TECHNISCHE UNIVERSITÄT WIEN

# DISSERTATION

# On adaptive FEM and BEM for indefinite and nonlinear problems

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften unter der Leitung von

**Univ.-Prof Dipl.-Math Dr. techn. Dirk Praetorius**
E101 – Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

**Dipl.-Ing. Alexander Haberl BSc.**
Matrikelnummer: 00925297
Bräuhausgasse 48/63
1050 Wien

Diese Dissertation haben begutachtet:

1. **Prof. Dr. Dirk Praetorius**
   Institut für Analysis and Scientific Computing, TU Wien
2. **Prof. Dr. Alex Bespalov**
   School of Mathematics, University of Birmingham
3. **Prof. Dr. Timo Betcke**
   Department of Mathematics, University College London

Wien, am 16. April 2018

# Kurzfassung

Das Ziel dieser Arbeit ist die Erweiterung der Analysis von adaptiven Algorithmen für Finite Elemente Methoden (FEM) und Randelementmethoden (BEM) von elliptischen Problemen im Rahmen des Lax–Milgram Lemmas zu elliptisch indefiniten und nichtlinearen partiellen Differentialgleichungen. Diese unterschiedlichen Problemklassen teilen die Arbeit in zwei Teile. Basierend auf einem *a posteriori* Fehlerschätzer formulieren wir jeweils einen adaptiven Algorithmus, welcher neben linearer Konvergenz auch zu optimalen Konvergenzverhalten des zugrunde liegenden Fehlerschätzers führt. Die Analysis bedient sich dabei eines komplett abstrakten Rahmens. Dieser erlaubt es, essentielle und hinreichende Eigenschaften des Fehlerschätzers und der zugrunde liegenden Netzverfeinerung zu bestimmen, welche in weiterer Folge optimale Konvergenzraten und optimale Komplexität garantieren.

Der Fokus des ersten Teils liegt auf kompakt gestörten Problemen. Diese Problemklasse beinhaltet allgemeine Diffusionsprobleme mit Konvektion und Reaktion, und im speziellen auch die Helmholtz-Gleichung. In bisherigen Resultaten für FEM und BEM mit stückweisen polynomiellen Ansatz- und Testräumen, wird mit Hilfe des dualen Problems die Existenz und Eindeutigkeit von diskreten Lösungen für hinreichend feine Netze garantiert. Wie jedoch der abstrakte Rahmen dieser Arbeit zeigt, ist diese pessimistische *a priori* Annahme bzw. Einschränkung nicht notwendig. Die adaptive Netzverfeinerung ist aufgrund von Stabilisierungseffekten in der Lage diese Startphase zu überwinden und liefert, unabhängig von der Netzweite des Startnetzes, asymptotisch optimales Abklingverhalten des Fehlerschätzers. Als Anwendung der abstrakten Analysis beweisen wir optimale Konvergenz von adaptiver FEM für kompakt gestörte elliptische Probleme.

Des Weiteren zeigen wir inverse Ungleichungen für alle fundamentalen Randintegraloperatoren der Helmholtz-Gleichung, welche bestehende Resultate für den Laplace Operator auf beliebige Wellenzahlen verallgemeinern. Mit Hilfe dieser Abschätzung, gibt die Arbeit einen ersten Beweis für die Optimalität der adaptiven Randelementmethode für die Helmholtz-Gleichung. Eine andere Stärke der BEM ist die Konvergenz des punktweisen Fehlers mit höherer Ordnung. Basierend auf Resultaten für elliptische Gleichungen zeigen wir zusätzlich optimales Konvergenzverhalten für eine berechenbare obere Schranke für den Punktfehler.

Im zweiten Teil betrachten wir nichtlineare Differentialgleichungen mit stark monotonen Operatoren. Im Kontrast zu bestehenden Arbeiten betrachtet der abstrakte Rahmen neben einer Picard-Iteration für das auftretende nichtlineare diskrete Problem, auch einen iterativen PCG-Löser für lineare Gleichungssysteme. Zusätzlich zu optimalem Konvergenzverhalten des Fehlerschätzers im Hinblick auf die Freiheitsgrade der verwendeten Diskretisierung zeigen wir auch Optimalität in Bezug auf den kumulativen Rechenaufwand des adaptiven Algorithmus.

# Abstract

The goal of this work is to generalize the analysis of adaptive algorithms for finite element methods (FEM) and boundary element methods (BEM) from elliptic problems, satisfying the setting of the Lax–Milgram theorem, to certain classes of elliptic indefinite and nonlinear problems. For each problem class, based on an *a posteriori* error estimator, we introduce an adaptive algorithm and prove that these algorithms do not only lead to linear convergence, but also guarantee optimal algebraic convergence behavior of the underlying error estimator. The thesis is split into two parts, where each part analyzes one specific problem class in an abstract framework. This general approach allows to formulate so-called *axioms of adaptivity* for the error estimator as well as the underlying mesh-refinement strategy, under which optimal algebraic convergence can be guaranteed.

First, we consider indefinite and compactly perturbed elliptic problems. This problem class covers general diffusion problems with convection and reaction and in particular, the Helmholtz equation. For a standard conforming FEM and BEM discretization by piecewise polynomials, usual duality arguments show that the underlying triangulation has to be sufficiently fine to ensure the existence and uniqueness of the Galerkin solution. Extending the abstract approach of existing works, we prove that adaptive mesh-refinement is capable of overcoming this preasymptotic behavior and eventually leads to convergence with optimal algebraic rates. Unlike previous works, one does not have to deal with the *a priori* assumption that the initial mesh is sufficiently fine. Due to stabilizing effects, the adaptive algorithm can, in particular, overcome possibly pessimistic restrictions on the meshes. As an application of the abstract framework, we prove optimal algebraic convergence rates for adaptive FEM.

Further, we show inverse estimates for the most important boundary integral operators associated with the Helmholtz equation, which generalizes the existing results for the Laplace equation to arbitrary wavenumbers. This allows us to give a first prove of optimal convergence rates for adaptive BEM for the Helmholtz equation. One particular strength of the boundary element methods is, that it allows for a higher-order point-wise approximation of the solution. As an application of the prior analysis, we generalize existing results for the elliptic case and prove optimal convergence behavior with respect to an *a posteriori* computable bound for the point error of the Helmholtz equation.

In the second part, we focus on nonlinear PDEs with strongly monotone operators. Unlike prior works, the analysis includes the iterative and inexact solution of the arising discrete nonlinear systems by means of the Picard iteration. We also consider an iterative PCG-solver for the invoked linear system in the computation of each Picard step. Using nested iteration, we show an improved linear convergence result as well as optimal algebraic convergence behavior of the underlying error estimator. Improving existing results, we also prove optimal convergence rates with respect to the cumulative computational costs of the adaptive algorithm.

# Danksagung

Diese Danksagung bietet mir Gelegenheit, jene Menschen zu würdigen, die mich seit Jahren begleiten und mir damit ermöglichen, meine Ziele zu verwirklichen.

Zuerst möchte ich mich bei meinem Betreuer Dirk Praetorius bedanken. Dieser hat mich während meiner Promotionszeit nicht nur als Mentor, sondern inbesondere auch als Freund, immer mit großem Eifer und Hingabe unterstützt.
Großer Dank gilt natürlich auch Alex Bespalov und Timo Betcke sowohl für die Begutachtung der Dissertation als auch für ihre Expertise und Bemühungen in der gemeinsamen Forschung. Besonders möchte ich noch folgenden Personen danken:

- Meinen Arbeitskollegen Alex, Bernhard, Carl, Gregor, Giovanni, Markus, Michael, Michele, Stefan und Thomas, dafür, dass die Universität zu einem zweiten Zuhause wurde und ich die Arbeitszeit mit Freunden verbringen durfte. Inbesondere auch Ursula Schweigler für Berge an Schokolade und Hilfe beim Überspringen so mancher bürokratischer Hürde.
- Allen Tarockierern und Freunden für unvergessliche Zeiten und Freundschaften seit dem Kindesalter.
- Meinen Studienkollegen Fabian, Stefan und Stefanie, für unzählige Lösungen und Erklärungen im Studium.
- Meiner Familie, Angelika, Daniela, Martin, Peter, für den Rückhalt, die Unterstützung und die Möglichkeiten, die sie mir gegeben haben.

Der größte Dank gilt meiner besseren Hälfte Anna. Ohne dich wäre ich wohl nie soweit gekommen. Danke für einfach Alles und noch vieles mehr.

Ich danke der TU Wien und dem FWF für die finanzielle Unterstützung meiner Forschungsarbeit im Zuge des Forschungsprojektes *Optimal adaptivity in BEM and FEM-BEM coupling (Grant P27005).*

# Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 16. April 2018

Alexander Haberl

# Contents

# 1 Introduction

## 1.1 Motivation

The *finite element method* (FEM) is one of the most important tools in numerical analysis to solve partial differential equations (PDEs). Over the last decades, it has shown its potential by providing solutions to a variety of problems, arising in applications in natural sciences as well as engineering. This overwhelming success gave a real impetus to the numerical analysis of FEM and lead to the development of various numerical schemes using the principal ideas of finite elements. The key ingredient in most of these methods is based on the discretization of the domain of interest by a mesh of polygons, which reduces the PDE to a finite dimensional linear system of equations and gives rise to a finite dimensional approximation of the unknown solution. The quality of the thereby computed approximation is controlled by the mesh-width of the underlying discretization. Hence, a simple and commonly used technique to guarantee convergence of the error to zero is to successively refine the corresponding mesh uniformly.

In general, geometry and data induced singularities of the unknown exact solution might reduce the possible order of convergence significantly and thus spoil the accuracy of the computed approximation. This leads to a dispensable increase of the underlying computational costs with respect to the quality of the achieved error. However, for many problems, the reduction of the order of convergence as well as the additional computational costs can be avoided by refining the mesh locally at these singularities. Doing this beforehand requires *a priori* information of the unknown exact solution, which is in general not available. This observation has led to the development of *adaptive finite element methods* (AFEM) and adaptive mesh-refinement. Based on an *a posteriori* error estimator which reflects the behavior of the approximation error, adaptive algorithms automatically steer the local refinement to recover optimal rates of convergence.

In recent years, the analysis of convergence and optimal convergence behavior of adaptive algorithms has matured. We refer to the seminal works [Dör96, MNS00, BDD04, Ste07, CKNS08, FFP14] for some milestones of AFEM for linear elliptic PDEs as well as to the works [FFK+14, FFK+15, FKMP13, AFF+17, Gan13] for adaptive boundary element methods (ABEM) in case of the Laplace equation.

The aim of this thesis is to extend the existing analysis on adaptive algorithms to certain classes of indefinite or nonlinear elliptic problems. To that end, we develop suitable adaptive algorithms and prove optimal algebraic convergence rates of the underlying error estimator.

## 1.2 State of the art and outline

### Chapter 2

First, we fix some basic notation and introduce Sobolev spaces on domains $\Omega \subset \mathbb{R}^d$ as well as on the corresponding boundaries $\partial\Omega$. Since these spaces are essential for the analysis of the succeeding chapters, we recall the definitions and summarize their most important properties.

### Chapter 3

One essential ingredient of the analysis of adaptive algorithms is the underlying mesh-refinement. In Chapter 3, we introduce meshes of domains $\Omega \subset \mathbb{R}^d$ and boundaries $\partial\Omega$. Section 3.3 formulates the so-called *axioms of refinement* in the spirit of [CFPP14]. Then, the *a posteriori* error analysis depends only on the refinement strategy properties (R3)–(R6). This allows to formulate the upcoming chapters in a completely abstract setting, independently of the actual refinement. Sections 3.4–3.5 recall that these axioms are satisfied for *newest vertex bisection* (NVB) and *extended bisection* (EB), which are used in specific settings for AFEM and ABEM later on.

### Chapter 4

In the prior mentioned references, only linear problems satisfying the setting of the Lax–Milgram theorem have been treated. In the more general case of compactly perturbed elliptic problems, those works require a sufficiently fine initial mesh [MN05, CN12] or strong monotonicity [FFP14] in order to guarantee optimal convergence. On the other hand, numerical experiments show that adaptive algorithms recover the optimal rate of convergence, independently of whether the initial mesh is sufficiently fine or not. Based on the own work [BHP17], we introduce an abstract framework in the style of [CFPP14], which is utilized to close this gap for conforming AFEM (Chapter 5) and enables a first optimality proof for ABEM for the Helmholtz equation in Chapter 6.

To that end, let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}^*$ denote its dual space. Given $f \in \mathcal{H}^*$, we consider variational formulations of the type

$$a(u,v) + \langle \mathcal{K}u, v\rangle = \langle f, v\rangle \quad \text{for all } v \in \mathcal{H}, \tag{1.1}$$

where $a(\cdot,\cdot)$ is an elliptic and symmetric bilinear form on $\mathcal{H}$ and $\mathcal{K} : \mathcal{H} \to \mathcal{H}^*$ is a continuous and compact linear operator. Given an initial triangulation $\mathcal{T}_0$, a typical adaptive algorithm consisting of the steps

$$\boxed{\texttt{SOLVE}} \implies \boxed{\texttt{ESTIMATE}} \implies \boxed{\texttt{MARK}} \implies \boxed{\texttt{REFINE}} \tag{1.2}$$

generates a sequence of refined meshes $\mathcal{T}_\ell$ with corresponding nested spaces $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{H}$ for all $\ell \geq 0$. We stress that unlike to the prior works [MN05, CN12, FFP14], Algorithm 4.4 in Section 4.4 will not be given any *a priori* information whether the mesh $\mathcal{T}_0$ is sufficiently fine. Instead, we introduce an additional step in the algorithm, which performs uniform refinement if the corresponding Galerkin solution of (1.1) does not exist. In Section 4.5, we define the *axioms of adaptivity* (E1)–(E4). These are a slight generalization of those of [CFPP14] and collect all estimator dependencies of the analysis. We introduce an additional axiom (E5). This ensures definiteness and therefore well-posedness of (1.1) on the "discrete" limit space. We emphasize that (E5) sets no additional limitation to the analysis and can be enforced by the adaptive algorithm using an expanded Dörfler marking strategy (Proposition 4.7). Then, Section 4.7 proves linear convergence of Algorithm 4.4 (Theorem 4.14), and also addresses the validity of the Céa lemma with optimal constant 1 (Proposition 4.16). Section 4.8.1 introduces approximation classes in the spirit of [CFPP14] and discusses their connection to other classes defined in [CKNS08]. Finally, we prove optimal algebraic convergence rates of Algorithm 4.4 in Theorem 4.21, independently of whether the initial mesh is sufficiently fine or not.

## Chapter 5

Chapter 5 applies the abstract framework of Chapter 4 to AFEM for general diffusion problems with convection and reaction; see Section 5.1. We consider piecewise polynomial ansatz and test spaces with arbitrary but fixed polynomial degree $p \geq 1$. Section 5.2 introduces the corresponding weighted-residual error estimator and proves that the axioms (E1)–(E4) are satisfied (Proposition 5.3–5.5). Then, Algorithm 5.2 with the expanded Dörfler marking additionally guarantees (E5). Utilizing the abstract analysis of Chapter 4, we obtain optimal convergence rates of the adaptive algorithm (Theorem 5.8). To underpin the theoretical findings, we conclude the chapter with numerical examples for the 2D Helmholtz equation.

## Chapter 6

Chapter 6 focuses on ABEM for the Helmholtz equation. First, Section 6.2 gives a brief introduction to BEM, where we fix the notation and recall some basic properties of the related boundary integral operators. This chapter generalizes existing results in [AFF$^+$17, FKMP13, FFK$^+$14, FFK$^+$15, Gan13] for the Laplace equation to general wavenumbers $k > 0$. To that end, we consider the weakly-singular integral equation: Given $f \in H^{1/2}(\Gamma)$, find $\phi \in \widetilde{H}^{-1/2}(\Gamma)$ such that

$$\langle \mathfrak{V}_k \, \phi \, , \, \psi \rangle = \langle f \, , \, \psi \rangle \quad \text{for all } \psi \in \widetilde{H}^{-1/2}(\Gamma), \tag{1.3}$$

where $\mathfrak{V}_k$ is the simple-layer operator for wavenumber $k > 0$. Building on a potential decomposition from [Mel12], Section 6.4 proves an inverse estimate for the most important boundary integral operators associated with the Helmholtz equation (Theorem 6.3). With the help of the inverse estimate, Section 6.5 shows the validity of the estimator axioms (E1)–(E4) for the weighted-residual error estimator corresponding to the model problem (1.3). The main result of this chapter is Theorem 6.11. It proves that the adaptive algorithm does not only lead to linear convergence, but also guarantees optimal algebraic convergence rates. We emphasize that Theorem 6.11 is independent of, whether direct or indirect BEM is used or the initial mesh is sufficiently fine. Section 6.7 transfers the main result to the hyper-singular integral equation (Theorem 6.14) and highlights the occurring differences in the proofs. We conclude this chapter with numerical experiments for 3D wave scattering problems in Section 6.8.

## Chapter 7

As an extension of Chapter 6, the goal of Chapter 7 is the optimal computation of point-values of solutions of the Helmholtz equation. Via the representation formula, BEM allows for a high-order pointwise approximation of the underlying PDE solution. Utilizing Chapter 4, we introduce two adaptive algorithms in Section 7.2 in the style of *goal oriented* AFEM from [MS09] and [BET11]. Further, both Algorithms 7.1–7.3 incorporate the expanded Dörfler marking from Proposition 4.7. Transferring ideas from [FGH$^+$16] to the Helmholtz setting, Theorem 7.5 proves optimal algebraic convergence of an *a posteriori* computable upper bound for the point error. We emphasize, that the analysis of this chapter covers in general also *goal-oriented* adaptivity for AFEM and thus transfers some results of [FPZ16] to the Helmholtz setting.

## Chapter 8

As for linear problems, the analysis of convergence and optimal convergence behavior of AFEM for nonlinear problems has been a fertile field for research on numerical analysis, cf. [Vee02, DK08, BDK12, GMZ12]. The interplay of adaptive mesh-refinement, optimal convergence rates, and inexact solvers has already been addressed and analyzed, e.g., in [Ste07, AGL13, ALMS13, CFPP14] for FEM for linear PDEs and in [CG12] for linear eigenvalue problems. The work [GMZ11] considers adaptive mesh-refinement in combination with a Kačanov-type iterative solver for strongly monotone operators. In the spirit of [MSV08, Sie11], the focus is on a plain convergence result of the overall strategy, while the proof of optimal convergence rates remains open. On the other hand, the influence of inexact solvers for nonlinear equations on optimal convergence has only recently been analyzed in our own work [GHPS17].

Chapter 8 considers elliptic nonlinear model problems with weak formulations of the following type: Given $F \in \mathcal{H}^*$, find $u^\star \in \mathcal{H}$ such that

$$\langle \mathfrak{A}u^\star \,,\, v\rangle = \langle F \,,\, v\rangle \quad \text{for all } v \in \mathcal{H}, \tag{1.4}$$

where $\mathfrak{A}$ is a strongly monotone and Lipschitz continuous operator. As in [GHPS17, CW17], we consider an inexact Picard iteration to compute the discrete solution of (1.4) (Section 8.3). Further, the computation of each Picard iteration requires to solve a discrete Laplace problem (Section 8.3.2). Unlike the prior works [GHPS17, CW17], we do not assume that the arising linear system is solved exactly. Instead, we consider an inexact iterative PCG solver. Section 8.5 introduces the adaptive strategy in Algorithm 8.7. Besides the normal adaptive loop in (1.2), the *a posteriori* error estimator steers also the Picard iteration as well as the PCG iteration. In the spirit of [CFPP14], the analysis is done in a completely abstract setting. To this end, Section 8.6 reformulates the estimator axioms, which slightly differ from Chapter 4. We first prove linear convergence of the adaptive algorithm (Theorem 8.20) and optimal convergence rates in Theorem 8.21 for sufficiently small, but independent control parameters $\lambda_{\mathrm{Pic}}$ and $\lambda_{\mathrm{PCG}}$. A more involved choice of the involved parameters leads to stronger linear convergence result (Theorem 8.30), i.e., contraction of the error estimator in each step of either PCG or Picard iteration. Although Algorithm 8.7 includes two nested iterative solvers, Theorem 8.32 proves that Algorithm 8.7 also guarantees optimal algebraic convergence with respect to the cumulative computational effort.

This guarantees optimal computational complexity and improves and generalizes the existing result from [GHPS17].

### Chapter 9

Chapter 9 is an application of Chapter 8. We consider AFEM for certain types of nonlinear boundary value problems similar to those of [GMZ11, GMZ12, BSF$^+$14, CW17] and show that these problems fit in the strongly monotone setting of Chapter 8. We prove that the corresponding weighted-residual error estimator satisfies the estimator axioms of Chapter 8. Then, with the help of the abstract framework, we prove optimal convergence rates for the underlying error estimator in Theorem 5.8. Finally, Section 9.4 underpins our theoretical findings with numerical experiments for lowest-order AFEM in $\mathbb{R}^2$.

# 2 Sobolev spaces and basic notation

In this chapter, we introduce some basic notations and definitions. Section 2.2 defines the necessary Sobolev spaces on domains $\Omega \subset \mathbb{R}^d$ as well as on boundaries $\partial\Omega$ and Section 2.2.3 introduces the corresponding dual spaces. Last, Section 2.2.4 recalls the definition and basic properties of the trace operators as well as of the conormal derivative. To abbreviate notation, we use the following convention.

**General notation.** Throughout all statements, all constants as well as their dependencies are explicitly given. In proofs, we may abbreviate the notation by use of the symbol $\lesssim$ which indicates $\leq$ up to some multiplicative constant which is clear from the context. Analogously, $\gtrsim$ indicates $\geq$ up to a multiplicative constant. Moreover, the symbol $\simeq$ states that both estimates $\lesssim$ and $\gtrsim$ hold.

## 2.1 Basic notation

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with piecewise $C^\infty$-boundary $\partial\Omega$ and exterior normal vector $\mathbf{n} = \mathbf{n}(y)$ for every $y \in \partial\Omega$; see, e.g., [SS11, Definition 2.2.10]. Further, let $\Omega^{\mathrm{ext}} := \mathbb{R}^d \setminus \overline{\Omega}$ denote the corresponding exterior domain. The Euclidean norm of two points $x \in \mathbb{R}^d$ is denoted by $|x|$. For measureable sets $S \subseteq \Omega$ or $S \subseteq \partial\Omega$ and if it is clear from the context, we use the same notation $|S|$ for the corresponding Lebesgue measure and the surface measure.

For $p > 0$, let $L^p(\Omega)$ be the usual Lebesgue spaces on $\Omega$ with corresponding norm $\|\cdot\|_{L^p(\Omega)}$. Analogously, Lebesgue spaces on the boundary $\partial\Omega$ are denoted by $L^p(\partial\Omega)$ with norm $\|\cdot\|_{L^p(\partial\Omega)}$. If a space $\mathcal{H}$ has additional Hilbert space structure, e.g., $\mathcal{H} = L^2(\Omega)$, the corresponding scalar product is denoted by $(\cdot\,,\,\cdot)_{\mathcal{H}}$.

Let $v : \Omega \subset \mathbb{R}^d \to \mathbb{R}$. The weak gradient and the divergence of $v$ are given by

$$\nabla v = \big(\partial_1 v, \dots, \partial_d v\big) \quad \text{and} \quad \operatorname{div} v = \sum_{j=1}^{d} \partial_j v_j,$$

where, $\partial_i := \frac{\partial}{\partial x_i}$ denotes the weak partial derivative (if it exists). This definition gives rise to the Laplace operator

$$\Delta v := \operatorname{div}(\nabla v) = \sum_{i=1}^{d} \partial_i^2 v.$$

Moreover, for $n \in \mathbb{N}$, let $\alpha \in \mathbb{N}_0^n$ be some multi index, where $|\alpha| := \sum_{i=1}^{n} \alpha_i$. Then, a function $v$ has a weak derivative $g := \partial^\alpha v$ of order $\alpha$ if

$$\int_\Omega g\, w \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega v\, \partial^\alpha w \,\mathrm{d}x \quad \text{for all } w \in C_0^\infty(\Omega),$$

where $C_0^\infty(\Omega)$ is the space of all infinitely differentiable functions with compact support. Note that if $\partial^\alpha v$ exists, e.g., $v \in L^2(\Omega)$ and $\partial^\alpha v \in L^2(\Omega)$, then the weak derivative is unique.

## 2.2 Sobolev spaces

In this section, we give a short introduction on Sobolev spaces on domains $\Omega$ as well as on their boundaries $\partial\Omega$. In contrast to spaces on domains $\Omega$, which are needed throughout the entire thesis, the usage of Sobolev spaces on boundaries is mostly restricted to the chapters on adaptive boundary element methods, i.e., Chapter 6 and Chapter 7. We emphasize that, although we recall just one definition, most spaces can be defined equivalently in different ways. Therefore, we refer to the monographs [McL00, SS11, Tar07, Tri83, Tri95] for further details.

### 2.2.1 Sobolev spaces on a domain $\Omega$

The Sobolev spaces on domains are defined in the usual sense; see, e.g., [McL00, p. 58] or [SS11, Section 2.3]. To that end, for $\ell \in \mathbb{N}_0$ we define $H^\ell(\Omega)$ by

$$H^\ell(\Omega) := \big\{ v \in L^2(\Omega) \,:\, \partial^\alpha v \in L^2(\Omega) \quad \text{exists in the weak sense for all } |\alpha| \leq \ell \big\}.$$

The corresponding scalar product $(\cdot\,,\,\cdot)_{H^\ell(\Omega)}$ is given by

$$(u\,,\,v)_{H^\ell(\Omega)} = \sum_{|\alpha|\leq\ell} (\partial^\alpha u\,,\,\partial^\alpha v)_{L^2(\Omega)} \quad \text{for all } u, v \in H^\ell(\Omega),$$

which induces the norm $\|v\|^2_{H^\ell(\Omega)} := (v\,,\,v)_{H^\ell(\Omega)}$. For $\ell = 1$, the latter definition simplifies to

$$H^1(\Omega) := \big\{ v \in L^2(\Omega) \,:\, \nabla v \in L^2(\Omega)^d \quad \text{exists in the weak sense} \big\}$$

with scalar product

$$(u\,,\,v)_{H^1(\Omega)} = \int_\Omega u\, v \;\mathrm{d}x + \int_\Omega \nabla u \cdot \nabla v \;\mathrm{d}x,$$

and norm $\|u\|^2_{H^1(\Omega)} = (u\,,\,u)_{H^1(\Omega)} = \|u\|^2_{L^2(\Omega)} + \|\nabla u\|^2_{L^2(\Omega)}$. Fractional-order Sobolev spaces are defined by the K-method of interpolation, i.e., for $\ell \in \mathbb{N}_0$ and $0 < s < 1$, we define $H^{\ell+s}(\Omega) := [H^\ell(\Omega), H^{\ell+1}(\Omega)]_{s,2}$; see, e.g., [Tri95, SS11]. Further, the analysis of ABEM for the Helmholtz equation additionally requires certain Besov spaces. Since these spaces are only needed in Chapter 6, we postpone their definition to Section 6.4.

### 2.2.2 Sobolev spaces on the boundary $\partial\Omega$

We emphasize that Sobolev spaces on the boundary can be defined in different ways; see, e.g., [SS11, McL00, HW08]. Suppose that $\Gamma = \partial\Omega$ or $\emptyset \neq \Gamma \subset \partial\Omega$ is a relative open set which stems from a Lipschitz dissection $\partial\Omega = \Gamma \cup \partial\Gamma \cup (\partial\Omega \setminus \Gamma)$ (see [McL00, p. 99]). Then, for $s \in \{-1/2, 0, 1/2\}$ the Sobolev spaces $H^{1/2+s}(\partial\Omega)$ are defined as in [McL00, p. 96-99] via Bessel-potentials and local Lipschitz parametrizations of $\partial\Omega$.

If $\Gamma \subsetneqq \partial\Omega$, let $E_{0,\Gamma}$ denote the extension operator which extends a function on $\Gamma$ to $\partial\Omega$ by zero. Then, the spaces $H^{1/2+s}(\Gamma)$ and $\widetilde{H}^{1/2+s}(\Gamma)$ are defined as in [AFF$^+$17] by

$$H^{1/2+s}(\Gamma) := \{v|_\Gamma : v \in H^{1/2+s}(\partial\Omega)\},$$
$$\widetilde{H}^{1/2+s}(\Gamma) := \{v : E_{0,\Gamma} v \in H^{1/2+s}(\partial\Omega)\}.$$

The corresponding norms are given by

$$\|v\|_{H^{1/2+s}(\Gamma)} := \inf_{w \in H^{1/2+s}(\partial\Omega)} \{\|w\|_{H^{1/2+s}(\partial\Omega)} : w|_\Gamma = v\},$$
$$\|v\|_{\widetilde{H}^{1/2+s}(\Gamma)} := \|E_{0,\Gamma}\, v\|_{H^{1/2+s}(\partial\Omega)}.$$

Next, we recap some important properties of $H^{1/2+s}(\Gamma)$ and $\widetilde{H}^{1/2+s}(\Gamma)$. For further details as well as a rigorous proof, we refer to [AFF$^+$17, Facts 2.1] or [SS11, Section 2.4]. Let $\nabla_\Gamma(\cdot) : H^1(\partial\Omega) \to L^2(\partial\Omega)$ denote the usual surface gradient, i.e., for sufficiently smooth functions $u$, it holds that $\nabla_\Gamma\, u = \nabla u - (\nabla u \cdot \mathbf{n})\mathbf{n}$.

- For $s = 1/2$, there hold the following equivalences
$$\|u\|^2_{H^1(\partial\Omega)} \simeq \|u\|^2_{L^2(\partial\Omega)} + \|\nabla_\Gamma\, u\|^2_{L^2(\partial\Omega)} \qquad \text{as well as}$$
$$\|u\|^2_{\widetilde{H}^1(\Gamma)} \simeq \|u\|^2_{L^2(\Gamma)} + \|\nabla_\Gamma\, u\|^2_{L^2(\Gamma)}.$$
- For $s = 0$, the norms $\|u\|_{H^{1/2}(\partial\Omega)}$ and $\|u\|_{\widetilde{H}^{1/2}(\Gamma)}$ can be equivalently described by the Aronstein-Slobodeckii norms of $u$ and $E_{0,\Gamma}\, u$.
- For $s = 0$, the spaces $H^{1/2}(\partial\Omega)$ and $\widetilde{H}^{1/2}(\Gamma)$ can equivalently be obtained by interpolation with the $K$-method, i.e.,
$$H^{1/2}(\partial\Omega) = \big[L^2(\partial\Omega), H^1(\partial\Omega)\big]_{1/2,2} \quad \text{and} \quad \widetilde{H}^{1/2}(\Gamma) = \big[L^2(\Gamma), \widetilde{H}^1(\Gamma)\big]_{1/2,2}.$$

To simplify notation and if it is clear from the context, we identify any $v \in \widetilde{H}^{1/2+s}(\Gamma)$ with its extension $E_{0,\Gamma} v \in H^{1/2+s}(\partial\Omega)$.

### 2.2.3 Dual spaces

Let $\langle \cdot\,, \cdot \rangle_\Omega$ and $\langle \cdot\,, \cdot \rangle_{\partial\Omega}$ denote the duality pairings which extend the $L^2(\Omega)$ and $L^2(\partial\Omega)$-scalar product. For $s \in \{-1/2, 0, 1/2\}$, the negative-order Sobolev spaces on the boundary are defined by duality as

$$H^{-(1/2+s)}(\partial\Omega) := H^{1/2+s}(\partial\Omega)',$$
$$\widetilde{H}^{-(1/2+s)}(\Gamma) := H^{1/2+s}(\Gamma)',$$
$$H^{-(1/2+s)}(\Gamma) := \widetilde{H}^{1/2+s}(\Gamma)';$$

see, e.g, [AFF$^+$17]. Note that for all $\psi \in L^2(\Gamma)$, it holds that $E_{0,\Gamma}\, \psi \in H^{-1/2}(\partial\Omega)$ as well as $\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} = \|E_{0,\Gamma}\, \psi\|_{H^{-1/2}(\partial\Omega)}$; see, e.g., [AFF$^+$17]. Further, we recall the continuous inclusions

$$\widetilde{H}^{\pm(1/2+s)}(\Gamma) \subseteq H^{\pm(1/2+s)}(\Gamma) \quad \text{and} \quad \widetilde{H}^{\pm(1/2+s)}(\partial\Omega) = H^{\pm(1/2+s)}(\partial\Omega).$$

### 2.2.4 Trace operators

In this section, we recall the definitions of the interior and exterior trace operator as well as the corresponding conormal derivative. To that end, let $\Omega$ be a Lipschitz domain and $1/2 < s < 3/2$. Then there exists a linear and continuous interior trace operator

$$\gamma_0^{\text{int}} \,:\, H^s(\Omega) \to H^{s-1/2}(\partial\Omega) \quad \text{such that} \quad \gamma_0^{\text{int}}\, w = w|_{\partial\Omega} \quad \text{for all } w \in C^0(\overline{\Omega});$$

see, e.g., [SS11, Theorem 2.6.8]. Let $u \in H^1_\Delta := \{u \in H^1(\Omega) : -\Delta u \in L^2(\Omega)\}$. We define the interior conormal derivative operator via the first Green's formula as

$$\begin{aligned} \gamma_1^{\text{int}} \,:\, H^1_\Delta(\Omega) &\to H^{-1/2}(\partial\Omega) && \text{such that,} \\ \langle \gamma_1^{\text{int}} u \,,\, \gamma_0^{\text{int}} v\rangle_{\partial\Omega} &= \langle \nabla u, \nabla v\rangle_\Omega - \langle -\Delta u, v\rangle_\Omega && \text{for all } v \in H^1(\Omega); \end{aligned}$$

see, e.g., [AFF+17]. The exterior counterparts $\gamma_0^{\text{ext}}$ and $\gamma_1^{\text{ext}}$ are defined analogously as follows. Let $U \subset \mathbb{R}^d$ be a bounded Lipschitz domain such that $\overline{\Omega} \subset U \subset \mathbb{R}^d$. Then, there exists a corresponding linear and continuous exterior trace operator

$$\begin{aligned} \gamma_0^{\text{ext}} \,:\, H^s(U \setminus \overline{\Omega}) &\to H^{s-1/2}(\partial\Omega) && \text{such that,} \\ \gamma_0^{\text{ext}}\, w = w|_{\partial\Omega} & && \text{for all } w \in C^0(\overline{U \setminus \overline{\Omega}}). \end{aligned}$$

The exterior conormal derivative operator $\gamma_1^{\text{ext}} : H^1_\Delta(U \setminus \overline{\Omega}) \to H^{-1/2}(\partial\Omega)$ is defined by

$$\langle \gamma_1^{\text{ext}}\, u \,,\, \gamma_0^{\text{ext}}\, v\rangle_{\partial\Omega} := \langle \nabla u \,,\, \nabla v\rangle_{U\setminus\Omega} - \langle -\Delta u \,,\, v\rangle_{U\setminus\Omega}$$

for all $v \in H^1(U \setminus \overline{\Omega})$ with $\gamma_0^{\text{ext}} v = 0$ on $\partial U$. For bounded $C^k$-domains with $k \in \mathbb{N} \cup \infty$, the range of the trace operator can even be expanded to $1/2 < s \le k$; see [SS11, Theorem 2.6.9] or [Néd01, Chapter 4]. Moreover, accoring to [SS11, Remark 3.1.18 (d)] for piecewise $C^\infty$-boundaries $\gamma_0^{\text{int}}$ is continuous for $-1/2 < s \le s_0$ with $s_0 > 1/2$. A proof is found in [BC01, Dau88].

The trace operators give rise to the following jump terms. If a function $u$ admits an interior and an exterior trace, we define the jump

$$[u]_0 := \gamma_0^{\text{ext}} u - \gamma_0^{\text{int}} u.$$

Analogously, if $u$ admits an interior and an exterior conormal derivative, the corresponding jump is given by

$$[u]_1 := \gamma_1^{\text{ext}} u - \gamma_1^{\text{int}} u.$$

For further details on trace operators, we refer to [SS11, Section 2.6–2.7]. The trace operator also allows to incorporate boundary conditions to the function spaces. To this end, the space of $H^1(\Omega)$-functions with zero boundary data is defined by

$$H^1_0(\Omega) := \big\{ v \in H^1(\Omega) \,:\, \gamma_0^{\text{int}} v = 0 \text{ for a.e. } x \in \partial\Omega \big\}.$$

The definitions of the corresponding discrete function spaces are given in Section 3.6.

# 3 Mesh-refinement

The analysis of adaptive algorithms heavily relies on the discretization as well as on the underlying mesh-refinement strategy. To this end, Section 3.1 and Section 3.2 introduce discretizations of domains $\Omega$ and boundaries $\partial\Omega$. Section 3.3 formulates some essential properties of the underlying refinement strategy, which are sufficient to prove optimal convergence rates for certain adaptive algorithms in an abstract setting. Depending on the dimension, we focus on (simplicial) triangulations for the discretization of $\Omega \subset \mathbb{R}^d$ with $d = 2$ or $d = 3$ as well as for boundaries $\partial\Omega$ with $d = 3$. In case of a one dimensional boundary, the mesh is simply given by a partition. In Section 3.4 and 3.5, we state that the refinement axioms of Section 3.3 are satisfied for *extended 1D bisection* (EB) and *newest vertex bisection* (NVB). Finally, the discretizations give rise to corresponding discrete subspaces, which are introduced in Section 3.6.

## 3.1 Triangulations of $\Omega$

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a polygonal or polyhedral Lipschitz domain. Further, let $\mathrm{conv}(S)$ denote the convex hull of a set $S \subset \mathbb{R}^d$. The following definition gives rise to conforming meshes $\mathcal{T}^\Omega$ on a domain $\Omega$.

**Definition 3.1.** *We call a set $\mathcal{T}^\Omega$ a conforming triangulation of $\Omega$, if the following conditions are satisfied:*

*i) Each element $T \in \mathcal{T}^\Omega$ is a $(d+1)$-simplex, i.e., there exist $d+1$ affinely independent points $x_1, \dots, x_{d+1} \in \overline{\Omega}$ such that*

$$T := \mathrm{conv}\{x_1, \dots, x_{d+1}\}.$$

*The set of vertices of an element $T$ is denoted by $\mathcal{N}(T) := \{x_1, \dots, x_{d+1}\}$.*

*ii) The intersection of two elements is either empty, a joint node, a joint edge ($d \geq 2$) or a joint facet ($d = 3$), i.e., for two elements $T, T' \in \mathcal{T}^\Omega$ it holds that*

$$T \cap T' = \mathrm{conv}\big(\mathcal{N}(T) \cap \mathcal{N}(T)\big).$$

*iii) The union of all elements cover $\Omega$, i.e.,*

$$\overline{\Omega} = \bigcup_{T \in \mathcal{T}^\Omega} T.$$

Note that Definition 3.1 ii) guarantees that the triangulations do not contain any hanging nodes. Further, we define the set of all nodes $\mathcal{N}$ of a triangulation $\mathcal{T}^\Omega$ by

$$\mathcal{N}_{\mathcal{T}^\Omega} := \mathcal{N}(\mathcal{T}^\Omega) := \bigcup_{T \in \mathcal{T}^\Omega} \mathcal{N}(T).$$

For an element $T \in \mathcal{T}^\Omega$ and a set of elements $\mathcal{U} \subseteq \mathcal{T}^\Omega$, the element patch is given by

$$\omega(T) := \bigcup \{T' \in \mathcal{T}^\Omega \,:\, T' \cap T \neq \emptyset\} \subseteq \overline{\Omega} \quad \text{and} \quad \omega(\mathcal{U}) := \bigcup_{T \in \mathcal{U}} \omega(T) \subseteq \overline{\Omega}.$$

Similar, for $x \in \mathcal{N}(\mathcal{T}^\Omega)$, we define the node patch by

$$\omega(x) := \omega(\{x\}) := \bigcup \{T' \in \mathcal{T}^\Omega \,:\, T' \cap \{x\} \neq \emptyset\} \subseteq \overline{\Omega}.$$

Further, it holds that $\omega(T) = \omega(\mathcal{N}(T))$. For each mesh $\mathcal{T}^\Omega$, the (local) mesh-size function $h_{\mathcal{T}^\Omega} \in L^\infty(\mathcal{T}^\Omega)$ is denoted by

$$h_{\mathcal{T}^\Omega}(T) := h_{\mathcal{T}^\Omega}|_T := |T|^{1/d} \quad \text{for all } T \in \mathcal{T}^\Omega,$$

where $|\cdot|$ denotes the volume ($d = 3$) or the area ($d = 2$) of an element. Next, we want to measure the degeneracy of a given triangulation. This can be done by the shape regularity constant given by

$$\sigma(T) := \frac{\operatorname{diam}(T)^d}{|T|} \quad \text{with} \quad \operatorname{diam}(T) := \sup_{x,y \in T} |x - y|.$$

We call $\mathcal{T}^\Omega$ a $\gamma$-shape regular triangulation if it holds that

$$\sigma(\mathcal{T}^\Omega) := \max_{T \in \mathcal{T}^\Omega} \sigma(T) \leq \gamma.$$

## 3.2 Triangulations of $\partial\Omega$

For boundary element methods, we additionally need to define regular triangulations of the boundary $\partial\Omega$. Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with piecewise $C^\infty$-boundary $\partial\Omega$; see, e.g., [SS11, Definition 2.2.10]. We suppose that $\Gamma = \partial\Omega$ or $\emptyset \neq \Gamma \subset \partial\Omega$ is a relative open set which stems from a Lipschitz dissection $\partial\Omega = \Gamma \cup \partial\Gamma \cup (\partial\Omega \setminus \Gamma)$; see, e.g., [McL00, p. 99]. Let $T_{\mathrm{ref}}$ denote the reference element defined by

$$T_{\mathrm{ref}} := \Big\{x \in \mathbb{R}^{d-1} : 0 \leq x_1, \dots, x_{d-1} \leq 1 \text{ and } \sum_{j=1}^{d-1} x_j \leq 1\Big\},$$

i.e., $T_{\mathrm{ref}} = (0, 1)$ is the open unit interval for $d = 2$ or $T_{\mathrm{ref}} = \operatorname{conv}\{(0, 0), (1, 0), (0, 1)\}$ is the open Kuhn simplex for $d = 3$. Analogously to Definition 3.1, we introduce conforming triangulations on the boundary $\Gamma$ as follows. A similar definition is also found in [SS11, Section 4.1.2].

