $\mathcal{T}_\bullet$ of Γ, let

$$\widetilde{\mathcal{S}}^1(\mathcal{T}_\bullet) := \left\{ v \in \widetilde{H}^{1/2}(\Gamma)\,:\, \forall T \in \mathcal{T}_\bullet \quad v|_T \text{ is affine} \right\}.$$

The Lax–Milgram theorem yields existence and uniqueness of $u_\bullet^\star \in \widetilde{\mathcal{S}}^1(\mathcal{T}_\bullet)$ such that

$$\langle u_\bullet^\star\,,\,v_\bullet\rangle = \langle f\,,\,v_\bullet\rangle \quad \text{for all } v_\bullet \in \widetilde{\mathcal{S}}^1(\mathcal{T}_\bullet).$$

With the corresponding weighted-residual error estimator, it holds that

$$\|\!|\,u^\star - u_\bullet^\star\,|\!\| \le C_{\text{rel}}\,\eta_\bullet(u_\bullet^\star) := \left(\sum_{T \in \mathcal{T}_\bullet} \eta_\bullet(T, u_\bullet^\star)^2 \right)^{1/2},$$
$$\text{where } \eta_\bullet(T, u_\bullet^\star)^2 := h_T\,\|f - Wu_\bullet^\star\|_{L^2(T)}^2;$$

see [10,17] for $d = 2$ resp. [14] for $d = 3$.

In [22,26], optimal additive Schwarz preconditioners are derived for this setting. Hence, Algorithm 5 can also be used in the present setting. We refer to [20, Section 3.3] for the fact that the *axioms of adaptivity* (A1)–(A4) from Proposition 16 remain valid for the hyper-singular integral equation. All other arguments in Section 6 rely only on general properties of the PCG algorithm (Sect. 6.2), the properties (A1)–(A4), and the Hilbert space setting of $\|\!|\,\cdot\,|\!\|$. Overall, this proves that our main results (Theorem 8 and Corollary 10) also cover the hyper-singular integral equation.

References

1. Aurada, M., Ebner, M., Feischl, M., Ferraz-Leite, S., Führer, T., Goldenits, P., Karkulik, M., Mayr, M., Praetorius, D.: A Matlab implementation of adaptive 2D-BEM. Numer. Algorithms **67**, 1–32 (2014)
2. Aurada, M., Feischl, M., Führer, T., Karkulik, M., Praetorius, D.: Efficiency and optimality of some weighted-residual error estimator for adaptive 2D boundary element methods. Comput. Methods Appl. Math. **13**(3), 305–332 (2013)
3. Aurada, M., Feischl, M., Führer, T., Karkulik, M., Melenk, J.M., Praetorius., D.: Local inverse estimates for non-local boundary integral operators. Math. Comput. **86**(308), 2651–2686 (2017)
4. Arioli, M., Georgoulis, E.H., Loghin, D.: Stopping criteria for adaptive finite element solvers. SIAM J. Sci. Comput. **35**(3), A1537–A1559 (2013)
5. Ainsworth, M., Guzmán, J., Sayas, F.-J.: Discrete extension operators for mixed finite element spaces on locally refined meshes. Math. Comput. **85**(302), 2639–2650 (2016)
6. Arioli, M., Liesen, J., Międlar, A., Strakoš, Z.: Interplay between discretization and algebraic computation in adaptive numerical solution of elliptic PDE problems. GAMM Mitt. **36**(1), 102–129 (2013)
7. Ainsworth, M., McLean, W., Tran, T.: The conditioning of boundary element equations on locally refined meshes and preconditioning by diagonal scaling. SIAM J. Numer. Anal. **36**(6), 1901–1932 (1999)
8. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rates. Numer. Math. **97**(2), 219–268 (2004)
9. Bespalov, A., Haberl, A., Praetorius, D.: Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. Comput. Methods Appl. Mech. Eng. **317**, 318–340 (2017)
10. Carstensen, C.: An a posteriori error estimate for a first-kind integral equation. Math. Comput. **66**(217), 139–155 (1997)
11. Carstensen, C., Feischl, M., Page, M., Praetorius, D.: Axioms of adaptivity. Comput. Math. Appl. **67**(6), 1195–1253 (2014)
12. Carstensen, C., Gedicke, J.: An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. SIAM J. Numer. Anal. **50**(3), 1029–1057 (2012)
13. Cascon, J.M., Kreuzer, C., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**(5), 2524–2550 (2008)
14. Carstensen, C., Maischak, M., Praetorius, D., Stephan, E.P.: Residual-based a posteriori error estimate for hypersingular equation on surfaces. Numer. Math. **97**(3), 397–425 (2004)
15. Carstensen, C., Maischak, M., Stephan, E.P.: A posteriori error estimate and h-adaptive algorithm on surfaces for Symm's integral equation. Numer. Math. **90**(2), 197–213 (2001)
16. Carstensen, C., Praetorius, D.: Averaging techniques for the effective numerical solution of Symm's integral equation of the first kind. SIAM J. Sci. Comput. **27**(4), 1226–1260 (2006)
17. Carstensen, C., Stephan, E.P.: A posteriori error estimates for boundary element methods. Math. Comput. **64**(210), 483–500 (1995)
18. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
19. Feischl, M., Führer, T., Karkulik, M., Melenk, J.M., Dirk, P.: Quasi-optimal convergence rates for adaptive boundary element methods with data approximation. Part I: weakly-singular integral equation. Calcolo **51**, 531–562 (2014)
20. Feischl, M., Führer, T., Karkulik, M., Melenk, J.M., Praetorius, D.: Quasi-optimal convergence rates for adaptive boundary element methods with data approximation. Part II: hyper-singular integral equation. Electron. Trans. Numer. Anal. **44**, 153–176 (2015)
21. Feischl, M., Führer, T., Praetorius, D.: Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. SIAM J. Numer. Anal. **52**(2), 601–625 (2014)
22. Feischl, M., Führer, T., Praetorius, D., Stephan, E.P.: Optimal additive Schwarz preconditioning for hypersingular integral equations on locally refined triangulations. Calcolo **54**(1), 367–399 (2017)
23. Feischl, M., Führer, T., Praetorius, D., Stephan, E.P.: Optimal preconditioning for the symmetric and nonsymmetric coupling of adaptive finite elements and boundary elements. Numer. Methods Partial Differ. Equ. **33**(3), 603–632 (2017)
24. Feischl, M., Karkulik, M., Melenk, J.M., Praetorius., D.: Quasi-optimal convergence rate for an adaptive boundary element method. SIAM J. Numer. Anal. **51**, 1327–1348 (2013)