**Definition 3.2.** *We call a set $\mathcal{T}^\Gamma$ a regular triangulation of $\Gamma$ if the following conditions are satisfied:*

*i) Each $T \in \mathcal{T}^\Gamma$ is a subset of $\Gamma$ and there exists a corresponding bijective element map $g_T \in C^\infty(T_{ref}, T)$ such that $g_T(T_{ref}) = T$. The set of nodes is given by $\mathcal{N}(T) := g_T(\mathcal{N}(T_{ref}))$.*

*ii) For all $T, T' \in \mathcal{T}^\Gamma$, the intersection $T \cap T'$ is either empty, a joint node ($d \geq 2$), or a joint facet ($d = 3$).*

*iii) The union of all elements cover $\Gamma$, i.e.,*

$$\overline{\Gamma} = \bigcup_{T \in \mathcal{T}^\Gamma} T.$$

*iv) In the case of $d = 3$, there holds the following: If $T \cap T'$ is a facet, there exist facets $f, f' \subseteq \partial T_{ref}$ of $T_{ref}$, such that $T \cap T' = g_T(f) = g_{T'}(f')$, and the composition $g_T^{-1} \circ g_{T'} : f' \to f$ is even affine.*

The definitions of the set of nodes and the element patches, are verbatim to Section 3.1. Similarly, the (local) mesh-size function $h_{\mathcal{T}^\Gamma} \in L^\infty(\mathcal{T})$ is given by

$$h_{\mathcal{T}^\Gamma}(T) := h_{\mathcal{T}^\Gamma}|_T := |T|^{1/(d-1)},$$

where $|\cdot|$ denotes the surface measure of an element. To introduce shape regularity, let $G_T(x) := D\, g_T(x)^\intercal D\, g_T(x) \in \mathbb{R}^{(d-1)\times(d-1)}$ be the symmetric Gramian matrix of $g_T$ and $\lambda_{\min}(G_T(x))$ and $\lambda_{\max}(G_T(x))$ its extremal eigenvalues. We call $\mathcal{T}^\Gamma$ a $\gamma$-shape regular triangulation, if the following holds:

- For all $T \in \mathcal{T}^\Gamma$, the corresponding element maps $g_T(\cdot)$ satisfy that

$$\sigma(T) := \sup_{x \in T_{\mathrm{ref}}} \Big( \frac{h_{\mathcal{T}^\Gamma}(T)^2}{\lambda_{\min}(G_T(x))} + \frac{\lambda_{\max}(G_T(x))}{h_{\mathcal{T}^\Gamma}(T)^2} \Big) \leq \gamma. \tag{3.1}$$

- If $d = 2$, it additionally holds that

$$\widetilde{\sigma}(\mathcal{T}^\Gamma) := \max_{\substack{T, T' \in \mathcal{T}^\Gamma \\ T' \cap T \neq \emptyset}} \frac{|T|}{|T'|} \leq \gamma. \tag{3.2}$$

Note that the Gramian matrix $G_T(x)$ is symmetric and positive definite. This implies that $0 \leq \lambda_{\min}(G_T) \leq \lambda_{\max}(G_T)$ and hence, $\sigma(T) \geq 0$. The additional assumption for $d = 2$ ensures that the mesh-size of neighboring elements remains comparable.

If $\Gamma$ is the union of $(d-1)$-dimensional hyperplanes, e.g., when $\Gamma \subseteq \partial\Omega$ with a polyhedral domain $\Omega \subset \mathbb{R}^d$, all element maps $g_T(\cdot)$ are affine. The corresponding Gramian matrix $G_T$ of each element map $g_T(\cdot)$ is constant and hence, the latter definition generalizes the concept of $\gamma$-shape regularity on $C^\infty$ boundaries $\Gamma$. For further details, we refer to [AFF+17, SS11].

## 3.3 Axioms for the mesh-refinement

The next lemma recaps some important properties of $\gamma$-shape regular meshes. A proof for boundary meshes is, e.g., found in [AFF$^+$17, Lemma 2.6].

**Lemma 3.3.** *Let $\mathcal{T} \in \{\mathcal{T}^\Omega, \mathcal{T}^\Gamma\}$ be a $\gamma$-shape regular triangulation. Then, there exists a constant $C_{\rm mesh} > 0$, depending only on $\gamma$ and in case of a boundary mesh, additionally on the Lipschitz parametrization of $\partial\Omega$, such that the following assertions i)–ii) hold:*

*i) For all $T, T' \in \mathcal{T}$ such that $T \cap T \neq \emptyset$, it holds that $h_\mathcal{T}(T) \le C_{\rm mesh} h_\mathcal{T}(T')$.*

*ii) The number of elements in an element patch is bounded by $C_{\rm mesh}$, i.e., $\#\omega(T) \le C_{\rm mesh}$ for all $T \in \mathcal{T}$.*

*iii) It holds that $\max_{T\in\mathcal{T}} \frac{\operatorname{diam}(T)}{h_\mathcal{T}} \le C_{\rm mesh}$.* □

To abbreviate the notation, we introduce the following convention. If (discrete) quantities are related to some triangulation, this is explicitly stated by use of appropriate indices, e.g., $h_\bullet$ is the local mesh-size function to the triangulation $\mathcal{T}_\bullet$, $v_\bullet$ is a generic discrete function in the corresponding discrete space $\mathcal{X}_\bullet$, and $\eta_\ell(\cdot)$ is the error estimator with respect to the triangulation $\mathcal{T}_\ell$.

From now on, suppose that $\mathcal{T}_\bullet \in \{\mathcal{T}^\Omega, \mathcal{T}^\Gamma\}$ is a given regular and $\gamma$-shape regular triangulation. Further, suppose that refine$(\cdot)$ is a fixed mesh-refinement strategy, such that given a conforming triangulation $\mathcal{T}_\bullet$ and $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$, the call $\mathcal{T}_\circ = \operatorname{refine}(\mathcal{T}_\bullet, \mathcal{M}_\bullet)$ returns the coarsest conforming refinement $\mathcal{T}_\circ$ of $\mathcal{T}_\bullet$ such that all $T \in \mathcal{M}_\bullet$ have been refined, i.e.,

- $\mathcal{T}_\circ$ is a conforming triangulation of $\Omega$ or $\Gamma$,
- for all $T \in \mathcal{T}_\bullet$, it holds $T = \bigcup \{T' \in \mathcal{T}_\circ \,:\, T' \subseteq T\}$,
- $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet \backslash \mathcal{T}_\circ$,
- the number of elements $\#\mathcal{T}_\circ$ is minimal amongst all other triangulations $\mathcal{T}'$ which share the three foregoing properties.

Furthermore, we write $\mathcal{T}_\circ \in \operatorname{refine}(\mathcal{T}_\bullet)$, if $\mathcal{T}_\circ$ is obtained by a finite number of refinement steps, i.e., there exists $n \in \mathbb{N}_0$ as well as a finite sequence $\mathcal{T}^{(0)}, \dots, \mathcal{T}^{(n)}$ of triangulations with corresponding sets $\mathcal{M}^{(j)} \subseteq \mathcal{T}^{(j)}$ such that

- $\mathcal{T}_\bullet = \mathcal{T}^{(0)}$,
- $\mathcal{T}^{(j+1)} = \operatorname{refine}(\mathcal{T}^{(j)}, \mathcal{M}^{(j)})$ for all $j = 0, \dots, n-1$,
- $\mathcal{T}_\circ = \mathcal{T}^{(n)}$.

In particular, it holds that $\mathcal{T}_\bullet \in \operatorname{refine}(\mathcal{T}_\bullet)$. Suppose that $\mathcal{T}_0$ is a given regular and $\gamma$-shape regular initial triangulation. To abbreviate notation, we define the set of all possible triangulations which can be obtained by refining $\mathcal{T}_0$ as $\mathbb{T} := \operatorname{refine}(\mathcal{T}_0)$. The analysis of optimal convergence rates for adaptive algorithms heavily relies on the underlying mesh-refinement strategy. To deal with this dependency, we formulate the following refinement axioms.

**R1) local mesh-size reduction:** Refinement yields a contraction of the local mesh-size function on refined elements, i.e., there exits $0 < q_{\rm mesh} < 1$ such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ implies that

$$h_\circ|_T \le q_{\rm mesh}\, h_\bullet|_T \quad \text{for all } T \in \mathcal{T}_\bullet \setminus \mathcal{T}_\circ. \tag{3.3}$$

**R2) uniform $\gamma$-shape regularity:** There exists a constant $\gamma_{\rm ref} > 0$ such that for all triangulations $\mathcal{T}_\bullet \in \mathbb{T}$ it holds that

$$\sigma(\mathcal{T}_\bullet) = \max_{T\in\mathcal{T}_\bullet} \sigma(T) \le \gamma_{\rm ref} \quad \text{and} \quad \widetilde{\sigma}(\mathcal{T}_\bullet) \le \gamma_{\rm ref} \quad \text{if } d = 2 \text{ and } \bigcup \mathcal{T}_\bullet \subseteq \Gamma. \tag{3.4}$$

**R3) splitting property:** Each refined element is split in at least 2 and at most $C_{\rm son} \ge 2$ many sons, i.e., for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$, the refined triangulation $\mathcal{T}_\circ = \mathrm{refine}(\mathcal{T}_\bullet, \mathcal{M}_\bullet)$ satisfies that

$$\#(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ) + \#\mathcal{T}_\bullet \le \#\mathcal{T}_\circ \le C_{\rm son}\, \#(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ) + \#(\mathcal{T}_\bullet \cap \mathcal{T}_\circ). \tag{3.5}$$

**R4) 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 $\mathcal{T}_\bullet \oplus \mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet) \cap \mathrm{refine}(\mathcal{T}_\circ) \subseteq \mathrm{refine}(\mathcal{T})$ which satisfies that

$$\#(\mathcal{T}_\bullet \oplus \mathcal{T}_\circ) \le \#\mathcal{T}_\bullet + \#\mathcal{T}_\circ - \#\mathcal{T}. \tag{3.6}$$

**R5) mesh-closure estimate:** There exists $C_{\rm mesh} > 0$ such that for all sequences $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of successively refined meshes, i.e., $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ with sets of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$, it holds that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 \le C_{\rm mesh} \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \quad \text{for all } \ell \in \mathbb{N}. \tag{3.7}$$

**R6) permutability of refinement steps:** For sequences $(\mathcal{M}_\ell)_{\ell=1}^n$ and $(\widehat{\mathcal{M}}_\ell)_{\ell=1}^m$ of marked elements with $\mathcal{T}_j = \mathrm{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1})$ for all $j = 1, \dots, n$ as well as $\widehat{\mathcal{T}}_j = \mathrm{refine}(\widehat{\mathcal{T}}_{j-1}, \widehat{\mathcal{M}}_{j-1})$ for all $j = 1, \dots, m$, it holds that

$$\bigcup_{j=0}^{n} \mathcal{M}_j = \bigcup_{j=0}^{m} \widehat{\mathcal{M}}_j \quad \Longrightarrow \quad \mathcal{T}_n = \widehat{\mathcal{T}}_m. \tag{3.8}$$

We emphasize that (R6) extends the refinement axioms in [CFPP14], but is necessary for the analysis of compactly perturbed problems in Chapter 4.

## 3.4 Extended 1D bisection

First, we consider extended 1D bisection (EB) for refining meshes on a 1-dimensional boundary $\Gamma \subseteq \partial\Omega$ with $\Omega \subset \mathbb{R}^2$. The algorithm is well known and used, e.g., in [FLP08, AFF$^+$13, EFLFP09, EFGP13]. For sake of completeness we include the formulation of the algorithm from [AFF$^+$13].

**Algorithm 3.4** (EB)**.** Input: *Mesh* $\mathcal{T}_\bullet$, *set of marked elements* $\mathcal{M}_\bullet^{(0)} := \mathcal{M}_\bullet \subseteq \mathcal{T}_\bullet$ *and* $i := 0$.

(i) *Define* $\mathcal{U}^{(i)} := \bigcup_{T \in \mathcal{M}_\bullet^{(i)}} \{T' \in \mathcal{T}_\bullet \setminus \mathcal{M}_\bullet^{(i)} \,:\, T \cap T' \neq \emptyset$ *and* $h_\bullet|_{T'} > \sigma(\mathcal{T}_0)\, h_\bullet|_T\}$.

(ii) `If` $\mathcal{U}^{(i)} \neq \emptyset$, *define* $\mathcal{M}_\bullet^{(i+1)} := \mathcal{M}_\bullet^{(i)} \cup \mathcal{U}^{(i)}$, *increase* $i \to i+1$, *and goto* (i).

(iii) `If` $\mathcal{U}^{(i)} = \emptyset$, *bisect all marked elements* $T \in \mathcal{M}_\bullet^{(i)}$ *to obtain* $\mathcal{T}_\circ$.

Output: *Refined mesh* $\mathcal{T}_\circ$, *with* $\mathcal{T}_\bullet \setminus \mathcal{T}_\circ = \mathcal{M}_\bullet^{(i)}$.

In addition to simple bisection of the marked elements, Step (i) and Step (ii) of Algorithm 3.4 guarantee that the ratio of the local mesh-size of neighboring elements remains bounded. In particular, this implies (R2).

### 3.4.1 Verification of the axioms

First, (R1) is shown in Lemma 3.5 (Section 3.5.1). The remaining properties of EB are proved in [AFF+13, Theorem 2.3]. [AFF+13, Theorem 2.3 (i)] guarantees that $\sigma(\mathcal{T}_\bullet) \leq 2\sigma(\mathcal{T}_0)$ for all $\mathcal{T}_\bullet \in \mathbb{T}$. This implies (R2) with $\gamma := 2\,\sigma(\mathcal{T}_0)$. Since EB uses bisection, (R3) follows directly from the definition of the refinement strategy with $C_{\rm son} = 2$. Further, the proofs of axiom (R4) and (R5) are found in of [AFF+13, Theorem 2.3 (ii)–(iii)]. Note that NVB is a binary refinement rule. Hence, the order of refinement does not matter which implies (R6).

## 3.5 Newest Vertex bisection

As second mesh-refinement strategy, we discuss *newest vertex bisection* (NVB); see, e.g., [Ste07] for $d = 2$ and [Ste08b] for $d = 3$. We use 2D NVB for refining triangulations of $\Omega \subset \mathbb{R}^2$ as well as for surface triangulations on $\Gamma \subseteq \partial\Omega$ with $\Omega \subset \mathbb{R}^3$. Further, 3D NVB is used for refinement of simplicial triangulations on $\Omega \subset \mathbb{R}^3$. A heuristic for refinement of an element with 2D NVB is illustrated in Figure 3.1. For an exact formulation we refer to [Ste07] or [Ste08b] in the 3D case.

### 3.5.1 Verification of the axioms

First, we prove (R1) for surface triangulations of $\Gamma \subseteq \partial\Omega$ where $\Omega \subseteq \mathbb{R}^d$ with $d = 2, 3$. Note that in case of meshes on the boundary, the bisection of an element with EB or NVB is understood in the following way. Let $T \in \mathcal{T}_\bullet$ be a marked element with element map $g_T(T_{\rm ref}) = T$. Bisection of the reference element $T_{\rm ref}$ (in case of a triangle according to 2D-NVB (Figure 3.1)) produces sons $T_{\rm ref}^1, \dots, T_{\rm ref}^k \subseteq T_{\rm ref}$. Then, with the element map $g_T$, we obtain two sons $T_1, \dots, T_n \subset T$ with $T = \bigcup_{i=1}^k T_i$ and $T_i = g_T(T_{\rm ref}^i)$ for all $i = 1, \dots, k$. We obtain the following lemma.

**Lemma 3.5.** *There exist* $0 < q_{\rm mesh} < 1$, *such that for all* $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$ *with* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ *it holds that* $h_\circ|_T \leq q_{\rm mesh}\, h_\bullet|_T$ *on all* $T \in \mathcal{T}_\bullet \setminus \mathcal{T}_\circ$. *In particular, there holds reduction of the local mesh-size* (R1).

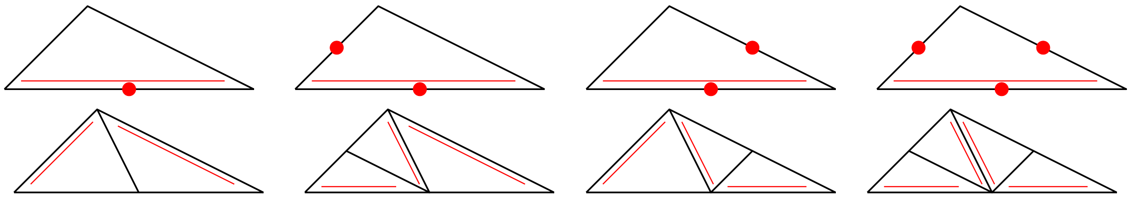


Figure 3.1: For each triangle $T \in \mathcal{T}_\bullet$, there is one fixed *reference edge*, indicated by the double line (top left). If $T$ is marked for refinement, we mark its reference edge. Then, refinement of $T$ is done by bisecting the reference edge, where its midpoint becomes a new vertex of the refined triangulation $\mathcal{T}_\circ$. The reference edges of the son triangles are opposite to this newest vertex (bottom left). To avoid hanging nodes, one proceeds as follows: We assume that certain edges of $T$, but at least the reference edge, are marked for refinement (top). Using iterated newest vertex bisection, the element is then split into 2, 3, or 4 son triangles (bottom).

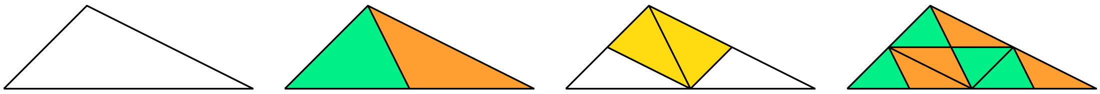


Figure 3.2: Newest vertex bisection does only lead (up to similarity) to a finite number of triangles. Above, the different colors represent similarity classes. Starting with a triangle (left), iterative use of NVB does only create (up to similarity) new triangles in the first two steps (mid left and mid right). Hence in following steps, no new similarity classes are generated.

*Proof.* To prove the lemma, we argue by contradiction. To that end, let $(\mathcal{T}_\bullet^n)_{n\in\mathbb{N}}, (\mathcal{T}_\circ^n)_{n\in\mathbb{N}} \subset \mathbb{T}$, be sequences of refinements with $\mathcal{T}_\circ^n \in \operatorname{refine}(\mathcal{T}_\bullet^n)$ and elements $T_\bullet^n \in \mathcal{T}_\bullet^n \setminus \mathcal{T}_\circ^n$ as well as $T_\circ^n \in \mathcal{T}_\circ^n \setminus \mathcal{T}_\bullet^n$ such that

$$T_\circ^n \subsetneqq T_\bullet^n \quad \text{as well as} \quad \frac{|T_\bullet^n|}{|T_\circ^n|} \to 1 \quad \text{as } n \to \infty.$$

This implies that $|T_\bullet^n \setminus T_\circ^n|/|T_\bullet^n| \to 0$ as $n \to \infty$. Further, for all $n \in \mathbb{N}$ there exists $T \in \mathcal{T}_0$ such $T_\circ^n \subsetneqq T_\bullet^n \subseteq T$. We obtain a corresponding sequence $\widetilde{T}_\circ^n \subsetneqq \widetilde{T}_\bullet^n \subseteq T_{\mathrm{ref}}$ with $g_T(\widetilde{T}_\bullet^n) = T_\bullet^n$ as well as $g_T(\widetilde{T}_\circ^n) = T_\circ^n$. Since bisection is done at first on the reference element, it holds that $|\widetilde{T}_\circ^n| \le 1/2\,|\widetilde{T}_\bullet^n|$ for all $n \in \mathbb{N}_0$. Then, $\gamma$-shape regularity implies that $\det G_T(x) \simeq (h_\bullet(T))^{2(d-1)} = |T|^2$ for all $x \in T_{\mathrm{ref}}$. This reveals the contradiction

$$\frac{1}{2} \le \frac{|\widetilde{T}_\bullet^n \setminus \widetilde{T}_\circ^n|}{|\widetilde{T}_\bullet^n|} \simeq \frac{\int_{\widetilde{T}_\bullet^n \setminus \widetilde{T}_\circ^n} |\det G_T(t)|^{1/2}\,\mathrm{d}t}{\int_{\widetilde{T}_\bullet^n} |\det G_T(t)|^{1/2}\,\mathrm{d}t} = \frac{|T_\bullet^n \setminus T_\circ^n|}{|T_\bullet^n|} \xrightarrow{n\to\infty} 0,$$

and concludes the proof. $\square$

In case of meshes on domains $\Omega$ for $d \ge 2$ there holds (3.3) with $q_{\mathrm{mesh}} = 2^{-1/d}$ and hence (R1). For a proof we refer to [CKNS08, Ste07]. Figure 3.2 illustrates for $d = 2$, that (up to similarity) only a finite number of triangles can be constructed during refinement of an initial mesh $\mathcal{T}_0$. A similar result for $d = 3$ and a rigorous proof of (R2) for arbitrary $d \geq 2$ can be found in [Ste08b].

The splitting property (R3) is proved in [GSS14] with $2 \leq C_{\rm son} < \infty$, where $C_{\rm son}$ depends only on $\mathcal{T}_0$ and $d$. For $d = 2$, it holds $C_{\rm son} = 4$ (see Figure 3.1).

The proof of the overlay estimate (R4) is found in [CKNS08, Ste07]. The mesh-closure estimate (R5) has first been proved for $d = 2$ in [BDD04] and later for $d \geq 2$ in [Ste08b]. Both works [BDD04, Ste08b] require an additional admissibility assumption on the initial mesh $\mathcal{T}_0$. While for $d = 2$, [BDD04, Section 2.2] gives a proof that every conforming mesh admits a labeling such that the admissibility condition is satisfied, a result for $d = 3$ is still missing. On the other hand, [KPP13] shows that the admissibility condition is unnecessary for $d = 2$.

Note that NVB is a binary refinement rule. Hence, the order of refinement does not matter which implies (R6), see [Ste08b].

### 3.5.2 Other refinement strategies

Red-green-blue refinement (see, e.g., [Ver13]), fails (R4) even for $d = 2$; see [Pav10, Satz 4.15] (in German) or [Fei15] for a counterexample. For red-refinement with first-order hanging nodes, the validity of (R3)–(R4) is shown in [BN10]. For mesh-refinement strategies in isogeometric analysis, we refer to [MP15] for T-splines and to [BGMP16, GHP17] for (truncated) hierarchical B-splines. A rigorous proof of (R1)–(R6) for (truncated) hierarchical B-splines is also found in [Gan17a]. For further details on mesh-refinement strategies which satisfy (R3)–(R4), we refer to [BN10, MP15, Fei15] and to the discussion in [CFPP14, Section 2.5].

## 3.6 Discrete function spaces

In this section, we introduce discrete function spaces on domains $\Omega$ and boundaries $\Gamma$. To this end, let $\mathcal{T}_\bullet^\Omega$ be a regular triangulation of $\Omega$. For some fixed polynomial degree $p \geq 1$, let

$$\mathcal{S}^p_\Omega(\mathcal{T}_\bullet^\Omega) := \big\{ V_\bullet \in C(\Omega) \,:\, \forall T \in \mathcal{T}_\bullet^\Omega \quad V_\bullet|_T \text{ is a polynomial of degree } \leq p \big\}$$

be the usual finite element space of globally continuous piecewise polynomials with the inclusion $\mathcal{S}^p_\Omega(\mathcal{T}_\bullet^\Omega) \subset H^1(\Omega)$. The corresponding conforming subspace of $H^1_0(\Omega)$ will be denoted by $\mathcal{S}^p_{\Omega,0}(\mathcal{T}_\bullet^\Omega) := \mathcal{S}^p_\Omega(\mathcal{T}_\bullet^\Omega) \cap H^1_0(\Omega)$.

In case of discrete spaces on the boundary, let $\mathcal{T}_\bullet^\Gamma$ be a regular triangulation of $\Gamma$. For a fixed $p \geq 0$, we define the space of (discontinuous) piecewise polynomials by

$$\mathcal{P}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) := \big\{ \Psi_\bullet \in L^\infty(\Gamma) : \ \forall T \in \mathcal{T}_\bullet^\Gamma, \quad \Psi_\bullet \circ g_T \text{ is a polynomial of degree } \leq p \big\}.$$

Further, let $\mathcal{S}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) := \mathcal{P}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) \cap H^1(\Gamma)$ resp. $\widetilde{\mathcal{S}}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) := \mathcal{P}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) \cap \widetilde{H}^1(\Gamma)$ be the space of continuous piecewise polynomials. We emphasize the following (compact) inclusions

$$\mathcal{P}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) \subset L^2(\Gamma) \subset \widetilde{H}^{-1/2}(\Gamma) \quad \text{and} \quad \widetilde{\mathcal{S}}^p_\Gamma(\mathcal{T}_\bullet^\Gamma) \subset \widetilde{H}^1(\Gamma) \subset \widetilde{H}^{1/2}(\Gamma); \tag{3.9}$$

see, e.g., [AFF+17]. In the case of $\Gamma = \partial\Omega$, it even holds that $\widetilde{\mathcal{S}}^p_\Gamma(\mathcal{T}^\Gamma_\bullet) = \mathcal{S}^p_\Gamma(\mathcal{T}^\Gamma_\bullet) \subset H^1(\Gamma)$. To shorten notation and if its clear from the context, we further omit the additional index and write $\mathcal{S}^p(\mathcal{T}_\bullet)$ instead of $\mathcal{S}^p_\Omega(\mathcal{T}^\Omega_\bullet)$ or $\mathcal{S}^p_\Gamma(\mathcal{T}^\Gamma_\bullet)$ as well as $\mathcal{S}^p_0(\mathcal{T}_\bullet)$ instead of $\mathcal{S}^p_{\Omega,0}(\mathcal{T}^\Omega_\bullet)$.

# 4 Abstract theory for compactly perturbed problems

## 4.1 State of the art and outline

In recent years, the analysis of convergence and optimal convergence behavior of adaptive finite element methods (AFEM) as well as adaptive boundary element methods (ABEM) has matured. We refer to [Dör96, MNS00, BDD04, Ste07, CKNS08, FFP14, CFPP14] for some milestones for AFEM and the works [FFK$^+$14, FFK$^+$15, FKMP13, AFF$^+$17, Gan13] for ABEM. In a more general case of compactly perturbed elliptic problems, existing results have the limitation that the initial mesh has to be sufficiently fine see, e.g., [MN05, CN12, FFP14] for AFEM. In the mentioned references, only problems satisfying the Lax–Milgram theorem have been treated. On the other hand, numerical examples in the engineering literature suggest that adaptive mesh-refinement for finite element methods performs well even if the initial mesh is coarse; see, e.g., [SH96, BI98, BI99] in the case AFEM of the Helmholtz equation.

In this chapter, we introduce a complete abstract setting in the spirit of [CFPP14]. This allows to formulate an adaptive Algorithm 4.4 and prove convergence with optimal algebraic rate, without limitations on the initial mesh. The abstract framework is based on some essential properties of the underlying error estimator and mesh-refinement, and covers in particular AFEM as well as ABEM.

Given an initial triangulation $\mathcal{T}_0$, a typical adaptive algorithm (1.2) generates a sequence of refined meshes $\mathcal{T}_\ell$ with corresponding nested spaces $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{H}$ for all $\ell \geq 0$. We stress that unlike the prior works [MN05, CN12, FFP14], Algorithm 4.4 will not be given any information on whether the current mesh is sufficiently fine to allow for a unique solution. In particular, we do not assume that the given initial mesh $\mathcal{T}_0$ and, in fact, any adaptive mesh $\mathcal{T}_\ell$ generated by Algorithm 4.4 is sufficiently fine. Independent of the missing *a priori* information, we derive similar results as for uniformly elliptic problems, see, e.g., [CKNS08, FFP14, CFPP14] and the references therein. Following some ideas of [FFP14], we prove convergence (Proposition 4.9), linear convergence (Theorem 4.14), and optimal algebraic convergence rates (Theorem 4.21). The main results of this chapter are based on the recent own article [BHP17].

**Outline of chapter.** Section 4.2 introduces the model problem and provides the functional analytic framework of the *a posteriori* analysis. The underlying error estimator and the precise formulation of the adaptive algorithm are given in Section 4.3–4.4. Section 4.5 adapts [CFPP14] to the present setting and formulates the necessary estimator axioms. Utilizing the estimator as well as the refinement axioms from Chapter 3, Section 4.6 gives a first (plain-) convergence result for the adaptive algorithm. Linear convergence is shown in Section 4.7, which also addresses the validity of the Céa lemma. In Section 4.8.1, we introduce approximation classes with respect to the error estimator in the spirit of [CFPP14] and discuss their connection to other definitions using the total error (see e.g., [CKNS08]). Finally, we conclude the chapter with the proof of optimal algebraic convergence rates in Section 4.8.

## 4.2 Abstract setting

The model problem is set in the following abstract framework. Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with polyhedral, or in case of boundary elements, piecewise $C^\infty$-boundary $\partial\Omega$. Let $\mathcal{H}$ denote a separable Hilbert space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with norm $\|\cdot\|_{\mathcal{H}}$. Further, suppose that $a(\cdot\,,\,\cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ is a hermitian, continuous, and elliptic sesquilinear form on $\mathcal{H}$, i.e., there exists some constant $\alpha > 0$ such that

$$\alpha \, \|v\|_{\mathcal{H}}^2 \le a(v\,,\,v) \quad \text{for all } v \in \mathcal{H}. \tag{4.1}$$

Since the sesquilinear form $a(\cdot\,,\,\cdot)$ is elliptic, the $a(\cdot\,,\,\cdot)$-induced energy norm $|\!|\!| \, v \, |\!|\!|^2 := a(v\,,\,v)$ is an equivalent norm on $\mathcal{H}$, i.e., $|\!|\!| \, v \, |\!|\!| \simeq \|v\|_{\mathcal{H}}$ for all $v \in \mathcal{H}$.

Let $\mathcal{H}^*$ denote the dual space of $\mathcal{H}$, and let $\langle \cdot\,,\,\cdot \rangle$ denote the corresponding duality pairing. Suppose that $\mathfrak{C} : \mathcal{H} \to \mathcal{H}^*$ is a compact linear operator and $f \in \mathcal{H}^*$. We consider the following weak model problem: Given $f \in \mathcal{H}^*$ find $u \in \mathcal{H}$ such that

$$b(u\,,\,v) := a(u\,,\,v) + \langle \mathfrak{C}u\,,\,v \rangle = \langle f\,,\,v \rangle \quad \text{for all } v \in \mathcal{H}. \tag{4.2}$$

We suppose that (4.2) admits a unique solution which is usually proved by the Fredholm alternative. Let $\mathcal{T}_0$ be a given regular initial mesh and suppose that refine$(\cdot)$ is a fixed refinement strategy satisfying the refinement axioms (R1)–(R6) of Chapter 3. For each triangulation $\mathcal{T}_\bullet \in \mathbb{T} := \text{refine}(\mathcal{T}_0)$, let $\mathcal{X}_\bullet \subseteq \mathcal{H}$ denote the corresponding conforming finite-dimensional subspace. Further, suppose that refinement of the underlying meshes $\mathcal{T}_\circ \in \text{refine}(\mathcal{T}_\bullet)$ leads to nestedness $\mathcal{X}_\bullet \subseteq \mathcal{X}_\circ$ of the corresponding subspaces. Then, the Galerkin formulation of (4.2) reads as: Given $f \in \mathcal{H}^*$, find $U_\bullet \in \mathcal{X}_\bullet$, such that

$$b(U_\bullet, V_\bullet) = \langle f\,,\,V_\bullet \rangle \quad \text{for all } V_\bullet \in \mathcal{X}_\bullet. \tag{4.3}$$

We additionally assume that iterated uniform mesh-refinement leads to a dense subspace of $\mathcal{H}$, i.e., for $\widehat{\mathcal{T}}_0 := \mathcal{T}_0$ and the inductively defined sequence $\widehat{\mathcal{T}}_{\ell+1} := \text{refine}(\widehat{\mathcal{T}}_\ell, \widehat{\mathcal{M}}_\ell)$ with $\widehat{\mathcal{M}}_\ell \subseteq \widehat{\mathcal{T}}_\ell$ for all $\ell \in \mathbb{N}_0$, it holds the following: If $\#\big\{\ell \in \mathbb{N}_0 \,:\, \widehat{\mathcal{M}}_\ell = \widehat{\mathcal{T}}_\ell\big\} = \infty$ (i.e., there are infinitely many steps that perform uniform refinement), then $\mathcal{H} = \overline{\bigcup_{\ell=0}^\infty \widehat{\mathcal{X}}_\ell}$. Note that the latter assumption is satisfied in most generic situations and can be guaranteed by use of suitable discrete spaces, see e.g., Chapter 5–6.

### 4.2.1 Existence of discrete solutions

In general, (4.3) may fail to allow for a (unique) solution $U_\bullet \in \mathcal{X}_\bullet$. However, existence and uniqueness are guaranteed if the corresponding mesh $\mathcal{T}_\bullet$ is sufficiently fine (see Corollary 4.2), e.g.,

$$\|h_\bullet\|_{L^\infty(\Omega)} \le H \ll 1.$$

Therefore, if (4.3) does not allow for a unique solution $U_\bullet \in \mathcal{X}_\bullet$, we employ one step of uniform refinement.

The next proposition is an improved version of [SS11, Theorem 4.2.9]. It guarantees the existence and uniqueness of discrete solutions $U_\bullet \in \mathcal{X}_\bullet$ for sufficiently fine meshes $\mathcal{T}_\bullet$. Even though the result appears to be well-known, we did not find the precise statement in the literature. We note that a similar result can be found in [BS08, Theorem 5.7.6], where additional regularity assumptions for the dual problem are required. Instead, our proof below proceeds without considering the dual problem, and hence no additional regularity assumptions are needed.

**Proposition 4.1.** *Suppose well-posedness of* (4.2)*, i.e.,*

$$\forall w \in \mathcal{H} \quad \big[w = 0 \quad \Longleftrightarrow \quad \big(\forall v \in \mathcal{H} \quad b(w\,,\,v) = 0\big)\big]. \tag{4.4}$$

*Suppose that $(\mathcal{X}_\ell)_{\ell\in\mathbb{N}_0}$ is a dense sequence of discrete subspaces $\mathcal{X}_\ell \subset \mathcal{H}$, i.e.,*

$$\lim_{\ell\to\infty} \min_{V_\ell\in\mathcal{X}_\ell} \|v - V_\ell\|_{\mathcal{H}} = 0 \quad \textit{for all } v \in \mathcal{H}. \tag{4.5}$$

*Then, there exists some index $\ell_\bullet \in \mathbb{N}_0$ such that for all discrete subspaces $\mathcal{X}_\bullet \subset \mathcal{H}$ with $\mathcal{X}_\bullet \supseteq \mathcal{X}_{\ell_\bullet}$, the following holds: There exists $\beta > 0$ which depends only on $\mathcal{X}_{\ell_\bullet}$, such that the* inf-sup *constant of $\mathcal{X}_\bullet$ is uniformly bounded from below, i.e.,*

$$\beta_\bullet := \inf_{W_\bullet\in\mathcal{X}_\bullet\backslash\{0\}} \sup_{V_\bullet\in\mathcal{X}_\bullet\backslash\{0\}} \frac{|b(W_\bullet\,,\,V_\bullet))|}{\|W_\bullet\|_{\mathcal{H}}\|V_\bullet\|_{\mathcal{H}}} \geq \beta > 0. \tag{4.6}$$

*In particular, the discrete formulation* (4.3) *admits a unique solution $U_\bullet \in \mathcal{X}_\bullet$. Moreover, there holds uniform validity of the Céa lemma, i.e., there is a constant $C > 0$ which depends only on $b(\cdot,\cdot)$ and $\beta$ but not on $\mathcal{X}_\bullet$, such that*

$$\|u - U_\bullet\|_{\mathcal{H}} \leq C \min_{V_\bullet\in\mathcal{X}_\bullet} \|u - V_\bullet\|_{\mathcal{H}}. \tag{4.7}$$

*If the spaces $\mathcal{X}_\ell$ are nested, i.e., $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ for all $\ell \in \mathbb{N}_0$, the latter guarantees convergence $\|u - U_\ell\|_{\mathcal{H}} \to 0$ as $\ell \to \infty$.*

*Proof.* Suppose well-posedness (4.4) of (4.2). Ellipticity (4.1) of $a(\cdot\,,\,\cdot)$ combined with the Fredholm alternative imply existence and uniqueness of a solution $u$ of (4.2) for all $f \in \mathcal{H}^*$; see [SS11, Theorem 4.2.7]. Let $\mathcal{X}_\bullet$ be an arbitrary discrete subspace of $\mathcal{H}$ with dual space $\mathcal{X}_\bullet^*$. Then, the bilinear form $b(\cdot\,,\,\cdot)$ induces the linear and continuous operator

$$\mathfrak{B}_\bullet : \mathcal{X}_\bullet \to \mathcal{X}_\bullet^*, \qquad \langle \mathfrak{B}_\bullet W_\bullet\,,\,V_\bullet\rangle \;:=\; b(W_\bullet\,,\,V_\bullet) \quad \text{for all } V_\bullet, W_\bullet \in \mathcal{X}_\bullet.$$

For convenience of the reader, we split the remainder of the proof into four steps.

**Step 1: Validity of (4.7).** Since $\mathcal{X}_\bullet$ is finite dimensional and since we use the same discrete ansatz and test space, well-posedness of the discrete problem (4.3) is equivalent to the discrete inf-sup condition

$$\beta_\bullet = \inf_{W_\bullet\in\mathcal{X}_\bullet\backslash\{0\}} \sup_{V_\bullet\in\mathcal{X}_\bullet\backslash\{0\}} \frac{|b(W_\bullet\,,\,V_\bullet)|}{\|W_\bullet\|_{\mathcal{H}}\|V_\bullet\|_{\mathcal{H}}} = \inf_{W_\bullet\in\mathcal{X}_\bullet\backslash\{0\}} \frac{\|\mathfrak{B}_\bullet W_\bullet\|_{\mathcal{X}_\bullet^*}}{\|W_\bullet\|_{\mathcal{H}}} > 0. \tag{4.8}$$

Note that (4.8) implies that $\mathfrak{B}_\bullet$ is injective. Then, the finite dimension of $\mathcal{X}_\bullet$, i.e., $\dim \mathcal{X}_\bullet = \dim \mathcal{X}_\bullet^* < \infty$ implies surjectivity. Moreover according to [Dem06, Theorem 3], there holds for arbitrary $W_\bullet \in \mathcal{X}_\bullet$ that

$$\begin{aligned}\beta_\bullet \, \|U_\bullet - W_\bullet\|_{\mathcal{H}} &\le \sup_{V_\bullet \in \mathcal{X}_\bullet \setminus \{0\}} \frac{|b(U_\bullet - W_\bullet \,,\, V_\bullet)|}{\|V_\bullet\|_{\mathcal{H}}} \\ &= \sup_{V_\bullet \in \mathcal{X}_\bullet \setminus \{0\}} \frac{|b(u - W_\bullet \,,\, V_\bullet)|}{\|V_\bullet\|_{\mathcal{H}}} \le \|\mathfrak{B}_\bullet\| \, \|u - W_\bullet\|_{\mathcal{H}}.\end{aligned}$$

The triangle inequality yields that

$$\|u - U_\bullet\|_{\mathcal{H}} \le \|u - W_\bullet\|_{\mathcal{H}} + \|W_\bullet - U_\bullet\|_{\mathcal{H}} \le \Big(1 + \frac{\|\mathfrak{B}_\bullet\|}{\beta_\bullet}\Big) \|u - W_\bullet\|_{\mathcal{H}}.$$

Since the latter estimate holds for arbitrary functions $W_\bullet \in \mathcal{X}_\bullet$, we obtain inequality (4.7) with

$$C := 1 + \frac{M}{\beta_\bullet}, \quad \text{where} \quad \|\mathfrak{B}_\bullet\| \le M := \sup_{\substack{v \in \mathcal{H} \setminus \{0\} \\ w \in \mathcal{H} \setminus \{0\}}} \frac{|b(w \,,\, v)|}{\|w\|_{\mathcal{H}} \|v\|_{\mathcal{H}}}.$$

**Step 2:** It remains to prove the following assertion:

$$\exists \beta > 0 \, \exists \ell_\bullet \in \mathbb{N}_0 \, \forall \mathcal{X}_\bullet \subset \mathcal{H} \text{ with } \mathcal{X}_\bullet \supseteq \mathcal{X}_{\ell_\bullet} \quad \inf_{W_\bullet \in \mathcal{X}_\bullet \setminus \{0\}} \frac{\|\mathfrak{B}_\bullet W_\bullet\|_{\mathcal{X}_\bullet^*}}{\|W_\bullet\|_{\mathcal{H}}} \ge \beta. \tag{4.9}$$

We will prove (4.9) by contradiction. Let us assume that (4.9) is wrong and hence

$$\forall \beta > 0 \;\; \forall \ell_\bullet \in \mathbb{N}_0 \;\; \exists \mathcal{X}_\bullet \subset \mathcal{H} \text{ with } \mathcal{X}_\bullet \supseteq \mathcal{X}_{\ell_\bullet} \quad \inf_{W_\bullet \in \mathcal{X}_\bullet \setminus \{0\}} \frac{\|\mathfrak{B}_\bullet W_\bullet\|_{\mathcal{X}_\bullet^*}}{\|W_\bullet\|_{\mathcal{H}}} < \beta. \tag{4.10}$$

For each $\ell_\bullet = \ell \ge 0$ and $\beta = 1/\ell$, we can thus find a discrete subspace $\widehat{\mathcal{X}}_\ell = \mathcal{X}_\bullet \subset \mathcal{H}$ as well as an element $\widehat{W}_\ell \in \widehat{\mathcal{X}}_\ell$ such that

$$\widehat{\mathcal{X}}_\ell \supseteq \mathcal{X}_\ell, \quad \|\widehat{W}_\ell\|_{\mathcal{H}} = 1, \quad \text{and} \quad \|\widehat{\mathfrak{B}}_\ell \widehat{W}_\ell\|_{\widehat{\mathcal{X}}_\ell^*} < 1/\ell. \tag{4.11}$$

Since the sequence $\widehat{W}_\ell$ is bounded, without loss of generality, we may assume weak convergence $\widehat{W}_\ell \rightharpoonup w \in \mathcal{H}$ as $\ell \to \infty$.

**Step 3: There holds $w = 0$.** Let $\widehat{P}_\ell : \mathcal{H} \to \widehat{\mathcal{X}}_\ell$ be the orthogonal projection onto $\widehat{\mathcal{X}}_\ell$ and $v \in \mathcal{H}$. Weak convergence $\widehat{W}_\ell \rightharpoonup w$ as well as $b(\cdot \,,\, v) \in \mathcal{H}^*$ imply that $b(\widehat{W}_\ell \,,\, v) \to b(w \,,\, v)$ as $\ell \to \infty$. Moreover, we employ $\|\widehat{W}_\ell\|_{\mathcal{H}} = 1$ and $\|\widehat{P}_\ell v\|_{\mathcal{H}} \le \|v\|_{\mathcal{H}}$ to estimate

$$\begin{aligned}|b(\widehat{W}_\ell \,,\, v)| \le |b(\widehat{W}_\ell \,,\, \widehat{P}_\ell v)| + |b(\widehat{W}_\ell \,,\, v - \widehat{P}_\ell v)| &\le |\langle \widehat{\mathfrak{B}}_\ell \widehat{W}_\ell \,,\, \widehat{P}_\ell v \rangle| + |b(\widehat{W}_\ell \,,\, v - \widehat{P}_\ell v)| \\ &\le \|\widehat{\mathfrak{B}}_\ell \widehat{W}_\ell\|_{\widehat{\mathcal{X}}_\ell^*} \, \|v\|_{\mathcal{H}} + M \, \|v - \widehat{P}_\ell v\|_{\mathcal{H}}.\end{aligned}$$

Recall density (4.5) and nestedness $\mathcal{X}_\ell \subseteq \widehat{\mathcal{X}}_\ell \subset \mathcal{H}$. With the orthogonal projection, this implies that

$$\|v - \widehat{P}_\ell v\|_{\mathcal{H}} = \min_{\widehat{V}_\ell \in \widehat{\mathcal{X}}_\ell} \|v - \widehat{V}_\ell\|_{\mathcal{H}} \le \min_{V_\ell \in \mathcal{X}_\ell} \|v - V_\ell\|_{\mathcal{H}} \xrightarrow{\ell \to \infty} 0.$$

Recall that $\|\widehat{\mathfrak{B}}_\ell \widehat{W}_\ell\|_{\widehat{\mathcal{X}}_\ell^*} \le 1/\ell$ form (4.11). With this, we thus conclude that $|b(\widehat{W}_\ell\,,\,v)| \to 0$ as $\ell \to \infty$. Altogether, we obtain that $b(w\,,\,v) = 0$ for all $v \in \mathcal{H}$. Hence, well-posedness (4.4) yields that $w = 0$.

**Step 4: Assumption (4.10) yields a contradiction so that (4.9) follows.** Recall that $\|\widehat{W}_\ell\|_{\mathcal{H}} = 1$. Ellipticity of $a(\cdot\,,\,\cdot)$ and the definition of $b(\cdot\,,\,\cdot)$ yield that

$$\|\widehat{W}_\ell\|_{\mathcal{H}}^2 \overset{(4.1)}{\lesssim} a(\widehat{W}_\ell\,,\,\widehat{W}_\ell) \le |b(\widehat{W}_\ell\,,\,\widehat{W}_\ell)| + |\langle \mathfrak{C}\widehat{W}_\ell\,,\,\widehat{W}_\ell\rangle| \le \|\widehat{\mathfrak{B}}_\ell \widehat{W}_\ell\|_{\widehat{\mathcal{X}}_\ell^*} + \|\mathfrak{C}\widehat{W}_\ell\|_{\mathcal{H}^*}.$$

Recall that compact operators turn weak convergence into strong convergence. Hence $\widehat{W}_\ell \rightharpoonup w = 0$ in $\mathcal{H}$ implies that $\|\mathfrak{C}\widehat{W}_\ell\|_{\mathcal{H}^*} \to 0$ as $\ell \to \infty$. Together with $\|\widehat{\mathfrak{B}}_\ell \widehat{W}_\ell\|_{\widehat{\mathcal{X}}_\ell^*} \le 1/\ell$, we thus obtain the contradiction $1 = \|\widehat{W}_\ell\|_{\mathcal{H}} \to 0$ as $\ell \to \infty$. □

We emphasize, that the proof of Proposition 4.1 heavily relies on definiteness of bilinear form $b(\cdot\,,\,\cdot)$ on the Hilbert space $\mathcal{H}$ (see assumption (4.4)). Usually, for adaptive algorithms there holds $\overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell} \neq \mathcal{H}$ and hence well-posedness of the discrete limit-space cannot be guaranteed. In order to overcome this difficulty, we have to additionally ensure the definiteness of the discrete limit-space (see Section 4.5.1 and Axiom (E5)), which can be guaranteed by modifying the marking strategy in Algorithm 4.4 (see Proposition 4.7).

Recall that uniform mesh-refinement leads to a sequence of dense subspaces $\mathcal{H} = \overline{\bigcup_{\ell=0}^\infty \widehat{\mathcal{X}}_\ell}$. Under the abstract assumptions, the following statement holds as an immediate consequence of Proposition 4.1.

**Corollary 4.2.** *Let $\widehat{\mathcal{T}}_0 := \mathcal{T}_0$ and $\widehat{\mathcal{T}}_{\ell+1} := \mathrm{refine}(\widehat{\mathcal{T}}_\ell, \widehat{\mathcal{M}}_\ell)$ with $\widehat{\mathcal{M}}_\ell \subseteq \widehat{\mathcal{T}}_\ell$ for all $\ell \in \mathbb{N}_0$. Suppose that $\#\{\ell \in \mathbb{N}_0 : \widehat{\mathcal{M}}_\ell = \widehat{\mathcal{T}}_\ell\} = \infty$, i.e., uniform refinement is performed infinitely many times. Then, there exists $m \in \mathbb{N}_0$ and $\beta > 0$ such that for all discrete spaces $\mathcal{X}_\bullet \subset \mathcal{H}$ with $\mathcal{X}_\bullet \supseteq \widehat{\mathcal{X}}_m$, it holds the following:*

- *The related* inf-sup *constant* (4.6) *satisfies $\beta_\bullet \ge \beta > 0$.*
- *$\mathcal{X}_\bullet$ admits a unique solution $U_\bullet \in \mathcal{X}_\bullet$ of* (4.3) *which is quasi-optimal in the sense of inequality* (4.7)*.*
- *For $\ell \ge m$, the Galerkin solutions $\widehat{U}_\ell \in \widehat{\mathcal{X}}_\ell$ yield convergence $\lim_{\ell \to \infty} \|u - \widehat{U}_\ell\|_{\mathcal{H}} = 0$.* □

### 4.2.2 Ellipticity of $a(\cdot\,,\,\cdot)$

The work [FFP14] considers problems, where the left-hand side of (4.2) is strongly elliptic on $\mathcal{H} = H_0^1(\Omega)$ in the following sense: There exists $\widetilde{\alpha} > 0$ such that

$$\widetilde{\alpha}\,\|v\|_{\mathcal{H}}^2 \le \mathrm{Re}\,\big(a(v, v) + \langle \mathfrak{C}\,v\,,\,v\rangle\big) \quad \text{for all } v \in \mathcal{H}. \tag{4.12}$$

The next lemma shows that the stronger assumption (4.12) already implies that $a(\cdot\,,\,\cdot)$ is elliptic in the sense of (4.1). Hence, the given abstract framework generalizes the analysis of [FFP14].

**Lemma 4.3.** *Let $\widetilde{\alpha} > 0$ and $a(\cdot\,,\,\cdot)$ such that $a(w\,,\,w) > 0$ for all $w \in \mathcal{H} \setminus \{0\}$ and (4.12) is satisfied. Then there exists a constant $\alpha > 0$ with*

$$\alpha\,\|v\|_{\mathcal{H}}^2 \le a(v\,,\,v) \quad \textit{for all } v \in \mathcal{H}.$$

*Proof.* We argue by contradiction, i.e., we assume the following: For all $\varepsilon > 0$, there is some $v \in \mathcal{H}$ with $|a(v\,,\,v)| < \varepsilon\,\|v\|_{\mathcal{H}}^2$. Choosing $\varepsilon = 1/n$, we obtain sequences $(v_n)_{n\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ with

$$|a(v_n\,,\,v_n)| < \frac{\|v_n\|_{\mathcal{H}}^2}{n} \quad \text{as well as} \quad w_n := \frac{v_n}{\|v_n\|_{\mathcal{H}}}.$$

By definition, $w_n$ is bounded. Without loss of generality, we may thus suppose weak convergence $w_n \rightharpoonup w$ in $\mathcal{H}$. Weakly lower semicontinuity yields that

$$|a(w,w)| \le \liminf_{n\to\infty} |a(w_n\,,\,w_n)| = 0$$

and hence $w = 0$. Compactness of the operator $\mathfrak{C}$ implies strong convergence $\|\mathfrak{C}w_n\|_{\mathcal{H}^*} \to 0$ as $n \to \infty$. Finally, ellipticity (4.12) gives

$$\widetilde{\alpha} = \widetilde{\alpha}\,\|w_n\|_{\mathcal{H}}^2 \le \operatorname{Re}\,\big(a(w_n\,,\,w_n) + \langle \mathfrak{C}w_n\,,\,w_n\rangle\big) < 1/n + \|\mathfrak{C}w_n\|_{\mathcal{H}^*} \xrightarrow{n\to\infty} 0.$$

This contradicts $\widetilde{\alpha} > 0$ and concludes the proof. □

## 4.3 A posteriori error estimator

In order to define a suitable *a posteriori* error estimator, we have to ensure unique solvability of the discrete problem (4.3) for the underlying meshes $\mathcal{T}_\bullet \in \mathbb{T}$ and corresponding discrete subspaces $\mathcal{X}_\bullet \subset \mathcal{H}$.

To this end, suppose that for all $T \in \mathcal{T}_\bullet \in \mathbb{T}$ such that a unique discrete solution to (4.3) exists, there exists an associated *refinement indicator* with $\eta_\bullet(\cdot) : \mathcal{T}_\bullet \to \mathbb{R}$ such that $\eta_\bullet(T) \ge 0$. The related *a posteriori* error estimator is given by

$$\begin{aligned} \eta_\bullet &:= \eta_\bullet(\mathcal{T}_\bullet), \quad \text{where} \\ \eta_\bullet(\mathcal{U}_\bullet) &:= \Big(\sum_{T\in\mathcal{U}_\bullet} \eta_\bullet(T)^2\Big)^{1/2} \quad \text{for all subsets } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet. \end{aligned} \tag{4.13}$$

In order to prove optimal rates of convergence of adaptive finite element or boundary element schemes, we have to ensure additional properties of the error estimator. These so called *axioms of adaptivity* from [CFPP14] play an essential role in the analysis later on and are discussed in Section 4.5. We also refer to the Chapters 5–7 for different applications of the weighted-residual error estimator concerning finite and boundary elements.

## 4.4 Adaptive algorithm

Based on the *a posteriori* error estimator from the previous section, we consider the following adaptive algorithm.

**Algorithm 4.4.** INPUT: *Parameters* $0 < \theta \leq 1$, $C_{\mathrm{mark}} \geq 1$, *initial triangulation* $\mathcal{T}_0$ *and parameters* $U_{-1} := 0 \in \mathcal{X}_0$ *and* $\eta_{-1} := 1$.

ADAPTIVE LOOP: *For all* $\ell = 0, 1, 2, \dots$, *iterate the following steps* (i)–(v)*:*

(i) `If` (4.3) *does not admit a unique solution in* $\mathcal{X}_\ell$*:*
   - *Define* $U_\ell := U_{\ell-1} \in \mathcal{X}_\ell$ *and* $\eta_\ell := \eta_{\ell-1}$,
   - *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ *be the uniform refinement of* $\mathcal{T}_\ell$,
   - *Increase* $\ell \to \ell + 1$, *and continue with Step* (i).

(ii) `Else` *compute the unique solution* $U_\ell \in \mathcal{X}_\ell$ *to* (4.3).

(iii) *Compute the corresponding indicators* $\eta_\ell(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iv) *Determine a set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative constant* $C_{\mathrm{mark}}$ *minimal cardinality such that*

$$\theta\, \eta_\ell^2 \leq \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2. \tag{4.14}$$

(v) *Compute* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, *increase* $\ell$ *by 1, and continue with Step* (i).

OUTPUT: *Sequences of successively refined triangulations* $\mathcal{T}_\ell$, *discrete solutions* $U_\ell$, *and corresponding estimators* $\eta_\ell$.

**Remark 4.5.**
- *Apart from Step* (i), *Algorithm 4.4 is the usual adaptive loop based on the Dörfler marking strategy [Dör96] in Step* (iv) *as used, e.g., in [CKNS08, FFP14, CFPP14] for AFEM and in [FFK⁺14, FFK⁺15, FKMP13] for ABEM.*
- *For* $C_{\mathrm{mark}} = 1$, *the algorithmic construction of a set* $\mathcal{M}_\ell'$ *with minimal cardinality which satisfies, for instance, the Dörfler criterion requires sorting of the refinement indicators and thus results in logarithmic-linear complexity. Instead, Stevenson [Ste07] proposes an approximate sorting based on binning. This allows the algorithmic construction of some set* $\mathcal{M}_\ell$ *in real linear complexity which satisfies the Dörfler criterion and has minimal cardinality up to the multiplicative factor* $C_{\mathrm{mark}} = 2$.

The following lemma exploits the validity of Proposition 4.1 for uniform mesh-refinement (Corollary 4.2) and recaps some important properties of the sequence of discrete solutions produced by Algorithm 4.4.

**Lemma 4.6.** *Let* $(U_\ell)_{\ell \in \mathbb{N}_0}$ *be the sequence of discrete solutions generated by Algorithm 4.4. Then, there exists a minimal index* $\ell_0 \in \mathbb{N}_0$ *such that the discrete model problem* (4.3)

- *does not admit a unique solution in* $\mathcal{X}_\ell$ *for* $0 \le \ell < \ell_0$,
- *but admits a unique solution* $U_{\ell_0} \in \mathcal{X}_{\ell_0}$.

*In particular, the corresponding mesh* $\mathcal{T}_{\ell_0}$ *is the* $\ell_0$*-times uniform refinement of* $\mathcal{T}_0$. *Furthermore, there exists* $\ell_1 \in \mathbb{N}_0$, *such that* (4.3) *admits a unique solution* $U_\ell \in \mathcal{X}_\ell$ *for all steps* $\ell \ge \ell_1$ *of Algorithm 4.4.*

*Proof.* Because of Corollary 4.2, the uniform refinement in Step (i) of Algorithm 4.4 will only be performed at most finitely many times. This concludes the proof. □

## 4.5 Axioms of adaptivity

To prove convergence with optimal algebraic rates for Algorithm 4.4, we rely on the following *axioms of adaptivity* which are slightly generalized when compared to those of [CFPP14], since we always have to suppose solvability of the related discrete problem (4.3).

**E1) stability on non-refined element domains:** There exists $C_{\rm stb} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, the following holds: Provided there exist unique discrete solutions $U_\bullet \in \mathcal{X}_\bullet$ and $U_\circ \in \mathcal{X}_\circ$, it holds that

$$\big|\eta_\circ(\mathcal{T}_\circ \cap \mathcal{T}_\bullet) - \eta_\bullet(\mathcal{T}_\circ \cap \mathcal{T}_\bullet)\big| \le C_{\rm stb}\, \|U_\circ - U_\bullet\|_{\mathcal{H}}.$$

**E2) reduction on refined element domains:** There exist $C_{\rm red} > 0$ and $0 < q_{\rm red} < 1$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, the following holds: Provided there exist unique discrete solutions $U_\bullet \in \mathcal{X}_\bullet$ and $U_\circ \in \mathcal{X}_\circ$, it holds that

$$\eta_\circ(\mathcal{T}_\circ\backslash\mathcal{T}_\bullet)^2 \le q_{\rm red}\, \eta_\bullet(\mathcal{T}_\bullet\backslash\mathcal{T}_\circ)^2 + C_{\rm red}^2\, \|U_\circ - U_\bullet\|_{\mathcal{H}}^2.$$

**E3) reliability:** There exists $\widetilde{C}_{\rm rel} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ the following holds: Provided there exists a unique discrete solutions $U_\bullet \in \mathcal{X}_\bullet$, it holds that

$$\|u - U_\bullet\|_{\mathcal{H}} \le \widetilde{C}_{\rm rel}\, \eta_\bullet.$$

**E4) discrete reliability:** There exists $C_{\rm rel} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, there exists a set $\mathcal{R}_{\bullet,\circ} \subseteq \mathcal{T}_\bullet$ such that the following holds: Provided there exist unique discrete solutions $U_\bullet \in \mathcal{X}_\bullet$ and $U_\circ \in \mathcal{X}_\circ$, it holds that

$$\|U_\circ - U_\bullet\|_{\mathcal{H}} \le C_{\rm rel}\, \beta_\circ^{-1}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ}) \quad \text{as well as} \quad \mathcal{T}_\bullet \setminus \mathcal{T}_\circ \subseteq \mathcal{R}_{\bullet,\circ},$$

with $\#\mathcal{R}_{\bullet,\circ} \le C_{\rm rel}\, \#(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)$, where $\beta_\circ > 0$ is the inf-sup constant (4.6) associated with $\mathcal{X}_\circ$.

### 4.5.1 Definiteness on the "discrete" limit space (E5)

We need an additional assumption (see (E5) below) which goes beyond the axioms of adaptivity in [CFPP14]. To this end, let us define the "discrete" limit space $\mathcal{X}_\infty := \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$. Because of nestedness $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ for all $\ell \geq 0$, $\mathcal{X}_\infty$ is a closed subspace of $\mathcal{H}$ and hence a Hilbert space.

**E5) definiteness of $b(\cdot,\cdot)$ on $\mathcal{X}_\infty$:** For all $w \in \mathcal{X}_\infty$, the following implication holds: If $b(w,v) = 0$ for all $v \in \mathcal{X}_\infty$, then $w = 0$.

While (E1)–(E4) rely only on the a posteriori error estimation strategy, the property (E5) involves the "discrete" limit space $\mathcal{X}_\infty = \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ generated by Algorithm 4.4 and is hence less accessible for the numerical analysis. Clearly, (E5) is satisfied if $b(\cdot,\cdot)$ is elliptic (4.12). Moreover, note that well-posedness (4.4) of (4.2) implies that (E5) is satisfied, if $\mathcal{X}_\infty = \mathcal{H}$. In many generic situations, the identity $\mathcal{X}_\infty = \mathcal{H}$ is automatically satisfied; see Section 5.2.1 in case of AFEM and Section 6.5.1 in case of ABEM.

The next proposition shows that $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ and hence (E5) with $\mathcal{X}_\infty = \mathcal{H}$ can also be guaranteed by employing an expanded Dörfler marking strategy in Step (iv) of Algorithm 4.4. We stress that this does not affect optimal convergence behavior in the sense of Theorem 4.21 below.

**Proposition 4.7** (expanded Dörfler marking)**.** *Suppose* $0 < \theta \leq 1$. *Employ the notation of Algorithm 4.4. Let* $C'_{\rm mark} > 0$. *For all* $\ell \in \mathbb{N}_0$, *we suppose that the set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *in Step* (iv) *of Algorithm 4.4 is selected as follows:*

- *Let* $\mathcal{M}'_\ell \subseteq \mathcal{T}_\ell$ *be a set of up to the multiplicative constant* $C'_{\rm mark}$ *minimal cardinality such that* $\theta\,\eta_\ell^2 \leq \eta_\ell(\mathcal{M}'_\ell)^2$.
- *Suppose that* $\mathcal{T}_\ell = \{T_1,\dots,T_N\}$ *is sorted such that* $|T_1| \geq |T_2| \geq \dots \geq |T_N|$.
- *With arbitrary* $1 \leq n \leq \#\mathcal{M}'_\ell$, *define* $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \{T_1,\dots,T_n\}$.

*Then,* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *is a set of up to the multiplicative constant* $C_{\rm mark} := 2C'_{\rm mark}$ *minimal cardinality such that the usual Dörfler marking criterion* $\theta\,\eta_\ell^2 \leq \eta_\ell(\mathcal{M}_\ell)^2$ *is satisfied. Moreover, Algorithm 4.4 with the expanded marking guarantees* $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ *as* $\ell \to \infty$. *In particular, assumption* (E5) *is satisfied with* $\mathcal{X}_\infty = \mathcal{H}$.

*Proof.* The claims on $\mathcal{M}_\ell$ are obvious. Recall that uniform refinement leads to a sequence of triangulations $(\widehat{\mathcal{T}}_\ell)_{\ell\in\mathbb{N}_0}$ with a corresponding dense sequence of discrete subspaces $(\widehat{\mathcal{X}}_\ell)_{\ell\in\mathbb{N}_0}$. Let $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ denote the sequence of meshes, generated by Algorithm 4.4 with the expanded marking. Axiom (3.3) guarantees that refinement of an element leads to a contraction of the local mesh-size, i.e., $h_{\ell+1}|_T \leq q_{\rm mesh}\, h_\ell|_T$ for all $T \in \mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$. Since each mesh $\mathcal{T}_\ell$ is a finite set and each step of the adaptive algorithm guarantees that (at least) the element $T \in \mathcal{T}_\ell$ with the largest size $|T| \simeq (h_\ell|_T)^d$ is refined, this implies necessarily $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$.

Further, (R6) implies that the order of refinement does not matter. Hence, for all $\widehat{\mathcal{T}}_j \in (\widehat{\mathcal{T}}_\ell)_{\ell\in\mathbb{N}_0}$ there exits a mesh $\mathcal{T}_k \in (\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ such that $\mathcal{T}_k \in \mathrm{refine}(\widehat{\mathcal{T}}_j)$. Nestedness of the corresponding subspaces implies $\mathcal{X}_k \supseteq \widehat{\mathcal{X}}_j$. Density of $(\widehat{\mathcal{X}}_\ell)_{\ell\in\mathbb{N}_0}$ implies that $\mathcal{H} = \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ and concludes the proof. ∎

The following technical lemma exploits the validity of (E5) and collects some essential properties of the meshes and solutions generated by Algorithm 4.4.

**Lemma 4.8.** *Suppose* (E1), (E2), (E4), *and* (E5). *Employ the notation of Algorithm 4.4 for* $0 < \theta \le 1$. *Then, there exists* $\ell_2 \in \mathbb{N}_0$ *and* $\beta > 0$ *such that for all* $\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}_{\ell_2})$ *with* $\mathcal{X}_\bullet \subseteq \mathcal{X}_\infty$, *the following assertion* (a) *holds:*

(a) *The corresponding* inf-sup *constant* (4.6) *is bounded from below by* $\beta_\bullet \ge \beta > 0$. *In particular, there exists a unique Galerkin solution* $U_\bullet \in \mathcal{X}_\bullet$ *to* (4.3) *which is quasi-optimal in the sense of inequality* (4.7).

*Moreover, let* $\mathcal{T}_\bullet \in \mathbb{T}$ *and* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet) \cap \mathrm{refine}(\mathcal{T}_{\ell_2})$ *and suppose that the Galerkin solution* $U_\bullet \in \mathcal{X}_\bullet$ *exists. Then, there hold the following assertions* (b)–(c) *with some additional constant* $C_{\mathrm{mon}} > 0$ *which depends only on* $C_{\mathrm{stb}}$, $C_{\mathrm{red}}$, $C_{\mathrm{rel}}$, *and* $\beta$*:*

(b) uniform discrete reliability, *i.e.,* $\|U_\circ - U_\bullet\|_{\mathcal{H}} \le C_{\mathrm{rel}}\, \beta^{-1}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ})$.

(c) quasi-monotonicity of error estimator, *i.e.,* $\eta_\circ \le C_{\mathrm{mon}}\, \eta_\bullet$.

*If in addition* $\mathcal{X}_\infty = \mathcal{H}$, *then the following assertion* (d) *holds:*

(d) *discrete reliability* (E4) *implies reliability* (E3), *i.e.,* $\|u - U_\bullet\|_{\mathcal{H}} \le C_{\mathrm{rel}}\, \beta^{-1}\, \eta_\bullet$.

*Proof.* **Step 1: Proof of (a) and (b).** Employ Proposition 4.1 with $\mathcal{H}$ replaced by $\mathcal{X}_\infty$. Axiom (E5) ensures well-posedness of (4.4) on the discrete limit space $\mathcal{X}_\infty$. This proves (a) and provides $\ell_2 \in \mathbb{N}_0$ and $\beta > 0$, such that the inf-sup constant (4.6) for all discrete subspaces $\mathcal{X}_\circ \subseteq \mathcal{X}_\infty$ with $\mathcal{X}_\circ \supseteq \mathcal{X}_{\ell_2}$ is uniformly bounded from below by $\beta_\circ \ge \beta > 0$. Together with (E4), this also proves (b).

**Step 2: Proof of (c).** To prove quasi-monotonicity (c), we follow the lines of [CFPP14, Lemma 3.5]. Stability (E1) and reduction (E2) imply that

$$\eta_\circ^2 \le q_{\mathrm{red}}\, \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 + 2\, \eta_\bullet(\mathcal{T}_\circ \cap \mathcal{T}_\bullet)^2 + (2C_{\mathrm{stb}}^2 + C_{\mathrm{red}}^2)\, \|U_\circ - U_\bullet\|_{\mathcal{H}}^2$$

With discrete reliability, we obtain that

$$\eta_\circ^2 \le 2\, \eta_\bullet^2 + (2C_{\mathrm{stb}}^2 + C_{\mathrm{red}}^2)\, C_{\mathrm{rel}}^2\, \beta_\circ^{-2}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2.$$

This concludes (c) with constant $C_{\mathrm{mon}} = \big(2 + (2C_{\mathrm{stb}}^2 + C_{\mathrm{red}}^2)\, C_{\mathrm{rel}}^2\, \beta^{-2}\big)^{1/2}$.

**Step 3: Proof of (d).** We follow the proof of [CFPP14, Lemma 3.4]. Recall that uniform refinement yields convergence (Corollary 4.2). Hence, given any $\varepsilon > 0$ and $\mathcal{T}_\bullet \in \mathbb{T}$, there exists a uniform refinement $\widehat{\mathcal{T}}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ with corresponding discrete solution $\widehat{U}_\circ \in \mathcal{X}(\widehat{\mathcal{T}}_\circ)$ such that $\|u - \widehat{U}_\circ\|_{\mathcal{H}} \le \varepsilon$. Discrete reliability (E4) implies that

$$\|u - U_\bullet\|_{\mathcal{H}} \le \|u - \widehat{U}\|_{\mathcal{H}} + \|\widehat{U}_\circ - U_\bullet\|_{\mathcal{H}} \le \varepsilon + C_{\mathrm{rel}}\, \beta_\circ^{-1}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ}) \le \varepsilon + C_{\mathrm{rel}}\, \beta^{-1}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ}).$$

Since $\varepsilon > 0$ is arbitrary, this implies (d) and concludes the proof. $\square$

## 4.6 Convergence

In this section, we show plain convergence $\|u - U_\ell\|_{\mathcal{H}} \to 0$ as $\ell \to \infty$ for Algorithm 4.4. For the proof of linear convergence and convergence with optimal algebraic rates, we refer to Section 4.7 resp. Section 4.8. The next proposition is the main result of this Section.

**Proposition 4.9.** *Suppose* (E1)–(E5) *and* $0 < \theta \leq 1$. *Employ the notation of Algorithm 4.4. Then, the "discrete" limit space* $\mathcal{X}_\infty = \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ *contains the exact solution to problem* (4.2)*, i.e.,* $u \in \mathcal{X}_\infty$. *Moreover, there holds*

$$\lim_{\ell\to\infty} \|u - U_\ell\|_{\mathcal{H}} = 0 = \lim_{\ell\to\infty} \eta_\ell.$$

The proof of Proposition 4.9 relies on the following estimator reduction which (in a weaker form) is first found also in [CKNS08]. We use a slightly generalized version, which is used in [FPZ16, Lemma 9] and follow ideas of [CFPP14, Lemma 4.7].

**Lemma 4.10** (generalized estimator reduction)**.** *Stability* (E1) *and reduction* (E2) *together with the Dörfler marking strategy from Step* (iv) *of Algorithm 4.4 imply the following perturbed contraction: For each* $\ell \in \mathbb{N}_0$ *and all* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_{\ell+1})$ *such that the discrete solutions* $U_\ell \in \mathcal{X}_\ell$ *and* $U_\circ \in \mathcal{X}_\circ$ *exist, it holds that*

$$\eta_\circ^2 \leq q_{\mathrm{est}}\, \eta_\ell^2 + C_{\mathrm{est}}\, \|U_\circ - U_\ell\|_{\mathcal{H}}^2.$$

*The constants* $C_{\mathrm{est}} > 0$ *and* $0 < q_{\mathrm{est}} < 1$ *depend only on* (E1)–(E2) *and on* $0 < \theta \leq 1$.

*Proof.* Let $\delta > 0$ and $C_{\mathrm{est}} := C_{\mathrm{red}}^2 + (1 + \delta^{-1})\, C_{\mathrm{stb}}^2$. The young inequality, stability (E1), and reduction (E2) imply that

$$\begin{aligned}
\eta_\circ^2 &= \eta_\circ(\mathcal{T}_\circ \setminus \mathcal{T}_\ell)^2 + \eta_\circ(\mathcal{T}_\circ \cap \mathcal{T}_\ell)^2 \\
&\leq q_{\mathrm{red}}\, \eta_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_\circ)^2 + C_{\mathrm{red}}^2\, \|U_\circ - U_\ell\|_{\mathcal{H}}^2 + \Big(\eta_\ell(\mathcal{T}_\circ \cap \mathcal{T}_\ell) + C_{\mathrm{stb}} \|U_\circ - U_\ell\|_{\mathcal{H}}\Big)^2 \\
&\leq q_{\mathrm{red}}\, \eta_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_\circ)^2 + (1 + \delta)\, \eta_\ell(\mathcal{T}_\circ \cap \mathcal{T}_\ell) + C_{\mathrm{est}}\, \|U_\circ - U_\ell\|_{\mathcal{H}}^2.
\end{aligned}$$

Note that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_{\ell+1})$ implies $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1} \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_\circ$. In combination with the Dörfler marking criterion, we obtain that

$$\begin{aligned}
\eta_\circ^2 &\leq (1 + \delta)\Big(\eta_\ell^2 - (1 - q_{\mathrm{red}})\, \eta_\ell^2(\mathcal{T}_\ell \setminus \mathcal{T}_\circ)^2\Big) + C_{\mathrm{est}}\, \|U_\circ - U_\ell\|_{\mathcal{H}}^2 \\
&\leq (1 + \delta)\big(1 - (1 - q_{\mathrm{red}})\theta\big)\, \eta_\ell^2 + C_{\mathrm{est}}\, \|U_\circ - U_\ell\|_{\mathcal{H}}^2.
\end{aligned}$$

Choosing $\delta > 0$ sufficient small, such that $q_{\mathrm{est}} := (1 + \delta)\big(1 - (1 - q_{\mathrm{red}})\, \theta\big) < 1$, we conclude the proof. ∎

With the estimator reduction of Lemma 4.10, we can proof plain convergence of Algorithm 4.4.

***Proof of Proposition 4.9.*** Let $\ell_2 \in \mathbb{N}_0$ be the index defined in Lemma 4.8. To simplify notation and without loss of generality, we may assume $\ell_2 = 0$ throughout the proof. In order to prove that $\eta_\ell \to 0$ as $\ell \to \infty$, we show that each subsequence $(\eta_{\ell_k})_{k\in\mathbb{N}_0}$ of the estimator sequence $(\eta_\ell)_{\ell\in\mathbb{N}_0}$ contains a further subsequence $(\eta_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ with $\eta_{\ell_{k_j}} \to 0$ as $j \to \infty$. According to basic calculus, this is in fact equivalent to $\eta_\ell \to 0$ as $\ell \to \infty$. We split the proof into several steps.

**Step 1: Boundedness of estimator sequence.** We apply Lemma 4.8 with $\ell_2 = 0$. Then, quasi-monotonicity of the error estimator proves $\eta_\ell \le C_{\rm mon}\,\eta_0$ for all $\ell \in \mathbb{N}_0$.

**Step 2: Weak convergence of discrete solutions (subsequence).** Recall that $a(\cdot\,,\,\cdot)$ is elliptic and induces an energy norm $|\!|\!|\,\cdot\,|\!|\!|$. Reliability (E3) in combination with Step 1 implies that

$$|\!|\!|\, U_\ell \,|\!|\!| \le |\!|\!|\, u \,|\!|\!| + |\!|\!|\, u - U_\ell \,|\!|\!| \lesssim |\!|\!|\, u \,|\!|\!| + \sup_{\ell\in\mathbb{N}_0} \eta_\ell < \infty,$$

i.e., the sequence of discrete solutions is uniformly bounded in $\mathcal{H}$. Let $(\eta_{\ell_k})_{k\in\mathbb{N}_0}$ be an arbitrary subsequence of $(\eta_\ell)_{\ell\in\mathbb{N}_0}$ with corresponding discrete solutions $U_{\ell_k}$. Since $U_{\ell_k} \in \mathcal{X}_{\ell_k} \subseteq \mathcal{X}_\infty$, there exists a subsequence $(U_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ of $(U_{\ell_k})_{k\in\mathbb{N}_0}$ and some limit $w \in \mathcal{H}$ such that $U_{\ell_{k_j}} \rightharpoonup w$ weakly in $\mathcal{H}$ as $j \to \infty$. According to Mazur's lemma (see, e.g., [Rud91, Theorem 3.12]), convexity and closedness imply that $\mathcal{X}_\infty$ is also closed with respect to the weak topology and hence $w \in \mathcal{X}_\infty$. Let $v \in \mathcal{X}_\infty$ and let $P_\ell : \mathcal{H} \to \mathcal{X}_\ell$ denote the orthogonal projection with respect to $|\!|\!|\,\cdot\,|\!|\!|$, i.e.,

$$|\!|\!|\, v - P_\ell v \,|\!|\!| = \min_{V_\ell\in\mathcal{X}_\ell} |\!|\!|\, v - V_\ell \,|\!|\!| \quad \text{for all } v \in \mathcal{H}.$$

By definition of $\mathcal{X}_\infty$, this also implies strong convergence $|\!|\!|\, v - P_\ell v \,|\!|\!| \to 0$ as $\ell \to \infty$. Recall that the product of a weakly convergent sequence and a strongly convergent sequence leads to convergence of the scalar product. Moreover, compact operators turn weak convergence into strong convergence, i.e., $\mathfrak{C}\,U_{\ell_{k_j}} \to \mathfrak{C}\,w$ strongly in $\mathcal{H}^*$ as $j \to \infty$. With these two observations, we derive that

$$0 \overset{(4.3)}{=} \langle f\,,\,P_{\ell_{k_j}} v\rangle - a(U_{\ell_{k_j}}\,,\,P_{\ell_{k_j}} v) - \langle \mathfrak{C}U_{\ell_{k_j}}\,,\,P_{\ell_{k_j}} v\rangle \xrightarrow{j\to\infty} \langle f\,,\,v\rangle - a(w,v) - \langle \mathfrak{C}w\,,\,v\rangle.$$

This proves that the weak limit $w \in \mathcal{X}_\infty$ solves the Galerkin formulation

$$a(w\,,\,v) + \langle \mathfrak{C}\,w\,,\,v\rangle = \langle f\,,\,v\rangle \quad \text{for all } v \in \mathcal{X}_\infty. \tag{4.15}$$

**Step 3: Strong convergence of discrete solutions (subsequence).** Note that $|\!|\!|\, w - U_{\ell_{k_j}} \,|\!|\!|^2 = |\!|\!|\, w \,|\!|\!|^2 - 2\,\mathrm{Re}\,a(w, U_{\ell_{k_j}}) + |\!|\!|\, U_{\ell_{k_j}} \,|\!|\!|^2$. Therefore, strong convergence $|\!|\!|\, w - U_{\ell_{k_j}} \,|\!|\!| \to 0$ is equivalent to weak convergence $U_{\ell_{k_j}} \rightharpoonup w$ plus convergence of the norm $|\!|\!|\, U_{\ell_{k_j}} \,|\!|\!| \to |\!|\!|\, w \,|\!|\!|$. It thus only remains to prove the latter. With the previous observations, it holds that

$$\begin{aligned} |\!|\!|\, U_{\ell_{k_j}} \,|\!|\!|^2 = a(U_{\ell_{k_j}}, U_{\ell_{k_j}}) &\overset{(4.3)}{=} \langle f\,,\,U_{\ell_{k_j}}\rangle - \langle \mathfrak{C}U_{\ell_{k_j}}\,,\,U_{\ell_{k_j}}\rangle \\ &\xrightarrow{j\to\infty} \langle f\,,\,w\rangle - \langle \mathfrak{C}w\,,\,w\rangle \overset{(4.15)}{=} a(w,w) = |\!|\!|\, w \,|\!|\!|^2. \end{aligned}$$

**Step 4: Estimator reduction principle (subsequence).** Let $(\eta_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ denote the estimator subsequence corresponding to $(U_{\ell_{k_j}})_{j\in\mathbb{N}_0}$. With $\mathcal{T}_{\ell_{k_{j+1}}} \in \mathrm{refine}(\mathcal{T}_{\ell_{k_j}+1})$ and

Lemma 4.10, it holds $\eta^2_{\ell_{k_{j+1}}} \le q_{\rm est}\,\eta^2_{\ell_{k_j}} + C_{\rm est}\,\|U_{\ell_{k_{j+1}}} - U_{\ell_{k_j}}\|^2_{\mathcal{H}}$. Moreover, Step 3 implies convergence $\|U_{\ell_{k_{j+1}}} - U_{\ell_{k_j}}\|_{\mathcal{H}} \simeq |\!|\!| U_{\ell_{k_{j+1}}} - U_{\ell_{k_j}} |\!|\!| \to 0$ as $j \to \infty$. Hence, the subsequence $(\eta_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ is contractive up to a sequence that converges to zero. Therefore, basic calculus (see [AFLP12, Lemma 2.3]) proves convergence $\eta_{\ell_{k_j}} \to 0$ as $j \to \infty$.