25. Führer, T., Melenk, J.M., Praetorius, D., Rieder, A.: Optimal additive Schwarz methods for the hp-BEM: the hypersingular integral operator in 3D on locally refined meshes. Comput. Math. Appl. **70**, 1583–1605 (2015)
26. Führer, T.: Zur Kopplung von finiten Elementen und Randelementen. Ph.D. thesis, TU Wien (2014)
27. Gantumur, T.: Adaptive boundary element methods with convergence rates. Numer. Math. **124**(3), 471–516 (2013)
28. Gantner, G., Haberl, A., Praetorius, D., Stiftner, B.: Rate optimal adaptive FEM with inexact solver for nonlinear operators. IMA J. Numer. Anal. **38**, 1797–1831 (2018)
29. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences, 4th edn. Johns Hopkins University Press, Baltimore (2013)
30. Hackbusch, W.: Hierarchical Matrices: Algorithms and Analysis, Volume 49 of Springer Series in Computational Mathematics. Springer, Heidelberg (2015)
31. Hiptmair, R., Jerez-Hanckes, C., Mao, S.: Extension by zero in discrete trace spaces: inverse estimates. Math. Comput. **84**(296), 2589–2615 (2015)
32. Hiptmair, R., Mao, S.: Stable multilevel splittings of boundary edge element spaces. BIT **52**(3), 661–685 (2012)
33. Hiptmair, R., Haijun, W., Zheng, W.: Uniform convergence of adaptive multigrid methods for elliptic problems and Maxwell's equations. Numer. Math. Theory Methods Appl. **5**(3), 297–332 (2012)
34. Hiptmair, R., Zheng, W.: Local multigrid in $H(\mathbf{curl})$. J. Comput. Math. **27**(5), 573–603 (2009)
35. Karkulik, M., Pavlicek, D., Praetorius, D.: On 2D newest vertex bisection: optimality of mesh-closure and H^1-stability of L_2-projection. Constr. Approx. **38**, 213–234 (2013)
36. McLean, W.: Strongly Elliptic Systems and Boundary Integral Equations. Cambridge University Press, Cambridge (2000)
37. Morin, P., Nochetto, R.H., Siebert, K.G.: Data oscillation and convergence of adaptive FEM. SIAM J. Numer. Anal. **38**(2), 466–488 (2000)
38. Oswald, P.: Multilevel finite element approximation. Teubner Skripten zur Numerik [Teubner Scripts on Numerical Mathematics]. Theory and Applications. B.G. Teubner, Stuttgart (1994)
39. Oswald, P.: Interface preconditioners and multilevel extension operators. In: Eleventh International Conference on Domain Decomposition Methods (London, 1998), pp. 97–104. DDM.org, Augsburg (1999)
40. Śmigaj, W., Betcke, T., Arridge, S., Phillips, J., Schweiger, M.: Solving Boundary Integral Problems with BEM++. ACM, New York (2013)
41. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**(2), 245–269 (2007)
42. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. Math. Comput. **77**(261), 227–241 (2008)
43. Stevenson, R., van Venetië, R.: Optimal preconditioning for problems of negative order. ArXiv e-prints, March 2018
44. Toselli, A., Widlund, O.: Domain Decomposition Methods-Algorithms and Theory, Volume 34 of Springer Series in Computational Mathematics. Springer, Berlin (2005)
45. Xu, J., Chen, L., Nochetto, R.H.: Optimal multilevel methods for $H(\text{grad})$, $H(\text{curl})$, and $H(\text{div})$ systems on graded and unstructured grids. In: DeVore, R., Kunoth, A. (eds.) Multiscale, Nonlinear and Adaptive Approximation, pp. 599–659. Springer, Berlin (2009)