**Step 5: Estimator convergence (full sequence).** We have shown that each subsequence $(\eta_{\ell_k})_{k\in\mathbb{N}_0}$ of $(\eta_\ell)_{\ell\in\mathbb{N}_0}$ has a further subsequence $(\eta_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ with $\eta_{\ell_{k_j}} \to 0$ as $j \to \infty$. As noted above, this already yields $\eta_\ell \to 0$ as $\ell \to \infty$.

**Step 6: Strong convergence of discrete solutions (full sequence).** Finally, reliability (E3) implies that $\|u - U_\ell\|_{\mathcal{H}} \lesssim \eta_\ell \to 0$ as $\ell \to \infty$. This concludes the proof. $\square$

**Remark 4.11.** *Note that the proof of Proposition 4.9 relies only on* (E4)–(E5) *to prove boundedness of the estimator sequence* $(\eta_\ell)_{\ell\in\mathbb{N}_0}$ *(see Step 1 of the proof). Instead, we can also modify the marking Step* (iv) *of Algorithm 4.4 so that the assertion of Proposition 4.9 remains true, if* (E1)–(E3) *still hold, while* (E4)–(E5) *fail. To this end, consider the following alternative criterion:*

(iv) *If* $\eta_\ell > \max_{j=0,\dots,\ell-1}\eta_j$*, define* $\mathcal{M}_\ell := \mathcal{T}_\ell$*. Otherwise, determine a set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative constant* $C_{\rm mark}$ *minimal cardinality such that* $\theta\eta_\ell^2 \le \eta_\ell(\mathcal{M}_\ell)^2$.

*To see that this new marking criterion ensures that* $(\eta_\ell)_{\ell\in\mathbb{N}_0}$ *is bounded, we argue as follows:*

**Case 1:** *Suppose that there exists an* $M \in \mathbb{N}$ *such that* $\eta_\ell \le \max_{j=0,\dots,\ell-1}\eta_j$ *for all* $\ell \ge M$*. Then, it even follows that* $\eta_\ell \le \max_{j=0,\dots,M-1}\eta_j$ *for all* $\ell \in \mathbb{N}_0$.

**Case 2:** *If the assumption of Case 1 fails, the alternative Step* (iv) *of Algorithm 4.4 enforces infinitely many steps of uniform refinement. Therefore, Corollary 4.2 applies and provides* $m \in \mathbb{N}_0$ *and* $C > 0$ *such that all discrete subspaces* $\mathcal{X}_\bullet \subseteq \mathcal{H}$ *with* $\mathcal{X}_\bullet \supseteq \mathcal{X}_m$ *admit a unique solution* $U_\bullet \in \mathcal{X}_\bullet$ *of* (4.3) *which is quasi-optimal in the sense of inequality* (4.7)*. Since* (E1)–(E3) *hold, [CFPP14, Lemma 3.5] applies and proves quasi-monotonicity of the estimator, i.e.,*

$$\eta_\circ \le C_{\rm mon}\,\eta_\bullet \quad \textit{for all } \mathcal{T}_\bullet \in \operatorname{refine}(\mathcal{T}_m) \textit{ and all } \mathcal{T}_\circ \in \operatorname{refine}(\mathcal{T}_\bullet).$$

*In particular, this implies* $\eta_\ell \le C_{\rm mon}\,\eta_m$ *for all* $\ell \ge m$*, and therefore*

$$\eta_\ell \le \max\{C_{\rm mon}, 1\} \max_{j=0,\dots,m} \eta_j \quad \textit{for all } \ell \in \mathbb{N}_0.$$

*Hence, Case 2 cannot happen.*

*Note that besides Step 1 all steps of the proof of Proposition 4.9 rely only on* (E1)–(E3)*. Therefore, we obtain* $\eta_\ell \to 0$ *as* $\ell \to \infty$*. In particular, this implies that Case 1 above is the generic case and that optimal convergence rates will not be affected by the new marking strategy.*

## 4.7 Linear convergence of adaptive algorithm

The analysis in this section adapts and extends some ideas from [FFP14]. We note that the latter work uses strong ellipticity (4.12) of $b(\cdot,\cdot)$, while we only rely on ellipticity (4.1) of $a(\cdot,\cdot)$ (see Section 4.2.2). The goal of this section is to prove linear convergence in Theorem 4.14.

**Lemma 4.12** ([FFP14, Lemma 3.5]). *Suppose* (E1)–(E5) *and* $0 < \theta \leq 1$. *Employ the notation of Algorithm 4.4. Then, the sequences* $(e_\ell)_{\ell\in\mathbb{N}}$ *and* $(E_\ell)_{\ell\in\mathbb{N}}$ *defined by*

$$e_\ell := \begin{cases} \frac{u-U_\ell}{\|u-U_\ell\|_{\mathcal{H}}} & \text{for } u \neq U_\ell, \\ 0 & \text{else}, \end{cases}$$

$$E_\ell := \begin{cases} \frac{U_{\ell+1}-U_\ell}{\|U_{\ell+1}-U_\ell\|_{\mathcal{H}}} & \text{for } U_{\ell+1} \neq U_\ell, \\ 0 & \text{else}, \end{cases}$$

*converge weakly to zero, i.e.,* $\lim_{\ell\to\infty} \langle \phi\,,\, e_\ell \rangle = 0 = \lim_{\ell\to\infty} \langle \phi\,,\, E_\ell \rangle$ *for all* $\phi \in \mathcal{H}^*$.

*Proof.* We consider the sequence $(e_\ell)_{\ell\in\mathbb{N}_0}$. The proof of the claim for $(E_\ell)_{\ell\in\mathbb{N}_0}$ follows along the same lines. To prove $e_\ell \rightharpoonup 0$ as $\ell \to \infty$, we show that each subsequence $(e_{\ell_k})_{k\in\mathbb{N}_0}$ admits a further subsequence $(e_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ such that $e_{\ell_{k_j}} \rightharpoonup 0$ as $j \to \infty$. Let $(e_{\ell_k})_{k\in\mathbb{N}_0}$ be a subsequence of $(e_\ell)_{\ell\in\mathbb{N}_0}$. Because of boundedness $\|e_{\ell_k}\|_{\mathcal{H}} \leq 1$, there exists a further weakly convergent subsequence $(e_{\ell_{k_j}})_{j\in\mathbb{N}_0}$ such that $e_{\ell_{k_j}} \rightharpoonup w \in \mathcal{H}$ as $j \to \infty$. It thus remains to show that $w = 0$.

Proposition 4.9 yields that $U_\ell, u \in \mathcal{X}_\infty$. This implies that $e_\ell \in \mathcal{X}_\infty$ and hence $w \in \mathcal{X}_\infty$. With Galerkin orthogonality we obtain that

$$0 = b(u - U_\bullet\,,\, V_\bullet) = a(u - U_\bullet\,,\, V_\bullet) + \langle \mathfrak{C}(u - U_\bullet)\,,\, V_\bullet \rangle \quad \text{for all } V_\bullet \in \mathcal{X}_\bullet. \tag{4.16}$$

Let $n \in \mathbb{N}$ be arbitrary and $V_n \in \mathcal{X}_n$. For $\ell_{k_j} \geq n$ and $e_{\ell_{k_j}} \neq 0$, the Galerkin orthogonality proves

$$b(e_{\ell_{k_j}}\,,\, V_n) = \frac{b(u - U_{\ell_{k_j}}\,,\, V_n)}{\|u - U_{\ell_{k_j}}\|_{\mathcal{H}}} = 0.$$

Hence, $b(e_{\ell_{k_j}}, V_n) = 0$ for all $\ell_{k_j} \geq n$. With weak convergence, this yields that

$$b(w\,,\, V_n) = \lim_{j\to\infty} b(e_{\ell_{k_j}}\,,\, V_n) = 0 \quad \text{for all } V_n \in \mathcal{X}_n \text{ and all } n \in \mathbb{N}_0.$$

Let $v \in \mathcal{X}_\infty$. By definition of $\mathcal{X}_\infty$, there exists a sequence $(V_n)_{n\in\mathbb{N}_0}$ with $V_n \in \mathcal{X}_n$ and $\|v - V_n\|_{\mathcal{H}} \to 0$ as $n \to \infty$. Therefore the latter identity implies that

$$b(w\,,\, v) = \lim_{n\to\infty} b(w\,,\, V_n) = 0 \quad \text{for all } v \in \mathcal{X}_\infty.$$

Finally, definiteness of the discrete limit space (E5) concludes $w = 0$. $\square$

The following quasi-orthogonality (4.17) is a consequence of Lemma 4.12 and the Galerkin orthogonality (4.16). For elliptic $b(\cdot\,,\,\cdot)$, it is proved in [FFP14, Proposition 3.6]. Our proof essentially follows those ideas, but since we lack ellipticity of the bilinear form $b(\cdot\,,\,\cdot)$ and since $b(\cdot\,,\,\cdot)$ does therefore not induce an equivalent quasi-norm, we instead use the energy norm $\|\!|\,\cdot\,|\!\|$ induced by $a(\cdot\,,\,\cdot)$.

**Lemma 4.13.** *Suppose* (E1)–(E5) *and* $0 < \theta \le 1$. *Employ the notation of Algorithm 4.4. Then, for any* $0 < \varepsilon < 1$, *there exists* $\ell_3 \in \mathbb{N}_0$ *such that*

$$\|\!|\, u - U_{\ell+1} \,|\!\|^2 + \|\!|\, U_{\ell+1} - U_\ell \,|\!\|^2 \le \frac{1}{1-\varepsilon} \|\!|\, u - U_\ell \,|\!\|^2 \quad \text{for all } \ell \ge \ell_3. \tag{4.17}$$

*Note that,* $\ell_3 = \ell_3(\varepsilon)$ *does depend on the given parameter* $\varepsilon$.

*Proof.* Let $\varepsilon > 0$. Further, let $\delta > 0$ be a free parameter which is fixed later. Consider the sequences $(e_\ell)_{\ell\in\mathbb{N}_0}$ and $(E_\ell)_{\ell\in\mathbb{N}_0}$ of Lemma 4.12. Recall that the compact operator $\mathfrak{C}$ turns weak convergence $e_\ell, E_\ell \rightharpoonup 0$ in $\mathcal{H}$ into strong convergence $\mathfrak{C}\,e_\ell, \mathfrak{C}\,E_\ell \to 0$ in $\mathcal{H}^*$ as $\ell \to \infty$. For any $\delta > 0$, this provides some $\ell_3 \in \mathbb{N}$ such that

$$\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*} + \|\mathfrak{C}\,E_\ell\|_{\mathcal{H}^*} \le \delta \quad \text{for all } \ell \ge \ell_3.$$

For any $w \in \mathcal{H}$, this gives

$$|\langle \mathfrak{C}(u - U_\ell)\,,\, w\rangle| = |\langle \mathfrak{C}e_\ell\,,\, w\rangle|\, \|u - U_\ell\|_{\mathcal{H}} \le \delta\, \|u - U_\ell\|_{\mathcal{H}} \|w\|_{\mathcal{H}},$$

as well as

$$|\langle \mathfrak{C}\,(U_{\ell+1} - U_\ell)\,,\, w\rangle| = |\langle \mathfrak{C}\,E_\ell\,,\, w\rangle|\, \|U_{\ell+1} - U_\ell\|_{\mathcal{H}} \le \delta\, \|U_{\ell+1} - U_\ell\|_{\mathcal{H}} \|w\|_{\mathcal{H}}.$$

Some basic computations in combination with the Galerkin orthogonality (4.16) show that

$$\begin{aligned}
b(u-U_{\ell+1}\,, u-U_{\ell+1}) &+ b(U_{\ell+1}-U_\ell\,,\, U_{\ell+1}-U_\ell) + b(U_{\ell+1}-U_\ell\,,\, u-U_{\ell+1}) \\
&= b(u-U_\ell\,,\, u-U_{\ell+1}) + b(U_{\ell+1}-U_\ell\,,\, U_{\ell+1}-U_\ell) \\
&= b(u-U_\ell\,,\, u-U_{\ell+1} - U_\ell + U_\ell) + b(U_{\ell+1}-U_\ell\,,\, U_{\ell+1}-U_\ell) \\
&= b(u-U_\ell\,,\, u-U_\ell) - b(u-U_\ell\,,\, U_{\ell+1}-U_\ell) + b(U_{\ell+1}-U_\ell\,,\, U_{\ell+1}-U_\ell) \\
&= b(u-U_\ell\,,\, u-U_\ell) + b(U_{\ell+1}-u\,,\, U_{\ell+1}-U_\ell) \\
&\overset{(4.16)}{=} b(u-U_\ell\,,\, u-U_\ell).
\end{aligned}$$

Recall that $\|\!|\, v \,|\!\|^2 = a(v\,,\,v) = b(v\,,\,v) - \langle \mathfrak{C}\,v\,,\,v\rangle$ for all $v \in \mathcal{H}$. Then, the latter equality is equivalent to

$$\begin{aligned}
\|\!|\, u - U_{\ell+1} \,|\!\|^2 &+ \|\!|\, U_{\ell+1} - U_\ell \,|\!\|^2 + \langle \mathfrak{C}\,(u - U_{\ell+1})\,,\, u - U_{\ell+1}\rangle + \langle \mathfrak{C}\,(U_{\ell+1} - U_\ell)\,,\, U_{\ell+1} - U_\ell\rangle \\
&+ b(U_{\ell+1} - U_\ell\,,\, u - U_{\ell+1}) = \|\!|\, u - U_\ell \,|\!\|^2 + \langle \mathfrak{C}\,(u - U_\ell)\,,\, u - U_\ell\rangle.
\end{aligned}$$

The remaining term with the bilinear form $b(U_{\ell+1} - U_\ell\,,\, u - U_{\ell+1})$ is estimated as follows:

$$\begin{aligned}
|b(U_{\ell+1} - U_\ell, u - U_{\ell+1})| &= |\overline{a(u - U_{\ell+1}\,,\, U_{\ell+1} - U_\ell)} + \langle \mathfrak{C}(U_{\ell+1} - U_\ell)\,,\, u - U_{\ell+1}\rangle| \\
&\overset{(4.16)}{=} |-\overline{\langle \mathfrak{C}(u - U_{\ell+1})\,,\, U_{\ell+1} - U_\ell\rangle} + \langle \mathfrak{C}(U_{\ell+1} - U_\ell)\,,\, u - U_{\ell+1}\rangle| \\
&\le 2\delta\, \|u - U_{\ell+1}\|_{\mathcal{H}} \|U_{\ell+1} - U_\ell\|_{\mathcal{H}}.
\end{aligned}$$

With norm equivalence $\|v\|_{\mathcal{H}}^2 \le C \,|\!|\!|\, v \,|\!|\!|^2$ for all $v \in \mathcal{H}$, we thus see that

$$\begin{aligned}(1-\delta C)\,|\!|\!|\, u - U_{\ell+1} \,|\!|\!|^2 + (1-\delta C)\,|\!|\!|\, U_{\ell+1} - U_\ell \,|\!|\!|^2 \\ \le (1+\delta C)\,|\!|\!|\, u - U_\ell \,|\!|\!|^2 + 2\delta C \,|\!|\!|\, u - U_{\ell+1} \,|\!|\!|\,|\!|\!|\, U_{\ell+1} - U_\ell \,|\!|\!|.\end{aligned}$$

Finally, recall the Young inequality $2cab \le ca^2 + cb^2$ for all $a, b, c \ge 0$. This yields that

$$(1-2\delta C)\,|\!|\!|\, u - U_{\ell+1} \,|\!|\!|^2 + (1-2\delta C)\,|\!|\!|\, U_{\ell+1} - U_\ell \,|\!|\!|^2 \le (1+\delta C)\,|\!|\!|\, u - U_\ell \,|\!|\!|^2.$$

For sufficiently small $\delta > 0$ and $\dfrac{1+\delta C}{1-2\delta C} \le \dfrac{1}{1-\varepsilon}$, this proves (4.17). □

The following result is found in [FFP14] for strongly elliptic problems (4.12). Our proof follows the ideas of [CKNS08] and generalizes [FFP14] to a more general class of compactly perturbed problems.

**Theorem 4.14.** *Suppose* (E1)–(E5) *and* $0 < \theta \le 1$. *Then, there exist constants* $0 < q_{\rm lin} < 1$ *and* $C_{\rm lin} > 0$ *such that the output of Algorithm 4.4 satisfies that*

$$\eta_{\ell+n} \le C_{\rm lin}\, q_{\rm lin}^n\, \eta_\ell \quad \textit{for all } \ell, n \in \mathbb{N}_0 \textit{ with } \ell \ge \ell_3, \tag{4.18}$$

*where* $\ell_3 \in \mathbb{N}_0$ *is the index from Lemma 4.13.*

In the proof of Theorem 4.14 we use the following generalized contraction property which is inspired by [CKNS08, Theorem 4.1].

**Lemma 4.15** (Generalized contraction)**.** *Suppose* (E1)–(E5) *and* $0 < \theta \le 1$. *Let* $\ell_3 \in \mathbb{N}_0$ *be the index from Lemma 4.13. Further, let* $\mathcal{T}_\ell, \mathcal{T}_\circ \in \mathbb{T}$ *with* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\ell)$ *such that the corresponding discrete solutions* $U_\ell, U_\circ$ *exist and the Dörfler marking criterion*

$$\theta\, \eta_\ell^2 \le \eta_\ell(\mathcal{T}_\ell \setminus \mathcal{T}_\circ)^2$$

*is satisfied. Then, there exist* $0 < q_{\rm lin}, \lambda < 1$ *such that for all* $\ell \ge \ell_3$ *it holds*

$$\Delta_\circ \le q_{\rm lin}\, \Delta_\ell \quad \textit{where} \quad \Delta_\bullet^2 := |\!|\!|\, u - U_\bullet \,|\!|\!|^2 + \lambda\, \eta_\bullet^2. \tag{4.19}$$

*Proof.* Recall norm equivalence $\|\cdot\|_{\mathcal{H}} \simeq |\!|\!|\cdot|\!|\!|$. This guarantees that reliability (E3) and estimator reduction (Lemma 4.10) also hold (up to a different constant) with respect to the $a(\cdot\,,\,\cdot)$-induced energy norm $|\!|\!|\cdot|\!|\!|$. To simplify the notation and without loss of generality, we therefore suppose that $\|\cdot\|_{\mathcal{H}} = |\!|\!|\cdot|\!|\!|$ throughout the proof.

Let $\varepsilon, \lambda > 0$ be free parameters which are fixed later. With estimator reduction from Lemma 4.10 and Lemma 4.13, we obtain that, for all $\ell \ge \ell_3 = \ell_3(\varepsilon)$,

$$\Delta_\circ^2 = |\!|\!|\, u - U_\circ \,|\!|\!|^2 + \lambda\, \eta_\circ^2 \le \frac{1}{1-\varepsilon}\,|\!|\!|\, u - U_\ell \,|\!|\!|^2 + \lambda\, q_{\rm est}\, \eta_\ell^2 + (\lambda C_{\rm est} - 1)\,|\!|\!|\, U_\circ - U_\ell \,|\!|\!|^2.$$

For sufficiently small $\lambda$, i.e., $\lambda C_{\rm est} \le 1$, and an additional free parameter $\delta > 0$, reliability (E3) yields that

$$\begin{aligned}\Delta_\circ^2 \le \frac{1}{1-\varepsilon}\,|\!|\!|\, u - U_\ell \,|\!|\!|^2 + \lambda\, q_{\rm est}\, \eta_\ell^2 &\le \Big(\frac{1}{1-\varepsilon} - \delta\lambda\Big)\,|\!|\!|\, u - U_\ell \,|\!|\!|^2 + \lambda(q_{\rm est} + \widetilde{C}_{\rm rel}^2 \delta)\, \eta_\ell^2 \\ &\le \max\Big\{\frac{1}{1-\varepsilon} - \delta\lambda\,,\, q_{\rm est} + \widetilde{C}_{\rm rel}^2 \delta\Big\}\, \Delta_\ell^2.\end{aligned}$$

Since $0 < q_{\rm est} < 1$, we may choose $\delta > 0$ sufficiently small such that $0 < q_{\rm est} + \widetilde{C}_{\rm rel}^2\,\delta < 1$. Finally choose $\varepsilon > 0$ sufficiently small such that $0 < 1/(1-\varepsilon) - \delta\lambda < 1$. This concludes the proof. $\square$

***Proof of Theorem 4.14.*** We employ the notation of Lemma 4.15 with $\mathcal{T}_\circ := \mathcal{T}_{\ell+1}$. This implies $\Delta_{\ell+1} \le q_{\rm lin}\,\Delta_\ell$. Induction on $n$ proves $\Delta_{\ell+n} \le q_{\rm lin}^n\,\Delta_\ell$ for all $\ell \ge \ell_3$ and all $n \in \mathbb{N}_0$. Note that reliability (E3) yields that $\eta_\bullet^2 \simeq \Delta_\bullet^2$. Combining these two observations, we conclude the proof. $\square$

### 4.7.1 Validity of the Céa lemma

In this section, we show that the discrete solutions $U_\ell$ computed by Algorithm 4.4 are quasi-optimal in the sense of the Céa lemma. Additionally, we obtain that the involved constants converge to 1 as $\ell \to \infty$.

**Theorem 4.16.** *Suppose* (E1)–(E5) *and* $0 < \theta \le 1$. *Then, there exist* $C_\ell \ge 1$ *with* $\lim\limits_{\ell\to\infty} C_\ell = 1$, *and* $\ell_4 > 0$ *such that the output of Algorithm 4.4 satisfies that*

$$\|\!| u - U_\ell \|\!| \le C_\ell \min_{V_\ell \in \mathcal{X}_\ell} \|\!| u - V_\ell \|\!| \quad \textit{for all } \ell \ge \ell_4. \tag{4.20}$$

*Proof.* Consider the sequences $(e_\ell)$ and $(E_\ell)$ from Lemma 4.12. We follow the arguments of the proof of Lemma 4.13. To this end, let $V_\ell \in \mathcal{X}_\ell$ be arbitrary. Then, Galerkin orthogonality (4.16) proves that

$$\begin{aligned}
\|\!| u - U_\ell \|\!|^2 &= b(u - U_\ell\,,\, u - U_\ell) - \langle \mathfrak{C}\,(u - U_\ell)\,,\, u - U_\ell \rangle \\
&\overset{(4.16)}{=} b(u - U_\ell\,,\, u - V_\ell) - \langle \mathfrak{C}\,(u - U_\ell)\,,\, u - U_\ell \rangle \\
&= a(u - U_\ell\,,\, u - V_\ell) + \langle \mathfrak{C}\,(u - U_\ell)\,,\, u - V_\ell \rangle - \langle \mathfrak{C}\,(u - U_\ell)\,,\, u - U_\ell \rangle \\
&\le \|\!| u - U_\ell \|\!| \|\!| u - V_\ell \|\!| + \|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*} \|u - U_\ell\|_{\mathcal{H}} \|u - V_\ell\|_{\mathcal{H}} + \|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*} \|u - U_\ell\|_{\mathcal{H}}^2.
\end{aligned}$$

Recall norm equivalence $\|v\|_{\mathcal{H}}^2 \le C\,\|\!| v \|\!|^2$ for all $v \in \mathcal{H}$. This implies that

$$\|\!| u - U_\ell \|\!| \le (1 + C\,\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*})\,\|\!| u - V_\ell \|\!| + C\,\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*} \|\!| u - U_\ell \|\!|.$$

Rearranging the terms in the latter estimate, we prove that

$$\|\!| u - U_\ell \|\!| \le \frac{1 + C\,\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*}}{1 - C\,\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*}}\,\|\!| u - V_\ell \|\!|.$$

This concludes (4.20). Lemma 4.12 yields weak convergence $e_\ell \rightharpoonup 0$. Compactness of $\mathfrak{C}$ implies $\|\mathfrak{C}\,e_\ell\|_{\mathcal{H}^*} \to 0$ as $\ell \to \infty$ and concludes the proof. $\square$

## 4.8 Optimal convergence rates

In this section we prove optimal algebraic convergence rates for the sequence of estimators generated by Algorithm 4.4. First, in order to quantify the optimal algebraic rate of convergence, we introduce so-called approximation classes in the spirit of [CFPP14]. Further, we discuss the incorporation of data oscillations (see, e.g., [CKNS08]) and show equivalence to approximation classes defined with respect to the total error (see [CKNS08]). Secondly, the main Theorem 4.21 is proved in Section 4.8.3. To this end, the next lemma recaps some important properties regarding sequences of successively uniformed refined meshes.

**Lemma 4.17.** *The mesh-refinement strategy guarantees the following properties* (a)–(c) *which are exploited in our analysis of optimal convergence rates:*

(a) *There exists $m_{\rm unif} \in \mathbb{N}$ such that the $m_{\rm unif}$-times uniform refinement $\widehat{\mathcal{T}}_0$ of $\mathcal{T}_0$ satisfies the assertions of Lemma 4.8 (with $\mathcal{T}_{\ell_2}$ replaced by $\widehat{\mathcal{T}}_0$). In particular, there holds the quasi-monotonicity of the estimator, i.e., there exists an independent constant $C_{\rm mon} > 0$ such that*

$$\eta_\circ \le C_{\rm mon}\, \eta_\bullet \quad \text{for all } \mathcal{T}_\bullet \in \mathbb{T} \text{ and all } \mathcal{T}_\circ \in \operatorname{refine}(\widehat{\mathcal{T}}_0) \cap \operatorname{refine}(\mathcal{T}_\bullet),$$

*provided that the Galerkin solution $U_\bullet \in \mathcal{X}_\bullet$ exists.*

(b) *Moreover, for all $\mathcal{T}_\bullet \in \mathbb{T}$, the $m_{\rm unif}$-times uniform refinement $\widehat{\mathcal{T}}_\bullet$ of $\mathcal{T}_\bullet$ guarantees $\widehat{\mathcal{T}}_\bullet \in \operatorname{refine}(\widehat{\mathcal{T}}_0)$ and $\#\widehat{\mathcal{T}}_\bullet \le C_{\rm son}^{m_{\rm unif}} \#\mathcal{T}_\bullet$.*

(c) *Suppose that $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ for $\ell \to \infty$ (e.g., the expanded Dörfler marking strategy from Proposition 4.7 is used). Then, there exists an index $\ell_5 \in \mathbb{N}_0$ such that $\mathcal{T}_\ell \in \operatorname{refine}(\widehat{\mathcal{T}}_0)$ for all $\ell \ge \ell_5$.*

*Proof.* Assertion (a) is a direct consequence of Corollary 4.2 resp. Proposition 4.1, if we argue as in the proof of Lemma 4.8.

Assertions (b) follows from the refinement axioms. For $\mathcal{T}_\bullet \in \mathbb{T}$ with the $m_{\rm unif}$-times uniform refinement $\widehat{\mathcal{T}}_\bullet$, (R6) implies $\widehat{\mathcal{T}}_\bullet \in \operatorname{refine}(\widehat{\mathcal{T}}_0)$ and (3.5) gives $\#\widehat{\mathcal{T}}_\bullet \le C_{\rm son}^{m_{\rm unif}} \#\mathcal{T}_\bullet$.

For Assertions (c), note that $h_\ell \to 0$ for $\ell \to \infty$ guarantees an index $\ell_5$ such that $\|h_{\ell_5}\|_{L^\infty(\Omega)} \le \|h_{\widehat{\mathcal{T}}_0}\|_{L^\infty(\Omega)}$. Hence, (R6) implies $\mathcal{T}_{\ell_5} \in \operatorname{refine}(\widehat{\mathcal{T}}_0)$ and concludes the proof. $\square$

### 4.8.1 Approximation classes

For $N \in \mathbb{N}_0$ and $\mathcal{T} \in \mathbb{T}$, we define the set of all possible refinements which contain at most $N$ elements more than $\mathcal{T}$ by

$$\mathbb{T}_N(\mathcal{T}) := \big\{\mathcal{T}_\bullet \in \operatorname{refine}(\mathcal{T}) \,:\, \#\mathcal{T}_\bullet - \#\mathcal{T} \le N \text{ and solution } U_\bullet \in \mathcal{X}_\bullet \text{ to (4.3) exists}\big\}.$$

We note that $\mathbb{T}_N(\mathcal{T})$ is finite, but may be empty. On the other hand, Lemma 4.17 guarantees $\mathbb{T}_N(\mathcal{T}) \neq \emptyset$ for sufficiently large $N$, e.g., $N \ge C_{\rm son}^{m_{\rm unif}} \#\mathcal{T}$.

We use the convention $\min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet = 0$, if $\mathbb{T}_N(\mathcal{T}) = \emptyset$. For $s > 0$, the corresponding approximation class is given by

$$\|u\|_{\mathbb{A}_s(\mathcal{T})} := \sup_{N \in \mathbb{N}_0} \Big((N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet\Big), \tag{4.21}$$

where $\eta_\bullet$ is the error estimator corresponding to the optimal triangulation $\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})$. Note that $\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty$ means that starting from a mesh $\mathcal{T}$, a convergence behavior of

$\eta_\bullet = \mathcal{O}\big((\#\mathcal{T}_\bullet)^{-s}\big)$ is possible, if the optimal meshes are chosen. To abbreviate notation, we let

$$\mathbb{T}_N := \mathbb{T}_N(\mathcal{T}_0) \quad \text{and} \quad \|u\|_{\mathbb{A}_s} := \|u\|_{\mathbb{A}_s(\mathcal{T}_0)}. \tag{4.22}$$

The next lemma examines the relation between approximation classes $\|u\|_{\mathbb{A}_s(\mathcal{T})}$ and $\|u\|_{\mathbb{A}_s}$. It essentially shows that, if an algebraic rate of convergence is possible starting with a mesh $\mathcal{T} \in \mathrm{refine}(\mathcal{T}_0)$, then the same rate can also be realized by starting with the coarser initial mesh $\mathcal{T}_0$.

**Lemma 4.18.** *There exists $C_{\mathbb{A}} > 0$ which depends only on $C_{\mathrm{son}}$, $m_{\mathrm{unif}}$ from Lemma 4.17, and $\mathcal{T}_0$, such that for all $s > 0$ and all $\mathcal{T} \in \mathbb{T}$, it holds that*

$$\sup_{N \geq C_{\mathbb{A}}\,\#\mathcal{T}} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet \Big) \leq 2^s\, \|u\|_{\mathbb{A}_s(\mathcal{T})}, \tag{4.23}$$

*as well as*

$$\sup_{N \geq C_{\mathbb{A}}\,\#\mathcal{T}} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet \Big) \leq C_{\mathrm{mon}}\, 2^s\, \|u\|_{\mathbb{A}_s}. \tag{4.24}$$

*In particular, there holds equivalence*

$$\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty \quad \Longleftrightarrow \quad \|u\|_{\mathbb{A}_s} < \infty. \tag{4.25}$$

In the proof of Lemma 4.18 we exploit the following elementary observation.

**Lemma 4.19.** *For all $\mathcal{T}_\bullet \in \mathbb{T}$ and $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, it holds*

$$\#\mathcal{T}_\circ - \#\mathcal{T}_\bullet + 1 \leq \#\mathcal{T}_\circ \leq \#\mathcal{T}_\bullet\,(\#\mathcal{T}_\circ - \#\mathcal{T}_\bullet + 1). \tag{4.26}$$

*Proof.* Note that $\big(\#\mathcal{T}_\circ - \#\mathcal{T}_\bullet + 1\big) - \#\mathcal{T}_\circ/\#\mathcal{T}_\bullet = \big(\#\mathcal{T}_\circ - \#\mathcal{T}_\bullet\big)\big(1 - 1/\#\mathcal{T}_\bullet\big) \geq 0$. Rearranging the terms, we conclude the upper bound in (4.26), while the lower bound is obvious. □

***Proof of Lemma 4.18 .*** We split the proof into three steps.

**Step 1: The estimates (4.23)–(4.24) imply (4.25).** For any $M > 0$, the sets $\bigcup_{N=0}^{M} \mathbb{T}_N$ and $\bigcup_{N=0}^{M} \mathbb{T}_N(\mathcal{T})$ are finite. The estimate (4.23) implies that

$$\begin{aligned}
\|u\|_{\mathbb{A}_s} &= \sup_{N \geq 0} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet \Big) \\
&\leq \max_{N \leq C_{\mathbb{A}}\,\#\mathcal{T}} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet \Big) + \sup_{N \geq C_{\mathbb{A}}\,\#\mathcal{T}} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet \Big) \\
&\leq \max_{N \leq C_{\mathbb{A}}\,\#\mathcal{T}} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet \Big) + 2^s\, \|u\|_{\mathbb{A}_s(\mathcal{T})}.
\end{aligned}$$

This provides an upper bound to $\|u\|_{\mathbb{A}_s}$ in terms of $\|u\|_{\mathbb{A}_s(\mathcal{T})}$, up to some finite summand. Therefore, $\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty$ implies $\|u\|_{\mathbb{A}_s} < \infty$. Using (4.24) instead of (4.23), the converse implication follows analogously.

**Step 2: Verification of (4.23).** Let $\mathcal{T} \in \mathbb{T}$ be arbitrary and let $\widetilde{N} \geq 0$. Apply Lemma 4.17 to see that the $m_{\mathrm{unif}}$-times uniform refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$ satisfies $\#\mathcal{T} \leq \#\widehat{\mathcal{T}} \leq C_{\mathrm{son}}^{m_{\mathrm{unif}}} \#\mathcal{T} =: C$. This implies that $\widehat{\mathcal{T}} \in \mathbb{T}_C(\mathcal{T}) \subseteq \mathbb{T}_{C+\widetilde{N}}(\mathcal{T})$ and hence, $\mathbb{T}_{C+\widetilde{N}}(\mathcal{T}) \neq \emptyset$. Choose the optimal mesh $\mathcal{T}_\circ \in \mathbb{T}_{C+\widetilde{N}}(\mathcal{T})$ with $\eta_\circ = \min_{\mathcal{T}_\bullet \in \mathbb{T}_{C+\widetilde{N}}(\mathcal{T})} \eta_\bullet > 0$. Then, we estimate

$$\#\mathcal{T}_\circ - \#\mathcal{T}_0 = (\#\mathcal{T}_\circ - \#\mathcal{T}) + (\#\mathcal{T} - \#\mathcal{T}_0) \leq (C + \widetilde{N}) + \#\mathcal{T} \leq 2C + \widetilde{N},$$

i.e., $\mathcal{T}_\circ \in \mathbb{T}_{2C+\widetilde{N}}$. By choice of $\mathcal{T}_\circ \in \mathbb{T}_{C+\widetilde{N}}(\mathcal{T})$ and the definition of $\|u\|_{\mathbb{A}_s(\mathcal{T})}$, it follows

$$\begin{aligned}(2C + \widetilde{N} + 1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{2C+\widetilde{N}}} \eta_\bullet &\leq \Big(\frac{2C + \widetilde{N} + 1}{C + \widetilde{N} + 1}\Big)^s (C + \widetilde{N} + 1)^s \eta_\circ \\ &\leq 2^s (C + \widetilde{N} + 1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{C+\widetilde{N}}(\mathcal{T})} \eta_\bullet \leq 2^s \|u\|_{\mathbb{A}_s(\mathcal{T})}.\end{aligned} \tag{4.27}$$

Define $C_{\mathbb{A}} := 2C_{\mathrm{son}}^{m_{\mathrm{unif}}}$. Since this estimate holds for all $\widetilde{N} \geq 0$, we obtain with $C_{\mathbb{A}}\#\mathcal{T} = 2C$ that

$$\sup_{N \geq C_{\mathbb{A}}\#\mathcal{T}} \Big((N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N} \eta_\bullet\Big) = \sup_{\widetilde{N} \geq 0} \Big((2C + \widetilde{N} + 1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{2C+\widetilde{N}}} \eta_\bullet\Big) \overset{(4.27)}{\leq} 2^s \|u\|_{\mathbb{A}_s(\mathcal{T})}.$$

This concludes the proof of (4.23) with $C_{\mathbb{A}} = 2C_{\mathrm{son}}^{m_{\mathrm{unif}}}$.

**Step 3: Verification of (4.24).** Let $\widetilde{N} \geq 0$ and let $\widehat{\mathcal{T}} \in \mathbb{T}$ be the $m_{\mathrm{unif}}$-times uniform refinement of $\mathcal{T}_0$. Adopt the notation from Step 2 and recall that $\widehat{\mathcal{T}} \in \mathbb{T}_C \subseteq \mathbb{T}_{C+\widetilde{N}}$. Choose $\mathcal{T}_\circ \in \mathbb{T}_{C+\widetilde{N}}$ with $\eta_\circ = \min_{\mathcal{T}_\bullet \in \mathbb{T}_{C+\widetilde{N}}} \eta_\bullet$. Define $\mathcal{T}_+ := \widehat{\mathcal{T}} \oplus \mathcal{T}_\circ$ to ensure that $\mathcal{T}_+ \in \mathrm{refine}(\widehat{\mathcal{T}})$ and that the discrete solution $U_+ \in \mathcal{X}_\circ$ exists. Then, it holds that

$$\#\mathcal{T}_+ - \#\mathcal{T} \overset{(\mathrm{R4})}{\leq} (\#\widehat{\mathcal{T}} + \#\mathcal{T}_\circ - \#\mathcal{T}_0) - \#\mathcal{T} \leq \#\widehat{\mathcal{T}} + C + \widetilde{N} \leq 2C + \widetilde{N},$$

i.e., $\mathcal{T}_+ \in \mathbb{T}_{2C+\widetilde{N}}(\mathcal{T})$. Moreover, quasi-monotonicity of the estimator (Lemma 4.17) and analogous argumentation as in Step 1 yield that

$$\begin{aligned}(2C + \widetilde{N} + 1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{2C+\widetilde{N}}(\mathcal{T})} \eta_\bullet &\leq (2C + \widetilde{N} + 1)^s \eta_+ \\ &\leq C_{\mathrm{mon}} \Big(\frac{2C + \widetilde{N} + 1}{C + \widetilde{N} + 1}\Big)^s (C + \widetilde{N} + 1)^s \eta_\circ \leq C_{\mathrm{mon}}\, 2^s \|u\|_{\mathbb{A}_s}.\end{aligned} \tag{4.28}$$

Since this estimate holds for all $\widetilde{N} \geq 0$, we again obtain with $C_{\mathbb{A}} = 2\,C_{\mathrm{son}}^{m_{\mathrm{unif}}}$ that

$$\begin{aligned}\sup_{N \geq C_{\mathbb{A}}\#\mathcal{T}} \Big((N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet\Big) &= \sup_{\widetilde{N} \geq 0} \Big((2C + \widetilde{N} + 1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{2C+\widetilde{N}}(\mathcal{T})} \eta_\bullet\Big) \\ &\overset{(4.28)}{\leq} C_{\mathrm{mon}}\, 2^s \|u\|_{\mathbb{A}_s}.\end{aligned}$$

This concludes the proof.

$\square$

### 4.8.2 Data oscillations

In this section we recap the definition of the approximation classes in the spirit of [CKNS08]. To this end, one can also consider approximation classes based on the so-called *total error*, e.g., error plus some suitable data-oscillations. Suppose that the Galerkin solution $U_\bullet \in \mathcal{X}_\bullet$ of (4.3) exists. Further, suppose that $\mathrm{osc}_\bullet : \mathcal{X}_\bullet \to \mathbb{R}$ are suitable oscillation terms such that the error estimator is reliable and efficient in the sense of

$$C_{\mathrm{rel}}^{-1} \, \|u - U_\bullet\|_{\mathcal{H}} \le \eta_\bullet \le C_{\mathrm{eff}} \big( \|u - U_\bullet\|_{\mathcal{H}} + \mathrm{osc}_\bullet(U_\bullet) \big). \tag{4.29}$$

Then, for $\mathcal{T} \in \mathbb{T}$, the work [CKNS08] considers approximation classes defined by

$$\|u\|_{\mathbb{E}_s(\mathcal{T})} := \sup_{N \in \mathbb{N}_0} \Big( (N+1)^s \min_{\substack{\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}) \\ \#\mathcal{T}_\bullet - \#\mathcal{T} \le N}} \inf_{V_\bullet \in \mathcal{X}_\bullet} \big( \|u - V_\bullet\|_{\mathcal{H}} + \mathrm{osc}_\bullet(V_\bullet) \big) \Big). \tag{4.30}$$

Note that the definition of $\|u\|_{\mathbb{E}_s(\mathcal{T})}$ uses $\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T})$ with $\#\mathcal{T}_\bullet - \#\mathcal{T} \le N$ and hence, $\|u\|_{\mathbb{E}_s(\mathcal{T})}$ also involves meshes for which the existence of the discrete solution may fail. The following lemma is an adaptation of [CFPP14, Theorem 4.4]. Starting from some arbitrary initial mesh $\mathcal{T}$, it shows that under the additional assumption of efficiency (4.29), the total error converges with the same algebraic rate as the error estimator.

**Lemma 4.20.** *Let* $\mathrm{osc}_\bullet : \mathcal{X}_\bullet \to \mathbb{R}$ *satisfy* (4.29)*. Suppose that there exists* $C_{\mathrm{osc}} > 0$*, such that for all* $\mathcal{T}_\bullet \in \mathbb{T}$ *for which the discrete solution* $U_\bullet \in \mathcal{X}_\bullet$ *of* (4.3) *exists, it holds the following:*

- $\mathrm{osc}_\bullet := \mathrm{osc}_\bullet(U_\bullet) \le C_{\mathrm{osc}} \, \eta_\bullet$,
- $C_{\mathrm{osc}}^{-1} \, \mathrm{osc}_\bullet(V_\bullet) \le \mathrm{osc}_\bullet(W_\bullet) + \|V_\bullet - W_\bullet\|_{\mathcal{H}}$ *for all* $V_\bullet, W_\bullet \in \mathcal{X}_\bullet$.

*Then, for all* $s > 0$ *and all* $\mathcal{T} \in \mathbb{T}$*, it holds that*

$$\|u\|_{\mathbb{E}_s(\mathcal{T})} < \infty \quad \Longleftrightarrow \quad \|u\|_{\mathbb{A}_s} < \infty.$$

*Proof.* We show that $\|u\|_{\mathbb{E}_s(\mathcal{T})} < \infty$ if and only if $\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty$. Then, (4.25) from Lemma 4.18 will conclude the proof. We split the proof into several steps.

**Step 1:** Let $\mathcal{T} \in \mathbb{T}$ and $\widehat{\mathcal{T}}_0 \in \mathbb{T}$ be the $m_{\mathrm{unif}}$-times uniform refinement of $\mathcal{T}_0$ from Lemma 4.17. With $C := (C_{\mathrm{son}}^{m_{\mathrm{unif}}} - 1)\#\mathcal{T}_0$, the triangulation $\mathcal{T}_+ := \mathcal{T} \oplus \widehat{\mathcal{T}}_0$ satisfies that

$$\#\mathcal{T}_+ \le \#\mathcal{T} + \#\widehat{\mathcal{T}}_0 - \#\mathcal{T}_0 \le \#\mathcal{T} + C,$$

and hence $\mathcal{T}_+ \in \mathbb{T}_C(\mathcal{T})$. This proves that $\mathbb{T}_N(\mathcal{T}) \neq \emptyset$ for $N \ge C$.

**Step 2:** We prove that $\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty$ implies that $\|u\|_{\mathbb{E}_s(\mathcal{T})} < \infty$ by showing that

$$\sup_{N \ge C} \Big( (N+1)^s \min_{\substack{\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}) \\ \#\mathcal{T}_\bullet - \#\mathcal{T} \le N}} \inf_{V_\bullet \in \mathcal{X}_\bullet} \big( \|u - V_\bullet\|_{\mathcal{H}} + \mathrm{osc}_\bullet(V_\bullet) \big) \Big) \lesssim \|u\|_{\mathbb{A}_s(\mathcal{T})}. \tag{4.31}$$

For $N \geq C$, note that Step 1 guarantees $\mathbb{T}_N(\mathcal{T}) \neq \emptyset$. Hence, there exists $\mathcal{T}_\circ \in \mathbb{T}_N(\mathcal{T})$ with discrete solution $U_\circ \in \mathcal{X}_\circ$ and $\eta_\circ = \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet$. With (4.29), we obtain that

$$\min_{\substack{\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}) \\ \#\mathcal{T}_\bullet - \#\mathcal{T} \leq N}} \inf_{V_\bullet \in \mathcal{X}_\bullet} \big(\|u - V_\bullet\|_{\mathcal{H}} + \mathrm{osc}_\bullet(V_\bullet)\big) \leq \|u - U_\circ\|_{\mathcal{H}} + \mathrm{osc}_\circ(U_\circ) \overset{(4.29)}{\simeq} \eta_\circ = \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet.$$

This proves (4.31).

**Step 3:** We prove that $\|u\|_{\mathbb{E}_s(\mathcal{T})} < \infty$ implies that $\|u\|_{\mathbb{A}_s(\mathcal{T})} < \infty$ by showing that

$$\sup_{N \geq C} \Big((N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_\bullet\Big) \leq (C+1)^s \, \|u\|_{\mathbb{E}_s}. \tag{4.32}$$

Let $N \geq 0$. Choose $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T})$ with $\#\mathcal{T}_\circ - \#\mathcal{T} \leq N$ and

$$\big(\|u - V_\circ\|_{\mathcal{H}} + \mathrm{osc}_\circ(V_\circ)\big) = \inf_{V_\bullet \in \mathcal{X}_\bullet} \big(\|u - V_\bullet\|_{\mathcal{H}} + \mathrm{osc}_\bullet(V_\bullet)\big).$$

Define $\mathcal{T}_+ := \mathcal{T}_\circ \oplus \widehat{\mathcal{T}}_0$ and note that $\mathcal{T}_+ \in \mathcal{T}_{N+C}(\mathcal{T})$. Combining (4.29) with the Céa lemma (4.7) and our assumptions on the data oscillations, we obtain that, for all $V_+ \in \mathcal{X}_+$,

$$\begin{aligned} \eta_+ \overset{(4.29)}{\simeq} \|u - U_+\|_{\mathcal{H}} + \mathrm{osc}_+(U_+) &\lesssim \|u - U_+\|_{\mathcal{H}} + \mathrm{osc}_+(V_+) + \|U_+ - V_+\|_{\mathcal{H}} \\ &\lesssim \|u - U_+\|_{\mathcal{H}} + \mathrm{osc}_\circ(V_+) + \|u - V_+\|_{\mathcal{H}} \\ &\overset{(4.7)}{\lesssim} \|u - V_+\|_{\mathcal{H}} + \mathrm{osc}_+(V_+). \end{aligned}$$

This reveals that $\eta_+ \simeq \inf_{V_+ \in \mathcal{X}_+} \big(\|u - V_+\|_{\mathcal{H}} + \mathrm{osc}_+(V_+)\big)$. Then, $\mathcal{T}_+ \in \mathcal{T}_{N+C}(\mathcal{T})$ together with $\mathcal{X}_+ \supseteq \mathcal{X}_\circ$, reveals that

$$\begin{aligned} (N+C+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_{N+C}(\mathcal{T})} \eta_\bullet &\leq (N+C+1)^s \, \eta_+ \\ &\simeq (N+C+1)^s \inf_{V_+ \in \mathcal{X}_+} \big(\|u - V_+\|_{\mathcal{H}} + \mathrm{osc}_+(V_+)\big) \\ &\leq \Big(\frac{N+C+1}{N+1}\Big)^s (N+1)^s \inf_{V_\circ \in \mathcal{X}_\circ} \big(\|u - V_\circ\|_{\mathcal{H}} + \mathrm{osc}_\circ(V_\circ)\big) \\ &\leq (C+1)^s \, \|u\|_{\mathbb{E}_s(\mathcal{T})}. \end{aligned}$$

This proves (4.32) and concludes the proof. □

### 4.8.3 Main result

The following theorem is the main result of this chapter. It states that Algorithm 4.4 does not only guarantee (linear) convergence, but also realizes the best possible algebraic convergence rate for the error estimator.

To that end, suppose that $\|u\|_{\mathbb{A}_s} < \infty$ for some $s > 0$. By definition (4.21) of the approximation class, there exists a sequence of "optimal" meshes $\widehat{\mathcal{T}}_\ell \in \mathbb{T} = \mathrm{refine}(\mathcal{T}_0)$ as well as corresponding error estimators $\widehat{\eta}_\ell$ such that $\widehat{\eta}_\ell \lesssim \big(\#\widehat{\mathcal{T}}_\ell - \#\mathcal{T}_0 + 1\big)^{-s}$ for all $\ell \in \mathbb{N}_0$. Note that these "optimal" triangulations are not necessarily successive refinements but in general even totally unrelated. Therefore, the important implication of the following theorem is that indeed the adaptively generated triangulations $\mathcal{T}_\ell$ yield the same algebraic decay $s > 0$ if the marking parameter $0 < \theta \ll 1$ is sufficiently small. Overall, Algorithm 4.4 thus guarantees that the error estimator decays asymptotically with any possible algebraic rate $s > 0$.

**Theorem 4.21.** *Suppose* (E1)–(E5) *with $\mathcal{X}_\infty = \mathcal{H}$. Employ the notation of Algorithm 4.4. Let $\widehat{\beta}_0 > 0$ be the lower-bound of the* inf-sup *constant* (4.6) *for the uniform refinement $\widehat{\mathcal{T}}_0$ from Lemma 4.17. Let $\ell_3, \ell_5 \in \mathbb{N}_0$ be the indices from Lemma 4.13 and Lemma 4.17 respectively. Define $\ell_6 := \max\{\ell_3, \ell_5\}$. Let $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$. Then, for all $s > 0$, there exists a constant $C_{\rm opt} > 0$ such that*

$$\|u\|_{\mathbb{A}_s} < \infty \quad \Longleftrightarrow \quad \forall \ell \geq \ell_6 \quad \eta_\ell \leq C_{\rm opt} \left(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\right)^{-s}. \tag{4.33}$$

*The constant $C_{\rm opt}$ depends only on $\#\mathcal{T}_{\ell_6}$, $\mathcal{T}_0$, $\theta$, $s$, $\|u\|_{\mathbb{A}_s}$, and validity of* (E1)–(E5)*.*

We emphasize that Axiom (E5) as well as $\mathcal{X}_\infty = \mathcal{H}$ can be enforced by the expanded Dörfler marking strategy from Proposition 4.7. Hence, by using the expanded Dörfler marking, Theorem 4.21 relies only on (E1)–(E4).

The proof of Theorem 4.21 requires the following two technical lemmas, which are adapted versions of [CFPP14, Proposition 4.12] and [CFPP14, Proposition 4.14].

**Lemma 4.22** (optimality of Dörfler marking)**.** *Under the assumptions of Theorem 4.21 and for all $0 < \theta < \theta_{\rm opt}$, there exists some $0 < \kappa_{\rm opt} < 1$, such that for all $\mathcal{T}_\bullet \in$* refine$(\mathcal{T}_{\ell_5})$ *and all $\mathcal{T}_\circ \in$* refine$(\mathcal{T}_\bullet)$*, it holds that*

$$\eta_\circ \leq \kappa_{\rm opt}\, \eta_\bullet \quad \Longrightarrow \quad \theta\, \eta_\bullet^2 \leq \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2, \tag{4.34}$$

*where $\mathcal{R}_{\bullet,\circ}$ is the (enlarged) set of refined elements from* (E4)*.*

*Proof.* Let $\mathcal{T}_\bullet \in$ refine$(\mathcal{T}_{\ell_5})$ and $\mathcal{T}_\circ \in$ refine$(\mathcal{T}_\bullet)$. Then, Lemma 4.17 guarantees that the discrete solutions $U_\bullet \in \mathcal{X}_\bullet$ and $U_\circ \in \mathcal{X}_\circ$ exist. With stability on non-refined elements (E1) and the Young inequality we obtain that, for all $\delta > 0$,

$$\begin{aligned}
\eta_\bullet^2 &= \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 + \eta_\bullet(\mathcal{T}_\bullet \cap \mathcal{T}_\circ)^2 \\
&\leq \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 + (1 + \delta^{-1})\, \eta_\circ(\mathcal{T}_\bullet \cap \mathcal{T}_\circ)^2 + (1+\delta) C_{\rm stb}^2\, \|U_\circ - U_\bullet\|_{\mathcal{H}}^2.
\end{aligned}$$

Recall that discrete reliability (E4) holds with the uniform constant $C_{\rm rel}/\widehat{\beta}_0$. Together with the assumption $\eta_\circ \leq \kappa_{\rm opt}\, \eta_\bullet$ and $\mathcal{T}_\bullet \setminus \mathcal{T}_\circ \subseteq \mathcal{R}_{\bullet,\circ}$, this yields that

$$\begin{aligned}
\eta_\bullet^2 &\leq \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2 + (1 + \delta^{-1}) \kappa_{\rm opt}\, \eta_\bullet^2 + (1+\delta)\, C_{\rm stb}^2\, C_{\rm rel}^2\, \widehat{\beta}_0^{-2}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2 \\
&= (1 + \delta^{-1})\, \kappa_{\rm opt}\, \eta_\bullet^2 + \big(1 + (1+\delta)\, C_{\rm stb}^2\, C_{\rm rel}^2\, \widehat{\beta}_0^{-2}\big)\, \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2.
\end{aligned}$$

Rearranging the terms in the latter estimate, we obtain that

$$\widetilde{\theta}\, \eta_\bullet^2 := \frac{1 - (1+\delta^{-1})\, \kappa_{\rm opt}}{1 + (1+\delta)\, C_{\rm stb}^2\, C_{\rm rel}^2\, \widehat{\beta}_0^{-2}}\, \eta_\bullet^2 \leq \eta_\bullet(\mathcal{R}_{\bullet,\circ})^2. \tag{4.35}$$

It remains to prove that $\widetilde{\theta} \geq \theta$ and hence, (4.35) implies (4.34). To this end, choose $\delta > 0$ sufficiently small and determine $0 < \kappa_{\rm opt} < 1$ such that

$$\theta \leq \widetilde{\theta} := \frac{1 - (1+\delta^{-1})\,\kappa_{\rm opt}}{1 + (1+\delta)\,C_{\rm stb}^2\,C_{\rm rel}^2\,\widehat{\beta}_0^{-2}} < \frac{1}{1+(1+\delta)C_{\rm stb}^2 C_{\rm rel}^2 \widehat{\beta}_0^{-2}} < \frac{1}{1 + C_{\rm stb}^2\,C_{\rm rel}^2\,\widehat{\beta}_0^{-2}} = \theta_{\rm opt}.$$

This concludes the proof. □

**Lemma 4.23.** *Suppose the assumptions of Theorem 4.21 and let $0 < \theta < \theta_{\rm opt}$. There exist constants $C_1, C_2 > 0$ such that for all $\ell \geq \ell_5$, there exists a corresponding set $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$ such that the following holds: For all $s > 0$ with $\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} < \infty$, it holds that*

$$\#\mathcal{R}_\ell \leq C_1\,(C_2\,\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})})^{1/s}\,\eta_\ell^{-1/s}, \tag{4.36}$$

*as well as the Dörfler marking criterion*

$$\theta\,\eta_\ell^2 \leq \eta_\ell(\mathcal{R}_\ell)^2. \tag{4.37}$$

*The constants $C_1, C_2$ depends only on $\theta$, $\widehat{\beta}_0$, and* (E1)–(E4).

*Proof.* If $\eta_\ell = 0$, the claim (4.36)–(4.37) is satisfied with $\mathcal{R}_\ell := \mathcal{T}_\ell$. Thus, we may suppose that $\eta_\ell > 0$. We split the remainder of the proof into three steps.

**Step 1: Construction of mesh $\mathcal{T}_\bullet$ and $\mathcal{R}_\ell := \mathcal{R}_{\ell,\bullet}$.** Let $\ell \geq \ell_5$ and $\varepsilon := C_{\rm mon}^{-1}\,\kappa_{\rm opt}\,\eta_\ell > 0$. Quasi-monotonicity of the estimator (Lemma 4.17) yields that

$$\varepsilon \leq \kappa_{\rm opt}\,\eta_{\ell_5} < \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} < \infty.$$

Choose the minimal $N \in \mathbb{N}_0$ such that $\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} \leq \varepsilon\,(N+1)^s$. This implies that $\varepsilon < \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} \leq \varepsilon(N+1)^s$ and hence $N \geq 1$. Note that $\mathcal{T}_{\ell_5} \in \mathbb{T}_N(\mathcal{T}_{\ell_5})$ and hence $\mathbb{T}_N(\mathcal{T}_{\ell_5}) \neq \emptyset$.

Choose $\mathcal{T}_\varepsilon \in \mathbb{T}_N(\mathcal{T}_{\ell_5})$ with $\eta_\varepsilon = \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_\bullet$. Define $\mathcal{T}_\bullet := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$. Recall that all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_{\ell_5})$ and corresponding spaces $\mathcal{X}_\circ \supseteq \mathcal{X}_{\ell_5}$ provide unique solutions of the discrete formulation (4.3). Hence, we obtain existence of the Galerkin solution $U_\bullet \in \mathcal{X}_\bullet$. Last, we define $\mathcal{R}_\ell := \mathcal{R}_{\ell,\bullet}$ as the set provided by discrete reliability (E4).

**Step 2: Optimality of Dörfler marking yields (4.37).** With the quasi-monotonicity of the estimator (Lemma 4.17) and the definition (4.21) of the approximation class, the choice of $N$ in Step 1 yields that

$$\eta_\bullet \leq C_{\rm mon}\,\eta_\varepsilon = C_{\rm mon} \min_{\mathcal{T}_\circ \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_\circ \overset{(4.21)}{\leq} C_{\rm mon}\,(N+1)^{-s}\,\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} \leq C_{\rm mon}\,\varepsilon = \kappa_{\rm opt}\,\eta_\ell.$$

This implies $\eta_\bullet \leq \kappa_{\rm opt}\,\eta_\ell$ and hence Lemma 4.22 proves (4.37).

**Step 3: Verification of (4.36).** Recall that $\mathcal{T}_\ell, \mathcal{T}_\varepsilon \in \mathrm{refine}(\mathcal{T}_{\ell_5})$ and $\mathcal{T}_\varepsilon \in \mathbb{T}_N(\mathcal{T}_{\ell_5})$. The definition $\mathcal{R}_\ell = \mathcal{R}_{\ell,\bullet}$ together with discrete reliability (E4), splitting property (R3), and the overlay estimate (R4) yields that

$$\#\mathcal{R}_\ell \overset{(E4)}{\leq} C_{\rm rel}\,\#(\mathcal{T}_\ell \setminus \mathcal{T}_\bullet) \overset{(R3)}{\leq} C_{\rm rel}\,(\#\mathcal{T}_\bullet - \#\mathcal{T}_\ell) \overset{(R4)}{\leq} C_{\rm rel}\,(\#\mathcal{T}_\varepsilon - \#\mathcal{T}_{\ell_5}) \leq C_{\rm rel}\,N. \tag{4.38}$$

Finally, minimality of $N$ in Step 1 implies that $\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} > \varepsilon N^s$ and hence

$$N < \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}^{1/s} \varepsilon^{-1/s} = C_3 \, \eta_\ell^{-1/s},$$

with $C_3 := \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}^{1/s} (C_{\mathrm{mon}}^{-1} \kappa_{\mathrm{opt}})^{-1/s} = (C_{\mathrm{mon}} \kappa_{\mathrm{opt}}^{-1} \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})})^{1/s}$. Altogether, we thus see

$$\#\mathcal{R}_\ell \overset{(4.38)}{\leq} C_{\mathrm{rel}} \, N < C_{\mathrm{rel}} \, C_3 \, \eta_\ell^{-1/s}.$$

This proves (4.36) with $C_1 = C_{\mathrm{rel}}$ and $C_2 = C_{\mathrm{mon}} \, \kappa_{\mathrm{opt}}^{-1}$. □

With optimality of Dörfler marking and Lemma 4.23 at hand, we can prove optimal algebraic convergence rates for Algorithm 4.4. The proof follows arguments from [CFPP14, Proposition 4.15] and corrects a small bug in the proof of [BHP17, Theorem 26].

***Proof of Theorem 4.21.*** We split the proof into two steps.

**Step 1: Implication "$\Longleftarrow$".** Suppose $\eta_\ell \leq C_{\mathrm{opt}} \big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\big)^{-s}$ for all $\ell \geq \ell_6$. According to Lemma 4.18 it is sufficient to prove $\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_6})} < \infty$. For $N \in \mathbb{N}$ with $N \geq \#\mathcal{T}_{\ell_6}$, choose the largest $\ell_N$ such that $\#\mathcal{T}_{\ell_N} - \#\mathcal{T}_{\ell_6} \leq N$. Due to maximality of $N$, we obtain with the splitting property (R3) that

$$N + 1 \leq \#\mathcal{T}_{\ell_N+1} - \#\mathcal{T}_{\ell_6} + 1 \overset{\text{(R3)}}{\lesssim} \#\mathcal{T}_{\ell_N} - \#\mathcal{T}_{\ell_6} + 1. \tag{4.39}$$

Note that $\#\mathcal{T}_{\ell_N} - \#\mathcal{T}_{\ell_6} \leq N$ directly implies $\mathcal{T}_{\ell_N} \in \mathbb{T}_N(\mathcal{T}_{\ell_6})$ and $\ell_N \geq \ell_6$. Together with the assumption, this yields that

$$\begin{aligned}\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_6})} = \sup_{N \geq 0} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T}_{\ell_6})} \eta_\bullet \Big) &\overset{(4.39)}{\lesssim} \sup_{N \geq 0} \Big( (\#\mathcal{T}_{\ell_N} - \#\mathcal{T}_{\ell_6} + 1)^s \, \eta_{\ell_N} \Big) \\ &\leq \sup_{N \geq 0} \Big( (\#\mathcal{T}_{\ell_N} - \#\mathcal{T}_{\ell_0} + 1)^s \, \eta_{\ell_N} \Big) \lesssim 1.\end{aligned}$$

This concludes Step 1.

**Step 2: Implication "$\Longrightarrow$".** To this end, suppose that $\|u\|_{\mathbb{A}_s} < \infty$. Lemma 4.18 then implies $\|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} < \infty$. For all $\ell \geq \ell_6 = \max\{\ell_3, \ell_5\}$, let $\mathcal{M}_\ell$ be the set of marked elements in the $\ell$-th step of Algorithm 4.4. According to Lemma 4.23, there exists $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$ with (4.36)–(4.37). Because of the minimal cardinality of $\mathcal{M}_\ell$ (cf. Step (iv) in Algorithm 4.4), it follows that

$$\#\mathcal{M}_\ell \overset{(4.37)}{\leq} C_{\mathrm{mark}} \, \#\mathcal{R}_\ell \overset{(4.36)}{\leq} C_{\mathrm{mark}} \, C_1 \, (C_2 \, \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})})^{1/s} \, \eta_\ell^{-1/s}.$$

The mesh-closure estimate (R5) yields that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_{\ell_0} + 1 \leq C_{\mathrm{mesh}} \sum_{j=\ell_0}^{\ell-1} \#\mathcal{M}_j. \tag{4.40}$$

Together with $C := \max_{j=0,\dots,\ell_6} \frac{\#\mathcal{M}_j}{\#\mathcal{M}_{\ell_6}}$, we obtain that

$$
\begin{aligned}
\#\mathcal{T}_\ell - \#\mathcal{T}_{\ell_0} + 1 &\leq C_{\mathrm{mesh}} \sum_{j=\ell_0}^{\ell-1} \#\mathcal{M}_j \\
&\leq C_{\mathrm{mesh}} \Big( \sum_{j=\ell_0}^{\ell_6} \#\mathcal{M}_j + \sum_{j=\ell_6}^{\ell-1} \#\mathcal{M}_j \Big) \\
&\leq C_{\mathrm{mesh}} \,(\ell_6\, C + 1) \sum_{j=\ell_6}^{\ell-1} \#\mathcal{M}_j \\
&\leq C_{\mathrm{mesh}} \,(\ell_6\, C + 1)\, C_{\mathrm{mark}}\, C_1 \,(C_2 \, \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})})^{1/s} \sum_{j=\ell_6}^{\ell} \eta_j^{-1/s}.
\end{aligned}
\tag{4.41}
$$

The linear convergence from Theorem 4.14 reads

$$
\eta_\ell \leq C_{\mathrm{lin}}\, q_{\mathrm{lin}}^{\ell-j}\, \eta_j \quad \text{for all } \ \ell_3 \leq j \leq \ell. \tag{4.42}
$$

Hence, this implies that

$$
\eta_j^{-1/s} \leq C_{\mathrm{lin}}^{1/s}\, q_{\mathrm{lin}}^{(\ell-j)/s}\, \eta_\ell^{-1/s} \quad \text{for all } \ \ell_3 \leq j \leq \ell.
$$

Since there holds $0 < q := q_{\mathrm{lin}}^{1/s} < 1$, the geometric series applies and the sum in (4.41) can be estimated by

$$
\sum_{j=\ell_6}^{\ell} \eta_j^{-1/s} \leq C_{\mathrm{lin}}^{1/s}\, \eta_\ell^{-1/s} \sum_{j=\ell_6}^{\ell} q^{(\ell-j)} \leq \frac{C_{\mathrm{lin}}^{1/s}}{1 - q_{\mathrm{lin}}^{1/s}}\, \eta_\ell^{-1/s}.
$$

Combining this estimate with (4.41), we derive

$$
\#\mathcal{T}_\ell - \#\mathcal{T}_{\ell_0} + 1 \leq \frac{C_{\mathrm{mesh}}\,(\ell_6\, C + 1)\, C_{\mathrm{mark}}\, C_1}{1 - q_{\mathrm{lin}}^{1/s}} (C_{\mathrm{lin}}\, C_2 \, \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})})^{1/s}\, \eta_\ell^{-1/s}.
$$

Rearranging these terms, we see that $\eta_\ell \lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_{\ell_0} + 1)^{-s}$. Using the definitions of $C_1, C_2 > 0$, this implies (4.33) with

$$
C_{\mathrm{opt}} := \left( \frac{\#\mathcal{T}_{\ell_6}\, C_{\mathrm{mesh}}\,(\ell_6\, C + 1)\, C_{\mathrm{mark}}\, C_{\mathrm{rel}}}{(1 - q_{\mathrm{lin}}^{1/s})} \right)^s C_{\mathrm{lin}}\, C_{\mathrm{mon}}\, \kappa_{\mathrm{opt}}^{-1} \, \|u\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}.
$$

This concludes the proof. $\square$

As a consequence of Section 4.8.2, Theorem 4.21 transfers into the setting of [CKNS08]. To that end, suppose $\mathrm{osc}_\bullet : \mathcal{X}_\bullet \to \mathbb{R}$ such that (4.29) and the assumptions of Lemma 4.20 are satisfied. Then, Lemma 4.20 implies that $\|u\|_{\mathbb{A}_s} < \infty \Longleftrightarrow \|u\|_{\mathbb{E}_s(\mathcal{T})} < \infty$ for all $\mathcal{T} \in \mathbb{T}$. This gives rise to the following remark.

**Remark 4.24.** *Under the assumptions of Theorem 4.21 and the assumptions on the data oscillations in Lemma 4.20, there holds the following: For all* $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$ *and for all* $s > 0$*, there exists a constant* $\widetilde{C_{\rm opt}} > 0$ *such that*

$$\|u\|_{\mathbb{E}_s(\mathcal{T}_0)} < \infty \quad \Longleftrightarrow \quad \forall \ell \geq \ell_6 \quad \eta_\ell \leq \widetilde{C_{\rm opt}} \big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\big)^{-s}. \tag{4.43}$$

*The constant* $\widetilde{C_{\rm opt}}$ *depends only on* $\#\mathcal{T}_{\ell_6}$, $\mathcal{T}_0$, $\theta$, $s$, $\|u\|_{\mathbb{E}_s(\mathcal{T}_0)}$, *validity of* (4.29), *and* (E1)–(E5).

# 5 Adaptive FEM for compactly perturbed problems

In this chapter, we apply the abstract framework of Chapter 4 to adaptive finite elements. To that end, we introduce the model problem in Section 5.1, which consists of a general diffusion problem with convection and reaction. Further, Section 5.1 and Section 5.2 ensure that the model problem fits in the abstract framework and that the corresponding error estimator satisfies the estimator axioms (E1)–(E4). Further, Section 5.3 recaps the main result of Chapter 4 in the current setting. At the end of this chapter (Section 5.4), we underpin our theoretical findings with some numerical experiment for the two dimensional Helmholtz equation.

This chapter is based on the work [BHP17], where besides the abstract framework of Chapter 4, optimal algebraic rates of convergence for adaptive finite elements for compactly perturbed problems are shown.

## 5.1 Model problem

Let $\Omega \subset \mathbb{R}^d$ with $d = 2$ or $d = 3$ be a polygonal resp. polyhedral Lipschitz domain with boundary $\Gamma := \partial\Omega$. Recall that $L^2(\Omega)$, $H^1_0(\Omega)$ denote the usual Lebesgue and Sobolev spaces and $(f\,,\,g) := \int_\Omega f\,g\,dx$ denotes the $L^2(\Omega)$ scalar product; see Chapter 2. We consider general diffusion problems with convection and reaction of the following type.

Let $c \in L^\infty(\Omega)$, $b \in L^\infty(\Omega)^d$, and $A \in L^\infty(\Omega)^{d\times d}$ be given coefficients such that $A(x) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is symmetric and uniformly positive definite, i.e., it holds that

$$\operatorname*{ess\,inf}_{x\in\Omega} \min_{\xi\in\mathbb{R}^d\setminus\{0\}} \frac{A\xi\cdot\xi}{|\xi|^2} \geq \alpha > 0. \tag{5.1}$$

Then, the model problem reads as follows: Given $f \in L^2(\Omega)$, find $u \in H^1_0(\Omega)$ such that

$$\begin{aligned} -\operatorname{div}(A\,\nabla u) + b\cdot\nabla u + cu &= f \qquad \text{in } \Omega,\\ u &= 0 \qquad \text{on } \Gamma. \end{aligned} \tag{5.2}$$

Possible examples include the weak formulation of the Helmholtz equation

$$-\Delta u - k^2 u = f \text{ in } \Omega \quad \text{subject to} \quad u = 0 \text{ on } \Gamma, \tag{5.3}$$

where $k > 0$ denotes the wavenumber. We emphasize that homogeneous Dirichlet conditions are only considered for the ease of presentation, while (inhomogeneous) mixed Dirichlet–Neumann–Robin boundary conditions can be included as in [FPP14, AFK$^+$13, CFPP14].

The weak formulation of (5.2) reads as follows: Given $f \in L^2(\Omega)$, find $u \in H^1_0(\Omega)$ such that

$$b(u\,,\,v) := (A\,\nabla u\,,\,\nabla v) + (b\cdot\nabla u + cu\,,\,v) = \langle f\,,\,v\rangle \quad \text{for all } v \in H^1_0(\Omega). \tag{5.4}$$

To guarantee unique solvability of the weak form (5.4), we suppose that (5.4) is well-posed in the sense of (4.4). In order to fit in the abstract framework of Chapter 4, we define the bilinear form $a(\cdot\,,\,\cdot) : H^1_0(\Omega)\times H^1_0(\Omega) \to \mathbb{R}$ by

$$a(u\,,\,v) := (A\nabla u\,,\,\nabla v) \quad \text{for all } u,v \in H^1_0(\Omega). \tag{5.5}$$

Furthermore, the linear operator $\mathfrak{C} : H^1_0(\Omega) \to H^{-1}(\Omega)$ is given by

$$\mathfrak{C}\,u := b\cdot\nabla u + cu \quad \text{for all } u \in H^1_0(\Omega). \tag{5.6}$$

The next proposition recaps some properties of the induced bilinear forms and the operator $\mathfrak{C}$. The proof follows the ideas of [FFP14, Lemma 3.4].

**Proposition 5.1.** *The bilinear forms $b(\cdot\,,\,\cdot)$, $a(\cdot\,,\,\cdot)$ and the linear operator $\mathfrak{C}$ satisfy the following properties:*

(a) *The bilinear form $b(\cdot\,,\,\cdot)$ is well defined and bounded with*

$$|b(w\,,\,v)| \le C_{\rm cont}\,\|\nabla w\|_{L^2(\Omega)}\,\|\nabla v\|_{L^2(\Omega)} \quad \textit{for all } w,v \in H^1_0(\Omega),$$

*where $C_{\rm cont} > 0$ just depends on the coefficients $A, b, c$ as well the Poincaré constant of $\Omega$.*

(b) *The bilinear $a(\cdot\,,\,\cdot)$ is symmetric, continuous, and elliptic with*

$$|a(w\,,\,v)| \le C_{\rm cont}\,\|\nabla w\|_{L^2(\Omega)}\,\|\nabla v\|_{L^2(\Omega)} \quad \textit{and} \quad \alpha\,\|v\|_{H^1_0(\Omega)} \le a(v\,,\,v)$$

*for all $u,v \in H^1_0(\Omega)$. The constants $C_{\rm cont} > 0$ and $\alpha$ just depend on $A$.*

(c) *The linear operator $\mathfrak{C} : H^1_0(\Omega) \to H^{-1}(\Omega)$ is bounded and compact.*

*Proof.* We split the proof into two steps.

**Step 1: Proof of** (a) **and** (b). Let $w,v \in H^1_0(\Omega)$. We estimate with the Cauchy–Schwarz inequality and the Poincaré inequality that

$$\begin{aligned}|b(w\,,\,v)| &\le \|A\nabla w\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} + \|b\cdot\nabla w\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + \|c\,w\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}\\ &\le \big(\|A\|_{L^\infty(\Omega)} + C_\Omega\,\|b\|_{L^\infty(\Omega)} + C_\Omega^2\,\|c\|_{L^\infty(\Omega)}\big)\|\nabla w\|_{L^2(\Omega)}\,\|\nabla v\|_{L^2(\Omega)},\end{aligned}$$

where $C_\Omega > 0$ denotes the Poincaré constant of $\Omega$. This implies (a).

Recall that, $A(x) \in \mathbb{R}^{d\times d}_{\rm sym}$ is symmetric and uniformly positive definite, hence $a(\cdot\,,\,\cdot)$ is elliptic in the sense of (4.1) resp. (4.12) (see Section 4.2.2). Continuity of $a(\cdot\,,\,\cdot)$ follows analogously to $b(\cdot\,,\,\cdot)$ with constant $C_{\rm cont} = \|A\|_{L^\infty(\Omega)}$ and concludes (b).

**Step 2: Proof of** (c). We define the operator $\widetilde{\mathfrak{C}} : H^1_0(\Omega) \to L^2(\Omega)$ by $\widetilde{\mathfrak{C}}\,w := \mathfrak{C}\,w$. Clearly $\widetilde{\mathfrak{C}}$ is linear and continuous. According to the Rellich compactness theorem, the embedding $\iota : H^1_0(\Omega) \to L^2(\Omega)$ is a compact operator. Schauder's theorem (see, e.g., [Rud91, Theorem 4.19]) implies that the adjoint operator $\iota^\star : L^2(\Omega) \to H^{-1}(\Omega)$ is also compact and coincides with the natural embedding. Therefore, we may write $\mathfrak{C}$ as

$$\mathfrak{C} = \iota^\star \circ \widetilde{\mathfrak{C}} : H^1_0(\Omega) \to L^2(\Omega) \to H^{-1}(\Omega).$$

Recall that the composition of bounded operator and a compact operator is compact. This concludes the proof. □

For the discretization, we consider standard finite element spaces based on regular triangulations (see Section 3.1). For mesh-refinement, we employ NVB (see Section 3.5), which satisfies the refinement axioms (R1)–(R6).

Recall that, for a given initial mesh $\mathcal{T}_0$, we denote the sets of all possible refinements by $\mathbb{T}$. Further, for a given $\mathcal{T}_\bullet \in \mathbb{T}$ , let $\mathcal{S}^p(\mathcal{T}_\bullet)$ be the space of globally continuous piecewise polynomials. To abbreviate notation let $\mathcal{S}^p_0(\mathcal{T}_\bullet) := \mathcal{S}^p(\mathcal{T}_\bullet) \cap H^1_0(\Omega)$ denote the space of continuous piecewise polynomials which vanish on the boundary. Then, the Galerkin formulation of (5.2) reads as follows: Given $f \in L^2(\Omega)$, find $U_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\bullet)$ such that

$$b(U_\bullet\,,\, V_\bullet) = \langle f\,,\, V_\bullet\rangle \quad \text{for all } V_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\bullet). \tag{5.7}$$

Proposition 5.1 guarantees that the model problem (5.2) fits in the abstract framework of Chapter 4 with $\mathcal{H} := H^1_0(\Omega)$ and $\mathcal{X}_\bullet := \mathcal{S}^p_0(\mathcal{T}_\bullet)$. Further, NVB guarantees nestedness of the discrete spaces $\mathcal{S}^p_0(\mathcal{T}_\bullet) \subseteq \mathcal{S}^p_0(\mathcal{T}_\circ)$ for all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$. We emphasize that iterated uniform mesh-refinement guarantees $\|h_\bullet\|_{L^\infty(\Omega)} \to 0$ and hence, leads to a dense sequence of subspaces. A rigorous proof is given in Lemma 5.6. Then, existence of solutions of (5.2) is guaranteed by Proposition 4.1.

In order to utilize the analysis and hence, obtain optimal algebraic convergence rates, it still remains to define a suitable *a posteriori* error estimator and validate the estimator axioms (E1)–(E5).

### 5.1.1 Weighted-residual error estimator

According to Section 4.3, we define the local contributions of the usual weighted-residual error estimator for the general diffusion problem (5.2) as follows. Suppose that $A$ is a piecewise Lipschitz diffusion coefficient with $A|_{T_0} \in W^{1,\infty}(T_0)$ for all $T_0 \in \mathcal{T}_0$. The space $W^{1,\infty}(T_0)$ is given by

$$W^{1,\infty}(T_0) := \big\{ f \in L^\infty(T_0) \,:\, \nabla f \in L^\infty(T_0) \text{ exists in the weak sense} \big\}.$$

For all $\mathcal{T}_\bullet \in \mathbb{T}$ and $T \in \mathcal{T}_\bullet$ the element contributions are given by

$$\eta_\bullet(T)^2 = h_T^2\, \|f + \mathrm{div}(A\nabla U_\bullet) - b\cdot\nabla U_\bullet - c\,U_\bullet\|^2_{L^2(T)} + h_T\, \|[(A\nabla U_\bullet)\cdot\mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)}, \tag{5.8}$$

where $[(\cdot)\cdot\mathbf{n}]$ denotes the normal jump over interior facets and $h_T := |T|^{1/d} \simeq \mathrm{diam}(T)$ denotes the local mesh size (see Chapters 2–3). For the Helmholtz problem (5.3), these local contributions simplify to

$$\eta_\bullet(T)^2 = h_T^2\, \|f + \Delta U_\bullet + k^2\, U_\bullet\|^2_{L^2(T)} + h_T\, \|[\nabla U_\bullet\cdot\mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)}. \tag{5.9}$$

Moreover, the related error estimator is given by

$$\eta_\bullet(\mathcal{U}_\bullet) := \Big( \sum_{T\in\mathcal{U}_\bullet} \eta_\bullet(T)^2 \Big)^{1/2} \quad \text{for all subsets } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet.$$

### 5.1.2 Adaptive algorithm

In this section, we recap the adaptive algorithm in the current setting. It combines Algorithm 4.4 with the expanded Dörfler marking strategy.

**Algorithm 5.2.** INPUT: *Parameters $0 < \theta \le 1$ and $C_{\rm mark} \ge 1$ as well as the initial triangulation $\mathcal{T}_0$ with $U_{-1} := 0 \in \mathcal{S}_0^p(\mathcal{T}_0)$ and $\eta_{-1} := 1$.*
ADAPTIVE LOOP: *For all $\ell = 0, 1, 2, \dots$, iterate the following steps* (i)–(vi)*:*

(i) `If` (5.7) *does not admit a unique solution in $\mathcal{S}_0^p(\mathcal{T}_\ell)$:*
   - *Define $U_\ell := U_{\ell-1} \in \mathcal{S}_0^p(\mathcal{T}_0)$ and $\eta_\ell := \eta_{\ell-1}$,*
   - *Let $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ be the uniform refinement of $\mathcal{T}_\ell$,*
   - *Increase $\ell \to \ell + 1$, and continue with Step* (i).

(ii) `Else` *compute the unique solution $U_\ell \in \mathcal{S}_0^p(\mathcal{T}_\ell)$ to* (5.7).

(iii) *Compute the corresponding indicators $\eta_\ell(T)$ for all $T \in \mathcal{T}_\ell$.*

(iv) *Determine a set $\mathcal{M}'_\ell \subseteq \mathcal{T}_\ell$ of up to the multiplicative factor $C_{\rm mark}$ minimal cardinality such that $\theta\, \eta_\ell^2 \le \eta_\ell(\mathcal{M}'_\ell)^2$.*

(v) *Find $\mathcal{M}''_\ell \subseteq \mathcal{T}_\ell$ such that $\#\mathcal{M}''_\ell = \#\mathcal{M}'_\ell$ as well as $h_\ell(T) \ge h_\ell(T')$ for all $T \in \mathcal{M}''_\ell$ and $T' \in \mathcal{T}_\ell \setminus \mathcal{M}''_\ell$. Define $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \mathcal{M}''_\ell$.*

(vi) *Generate $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, increase $\ell \to \ell + 1$, and continue with Step* (i).

OUTPUT: *Sequences of successively refined triangulations $\mathcal{T}_\ell$, discrete solutions $U_\ell$, and corresponding estimators $\eta_\ell$.*

Apart from Step (iv) and Step (v), Algorithm 5.2 coincides with Algorithm 4.4 of Chapter 4. This step realize the expanded Dörfler marking of from Proposition 4.7 which guarantees (E5).

## 5.2 Verification of the axioms

In this section, we prove that the weighted-residual error estimator from (5.8) satisfies the estimator axioms (E1)–(E4). Further, we show that Algorithm 5.2 guarantees (E5) and hence all assumptions of the abstract framework of Chapter 4 are met. We note that the proofs of (E1)–(E4) are well known in the literature; see, e.g., [CKNS08, CFPP14, FFP14]. For the sake of completeness, we sketch the most important steps.

**Proposition 5.3** (stability on non-refined element domains)**.** *There exists $C_{\rm stb} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, the following holds: Provided there exist unique discrete solutions $U_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$ and $U_\circ \in \mathcal{S}_0^p(\mathcal{T}_\circ)$, it holds that*

$$\big|\eta_\circ(\mathcal{U}_\bullet) - \eta_\bullet(\mathcal{U}_\bullet)\big| \le C_{\rm stb}\, \|U_\circ - U_\bullet\|_{H^1(\Omega)} \quad \text{for all } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet \cap \mathcal{T}_\circ.$$

*In particular, there holds* (E1)*. The constant $C_{\rm stb} \ge 1$ just depends on the given data, the polynomial degree, and on $\gamma$-shape regularity.*

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$ such that $\mathcal{T}_\circ \in \operatorname{refine}(\mathcal{T}_\bullet)$. Suppose that the corresponding discrete solutions $U_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$ and $U_\circ \in \mathcal{S}_0^p(\mathcal{T}_\circ)$ exist. Let $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet \cap \mathcal{T}_\circ$. The triangle inequality reveals that

$$\Big|\eta_\circ(\mathcal{U}_\bullet) - \eta_\bullet(\mathcal{U}_\bullet))\Big| = \Big|\Big(\sum_{T\in\mathcal{U}_\bullet} \eta_\circ(\mathcal{U}_\bullet)\Big)^{1/2} - \Big(\sum_{T\in\mathcal{U}_\bullet} \eta_\bullet(\mathcal{U}_\bullet)\Big)^{1/2}\Big| \leq \Big(\sum_{T\in\mathcal{U}_\bullet} R_\bullet(T)^2\Big)^{1/2},$$

where $R_\bullet(T)$ is given by

$$\begin{aligned} R_\bullet(T) &= h_T \,\|\operatorname{div}(A\,\nabla(U_\circ - U_\bullet)) - b\cdot\nabla(U_\circ - U_\bullet) - c(U_\circ - U_\bullet)\|_{L^2(T)} \\ &\qquad + h_T^{1/2}\,\|[A\,\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|_{L^2(\partial T\cap\Omega)} \\ &\leq h_T\,\|\operatorname{div}(A\,\nabla(U_\circ - U_\bullet))\|_{L^2(T)} + h_T\,\big(\|b\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\big)\|U_\circ - U_\bullet\|_{H^1(T)} \\ &\qquad + h_T^{1/2}\,\|[A\,\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|_{L^2(\partial T\cap\Omega)}. \end{aligned} \tag{5.10}$$

The product rule as well as the inverse estimate imply for the first term on the right-hand side that

$$\begin{aligned} \|\operatorname{div}(A\nabla(U_\circ - U_\bullet))\|_{L^2(T)} &\leq \|\nabla A\|_{L^\infty(\Omega)}\|\nabla(U_\circ - U_\bullet)\|_{L^2(T)} + \|A\|_{L^\infty(\Omega)}\|\Delta(U_\circ - U_\bullet)\|_{L^2(T)} \\ &\leq \big(\|\nabla A\|_{L^\infty(\Omega)} + C_{\mathrm{inv}}\,h_T^{-1}\,\|A\|_{L^\infty(\Omega)}\big)\,\|\nabla(U_\circ - U_\bullet)\|_{L^2(T)}. \end{aligned}$$

To estimate the jump term, we emphasize that each interior hyperface $E \subseteq \partial\mathcal{T}\cap\Omega$ is the intersection of two elements $T_E^1, T_E^2 \in \mathcal{T}_\bullet$ with $T_E^1 \cap T_E^2 = E$. This leads to

$$\begin{aligned} \|[A\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)} &= \sum_{E\subset\partial T\cap\Omega} \|[A\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|^2_{L^2(E)} \\ &= \sum_{E\subset\partial T\cap\Omega} \|[A\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|^2_{L^2(T_E^1\cap T_E^2)}. \end{aligned}$$

With the Young inequality and a scaling argument, we estimate for each hyperface $E$ that

$$\begin{aligned} &\|[A\,\nabla(U_\circ - U_\bullet)\cdot\mathbf{n}]\|^2_{L^2(T_E^1\cap T_E^2)} \\ &\quad\leq \|A\|^2_{L^\infty(\Omega)}\Big(\|\nabla(U_\circ - U_\bullet)\|^2_{L^2(\partial T_E^1)} + \|\nabla(U_\circ - U_\bullet)\|^2_{L^2(\partial T_E^2)}\Big) \\ &\quad\leq C\,\|A\|^2_{L^\infty(\Omega)}\,\max\Big\{\frac{|E|}{|T_E^1|}, \frac{|E|}{|T_E^2|}\Big\}\|\nabla(U_\circ - U_\bullet)\|^2_{L^2(T_E^1\cup T_E^2)} \\ &\quad\leq C\,\|A\|^2_{L^\infty(\Omega)}\max\{h^{-1}_{T_E^1}, h^{-1}_{T_E^2}\}\,\|U_\circ - U_\bullet\|^2_{H^1(T_E^1\cup T_E^2)}, \end{aligned} \tag{5.11}$$

where the constant $C > 0$ depends only on $\gamma$. The $\gamma$-shape regularity implies $|E|/|T_E^2| \simeq h_T^{-1}$ as well as $\|h_\bullet\|_{L^\infty(\Omega)} \leq \operatorname{diam}(\Omega)$. Moreover, it guarantees $|T| \simeq |T_E^1| \simeq |T_E^2|$ and hence, $h_T \simeq h_{T_E^1} \simeq h_{T_E^2}$ for all hyperfaces $E \subset \partial T\cap\Omega$. Combining the latter estimates, we obtain that

$$\begin{aligned}
R_\bullet(T) &\le h_T \big(\|\nabla A\|_{L^\infty(\Omega)} + C_{\rm inv} h_T^{-1} \|A\|_{L^\infty(\Omega)}\big) \|\nabla(U_\circ - U_\bullet)\|_{L^2(T)} \\
&\quad + h_T \big(\|b\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\big) \|U_\circ - U_\bullet\|_{H^1(T)} \\
&\quad + C^{1/2} \|A\|_{L^\infty(\Omega)} h_T^{1/2} \Big( \sum_{E \subset \partial T \cap \Omega} C' h_T \|U_\circ - U_\bullet\|^2_{H^1(T_E^1 \cup T_E^2)} \Big)^{1/2} \\
&\le (C'' + h_T) \big(\|A\|_{W^{1,\infty}(\Omega)} + \|b\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\big) \|U_\circ - U_\bullet\|_{H^1(\omega_T)},
\end{aligned}$$

where $\omega_T$ denotes the element patch of $T$ and $C''$ depends only on the $\gamma$-shape regularity. Recall that the number of elements in $\omega_T$ is uniformly bounded. Hence, summing over all elements $T \in \mathcal{U}_\bullet$ reveals that

$$\Big|\eta_\circ(\mathcal{U}_\bullet) - \eta_\bullet(\mathcal{U}_\bullet))\Big| \le C''' \big(C'' + \mathrm{diam}(\Omega)\big) \|U_\circ - U_\bullet\|_{H^1(\Omega)}.$$

This concludes the proof. □

**Proposition 5.4** (reduction on refined element domains). *There exist* $C_{\rm red} > 0$ *and* $0 < q_{\rm red} < 1$ *such that for all* $\mathcal{T}_\bullet \in \mathbb{T}$ *and all* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ *the following holds: Provided there exist unique discrete solutions* $U_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$ *and* $U_\circ \in \mathcal{S}_0^p(\mathcal{T}_\circ)$*, it holds that*

$$\eta_\circ(\mathcal{T}_\circ \backslash \mathcal{T}_\bullet)^2 \le q_{\rm red}\, \eta_\bullet(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ)^2 + C_{\rm red}^2 \|U_\circ - U_\bullet\|^2_{H^1(\Omega)}.$$

*In particular, there holds* (E2)*. The constants* $0 < q_{\rm red} < 1$ *and* $C_{\rm red} \ge 1$ *just depend on the given data, the polynomial degree, on* $\gamma$*-shape regularity and on* $q_{\rm mesh}$ *from* (R1)*.*

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$, such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ and the corresponding discrete solutions $U_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$ and $U_\circ \in \mathcal{S}_0^p(\mathcal{T}_\circ)$ exist. Recall the notation of Proposition 5.3. Once again, the triangle inequality implies that

$$\begin{aligned}
\eta_\circ(\mathcal{T}_\circ \setminus \mathcal{T}_\bullet) \le \Big( \sum_{T \in \mathcal{T}_\circ \backslash \mathcal{T}_\bullet} & h_T^2 \|f + \mathrm{div}(A \nabla U_\bullet) - b \cdot \nabla U_\bullet - c\, U_\bullet\|^2_{L^2(T)} \\
&+ h_T \|[A \nabla U_\bullet \cdot \mathbf{n}]\|^2_{L^2(\partial T \cap \Omega)} \Big)^{1/2} + \Big( \sum_{T \in \mathcal{T}_\circ \backslash \mathcal{T}_\bullet} R_\circ(T)^2 \Big)^{1/2},
\end{aligned} \tag{5.12}$$

where $R_\circ(T)$ is defined in (5.10). Analogously to the proof of Proposition 5.3, there holds

$$\Big( \sum_{T \in \mathcal{T}_\circ \backslash \mathcal{T}_\bullet} R_\circ(T)^2 \Big)^{1/2} \le C \|U_\circ - U_\bullet\|_{H^1(\Omega)}.$$

The first sum on the right-hand side of (5.12) can be treated as follows: Reduction of the local mesh-size (R1) and the splitting property (R3) imply for the volume residual that

$$\begin{aligned}
&\sum_{T \in \mathcal{T}_\circ \backslash \mathcal{T}_\bullet} h_T^2 \|f + \mathrm{div}(A \nabla U_\bullet) - b \cdot \nabla U_\bullet - c\, U_\bullet\|^2_{L^2(T)} \\
&\qquad \le \sum_{T' \in \mathcal{T}_\bullet \backslash \mathcal{T}_\circ} q_{\rm mesh}^2 h_{T'}^2 \sum_{T \subseteq T'} \|f + \mathrm{div}(A \nabla U_\bullet) - b \cdot \nabla U_\bullet - c U_\bullet\|^2_{L^2(T)} \\
&\qquad \le q_{\rm mesh}^2 \sum_{T' \in \mathcal{T}_\bullet \backslash \mathcal{T}_\circ} h_{T'}^2 \|f + \mathrm{div}(A \nabla U_\bullet) - b \cdot \nabla U_\bullet - c\, U_\bullet\|^2_{L^2(T')}.
\end{aligned}$$

For the jump term, we additionally emphasize that $A\nabla U_\bullet$ is continuous inside an element $T' \in \mathcal{T}_\bullet$. Hence for all $T' \in \mathcal{T}_\bullet \setminus \mathcal{T}_\circ$ with sons $T \subset T'$ there holds $[A\nabla U_\bullet \cdot \mathbf{n}] = 0$ on $\partial T \setminus \partial T'$. Similar to the latter estimate, we obtain that

$$\begin{aligned}\sum_{T\in\mathcal{T}_\circ\setminus\mathcal{T}_\bullet} h_T \,\|[A\nabla U_\bullet \cdot \mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)} &\le \sum_{T'\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} q_{\mathrm{mesh}}\, h_T \sum_{T\subseteq T'} \|[A\nabla U_\bullet \cdot \mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)} \\ &\le q_{\mathrm{mesh}} \sum_{T'\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} h_T \,\|[A\nabla U_\bullet \cdot \mathbf{n}]\|^2_{L^2(\partial T'\cap\Omega)}.\end{aligned}$$

Combining the latter estimates with (5.12), we see that

$$\eta_\circ(\mathcal{T}_\circ \setminus \mathcal{T}_\bullet) \le q_{\mathrm{mesh}}^{1/2}\, \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ) + C\, \|U_\circ - U_\bullet\|_{H^1(\Omega)}.$$

With the Young inequality for $\delta > 0$, we obtain that

$$\eta_\circ(\mathcal{T}_\circ \setminus \mathcal{T}_\bullet)^2 \le (1+\delta)\, q_{\mathrm{mesh}}\, \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 + (1+\delta^{-1})\, C^2\, \|U_\circ - U_\bullet\|^2_{H^1(\Omega)}.$$

Choosing $\delta > 0$ such that $q_{\mathrm{red}} := (1+\delta)\, q_{\mathrm{mesh}} < 1$, we conclude the proof. □

**Proposition 5.5** (discrete reliability). *There exists $C_{\mathrm{rel}} > 0$, such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, there exists a set $\mathcal{R}_{\bullet,\circ} \subseteq \mathcal{T}_\bullet$ such that the following holds: Provided there exist unique discrete solutions $U_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\bullet)$ and $U_\circ \in \mathcal{S}^p_0(\mathcal{T}_\circ)$, it holds that*

$$\|U_\circ - U_\bullet\|_{H^1(\Omega)} \le C_{\mathrm{rel}}\, \beta_\circ^{-1}\, \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)$$

*where $\beta_\circ > 0$ is the* inf-sup *constant* (4.6) *associated with $\mathcal{S}^p_0(\mathcal{T}_\circ)$. In particular, there holds* (E4). *The constant $C_{\mathrm{rel}} \ge 1$ depends only on the given data, the initial mesh $\mathcal{T}_0$the polynomial degree, and $\gamma$-shape regularity of $\mathcal{T}_\bullet$.*

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$, such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$. Suppose that the corresponding discrete solutions $U_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\bullet)$ and $U_\circ \in \mathcal{S}^p_0(\mathcal{T}_\circ)$ of (5.7) exist. Recall the discrete inf-sup condition on $\mathcal{S}^p_0(\mathcal{T}_\circ)$ from Proposition 4.1

$$\beta_\circ = \inf_{W_\circ\in\mathcal{S}^p_0(\mathcal{T}_\circ)} \sup_{V_\circ\in\mathcal{S}^p_0(\mathcal{T}_\circ)} \frac{|b(W_\circ\,,\,V_\circ)|}{\|W_\circ\|_{H^1(\Omega)}\|V_\circ\|_{H^1(\Omega)}}.$$

Choosing $W_\circ := U_\circ - U_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\circ)$ gives

$$\beta_\circ\, \|U_\circ - U_\bullet\|_{H^1(\Omega)} \le \sup_{V_\circ\in\mathcal{S}^p_0(\mathcal{T}_\circ)} \frac{|b(U_\circ - U_\bullet\,,\,V_\circ)|}{\|V_\circ\|_{H^1(\Omega)}}. \tag{5.13}$$

Now, let $V_\circ \in \mathcal{S}^p_0(\mathcal{T}_\circ)$ be arbitrary but fixed. Galerkin orthogonality implies for all $V_\bullet \in \mathcal{S}^p_0(\mathcal{T}_\bullet)$ that

$$b(U_\circ - U_\bullet\,,\,V_\circ) = b(U_\circ - U_\bullet\,,\,V_\circ - V_\bullet).$$

To estimate the right-hand side, define $\widetilde{\Omega} := \mathrm{interior}\Big(\bigcup(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)\Big)$ with shape regular triangulation $\widetilde{\mathcal{T}} := \mathcal{T}_\bullet|_{\widetilde{\Omega}}$. Let $J_{\widetilde{\mathcal{T}}} : H^1(\widetilde{\Omega}) \to \mathcal{S}^p(\widetilde{\mathcal{T}})$ denote the Scott–Zhang projection, see e.g., [SZ90]. For a deeper discussion on the properties of the Scott–Zhang projection we refer to [AFK$^+$13, Section 3]. We define

$$V_\bullet := \begin{cases} V_\circ & \text{on } \Omega \setminus \widetilde{\Omega} \\ J_{\widetilde{\mathcal{T}}}(V_\circ) & \text{on } \widetilde{\Omega}. \end{cases}$$

Since $V_\circ \in \mathcal{S}_0^p(\mathcal{T}_\circ)$, there holds $V_\bullet|_{\Gamma\setminus\partial\widetilde{\Omega}} = V_\circ|_{\Gamma\setminus\partial\widetilde{\Omega}} = 0$. Note that the Scott–Zhang projection preserves discrete boundary data on $\partial\widetilde{\Omega}$ (see [AFK$^+$13, Section 3.1]). This implies that $V_\circ - V_\bullet = 0$ on $\Omega \setminus \widetilde{\Omega}$ and hence $V_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$. Integration by parts and the Cauchy inequality yield that

$$\begin{aligned}
b(U_\circ - U_\bullet\,,\, V_\circ - V_\bullet) &= \langle f\,,\, V_\circ - V_\bullet\rangle - b(U_\bullet\,,\, V_\circ - V_\bullet) \\
&= \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} \Big( \int_T f(V_\circ - V_\bullet)\,dx - \int_T A\nabla U_\bullet \cdot \nabla(V_\circ - V_\bullet)\,dx \\
&\qquad - \int_T \big(b\cdot\nabla U_\bullet + c\,U_\bullet\big)(V_\circ - V_\bullet)\,dx\Big) \\
&= \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} \Big( \int_T \big(f + \operatorname{div}(A\nabla U_\bullet) - b\cdot\nabla U_\bullet - c\,U_\bullet\big)(V_\circ - V_\bullet)\,dx \\
&\qquad + \int_{\partial T} A\nabla U_\bullet \cdot \mathbf{n}\,(V_\circ - V_\bullet)\,ds\Big) \\
&\le \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} \Big( \|h_T(f + \operatorname{div}(A\nabla U_\bullet) - b\cdot\nabla U_\bullet - c\,U_\bullet)\|_{L^2(T)}\,\|h_T^{-1}(V_\circ - V_\bullet)\|_{L^2(T)} \\
&\qquad + \|h^{1/2}[A\nabla U_\bullet\cdot\mathbf{n}]\|_{L^2(\partial T\cap\Omega)}\,\|h_T^{-1/2}(V_\circ - V_\bullet)\|_{L^2(\partial T\cap\Omega)}\Big).
\end{aligned} \tag{5.14}$$

We recall some standard properties of the Scott–Zhang projection; see [SZ90]. Let $\widetilde{\omega}_T$ denote the patch of $T$ in $\widetilde{\mathcal{T}}$. For an element $T \in \mathcal{T}_\bullet \setminus \mathcal{T}_\circ$, there holds that

$$\|V_\circ - V_\bullet\|_{L^2(T)} = \|(1 - J_{\widetilde{\mathcal{T}}})V_\circ\|_{L^2(T)} \lesssim h_T\,\|\nabla V_\circ\|_{L^2(\widetilde{\omega}_T)}. \tag{5.15}$$

In combination with the trace inequality, we obtain that

$$\begin{aligned}
\|V_\circ - V_\bullet\|_{L^2(\partial T)}^2 &= \|(1 - J_{\widetilde{\mathcal{T}}})V_\circ\|_{L^2(\partial T)}^2 \\
&\lesssim h_T^{-1}\,\|(1 - J_{\widetilde{\mathcal{T}}})V_\circ\|_{L^2(T)}^2 + h_T\,\|\nabla(1 - J_{\widetilde{\mathcal{T}}})V_\circ\|_{L^2(T)}^2 \\
&\lesssim h_T\,\|\nabla V_\circ\|_{L^2(\widetilde{\omega}_T)}^2.
\end{aligned} \tag{5.16}$$

We emphasize that the constants of (5.15) and (5.16) just depend on the $\gamma$-shape regularity, the polynomial degree $p$ and the dimension. Combining the latter estimates with (5.14), the Cauchy inequality yields that

$$\begin{aligned} b(U_\circ - U_\bullet\,,\, V_\circ - V_\bullet) \lesssim \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} \Big( &\|h_T(f + \operatorname{div}(A\nabla U_\bullet) - b\cdot\nabla U_\bullet - c\,U_\bullet)\|_{L^2(T)}\, \|\nabla V_\circ\|_{L^2(\omega_T)} \\ &+ \|h^{1/2}[A\nabla U_\bullet\cdot \mathbf{n}]\|_{L^2(\partial T\cap\Omega)}\, \|\nabla V_\circ\|_{L^2(\omega_T)}\Big) \\ \le \Big( \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} &\|h_T(f + \operatorname{div}(A\nabla U_\bullet) - b\cdot\nabla U_\bullet - c\,U_\bullet)\|^2_{L^2(T)} \\ &+ \|h^{1/2}[A\nabla U_\bullet\cdot \mathbf{n}]\|^2_{L^2(\partial T\cap\Omega)}\Big)^{1/2} \times \Big( \sum_{T\in\mathcal{T}_\bullet\setminus\mathcal{T}_\circ} \|\nabla V_\circ\|^2_{L^2(\omega_T)}\Big)^{1/2}. \end{aligned}$$

Recall that every element patch $\omega_T$ consists only of finitely many elements. Since the latter estimate holds for arbitrary $V_\circ \in \mathcal{S}^p_0(\mathcal{T}_\circ)$, we obtain with (5.13) that

$$\begin{aligned} \beta_\circ\, \|U_\circ - U_\bullet\|_{H^1(\Omega)} &\le \sup_{V_\circ\in\mathcal{S}^p_0(\mathcal{T}_\circ)} \frac{|b(U_\circ - U_\bullet\,,\, V_\circ)|}{\|V_\circ\|_{H^1(\Omega)}} \\ &\lesssim \sup_{V_\circ\in\mathcal{S}^p_0(\mathcal{T}_\circ)} \frac{\eta_\bullet(\mathcal{T}_\bullet\setminus\mathcal{T}_\circ)\, \|V_\circ\|_{H^1(\Omega)}}{\|V_\circ\|_{H^1(\Omega)}} = \eta_\bullet(\mathcal{T}_\bullet\setminus\mathcal{T}_\circ). \end{aligned}$$

This concludes the proof of (E4) with $\mathcal{R}_{\bullet,\circ} = \mathcal{T}_\bullet \setminus \mathcal{T}_\circ$. □

It remains to prove reliability (E3). Since uniform refinement leads to convergence, discrete reliability (E4) implies reliability. The proof follows analogously to Step 3 of the proof of Lemma 4.8.

We emphasize that the error estimator can be extended to mixed Dirichlet–Neumann–Robin boundary conditions, where inhomogeneous Dirichlet conditions are discretized by nodal interpolation for $d = 2$ and $p = 1$, see [FPP14], or by Scott–Zhang interpolation for $d \ge 2$ and $p \ge 1$, see [CFPP14]. In any case (E1)–(E4) remain valid [FPP14, CFPP14], but $\mathcal{R}_{\bullet,\circ}$ consists of a fixed patch of $\mathcal{T}_\bullet\setminus\mathcal{T}_\circ$ [AFK+13, CFPP14].

### 5.2.1 Definiteness on the "discrete" limit space (E5)

It remains to prove validity of (E5). Recall that $\mathcal{H} = H^1_0(\Omega)$ and $\mathcal{X}_\ell = \mathcal{S}^p_0(\mathcal{T}_\ell)$ in the sense of Section 4.5.1. We define the discrete limit space $\mathcal{X}_\infty := \overline{\bigcup_{\ell=0}^\infty \mathcal{S}^p_0(\mathcal{T}_\ell)}$ and obtain the following lemma.

**Lemma 5.6.** *Let $p \ge 1$. The triangulations $\mathcal{T}_\ell$ generated by Algorithm 5.2 are uniformly $\gamma$-shape regular with $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$. Further, there holds $\mathcal{X}_\infty = H^1_0(\Omega)$ and hence assumption* (E5) *is satisfied.*

*Proof.* Recall the notation of Algorithm 5.2. NVB guarantees $\gamma$-shape regularity (3.4) of the generated meshes $\mathcal{T}_\ell$. Further, $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$ is enforced by the expanded marking strategy in Step (iv) and Step (v) of Algorithm 5.2 (see Proposition 4.7).

For $w \in \mathcal{D} := H^2(\Omega) \cap H^1_0(\Omega)$, recall the approximation property $\inf_{V_\ell\in\mathcal{X}_\ell} \|w - V_\ell\|_{\mathcal{H}} \lesssim \|h_\ell\|_{L^\infty(\Omega)}\|D^2 w\|_{L^2(\Omega)}$ from, e.g., [BS08, Chapter 4]. This proves that

$$\lim_{\ell\to\infty} \inf_{V_\ell\in\mathcal{S}^p_0(\mathcal{T}_\ell)} \|w - V_\ell\|_{\mathcal{H}} = 0 \quad \text{for all } w \in \mathcal{D}. \tag{5.17}$$

Let $v \in H^1_0(\Omega)$ and $\varepsilon > 0$. Since $\mathcal{D}$ is dense within $H^1_0(\Omega)$, choose $w \in \mathcal{D}$ with $\|v - w\|_{\mathcal{H}} \le \varepsilon/2$. According to (5.17), there exists an index $\ell_\bullet \in \mathbb{N}_0$ such that $\inf_{V_\ell \in \mathcal{X}_\ell} \|w - V_\ell\|_{\mathcal{H}} \le \varepsilon/2$ for all $\ell \ge \ell_\bullet$. In particular, the triangle inequality gives

$$\inf_{V_\ell \in \mathcal{X}_\ell} \|v - V_\ell\|_{\mathcal{H}} \le \|v - w\|_{\mathcal{H}} + \inf_{V_\ell \in \mathcal{X}_\ell} \|w - V_\ell\|_{\mathcal{H}} \le \varepsilon \quad \text{for all } \ell \ge \ell_\bullet.$$

This proves $v \in \mathcal{X}_\infty = \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ and hence concludes $\mathcal{X}_\infty = H^1_0(\Omega)$. □

The next remark shows that definiteness on the "discrete" limit space (E5) is already satisfied for many generic situation, even without using the expanded Dörfler marking strategy.

**Remark 5.7.** *In many generic situations, $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ and hence* (E5) *with $\mathcal{X}_\infty = H^1_0(\Omega)$ is satisfied even without using the expanded Dörfler marking of Proposition 4.7. To see that, suppose $p \ge 1$ and $q \ge 0$ are polynomial degrees and $f \in L^2(\Omega)$. Let $f_\ell \in \mathcal{P}^q(\mathcal{T}_\ell)$ be the $L^2$-best approximation of $f$ in $\mathcal{P}^q(\mathcal{T}_\ell)$. Suppose further that the error estimator is even reliable in the sense of*

$$\|u - U_\ell\|_{H^1(\Omega)} + \|h_\ell(f - f_\ell)\|_{L^2(\Omega)} \le C_{\mathrm{rel}}\, \eta_\ell \quad \textit{for all } \ell \ge 0, \tag{5.18}$$

*where the constant $C_{\mathrm{rel}}$ is independent of $\ell$. Note that* (5.18) *is well-known for residual error estimators and elliptic PDEs with polynomial coefficients. Suppose that for all $\ell \in \mathbb{N}$ and all $T \in \mathcal{T}_\ell$, it holds $u|_T \notin \mathcal{P}^p(T)$ or $f|_T \notin \mathcal{P}^q(T)$, i.e., the continuous solution as well as the given data are not locally polynomial. Using* (5.18)*, a simply argument by contradiction shows that convergence $\eta_\ell \to 0$ as $\ell \to 0$ already implies $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$. Then, Lemma 5.6 implies* (E5) *with $\mathcal{X}_\infty = H^1_0(\Omega)$.*

## 5.3 Optimal convergence

In this section we show the main result of this chapter, which proves optimal algebraic convergence rates for adaptive FEM for general second-order diffusion problems with convection and reaction.

Section 5.2 proves that the model problem (5.2) fits in the abstract framework of Chapter 4. Utilizing the previous sections, we can apply the abstract analysis of Sections 4.6–4.8 to obtain convergence of the adaptive algorithm (Sections 4.6). Further, as a direct consequence of Theorem 4.14 and Theorem 4.21 we get the following result.

**Theorem 5.8.** *Employ the notation of Algorithm 5.2. Suppose $0 < \theta \le 1$. Then, there exist $\ell_{\mathrm{lin}} > 0$ and constants $0 < q_{\mathrm{lin}} < 1$ as well as $C_{\mathrm{lin}} > 0$, such that the output of Algorithm 5.2 satisfies*

$$\eta_{\ell+n} \le C_{\mathrm{lin}}\, q_{\mathrm{lin}}^n\, \eta_\ell \quad \textit{for all } \ell, n \in \mathbb{N}_0 \textit{ with } \ell \ge \ell_{\mathrm{lin}}. \tag{5.19}$$

*Further suppose $0 < \theta < \theta_{\mathrm{opt}} := (1 + C_{\mathrm{stb}}^2 C_{\mathrm{rel}}^2 / \widehat{\beta}_0^2)^{-1}$, where $\widehat{\beta}_0^2$ is the* inf-sup *constant of the $m_{\mathrm{unif}}$-times uniform refinement of $\mathcal{T}_0$; see Lemma 4.17. For all $s > 0$, it holds that*

$$\|u\|_{\mathbb{A}_s} < \infty \quad \Longleftrightarrow \quad \exists \ell_{\mathrm{opt}} > 0\ \exists C_{\mathrm{opt}} > 0\ \forall \ell \ge \ell_{\mathrm{opt}} \quad \eta_\ell \le C_{\mathrm{opt}} \big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\big)^{-s}. \tag{5.20}$$

*The constant $C_{\mathrm{opt}}$ depends only on $\#\mathcal{T}_{\ell_{\mathrm{opt}}}$, $\mathcal{T}_0$, $\theta$, $s$, and validity of* (E1)–(E5). □

Theorem 5.8 does not only prove linear convergence, but also optimal algebraic convergence rates for the sequence of estimators and solutions provided by Algorithm 5.2. We emphasize that in many generic situations, Theorem 5.8 even holds without using the expanded Dörfler marking in Algorithm 5.2 (see Remark 5.7).

To get optimal convergence in the spirit of [CKNS08] with approximation classes $\|\cdot\|_{\mathbb{E}_s(\mathcal{T}_\bullet)}$ (cf. see Section 4.8.2), one can define suitable oscillation indicators in the following way. For each element $T \in \mathcal{T}_\bullet \in \mathbb{T}$, let $\mathcal{F}_T$ denote the set of its facets (i.e., edges for $d = 2$). For arbitrary $q \geq p - 1$, the data oscillations corresponding to the error indicators from (5.8) are given by

$$\operatorname{osc}_\bullet(V_\bullet)^2 := \sum_{T \in \mathcal{T}_\bullet} \operatorname{osc}_\bullet(T, V_\bullet)^2. \tag{5.21a}$$

We define the element contributions for all $T \in \mathcal{T}_\bullet$ and $V_\bullet \in \mathcal{S}_0^p(\mathcal{T}_\bullet)$ by

$$\begin{aligned} \operatorname{osc}_\bullet(T, V_\bullet)^2 =& h_T^2 \min_{Q \in \mathcal{P}^q(T)} \|f + \operatorname{div}(A\nabla V_\bullet) - b \cdot \nabla V_\bullet - c\, V_\bullet - Q\|_{L^2(T)}^2 \\ &+ h_T \sum_{F \in \mathcal{F}_T} \min_{Q \in \mathcal{P}^q(T)} \|[(A\nabla V_\bullet) \cdot \mathbf{n}] - Q\|_{L^2(F \cap \Omega)}^2. \end{aligned} \tag{5.21b}$$

Then, $\operatorname{osc}_\bullet(\cdot)^2$ satisfies the assumptions of Lemma 4.20 in Section 4.8.2 as well as (4.29) with $\mathcal{X}_\bullet := \mathcal{S}_0^p(\mathcal{T}_\bullet) = \mathcal{S}^p(\mathcal{T}_\bullet) \cap H_0^1(\Omega)$ and estimator $\eta_\ell$ from(5.8). Further details as well as the exact proof is found in, e.g., [CKNS08, CN12, FFP14]. Note that the constant $C_{\mathrm{osc}}$ in Lemma 4.20 depends on $q$ and $p$. We emphasize that, if $A, b, c$ are piecewise polynomial and if $q$ is chosen sufficiently large, the local contributions simplify to the well-known data oscillations

$$\operatorname{osc}_\bullet(T, V_\bullet)^2 = h_T^2 \min_{f_T \in \mathcal{P}^q(T)} \|f - f_T\|_{L^2(T)}^2$$

as for the Laplace problem. In any case, Lemma 4.20 applies and yields for all $s > 0$ that

$$\|u\|_{\mathbb{E}_s} < \infty \quad \Longleftrightarrow \quad \|u\|_{\mathbb{A}_s} < \infty.$$

Hence, Algorithm 5.2 guarantees optimal algebraic convergence rates with respect to the total error.

## 5.4 Numerical experiments

In this section, we present two numerical experiments for the 2D Helmholtz equation (5.3) that underpin our theoretical finding from the previous sections. The numerical experiments are taken from [BHP17]. We use lowest-order FEM with $\mathcal{X}_\bullet := \mathcal{S}_0^1(\mathcal{T}_\bullet)$ and employ the weighted-residual error estimator. We also refer to [BISG97] for a first systematic a posteriori error analysis for finite elements for the Helmholtz equation and [OPD05] for an overview on the state of the art and available error estimation techniques for this problem. In the experiments, we compare the performance of Algorithm 5.2 with respect to

- different values of $k \in \{1, 2, 4, 8, 16\}$,
- different values of $\theta \in \{0.1, 0.2, \dots, 0.9\}$,
- standard Dörfler marking strategy (with $C_{\mathrm{mark}} = 1$) as well as the expanded Dörfler marking strategy of Proposition 4.7 with $C_{\mathrm{mark}} = 2$.

We consider domains $\Omega \subset \mathbb{R}^2$ with a single re-entrant corner and corresponding interior angle $\alpha > \pi$, cf. Figure 5.1. Note that elliptic regularity thus predicts a generic convergence order $\mathcal{O}(N^{-\delta/2})$ for the error on uniform meshes with $N$ elements, where $\delta = \pi/\alpha < 1$. On the other hand, the optimal convergence behavior for lowest-order elements is $\mathcal{O}(N^{-1/2})$ if the mesh is appropriately refined.

### 5.4.1 Experiment with unknown solution

We consider the Z-shaped domain $\Omega \subset \mathbb{R}^2$ from Figure 5.1 (upper left). The marked node has the coordinates $(-1, -t) = (-1, -0.5)$ and determines the angle $\alpha$ at the re-entrant corner $(0, 0)$. This choice leads to $\alpha = 2\pi - \arcsin\big(t/\sqrt{1+t^2}\big)$ and $\delta \approx 0.5398$. Consider the constant right-hand side $f \equiv 1$ in (5.3) so that the residual error estimator is equivalent to the actual error, i.e., $\eta_\star \simeq \|u - U_\star\|_{H^1(\Omega)}$.

For $k = 2$, Figure 5.1 shows a generically reduced convergence rate for the error estimator on uniform meshes, while on the other hand, Algorithm 5.2 with $\theta = 0.2$ regains the optimal convergence rate. Empirically, the results generated by employing the standard Dörfler marking lead to similar results as for the expanded Dörfler marking from Proposition 4.7. The same observation is made for all tested choices of $\theta$ (not displayed), so that we only consider the expanded Dörfler marking in the remaining plots. Figure 5.2 compares uniform vs. adaptive mesh-refinement for some fixed $\theta \in \{0.2, 0.5\}$ but various wavenumbers $k \in \{1, 2, 4, 8\}$. As expected, the preasymptotic phase increases with growing $k$. However, adaptive mesh-refinement results in asymptotically optimal convergence behavior. Figure 5.3 compares uniform vs. adaptive mesh-refinement for fixed $k \in \{2, 8\}$ but various $\theta \in \{0.1, \dots, 0.9\}$. Although Theorem 5.8 predicts optimal convergence rates only for small marking parameters $0 < \theta < \theta_{\mathrm{opt}} := (1 + C_{\mathrm{stb}}^2 C_{\mathrm{rel}}^2)^{-1}$, we observe that Algorithm 5.2 is stable in $\theta$, and any choice of $\theta \leq 0.9$ leads to the optimal convergence behavior. Even though this gap between analytical and computational results is well-known and typical for adaptive algorithms, this problem is still open. Finally, we observe that Algorithm 5.2 did never enforce uniform mesh-refinement in Step (i), i.e., throughout the resulting discrete linear systems were indefinite but regular.

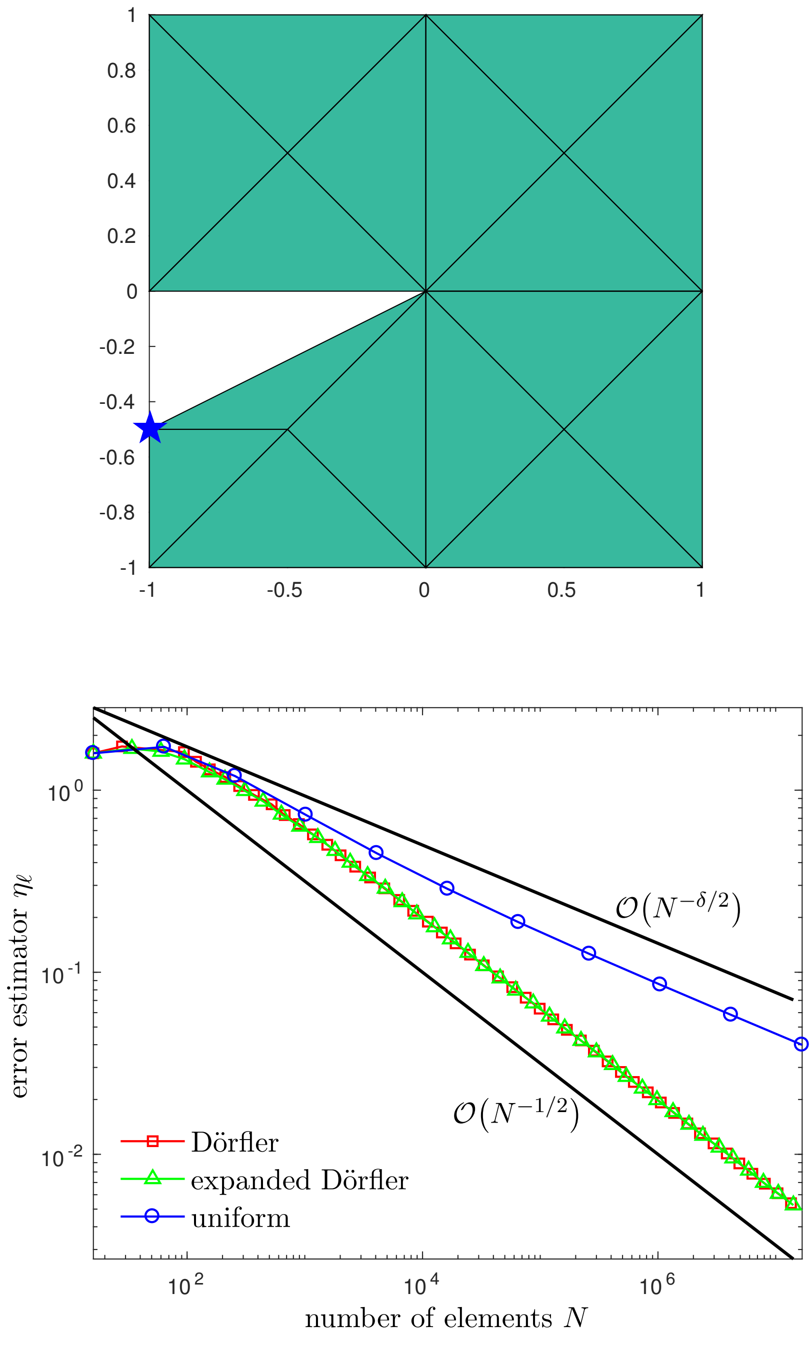
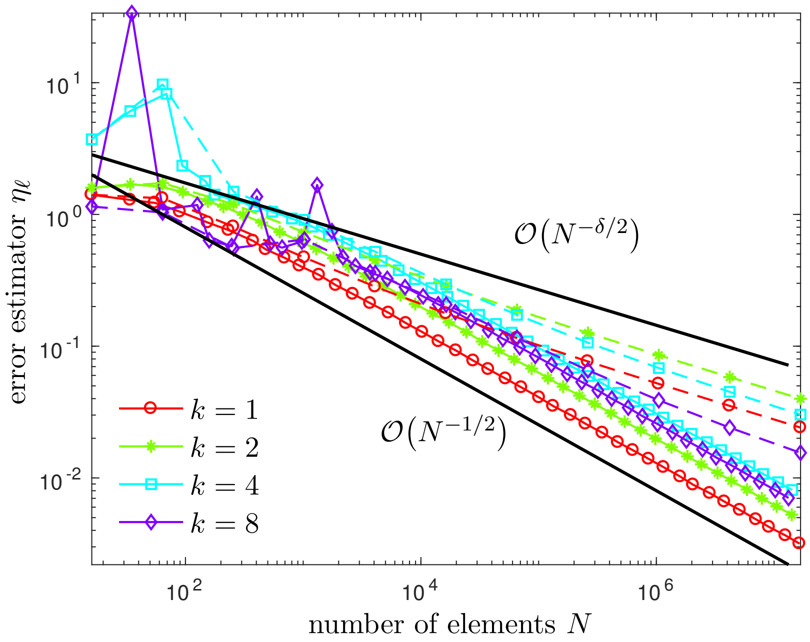



Figure 5.1: Geometry and initial partition $\mathcal{T}_0$ in the experiment from Section 5.4.1 (above). The blue star indicates the node $(-1,-t) := (-1,-0.5)$. Below, we compare the error estimator for uniform vs. adaptive refinement with $\theta = 0.2$ and $k = 2$. Uniform mesh-refinement leads to a suboptimal convergence rate, while Algorithm 5.2 with Dörfler marking and expanded Dörfler marking recovers the optimal convergence rate.



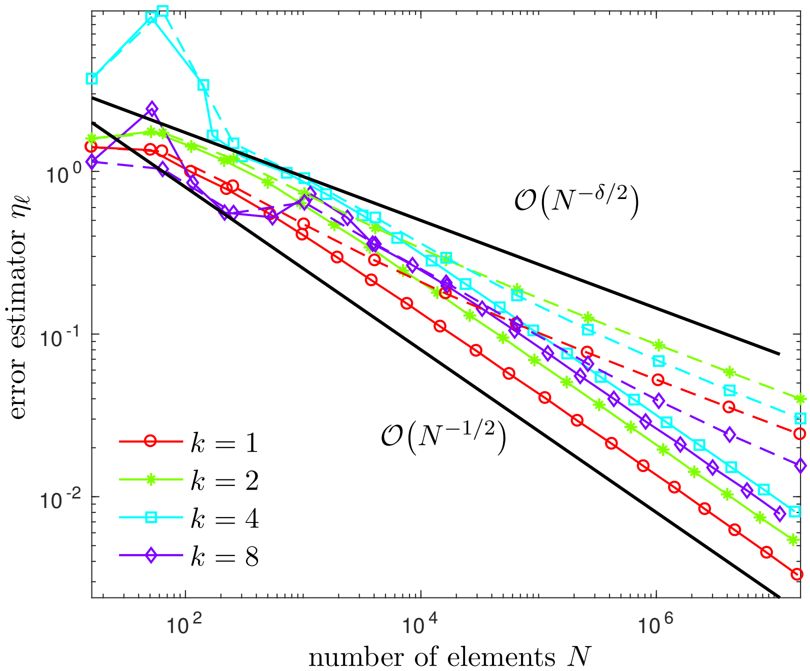




Figure 5.2: Convergence rates for $\eta_\ell$ in the experiment from Section 5.4.1 for different values of $k$. We used $\theta = 0.2$ (above) as well as $\theta = 0.5$ (below). Dashed lines mark uniform refinement, while solid lines mark the corresponding estimator $\eta_\ell$ of Algorithm 5.2 with expanded Dörfler marking. The latter recovers optimal convergence rates, while uniform refinement does not.

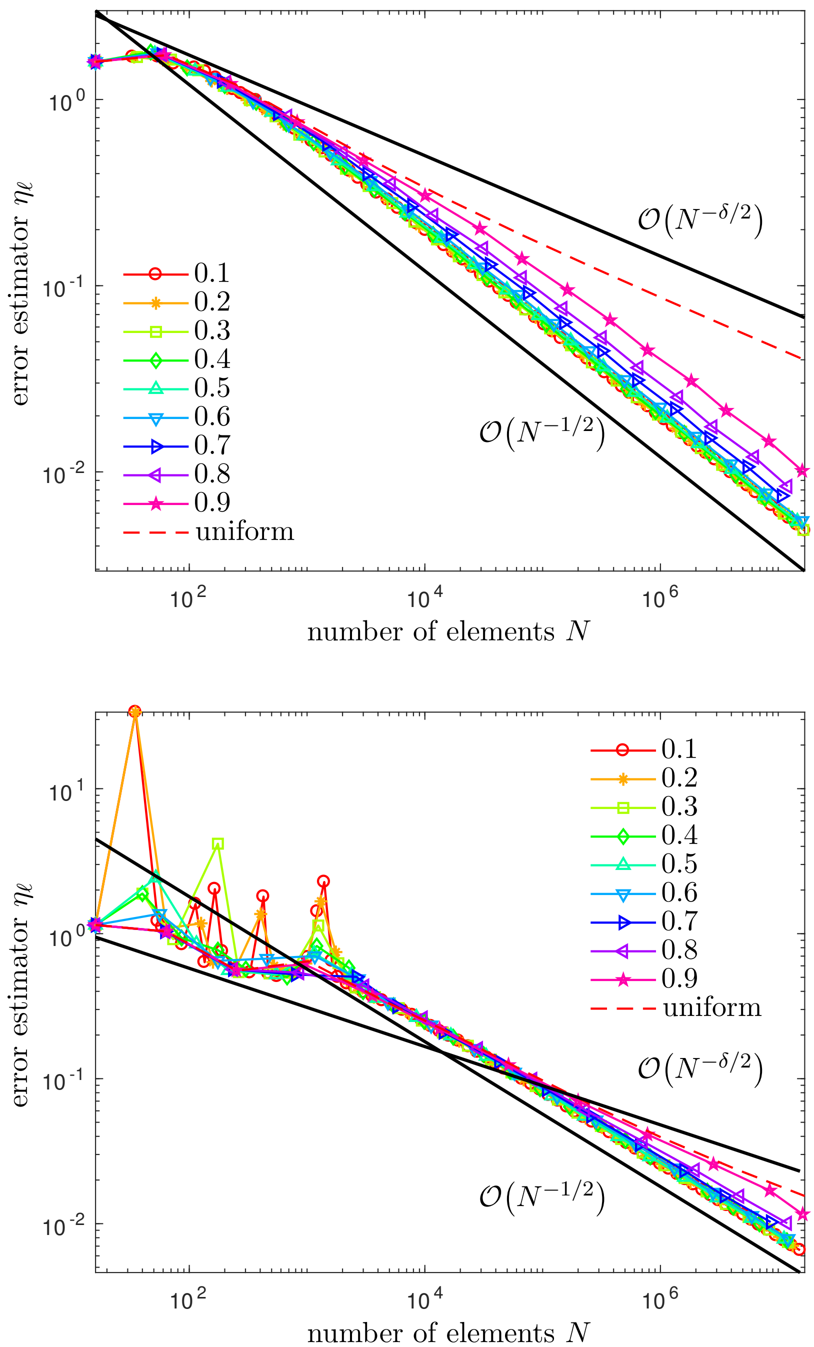



Figure 5.3: Convergence rates for $\eta_\ell$ in the experiment from Section 5.4.1. We compare uniform and adaptive mesh-refinement for different values of $\theta \in \{0.1, \dots, 0.9\}$ and $k = 2$ (above) as well as $k = 8$ (below). For all $\theta < 1$, adaptive mesh-refinement leads to optimal convergence behavior, while the preasymptotic behavior increases with $k$.

### 5.4.2 Experiment with mixed boundary conditions

We consider a Z-shaped domain with a symmetric opening at the re-entrant corner, see Figure 5.7. The marked nodes read $(-1, \pm t) = (-1, \pm 0.25)$. Analogously to the previous example, we expect a reduced convergence order $\mathcal{O}(N^{-\delta/2})$ for uniform mesh-refinement with $\delta \approx 0.5423$. We prescribe the exact solution of the Helmholtz equation in polar coordinates $(r, \phi)$ by

$$u(x, y) = r^{\delta} \cos\big(\delta\phi\big) \tag{5.22}$$

and define $f := -k^2 u$ in $\Omega$ and $g := \partial_{\mathbf{n}} u$ on $\Gamma$. Note that $u$ has a generic singularity at the re-entrant corner $(0, 0)$ of $\Omega$. Further, there holds $u|_{\Gamma_D} = 0$ on the Dirichlet boundary $\Gamma_D := \operatorname{conv}\{(-1, \pm t), (0, 0)\}$. The Neumann boundary is given by $\Gamma_N := \partial\Omega \backslash \Gamma_D$. Thus, $u$ is the unique weak solution of the mixed boundary value problem

$$\begin{aligned} -\Delta u - k^2\, u &= f && \text{in } \Omega \\ u &= 0 && \text{on } \Gamma_D \\ \partial_{\mathbf{n}} u &= g && \text{on } \Gamma_N. \end{aligned} \tag{5.23}$$

The weak formulation (5.23) can be written in the variational formulation (5.4) with Hilbert space

$$\mathcal{H} := H^1_D(\Omega) = \big\{v \in H^1(\Omega) \,:\, \gamma_0^{\text{int}}(v)|_{\Gamma_D} = 0\big\},$$

where $\gamma_0^{\text{int}}(\cdot)$ denotes the interior trace operator (cf. Section 2.2). Note that assumption (E5) is guaranteed by Remark 5.7 even for standard Dörfler marking. Moreover, since the exact solution $u$ is given, we can compute the error $\texttt{err}_\star := \|u - U_\star\|_{H^1(\Omega)}$ besides the corresponding error estimator $\eta_\star$.

Our empirical observations for mixed boundary value problem are similar to those of the previous experiment in Section 5.4.1; see, e.g., Figure 5.4–5.6. Uniform mesh-refinement leads to suboptimal convergence behavior with rate $\mathcal{O}(N^{-\delta/2})$ for both the error as well as the error estimator. On the other hand, adaptive mesh-refinement resolves the geometric singularity at the re-entrant corner (Figure 5.7) and recovers the optimal convergence rate $\mathcal{O}(N^{-1/2})$, see Figure 5.4.

As in the previous example, Algorithm 5.2 appears to be stable for all $\theta \in \{0.1, \dots, 0.9\}$ and always realizes the optimal rate (Figure 5.6). Different values for the wavenumber $k \in \{1, 2, 4, 8, 16\}$ only affect the preasymptotic phase (Figure 5.5). Finally Figure 5.4 shows that there is no empirical difference between the standard Dörfler marking and the expanded Dörfler marking, and therefore, both refinement strategies lead to optimal convergence behavior for the error as well as the error estimator.

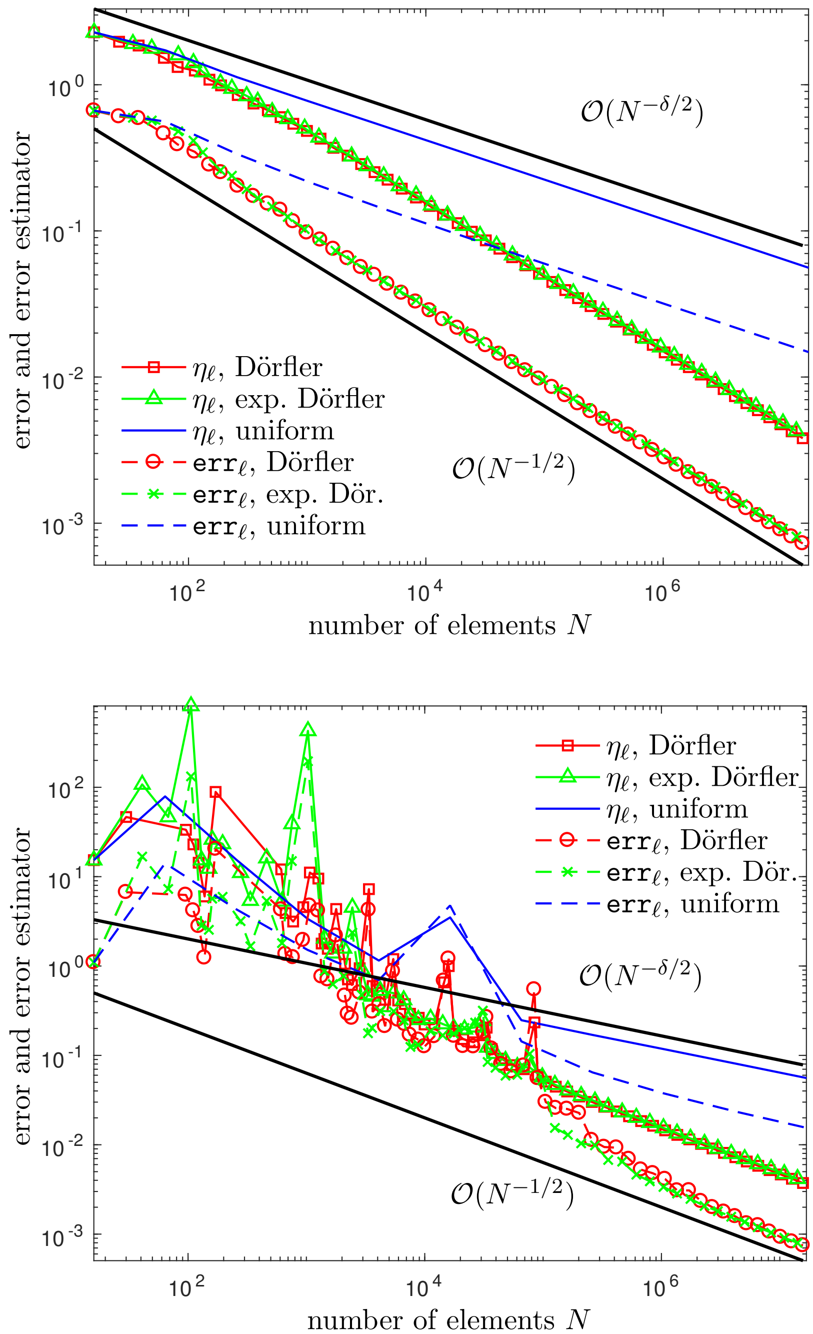



Figure 5.4: Convergence of $\texttt{err}_\ell$ and $\eta_\ell$ in the experiment of Section 5.4.2 for $k = 2$ (above) and $k = 16$ (below). We compare uniform vs. adaptive mesh-refinement using Dörfler marking as well as expanded Dörfler marking with $\theta = 0.2$.

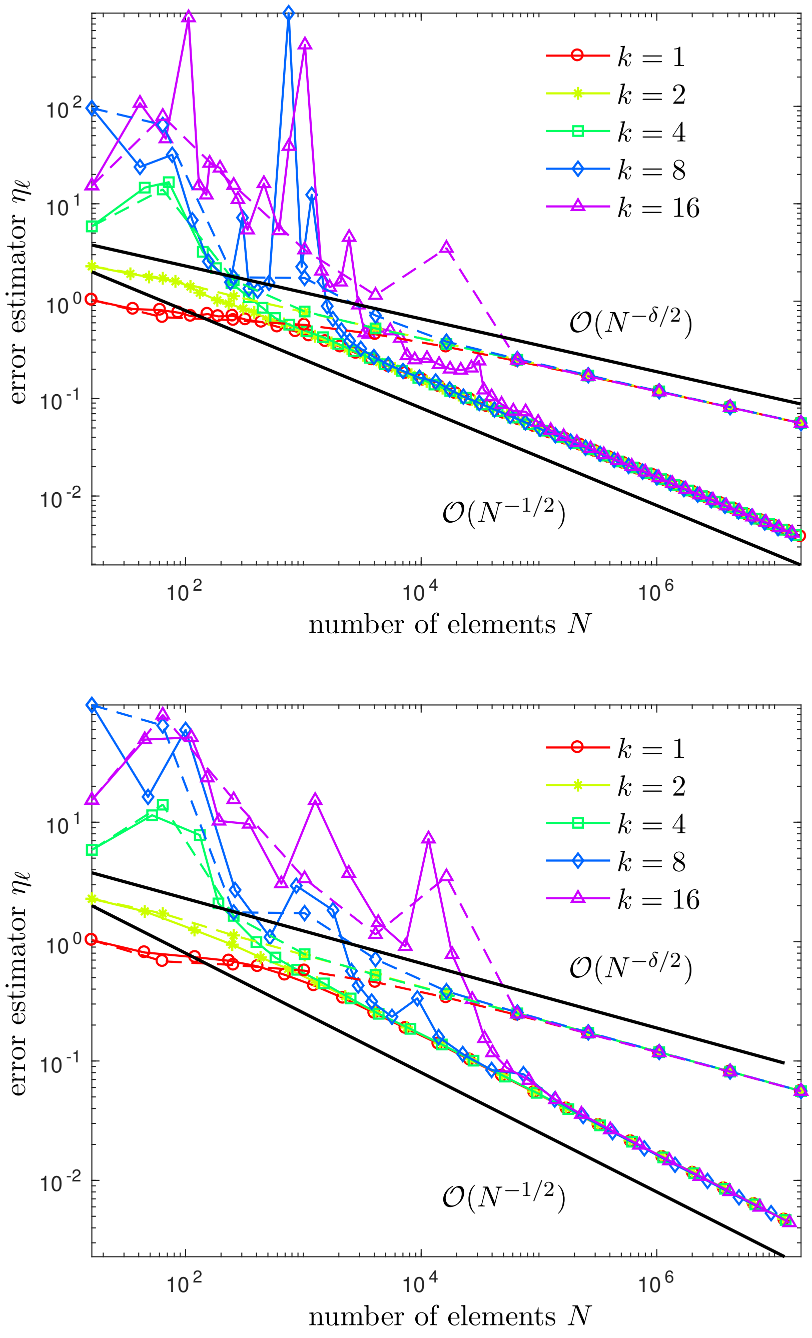



Figure 5.5: Convergence rates for $\eta_\ell$ in the experiment of Section 5.4.2 for different values of $k$ and marking parameters $\theta = 0.2$ (above) as well as $\theta = 0.5$ (below). Dashed lines indicate uniform refinement, while solid lines indicate the output of Algorithm 5.2 with expanded Dörfler marking. The latter recovers optimal convergence rates, while uniform refinement does not.

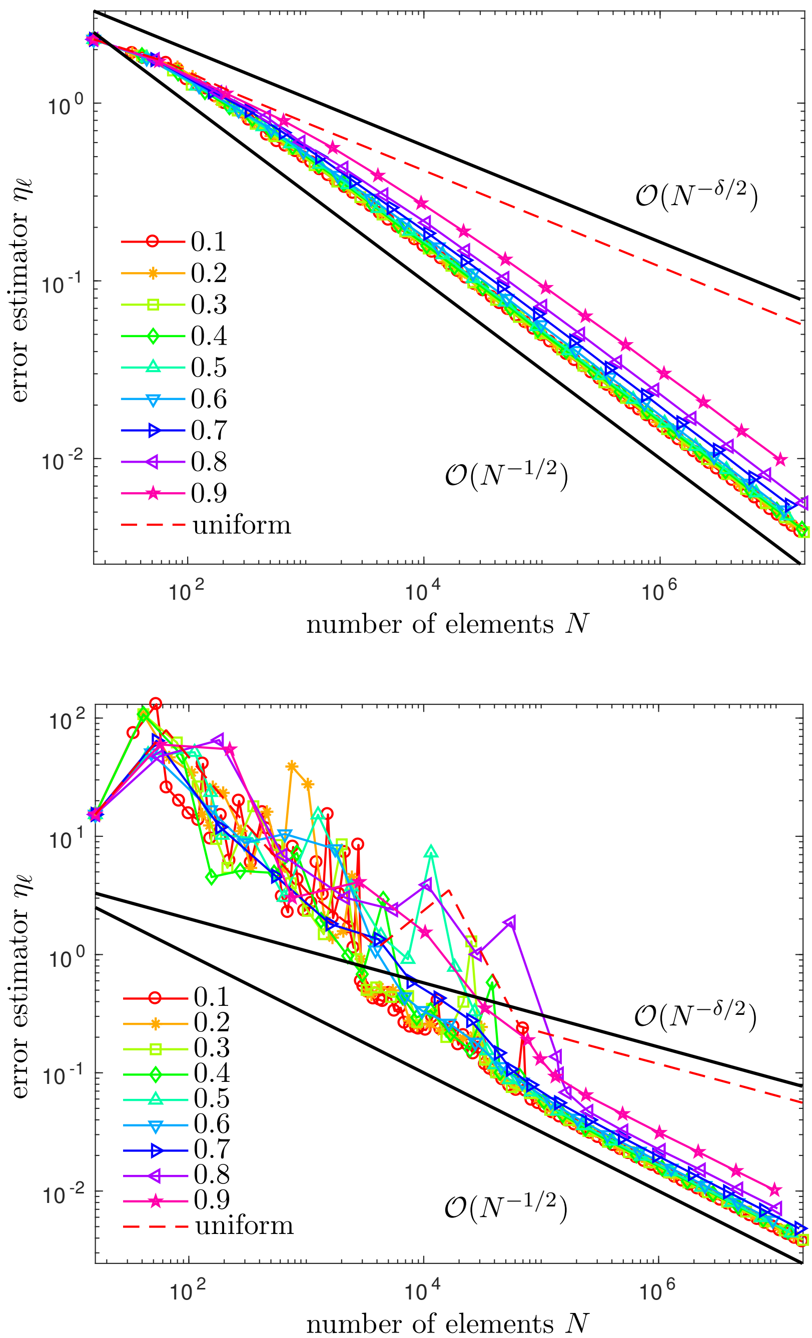
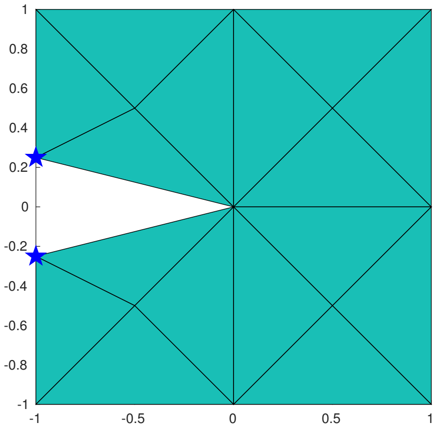
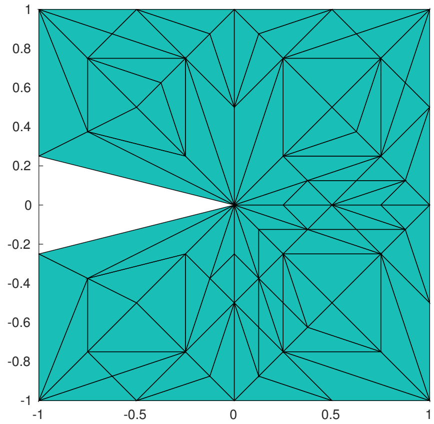


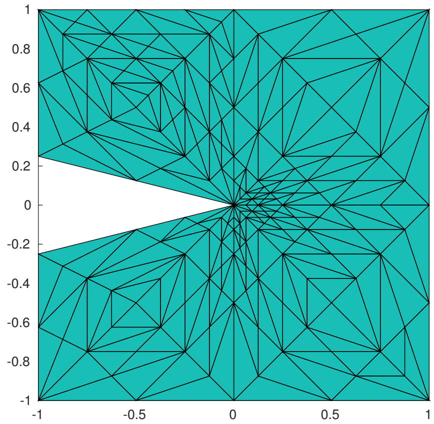

Figure 5.6: Convergence rates for $\eta_\ell$ in the experiment from Section 5.4.2 for uniform and adaptive refinement with different values of $\theta \in \{0.1, \dots, 0.9\}$ and $k = 2$ (above) as well as $k = 16$ (below). Independent of the choice of $\theta < 1$, Algorithm 5.2 leads to optimal convergence behavior. As in Figure 5.5, the preasymptotic phase increases with $k$.



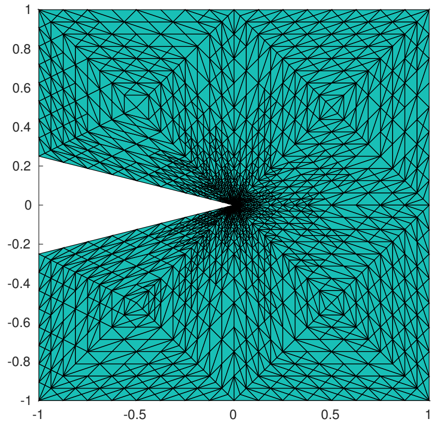

$\#\mathcal{T}_0 = 16$




$\#\mathcal{T}_4 = 82$




$\#\mathcal{T}_8 = 232$




$\#\mathcal{T}_{16} = 2368$

Figure 5.7: Initial mesh $\mathcal{T}_0$ (upper left) and adaptively generated meshes $\mathcal{T}_\ell$ in the experiment from Section 5.4.2 for $k = 2$ and $\theta = 0.2$. The nodes marked with blue stars are given be $(-1, \pm t) = (-1, \pm 0.25)$.

# 6 Adaptive BEM for the Helmholtz equation

## 6.1 State of the art and outline

Adaptive boundary element methods with (dis)continuous piecewise polynomials for second order elliptic problems are well understood if the boundary integral operator is strongly elliptic. In particular, optimal algebraic rates of convergence have been proved in [FFK+14, FFK+15, FKMP13] for polyhedral boundaries and in [Gan17b] for smooth boundaries. An abstract framework is also found in [CFPP14, Fei15]. With [AFF+17] these results can also be extended to piecewise smooth boundaries.

In recent years, isogeometric analysis has lead to a variety of works proving optimal rates for ABEM using spline basis functions; see, e.g., [FGP15, FGHP16, FGHP17] for the Laplace problem in two dimension as well as [Gan17a] for a generalization to second-order linear elliptic PDEs in three dimensions.

On the other hand, boundary element methods for the Helmholtz equation are very popular and used in many applications; see, e.g., [CWGLS12, CK83] for an overview of techniques in acoustic scattering. To our knowledge, there are no results concerning optimal convergence of ABEM for indefinite problems, even for sufficiently fine initial meshes.

As second application of the abstract framework presented in Chapter 4, we consider ABEM for the Dirichlet or Neumann boundary value problem for the Helmholtz equation, i.e.,

$$\begin{aligned} -\Delta u - k^2 u &= 0 \text{ in } \Omega \qquad && \text{subject to} \\ u &= g \text{ on } \Gamma && \text{or} \\ \partial_{\mathbf{n}} u &= \phi \text{ on } \Gamma. \end{aligned}$$

where $k \in \mathbb{R}$ denotes the wavenumber. Using a Potential decomposition from [Mel12], we generalize the inverse estimates in [AFF+17] for the Laplacian to the Helmholtz case. Using the inverse estimate, we prove the estimator axioms (E1)–(E5). Hence, we can apply the abstract framework of Chapter 4. Then, the latter provides the main results of this chapter (Theorem 6.11 and Theorem 6.14), where linear convergence as well as optimal algebraic convergence behavior is proved for the weakly-singular as well as the hyper-singular integral equation.

**Outline of chapter.** First, in Section 6.2.1 we give a short introduction into BEM, where we recap the most important properties of the involved integral operators. As model problem serves the weakly-singular integral equation, which is introduced in Section 6.3. There, we also discuss the weighted-residual error estimator and recall the adaptive Algorithm 4.4 in the current setting. In Section 6.4, we prove the inverse estimates in the style of [AFF+17] for the Helmholtz operators. To that end, we introduce Besov spaces in

Section 6.4.1 and decompositions of the corresponding potential operators in Section 6.4.2. Utilizing the inverse estimates, Section 6.5 verifies the estimator axioms. The main result (Theorem 6.11) of this chapter is given in Section 6.6. Further, in Section 6.7, we focus on the hyper-singular integral equation and prove optimal convergence rates also in this setting. In the last Section 6.8, we underpin our theoretical findings with some experiments.

## 6.2 Boundary element method for the Helmholtz equation

This section gives a short introduction into boundary element methods. Our focus is on the Helmholtz equation and the related integral operators. The interior Helmholtz equation reads

$$-\Delta u - k^2 u = 0 \quad \text{in } \Omega, \tag{6.1a}$$

where $k \in \mathbb{R}$ denotes the wavenumber. The corresponding outer problem reads

$$-\Delta u - k^2 u = 0 \quad \text{in } \mathbb{R}^d \setminus \Omega, \tag{6.1b}$$

where we additionally impose the Sommerfeld radiation condition

$$\left.\begin{aligned} |u(x)| &\le C\,|x|^{-1} \\ |\partial u/\partial r - iku| &\le C\,|x|^{-2} \end{aligned}\right\} \quad \text{for } |x| \to \infty. \tag{6.1c}$$

Here $\partial u/\partial r = x/|x| \cdot \nabla u$ denotes the radial derivative. Note that, (6.1c) is needed to guarantee unique solvability for the outer problem (6.1b). The next section introduces the corresponding integral operators.

### 6.2.1 Layer potential and boundary integral operators

Recall the notation of Chapter 2. Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with piecewise $C^\infty$-boundary $\partial\Omega$. For $d = 2$, we additionally assume that $\mathrm{diam}(\Omega) < 1$. Let $\Omega^{\mathrm{ext}} := \mathbb{R}^d \setminus \Omega$ be the exterior domain and $\mathbf{n}(y)$ denote the exterior normal vector for all $x \in \partial\Omega$. For $k > 0$, the Helmholtz kernel function and the fundamental solution of (6.1) are given by

$$G_k(x, y) := \begin{cases} \frac{i}{4} H_0^{(1)}(k|x-y|) & \text{for } d = 2, \\ \frac{1}{4\pi|x-y|}\, e^{ik|x-y|} & \text{for } d = 3, \end{cases} \tag{6.2}$$

where $H_0^{(1)}$ denotes the first-kind Hankel function of order zero; see, e.g., [Ste08a, Chapter 5.4]. For $k = 0$, the kernel $G_0(x, y)$ coincides with the fundamental solution of the Laplace operator and reads

$$G_0(x, y) := \begin{cases} -\frac{1}{2\pi} \log|x-y| & \text{for } d = 2, \\ \frac{1}{4\pi|x-y|} & \text{for } d = 3. \end{cases} \tag{6.3}$$

In case of a negative wavenumber $k < 0$, we define $G_k := \overline{G_{-k}}$. For smooth solutions $u \in C^2(\overline{\Omega})$ of equation (6.1a), there holds the representation formula

$$u(x) = \int_{\partial\Omega} G_k(x,y)\,\partial_{\mathbf{n}(y)} u(y)\,dy - \int_{\partial\Omega} \partial_{\mathbf{n}(y)} G_k(x,y)\,u(y)\,dy \quad \text{for all } x \in \Omega. \tag{6.4}$$

Hence, in oder to get the solution $u(x)$ in $\Omega$, depending on the given boundary conditions, one has to compute either $\partial_{\mathbf{n}}\, u$ or $u$ on the boundary. We emphasize that the representation formula (6.4) holds, up to a different sign, also for smooth solutions of the exterior problem (6.1b) with the Sommerfeld radiation condition (6.1c). Without (6.1c), one can prove a similar result with an additional Helmholtz-harmonic function on the right hand side; see, e.g., [SS11, Theorem 3.1.8].

The representation formula motivates the definition of the following potential operators. For all $k \in \mathbb{R}$, we define the corresponding simple-layer potential by

$$(\widetilde{\mathfrak{V}}_k \phi)(x) := \int_{\partial\Omega} G_k(x,y)\phi(y)\,dy \quad \text{for all } x \in \mathbb{R}^d \setminus \partial\Omega, \tag{6.5}$$

and double-layer potential by

$$(\widetilde{\mathfrak{K}}_k \phi)(x) := \int_{\partial\Omega} \partial_{\mathbf{n}(y)} G_k(x,y)\phi(y)\,dy \quad \text{for all } x \in \mathbb{R}^d \setminus \partial\Omega. \tag{6.6}$$

For $-1/2 < s < 1/2$, these potentials give rise to bounded linear operators

$$\widetilde{\mathfrak{V}}_k : H^{-1/2+s}(\partial\Omega) \to H^1_{\mathrm{loc}}(\mathbb{R}^d) \quad \text{and} \quad \widetilde{\mathfrak{K}}_k \in H^{1/2+s}(\partial\Omega) \to H^1_{\mathrm{loc}}(\mathbb{R}^d), \tag{6.7}$$

where $H^1_{\mathrm{loc}}(\mathbb{R}^d)$ denotes the space of $H^1$-functions with compact support; see [SS11, Theorem 3.1.16].

Recall the interior and exterior trace operators $\gamma_0^{\mathrm{int}}, \gamma_0^{\mathrm{ext}}$ as well as the conormal derivatives $\gamma_1^{\mathrm{int}}, \gamma_1^{\mathrm{ext}}$ from Section 2.2.4. For $-1/2 < s < 1/2$, application of this trace operators gives rise to the following linear and continuous boundary integral operators:

- simple-layer operator

$$\mathfrak{V}_k : H^{-1/2+s}(\partial\Omega) \to H^{1/2+s}(\partial\Omega) \quad \text{with} \quad \mathfrak{V}_k := \gamma_0^{\mathrm{int}} \widetilde{\mathfrak{V}}_k; \tag{6.8}$$

- double-layer operators

$$\mathfrak{K}_k^\sigma : H^{1/2+s}(\partial\Omega) \to H^{1/2+s}(\partial\Omega) \quad \text{with} \quad \mathfrak{K}_k^\sigma := \gamma_0^\sigma \widetilde{\mathfrak{K}}_k \quad \text{and } \sigma \in \{\mathrm{int}, \mathrm{ext}\},$$
$$\mathfrak{K}_k : H^{1/2+s}(\partial\Omega) \to H^{1/2+s}(\partial\Omega) \quad \text{with} \quad \mathfrak{K}_k := \frac{1}{2}\,(\mathfrak{K}_k^{\mathrm{int}} + \mathfrak{K}_k^{\mathrm{ext}});$$

- adjoint double-layer operator

$$\mathfrak{K}_k' : H^{-1/2+s}(\partial\Omega) \to H^{-1/2+s}(\partial\Omega) \quad \text{with} \quad \mathfrak{K}_k' := -\frac{1}{2}\,\mathrm{Id} + \gamma_1^{\mathrm{int}}\, \widetilde{\mathfrak{V}}_k, \tag{6.9}$$

- hyper-singular operator

$$\mathfrak{W}_k : H^{1/2+s}(\partial\Omega) \to H^{-1/2+s}(\partial\Omega) \quad \text{with} \quad \mathfrak{W}_k := -\gamma_1^{\mathrm{int}} \widetilde{\mathfrak{K}}_k \tag{6.10}$$

A proof of the stated mapping properties is given in [SS11, Theorem 3.1.16] resp. [SS11, Section 3.1.2]. We emphasize that the simple-layer operator $\mathfrak{V}_k$ is symmetric and even continuous for $s = 1/2$ (see Theorem 6.3). For $k = 0$, $\mathfrak{V}_0$ is a well-defined isomorphism for $-1/2 \le s \le 1/2$, and elliptic and symmetric for $s = 0$. Further, for $k \neq 0$, it is well-known that the simple-layer operator $\mathfrak{V}_k$ is invertible, if and only if $k^2$ is not an eigenvalue of the interior Dirichlet problem (IDP) for the Laplace operator, i.e., it holds that

$$\forall u \in H^1(\Omega) \quad \Big( -\Delta u = k^2 u \quad \text{with} \quad \gamma_0^{\mathrm{int}} u = 0 \quad \Longrightarrow \quad u = 0 \quad \text{in } \Omega \Big); \tag{IDP}$$

see, e.g., [SS11, Theorem 3.9.1]. Hence, to ensure solvability, we assume throughout this chapter that $k^2$ is not an eigenvalue of (IDP).

Similar to the simple-layer operator, it can be shown that the double-layer operator is continuous for $s = 1/2$ (see Theorem 6.3). In case of the Laplace equation, i.e., $k = 0$, the operators $\mathfrak{K}_0$, $\mathfrak{K}_0'$, as well as $\mathfrak{W}_0$ are even well defined for $s = \pm 1/2$; see [SS11, Remark 3.1.18]). For further properties on the hypersingular operator $\mathfrak{W}_k$, we refer to Section 6.7.

For ease of presentation, the main part of this chapter (Sections 6.3–6.6) focuses on the weakly-singular integral equation. The hyper-singular integral equation is discussed in Section 6.7.

According to the representation formula (6.4), the solution $u(x)$ in $\Omega$ is given in terms of the normal derivative $\partial_{\mathbf{n}} u$ and the trace $u$ on the boundary $\partial\Omega$. For the Dirichlet problem, i.e., (6.1a) subject to $u = g$ on $\partial\Omega$, the missing normal derivative $\phi = \partial_{\mathbf{n}} u$ is given by Symm's intgral equation

$$\mathfrak{V}_k \phi = \big(\mathfrak{K}_k + \frac{1}{2}\,\mathrm{Id}\big)\, g \quad \text{on } \partial\Omega. \tag{6.11}$$

For ease of presentation, we restrict ourself to an indirect formulation, where the solution $u$ of the Dirichlet problem is given in terms of the simple-layer potential

$$u = \widetilde{\mathfrak{V}}_k\, \phi, \qquad \text{where } \phi \text{ is the solution of} \qquad \mathfrak{V}_k\, \phi = g \quad \text{on } \partial\Omega. \tag{6.12}$$

Note that in the indirect formulation, the density $\phi$ has no direct physical meaning. However, we stress that all results of Sections (6.3)–(6.6) also holds for direct boundary element methods; cf. Remark 6.12. For the Neumann problem, i.e., (6.1a) subject to $\partial_{\mathbf{n}} u = \phi$ on $\partial\Omega$, we refer to Section 6.7.

## 6.3 Model problem

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with piecewise $C^\infty$-boundary $\partial\Omega$. For $d = 2$, we additionally assume that $\mathrm{diam}(\Omega) < 1$. Suppose $\Gamma = \partial\Omega$ or $\emptyset \neq \Gamma \subset \partial\Omega$ is a relative open set which stems from a Lipschitz dissection $\Gamma \cup \partial\Gamma \cap (\partial\Omega \setminus \Gamma)$.

To simplify notation, we make the following convention: If $\Gamma \subsetneqq \partial\Omega$, and it is clear from the context, we identify any $v \in \widetilde{H}^{1/2+s}(\Gamma)$ with its extension $E_{0,\Gamma}v \in H^{1/2+s}(\partial\Omega)$. Further, the operators $\widetilde{\mathfrak{V}}_k, \mathfrak{V}_k, \mathfrak{K}_k'$ are often applied to functions in $L^2(\Gamma)$, resp. $\widetilde{\mathfrak{K}}_k, \mathfrak{K}_k, \mathfrak{W}_k$ are applied to functions in $\widetilde{H}^{1/2}(\Gamma)$. To ease notation, we use the following convention: For $\psi \in L^2(\Gamma)$ and $v \in \widetilde{H}^{1/2}(\Gamma)$, we implicitly extend with zero, e.g., we write $\mathfrak{V}_k\psi$ instead of $\mathfrak{V}_k(E_{0,\Gamma}\psi)$ and $\mathfrak{K}_k v$ instead of $\mathfrak{K}_k(E_{0,\Gamma}v)$.

As model problem, we consider the weakly-singular integral equation: Given an $f \in H^{1/2}(\Gamma)$, find $\phi \in \widetilde{H}^{-1/2}(\Gamma)$ such that

$$\mathfrak{V}_k\,\phi = f \quad \text{on } \Gamma. \tag{6.13}$$

Recall the notation of Section 2.2, and let $\langle\cdot\,,\,\cdot\rangle$ denote the duality pairing which extends the $L^2(\partial\Omega)$-scalar product. The weak formualtion of (6.13) reads as. Given $f \in H^{1/2}(\Gamma)$, find $\phi \in \widetilde{H}^{-1/2}(\Gamma)$ such that

$$\langle \mathfrak{V}_k\,\phi\,,\,\psi\rangle = \langle f\,,\,\psi\rangle \quad \text{for all } \psi \in \widetilde{H}^{-1/2}(\Gamma). \tag{6.14}$$

In case of $k = 0$, $\mathfrak{V}_0$ corresponds with the simple-layer operator of the Laplace equation. Adaptive algorithms for the Laplace equations are well studied and understood. In this case, the proof of optimal algebraic rates of convergence for the weakly-singular integral equation (6.13) is found in [FKMP13, FFK+14, Gan13].

In order to fit in the abstract framework of Chapter 4, we recast the model problem (6.13) in the following functional analytic setting. The bilinear form $a(\cdot\,,\,\cdot) : \widetilde{H}^{-1/2}(\Gamma)\times \widetilde{H}^{-1/2}(\Gamma) \to \mathbb{R}$ is given by

$$a(\chi\,,\,\psi) := \langle \mathfrak{V}_0\,\chi\,,\,\psi\rangle \quad \text{for all } \chi,\psi \in \widetilde{H}^{-1/2}(\Gamma). \tag{6.15}$$

Further more, we define the linear operator

$$\mathfrak{C}_k := \mathfrak{V}_k - \mathfrak{V}_0 : \widetilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma). \tag{6.16}$$

Then, the model problem (6.13) can equivalently be reformulated as follows: Given $f \in H^{1/2}(\Gamma)$, find $\phi \in \widetilde{H}^{-1/2}(\Gamma)$ such that

$$\big(\mathfrak{V}_0 + \mathfrak{C}_k\big)\phi = f \quad \text{on } \Gamma. \tag{6.17}$$

The weak formulation of (6.17) thus reads as: Given $f \in H^{1/2}(\Gamma)$, find $\phi \in \widetilde{H}^{-1/2}(\Gamma)$ such that

$$b(\phi\,,\,\psi) := a(\phi\,,\,\psi) + \langle \mathfrak{C}_k\,\phi\,,\,\psi\rangle = \langle f\,,\,\psi\rangle \quad \text{for all } \psi \in \widetilde{H}^{-1/2}(\Gamma). \tag{6.18}$$

The following Proposition recaps some important properties of $a(\cdot\,,\,\cdot)$, and $\mathfrak{C}$ and ensures that (6.17) fits in the compactly perturbed setting of Chapter 4.

**Proposition 6.1.** *There exist constants $\alpha, C_{\mathrm{cont}} > 0$ such that the bilinear form $a(\cdot\,,\,\cdot)$ from* (6.15) *and $\mathfrak{C}_k$ from* (6.16) *satisfy:*

(i) $a(\cdot\,,\,\cdot)$ *is symmetric, continuous and elliptic with*

$$|a(\chi\,,\,\psi)| \le C_{\rm cont}\,\|\chi\|_{\widetilde{H}^{-1/2}(\Gamma)}\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} \quad \textit{and} \quad \alpha\,\|\psi\|^2_{\widetilde{H}^{-1/2}(\Gamma)} \le a(\psi\,,\,\psi),$$

*for all* $\chi,\psi \in \widetilde{H}^{-1/2}(\Gamma)$*. Hence,* $a(\cdot\,,\,\cdot)$ *defines a scalar product and induces an equivalent energy norm* $|\!|\!|\,\psi\,|\!|\!| := a(\psi\,,\,\psi)^{1/2} \simeq \|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}$ *on* $\widetilde{H}^{-1/2}(\Gamma)$*.*

(ii) *The operator* $\mathfrak{C}_k : \widetilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ *is bounded and compact.*

*Proof.* We split the proof into two steps.
**Step 1: Proof of (i).** Recall that $\mathfrak{V}_0$ is a symmetric isomorphism. This yields that

$$a(\chi\,,\,\psi) = \langle \mathfrak{V}_0\,\chi\,,\,\psi\rangle = \langle \chi\,,\,\mathfrak{V}_0\,\psi\rangle = a(\psi\,,\,\chi) \quad \text{for all } \chi,\psi \in \widetilde{H}^{-1/2}(\Gamma).$$

Hence, $a(\cdot\,,\,\cdot)$ is symmetric. Recall the extension operator $E_{0,\Gamma}$ from Section 2.2. Continuity of $\mathfrak{V}_0 : H^{-1/2+s}(\partial\Omega) \to H^{1/2+s}(\partial\Omega)$ further implies that

$$\begin{aligned}
|a(\chi\,,\,\psi)| = |\langle \mathfrak{V}_0\,\chi\,,\,\psi\rangle| &\le \|\mathfrak{V}_0\,\chi\|_{H^{1/2}(\Gamma)}\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}\\
&= \|\mathfrak{V}_0(E_{0,\Gamma}\chi)\|_{H^{1/2}(\Gamma)}\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}\\
&\le \|\mathfrak{V}_0\|\,\|E_{0,\Gamma}\chi\|_{H^{-1/2}(\partial\Omega)}\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}\\
&= \|\mathfrak{V}_0\|\,\|\chi\|_{\widetilde{H}^{-1/2}(\Gamma)}\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.
\end{aligned}$$

This proves continuity of $a(\cdot\,,\,\cdot)$. In case of $d = 3$, $\mathfrak{V}_0$ is elliptic and we directly obtain that

$$|a(\chi\,,\,\chi)| = |\langle \mathfrak{V}_0\,(E_{0,\Gamma}\chi)\,,\,E_{0,\Gamma}\chi\rangle| \gtrsim \|E_{0,\Gamma}\,\psi\|^2_{H^{-1/2}(\partial\Omega)} = \|\psi\|^2_{\widetilde{H}^{-1/2}(\Gamma)},$$

where the hidden constant depends only on $\partial\Omega$. For $d = 2$, note that $\mathrm{diam}(\Omega) < 1$ implies ellipticity of $\mathfrak{V}_0$; see, e.g., [Ste08a, Section 6.6]. This concludes the proof of (i).

**Step 2: Proof of (ii).** Boundedness of $\mathfrak{C}_k$ follows directly from the definition. On Lipschitz boundaries $\partial\Omega$, the operator $\mathfrak{C}_k := \mathfrak{V}_k - \mathfrak{V}_0 : H^{-1/2}(\partial\Omega) \to H^{1/2}(\partial\Omega)$ is compact; see, e.g., [SS11, Lemma 3.9.8] or [Ste08a, Section 6.9]. This implies compactness of $\mathfrak{C}_k : \widetilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$. $\square$

For mesh-refinement, we consider extended 1D bisection (see Section 3.4) in the case of $d = 2$. For $d = 3$, we use NVB (Section 3.5) on the two dimensional boundary $\Gamma$. Then, Section 3.4 and Section 3.5 prove the refinement axioms (R1)–(R6) for both strategies.

For discretization, we consider standard piecewise polynomial ansatz and test spaces, based on regular triangulations of $\Gamma$ (see Section 3.2). To that end, let $\mathcal{T}_0$ be a given regular and $\gamma$-shape regular initial triangulation on $\Gamma$. Further, let $p \in \mathbb{N}_0$ be an arbitrary but fixed polynomial degree. For each admissible mesh $\mathcal{T}_\bullet$, the corresponding $\mathcal{T}_\bullet$-piecewise polynomial space is denoted by $\mathcal{P}^p(\mathcal{T}_\bullet)$. Then, the Galerkin discretization of (6.17) reads as follows: Find $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ such that

$$a(\Phi_\bullet\,,\,\Psi_\bullet) + \langle \mathfrak{C}_k\,\Phi_\bullet\,,\,\Psi_\bullet\rangle = \langle f\,,\,\Psi_\bullet\rangle \quad \text{for all } \Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet). \tag{6.19}$$

It remains to prove that iterated uniform refinement leads to a dense subspace of $\widetilde{H}^{-1/2}(\Gamma)$. To that end, let $\Pi_\bullet : L^2(\Gamma) \to \mathcal{P}^0(\mathcal{T}_\bullet)$ denote the $L^2(\Gamma)$- orthogonal projection onto $\mathcal{P}^0(\mathcal{T}_\bullet)$. For all $\psi \in L^2(\Gamma)$, it holds that

$$\|(1 - \Pi_\bullet)\,\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} \lesssim \|h_\bullet^{1/2}\,(1 - \Pi_\bullet)\psi\|_{L^2(\Gamma)}, \tag{6.20}$$

where the hidden constant depends on the shape regularity of $\mathcal{T}_\bullet$; see, e.g., [AFF$^+$17, Corollary 3.3] or [SS11, Section 4.3.4]. Note that there also holds nestedness $\mathcal{P}^0(\mathcal{T}_\bullet) \subseteq \mathcal{P}^p(\mathcal{T}_\bullet)$ for $p \geq 0$. For $\psi \in L^2(\Gamma)$, we thus obtain that

$$\inf_{\Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)} \|\psi - \Psi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \overset{(6.20)}{\lesssim} \|h_\bullet^{1/2}\,(1 - \Pi_\bullet)\psi\|_{L^2(\Gamma)} \lesssim \|h_\bullet^{1/2}\|_{L^\infty(\Gamma)}\,\|\psi\|_{L^2(\Gamma)}.$$

Recall that the embedding $L^2(\Gamma) \subset \widetilde{H}^{-1/2}(\Gamma)$ is dense. Uniform mesh-refinement guarantees $\|h_\bullet\|_{L^\infty(\Omega)} \to 0$ and hence, leads to a dense sequence of subspaces. Note that the involved constants only depend on the shape regularity of $\mathcal{T}_\bullet$ and $\Gamma$. Thus, the above argumentation holds for every refinement strategy satisfying (R2).

Proposition 6.1 proves that model problem (6.13) as well as (6.17) and (6.19) fit in the compactly perturbed framework of Section 4.2 with $\mathcal{H} := \widetilde{H}^{-1/2}(\Gamma)$ and $\mathcal{X}_\bullet := \mathcal{P}^p(\mathcal{T}_\bullet)$. Hence, existence and uniqueness of solutions of (6.19) is guaranteed in the sense of Proposition 4.1.

### 6.3.1 Weighted-residual error estimator

In this subsection, we introduce the weighted-residual error estimator for the weakly-singular integral equation (6.13). To that end, suppose $f \in H^1(\Gamma)$ and $\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}_0)$, such that the discrete solution $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ of (6.19) exists. To guarantee well posedness of the estimator, we note that there holds $\mathfrak{V}_k : L^2(\Gamma) \to H^1(\Gamma)$ (see Theorem 6.3). For all $T \in \mathcal{T}_\bullet$, the local contributions are defined by

$$\eta_\bullet(T) := \|h_\bullet^{1/2}\,\nabla_\Gamma\,(\mathfrak{V}_k\Phi_\bullet - f)\|_{L^2(T)}.$$

The corresponding *a posteriori* error estimator is given by

$$\eta_\bullet := \eta_\bullet(\mathcal{T}_\bullet) \quad \text{with} \quad \eta_\bullet(\mathcal{U}_\bullet) := \Big(\sum_{T \in \mathcal{U}_\bullet} \eta_\bullet(T)^2\Big)^{1/2} \quad \text{for all } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet. \tag{6.21}$$

Further, for a set of elements $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet$, we define

$$\bigcup \mathcal{U}_\bullet := \{x \in \Gamma : \exists T \in \mathcal{T}_\bullet\ x \in T\}.$$

Then, it holds that $\eta_\bullet(\mathcal{U}_\bullet) = \|h^{1/2}\,\nabla_\Gamma\,(\mathfrak{V}_k\Phi_\bullet - f)\|_{L^2(\bigcup \mathcal{U}_\bullet)}$. The weighted-residual error estimator (6.21) has first been proposed for *a posteriori* BEM error control for the weakly-singular integral equation in 2D in [CS95, Car96] and later in 3D in [CMS01].

### 6.3.2 Adaptive algorithm

With the *a posteriori* error estimator $\eta_\bullet$ at hand, we consider the following adaptive algorithm which consists of Algorithm 4.4 combined with the expanded Dörfler marking from Proposition 4.7.

**Algorithm 6.2.** Input: *Parameters* $0 < \theta \leq 1$ *and* $C_{\mathrm{mark}} \geq 1$ *as well as initial triangulation* $\mathcal{T}_0$ *with* $\Phi_{-1} := 0 \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{-1} := 1$.
Adaptive loop: *For all* $\ell = 0, 1, 2, \dots,$ *iterate the following Steps* (i)–(vi)*:*

(i) `If` (6.19) *does not admit a unique solution in* $\mathcal{P}^p(\mathcal{T}_\ell)$*:*
   - *Define* $\Phi_\ell := \Phi_{\ell-1} \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_\ell := \eta_{\ell-1}$.
   - *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ *be the uniform refinement of* $\mathcal{T}_\ell$,
   - *Increase* $\ell \to \ell + 1$ *and continue with Step* (i).

(ii) `Else` *compute the unique solution* $\Phi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$ *to* (6.19).

(iii) *Compute the corresponding indicators* $\eta_\ell(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iv) *Determine a set* $\mathcal{M}'_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\mathrm{mark}}$ *minimal cardinality such that* $\theta\eta_\ell^2 \leq \eta_\ell(\mathcal{M}'_\ell)^2$.

(v) *Find* $\mathcal{M}''_\ell \subseteq \mathcal{T}_\ell$ *such that* $\#\mathcal{M}''_\ell = \#\mathcal{M}'_\ell$ *as well as* $h_\ell(T) \geq h_\ell(T')$ *for all* $T \in \mathcal{M}''_\ell$ *and* $T' \in \mathcal{T}_\ell \setminus \mathcal{M}''_\ell$. *Define* $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \mathcal{M}''_\ell$.

(vi) *Generate* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, *increase* $\ell \to \ell + 1$, *and continue with Step* (i).

Output: *Sequences of successively refined triangulations* $\mathcal{T}_\ell$, *discrete solutions* $\Phi_\ell$, *and corresponding estimators* $\eta_\ell$.

Apart from the model problem and involved discrete spaces, Algorithm 6.2 coincides with Algorithm 5.2 for adaptive finite elements.

## 6.4 Inverse estimates

The main result of this section is the following inverse-type estimate. In case of the Laplace equation ($k = 0$), similar estimates are shown in [FKMP13, Gan13] for polyhedral or smooth boundaries. In [AFF$^+$17], the analysis has been lifted to piecewise $C^1$-boundaries.

The proof of Theorem 6.3 uses ideas from [AFF$^+$17, Mel12] and generalizes the existing results in [FKMP13, Theorem 3.1] and in [AFF$^+$17, Theorem 3.1] from $k = 0$ to general $k > 0$. We note that, Theorem 6.3 is an essential tool to prove the estimator Axioms (E1)–(E4) in Section 6.5.

**Theorem 6.3.** *The simple-layer and double-layer operator satisfy*

$$\mathfrak{V}_k \in L\big(L^2(\Gamma), H^1(\Gamma)\big) \quad \textit{resp.} \quad \mathfrak{K}_k \in L\big(\widetilde{H}^1(\Gamma), H^1(\Gamma)\big). \tag{6.22}$$

*Additionally, let $\mathcal{T}_\bullet$ be a $\gamma$-shape regular triangulation of $\Gamma$. Then, there exists a constant $C_{\rm inv} > 0$ which depends only on $\Gamma$, $\Omega$, and $\gamma$, such that for all $k > 0$, it holds that*

$$C_{\rm inv}^{-1} \|h_\bullet^{1/2} \nabla_\Gamma \mathfrak{V}_k \psi\|_{L^2(\Gamma)} \leq (1 + k^3) \|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2} \psi\|_{L^2(\Gamma)}, \tag{6.23}$$

$$C_{\rm inv}^{-1} \|h_\bullet^{1/2} \nabla_\Gamma \mathfrak{K}_k v\|_{L^2(\Gamma)} \leq (1 + k^3) \|v\|_{\widetilde{H}^{1/2}(\Gamma)} + \|h_\bullet^{1/2} \nabla_\Gamma v\|_{L^2(\Gamma)}, \tag{6.24}$$

$$C_{\rm inv}^{-1} \|h_\bullet^{1/2} \mathfrak{K}'_k \psi\|_{L^2(\Gamma)} \leq (1 + k^3) \|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2} \psi\|_{L^2(\Gamma)}, \tag{6.25}$$

$$C_{\rm inv}^{-1} \|h_\bullet^{1/2} \mathfrak{W}_k v\|_{L^2(\Gamma)} \leq (1 + k^3) \|v\|_{\widetilde{H}^{1/2}(\Gamma)} + \|h_\bullet^{1/2} \nabla_\Gamma v\|_{L^2(\Gamma)}, \tag{6.26}$$

*for all $\psi \in L^2(\Gamma)$ and $v \in \widetilde{H}^1(\Gamma)$. Furthermore, there exists $\widetilde{C}_{\rm inv} > 0$ which depends only on $\Omega$, $\Gamma$, $\gamma$, and $p$, such that*

$$\|h_\bullet^{1/2} \nabla_\Gamma \mathfrak{V}_k \Psi_\bullet\|_{L^2(\Gamma)} + \|h_\bullet^{1/2} \mathfrak{K}'_k \Psi_\bullet\|_{L^2(\Gamma)} \leq \widetilde{C}_{\rm inv} (1 + k^3) \|\Psi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)}, \tag{6.27}$$

$$\|h_\bullet^{1/2} \nabla_\Gamma \mathfrak{K}_k V_\bullet\|_{L^2(\Gamma)} + \|h_\bullet^{1/2} \mathfrak{W}_k V_\bullet\|_{L^2(\Gamma)} \leq \widetilde{C}_{\rm inv} (1 + k^3) \|V_\bullet\|_{\widetilde{H}^{1/2}(\Gamma)}, \tag{6.28}$$

*for all $\Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ and $V_\bullet \in \widetilde{\mathcal{S}}^{p+1}(\mathcal{T}_\bullet)$. In particular, the constants $C_{\rm inv}, \widetilde{C}_{\rm inv}$ are independent of the wavenumber $k > 0$.*

### 6.4.1 Function spaces revisited

The proof of Theorem 6.3 involves certain Besov spaces on domains $\widetilde{\Omega}$. Therefore, we recap the definitions and some important properties. Besov spaces can be defined by the K-method of interpolation; see, e.g., [Tar07, Tri83, Tri92]. To that end, let $\widetilde{\Omega} \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $s \in \mathbb{N}_0$ and $s' \in (0,1)$. Then, the Besov space $B^{s+s'}_{2,\infty}(\widetilde{\Omega})$ is given by

$$B^{s+s'}_{2,\infty}(\widetilde{\Omega}) := \big[H^s(\widetilde{\Omega}), H^{s+1}(\widetilde{\Omega})\big]_{s',\infty}.$$

Moreover, according to [Tar07, Lemma 22.2], there holds the continuous embedding $H^{s+s'}(\widetilde{\Omega}) \subset B^{s+s'}_{2,\infty}(\widetilde{\Omega})$.

### 6.4.2 Potential decompositions

The proof of Theorem 6.3 is based on the decomposition of the layer potentials into a singular part, which consists of the layer potentials $\widetilde{\mathfrak{V}}_0$, resp., $\widetilde{\mathfrak{K}}_0$ of the Laplacian, and two smoothing operators. We employ the following notation

$$|\nabla^n \psi(x)|^2 := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| = n}} \frac{n!}{\alpha!} |D^\alpha \psi(x)|^2 \text{ with } \alpha! := \alpha_1! \cdot \alpha_2! \cdot \ldots \cdot \alpha_d! \quad \text{and} \quad |\nabla^0 \psi(x)|^2 := |\psi(x)|^2.$$

Lemma 6.4 provides a decomposition for the simple-layer potential, while Lemma 6.5 states a similar result for the double-layer potential. Both results are proved in [Mel12].

**Lemma 6.4** ([Mel12, Theorem 5.1.1]). *Let $R > 0$ with $\overline{\Omega} \subsetneqq B_R := \{x \in \mathbb{R}^d : |x| < R\}$. Let $0 < \rho < 1$. Then, it holds that*

$$\widetilde{\mathfrak{V}}_k = \widetilde{\mathfrak{V}}_0 + \widetilde{\mathfrak{S}}_{\mathfrak{V},k} + \widetilde{\mathfrak{A}}_{\mathfrak{V},k}, \tag{6.29}$$

*with linear potential operators $\widetilde{\mathfrak{S}}_{\mathfrak{V},k} : H^{-1/2+s}(\partial\Omega) \to H^{3+s}(B_R)$ and $\widetilde{\mathfrak{A}}_{\mathfrak{V},k} : H^{-1/2+s}(\partial\Omega) \to H^{3+s}(B_R) \cap C^\infty(B_R)$ for all $-1/2 < s < 1/2$. Moreover, there exist constants $C_1^V, C_2^V, C_3^V > 0$ such that*

$$\|\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{H^{s'}(B_R)} \le C_1^V\,\rho^2\,(\rho k^{-1})^{1+s-s'}\|\psi\|_{H^{-1/2+s}(\partial\Omega)} \quad \textit{for all } 0 \le s' \le 3+s, \tag{6.30}$$

$$\|\nabla^n\,\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{L^2(B_R)} \le C_2^V k^{n+1}\|\widetilde{\mathfrak{V}}_0\,\psi\|_{L^2(B_R)} \le C_3^V k^{n+1}\|\psi\|_{H^{-1}(\partial\Omega)} \quad \textit{for all } n \in \mathbb{N}_0. \tag{6.31}$$

*The constants $C_1^V, C_2^V$, and $C_3^V$ depend only on $\rho$, $R$, $\Omega$ , but not on the wavenumber $k$.* $\square$

Similar to the simple-layer potential, the double-layer potential can be split in the following way.

**Lemma 6.5** ([Mel12, Theorem 5.2]). *Let $R > 0$ with $\overline{\Omega} \subsetneqq B_R := \{x \in \mathbb{R}^d : |x| < R\}$. Let $0 < \rho < 1$. Then, it holds that*

$$\widetilde{\mathfrak{K}}_k = \widetilde{\mathfrak{K}}_0 + \widetilde{\mathfrak{S}}_{\mathfrak{K},k} + \widetilde{\mathfrak{A}}_{\mathfrak{K},k}, \tag{6.32}$$

*with linear potential operators $\widetilde{\mathfrak{S}}_{\mathfrak{K},k} : L^2(\partial\Omega) \to B^{5/2}_{2,\infty}(B_R)$ and $\widetilde{\mathfrak{A}}_{\mathfrak{K},k} : L^2(\partial\Omega) \to B^{5/2}_{2,\infty}(B_R) \cap C^\infty(B_R)$. Moreover, there exist constants $C_1^K, C_2^K, C_3^K > 0$, such that*

$$\|\widetilde{\mathfrak{S}}_{\mathfrak{K},k}\,v\|_{B^{5/2}_{2,\infty}(B_R)} \le C_1^K\,k\,\|v\|_{L^2(\partial\Omega)}, \tag{6.33}$$

$$\|\nabla^n\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{L^2(B_R)} \le C_2^K\,k^{n+1}\,\|\widetilde{\mathfrak{K}}_0 v\|_{L^2(B_R)} \le C_3^K\,k^{n+1}\,\|v\|_{L^2(\partial\Omega)} \quad \textit{for all } n \in \mathbb{N}_0. \tag{6.34}$$

*The constants $C_1^K, C_2^K$, and $C_3^K$ depend only on $\rho$, $R$, $\Omega$, but not on the wavenumber $k$.* $\square$

### 6.4.3 Proof of Theorem 6.3

With the potential decompositions of Lemma 6.4 and Lemma 6.5, we can prove the inverse estimate.

***Proof of Theorem 6.3.*** Let $k > 0$ and $R > 0$ with $B_R \supsetneqq \overline{\Omega}$. For convenience of the reader, we split the proof into several steps.

**Step 1: Proof of (6.22) for $\mathfrak{V}_k$.** Let $\psi \in L^2(\Gamma)$ and recall that $\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} = \|\psi\|_{H^{-1/2}(\partial\Omega)}$, where we identify $\psi$ identified with its extension $E_{0,\Gamma}\,\psi$. With Lemma 6.4 and the definition of $\mathfrak{V}_k := \gamma_0^{\mathrm{int}}\,\widetilde{\mathfrak{V}}_k$, we decompose $\mathfrak{V}_k = \mathfrak{V}_0 + \mathfrak{S}_{\mathfrak{V},k} + \mathfrak{A}_{\mathfrak{V},k}$, where

$$\mathfrak{S}_{\mathfrak{V},k} := \gamma_0^{\mathrm{int}}\,\widetilde{\mathfrak{S}}_{\mathfrak{V},k} \quad \text{and} \quad \mathfrak{A}_{\mathfrak{V},k} := \gamma_0^{\mathrm{int}}\,\widetilde{\mathfrak{A}}_{\mathfrak{V},k}.$$

For all $1/2 < s' \le 3 + s \le 3 + 1/2$, equation (6.30) implies that

$$\|\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{H^{s'}(B_R)} \overset{(6.30)}{\lesssim} \rho^2\,(\rho\,k^{-1})^{1+s-s'}\|\psi\|_{H^{-1/2+s}(\partial\Omega)}. \tag{6.35}$$

For $s' = 2$, this reveals that $\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi \in H^2(B_R)$. Further, stability of $\gamma_0^{\rm int}$ yields that

$$\begin{aligned}\|\mathfrak{S}_{\mathfrak{V},k}\psi\|_{H^1(\Gamma)} &\le \|\mathfrak{S}_{\mathfrak{V},k}\psi\|_{H^1(\partial\Omega)}\\ &\lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\psi\|_{H^{3/2}(B_R)} \lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\psi\|_{H^2(B_R)} \overset{(6.35)}{\lesssim} \rho k\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.\end{aligned} \tag{6.36}$$

Next, note that equation (6.31) proves that $\mathfrak{A}_{\mathfrak{V},k}\psi \in H^2(B_R)$. With the (compact) embedding $H^{-1/2}(\partial\Omega) \subset H^{-1}(\partial\Omega)$ with $\|\cdot\|_{H^{-1}(\partial\Omega)} \lesssim \|\cdot\|_{H^{-1/2}(\partial\Omega)}$, this yields that

$$\|\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{H^2(B_R)} \overset{(6.31)}{\lesssim} (k + k^2 + k^3)\,\|\psi\|_{H^{-1}(\partial\Omega)} \lesssim (1 + k^3)\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}. \tag{6.37}$$

Similarly to (6.36), continuity of the trace operator proves that

$$\begin{aligned}\|\mathfrak{A}_{\mathfrak{V},k}\,\psi\|_{H^1(\Gamma)} &\le \|\mathfrak{A}_{\mathfrak{V},k}\,\psi\|_{H^1(\partial\Omega)}\\ &\lesssim \|\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{H^{3/2}(B_R)} \le \|\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{H^2(B_R)} \overset{(6.37)}{\lesssim} (1 + k^3)\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.\end{aligned} \tag{6.38}$$

Combining the estimates (6.36) and (6.38) with the (compact) embedding $L^2(\Gamma) \subset \widetilde{H}^{-1/2}(\Gamma)$, we see that $\mathfrak{A}_{\mathfrak{V},k}, \mathfrak{S}_{\mathfrak{V},k} \in L\big(L^2(\Gamma), H^1(\Gamma)\big)$. With $\mathfrak{V}_0 \in L\big(L^2(\Gamma), H^1(\Gamma)\big)$, we conclude that $\mathfrak{V}_k = \mathfrak{V}_0 + \mathfrak{S}_{\mathfrak{V},k} + \mathfrak{A}_{\mathfrak{V},k} \in L\big(L^2(\Gamma), H^1(\Gamma)\big)$.

**Step 2: Proof of equation (6.23).** Recall that $\mathfrak{V}_k = \mathfrak{V}_0 + \mathfrak{S}_{\mathfrak{V},k} + \mathfrak{A}_{\mathfrak{V},k}$. This decomposition directly yields that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_k\psi\|_{L^2(\Gamma)} \le \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_0\,\psi\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{S}_{\mathfrak{V},k}\psi\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{A}_{\mathfrak{V},k}\psi\|_{L^2(\Gamma)}. \tag{6.39}$$

We treat each term on the right-hand side separately. First, [AFF+17, Theorem 3.1] yields that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_0\,\psi\|_{L^2(\Gamma)} \lesssim \|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\psi\|_{L^2(\Gamma)}.$$

Second, $\|h_\bullet\|_{L^\infty(\Gamma)} \lesssim \mathrm{diam}(\Omega) \lesssim 1$ and equation (6.36) imply that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{S}_{\mathfrak{V},k}\,\psi\|_{L^2(\Gamma)} \lesssim \|\mathfrak{S}_{\mathfrak{V},k}\,\psi\|_{H^1(\Gamma)} \overset{(6.36)}{\lesssim} k\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.$$

Third, we use equation (6.38) to estimate the last term on the right hand side of (6.39) by

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{A}_{\mathfrak{V},k}\psi\|_{L^2(\Gamma)} \lesssim \|\mathfrak{A}_{\mathfrak{V},k}\psi\|_{H^1(\Gamma)} \overset{(6.38)}{\lesssim} (1 + k^3)\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.$$

Combining the latter four estimates, we prove that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_k\psi\|_{L^2(\Gamma)} \lesssim (1 + k^3)\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\psi\|_{L^2(\Gamma)}.$$

This concludes the proof of (6.23).

**Step 3: Proof of equation (6.25).** Recall the definition of the adjoint double-layer operator. This gives rise to $\mathfrak{K}_k' = -\frac{1}{2}\mathrm{Id} + \gamma_1^{\mathrm{int}}\widetilde{\mathfrak{V}}_k = \mathfrak{K}_0' + \gamma_1^{\mathrm{int}}\widetilde{\mathfrak{S}}_{\mathfrak{V},k} + \gamma_1^{\mathrm{int}}\widetilde{\mathfrak{A}}_{\mathfrak{V},k}$ and hence implies that

$$\|h_\bullet^{1/2}\,\mathfrak{K}_k'\,\psi\|_{L^2(\Gamma)} \le \|h_\bullet^{1/2}\,\mathfrak{K}_0'\,\psi\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{L^2(\Gamma)}.$$

Again, we treat each term on the right-hand side separately. First, [AFF+17, Theorem 3.1] yields that

$$\|h_\bullet^{1/2}\,\mathfrak{K}_0'\,\psi\|_{L^2(\Gamma)} \lesssim \|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\psi\|_{L^2(\Gamma)}.$$

Recall from Step 1 that $\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi$, $\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi \in H^2(B_R)$. Therefore, [SS11, Remark 2.7.5] implies that $\gamma_1^{\mathrm{int}}\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\psi, \gamma_1^{\mathrm{int}}\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\psi \in H^{1/2}(\partial\Omega)$. With $\|h_\bullet\|_{L^\infty(\Gamma)} \lesssim \mathrm{diam}(\Omega) \lesssim 1$, the (compact) embedding $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$ and stability ([SS11, Remark 2.7.5]) of the conormal derivative yield that

$$\|h_\bullet^{1/2}\,\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{L^2(\Gamma)} \lesssim \|\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{H^{1/2}(\partial\Omega)} \lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{V},k}\,\psi\|_{H^2(B_R)} \overset{(6.36)}{\lesssim} k\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.$$

Third, we argue as before and prove that

$$\|h_\bullet^{1/2}\,\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\psi\|_{L^2(\Gamma)} \lesssim \|\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{H^{1/2}(\partial\Omega)} \lesssim \|\widetilde{\mathfrak{A}}_{\mathfrak{V},k}\,\psi\|_{H^2(B_R)} \overset{(6.37)}{\lesssim} (1+k^3)\,\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)}.$$

Combining all right-hand side estimates, we obtain that

$$\|h_\bullet^{1/2}\,\mathfrak{K}_k'\psi\|_{L^2(\Gamma)} \lesssim (1+k^3)\|\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\psi\|_{L^2(\Gamma)},$$

and conclude the proof of (6.25).

**Step 4: Proof of (6.22) for $\mathfrak{K}_k$.** Let $v \in \widetilde{H}^1(\Gamma)$. Analogously to Step 1, Lemma 6.5 gives rise to the decomposition $\mathfrak{K}_k = \mathfrak{K}_0 + \mathfrak{S}_{\mathfrak{K},k} + \mathfrak{A}_{\mathfrak{K},k}$, where $\mathfrak{S}_{\mathfrak{K},k} := \gamma_0^{\mathrm{int}}\,\widetilde{\mathfrak{S}}_{\mathfrak{K},k}$ and $\mathfrak{A}_{\mathfrak{K},k} := \gamma_0^{\mathrm{int}}\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}$.

For $-\infty < \sigma < s < \infty$, $0 < q < \infty$, and $0 < r, t \le \infty$, there holds the continuous embedding $B^s_{q,r}(B_R) \subset B^\sigma_{q,t}(B_R)$; see, e.g., [Tri92, Section 2.32]. This implies that $B^{5/2}_{2,\infty}(B_R) \subset B^2_{2,2}(B_R) = H^2(B_R)$ with $\|\cdot\|_{H^2(B_R)} \lesssim \|\cdot\|_{B^{5/2}_{2,\infty}(B_R)}$. Analogously to (6.36), continuity of the interior trace operator $\gamma_0^{\mathrm{int}}$ and inequality (6.33) reveal that

$$\begin{aligned}\|\mathfrak{S}_{\mathfrak{K},k}v\|_{H^1(\Gamma)} \le \|\mathfrak{S}_{\mathfrak{K},k}v\|_{H^1(\partial\Omega)} &\lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{K},k}v\|_{H^2(B_R)} \\ &\lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{K},k}v\|_{B^{5/2}_{2,\infty}(B_r)} \overset{(6.33)}{\lesssim} k\,\|v\|_{L^2(\partial\Omega)} = k\,\|v\|_{L^2(\Gamma)}.\end{aligned} \tag{6.40}$$

The operator $\mathfrak{A}_{\mathfrak{K},k}$ is treated analogously to Step 1 and hence satisfies that

$$\|\mathfrak{A}_{\mathfrak{K},k}\,v\|_{H^1(\Gamma)} \lesssim \|\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{H^{3/2}(B_R)} \le \|\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{H^2(B_R)} \lesssim (1+k^3)\,\|v\|_{L^2(\Gamma)}. \tag{6.41}$$

Then, the estimates (6.40) and (6.41) prove that $\mathfrak{S}_{\mathfrak{K},k}$, $\mathfrak{A}_{\mathfrak{K},k} \in L\big(L^2(\Gamma), H^1(\Gamma)\big)$. With $\mathfrak{K}_0 \in L\big(\widetilde{H}^1(\Gamma), H^1(\Gamma)\big)$, we conclude that $\mathfrak{K}_k = \mathfrak{K}_0 + \mathfrak{S}_{\mathfrak{K},k} + \mathfrak{A}_{\mathfrak{K},k} \in L\big(\widetilde{H}^1(\Gamma), H^1(\Gamma)\big)$.

**Step 5: Proof of equation (6.24).** Let $v \in \widetilde{H}^1(\Gamma)$. Analogously to Step 2, the decomposition $\mathfrak{K}_k = \mathfrak{K}_0 + \mathfrak{S}_{\mathfrak{K},k} + \mathfrak{A}_{\mathfrak{K},k}$ implies that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{K}_k v\|_{L^2(\Gamma)} \le \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{K}_0 v\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{S}_{\mathfrak{K},k} v\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{A}_{\mathfrak{K},k} v\|_{L^2(\Gamma)}.$$

We treat each term on the right-hand side separately. First, [AFF+17, Theorem 3.1] yields that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{K}_0\,v\|_{L^2(\Gamma)} \lesssim \|v\|_{\widetilde{H}^{1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,v\|_{L^2(\Gamma)}.$$

Second, $\|h_\bullet\|_{L^\infty(\Gamma)} \lesssim \operatorname{diam}(\Omega) \lesssim 1$ and equation (6.40) imply that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{S}_{\mathfrak{K},k}\,v\|_{L^2(\Gamma)} \lesssim \|\mathfrak{S}_{\mathfrak{K},k}\,v\|_{H^1(\Gamma)} \overset{(6.40)}{\lesssim} k\,\|v\|_{L^2(\Gamma)}.$$

Third, we use equation (6.41) to see that

$$\|h_\bullet^{1/2}\,\nabla_\Gamma\mathfrak{A}_{\mathfrak{K},k} v\|_{L^2(\Gamma)} \lesssim \|\mathfrak{A}_{\mathfrak{K},k}\,v\|_{H^1(\Gamma)} \overset{(6.41)}{\lesssim} (1+k^3)\,\|v\|_{L^2(\Gamma)}.$$

Combining the latter estimates, we obtain that

$$\begin{aligned}\|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{K}_k v\|_{L^2(\Gamma)} &\lesssim \|v\|_{\widetilde{H}^{1/2}(\Gamma)} + (1+k^3)\,\|v\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,v\|_{L^2(\Gamma)}\\ &\lesssim (1+k^3)\,\|v\|_{\widetilde{H}^{1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,v\|_{L^2(\Gamma)}.\end{aligned}$$

This concludes the proof of (6.24).

**Step 6: Proof of equation (6.26).** Recall the definition of $\mathfrak{W}_k$. With $\widetilde{\mathfrak{K}}_k = \widetilde{\mathfrak{K}}_0 + \widetilde{\mathfrak{S}}_{\mathfrak{K},k} + \widetilde{\mathfrak{A}}_{\mathfrak{K},k}$ there holds $\mathfrak{W}_k = -\gamma_1^{\rm int}\,\widetilde{\mathfrak{K}}_k = \mathfrak{W}_0 - \gamma_1^{\rm int}\,\widetilde{\mathfrak{S}}_{\mathfrak{K},k} - \gamma_1^{\rm int}\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}$. This yields that

$$\|h_\bullet^{1/2}\,\mathfrak{W}_k v\|_{L^2(\Gamma)} \le \|h_\bullet^{1/2}\,\mathfrak{W}_0\,v\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\gamma_1^{\rm int}\widetilde{\mathfrak{S}}_{\mathfrak{K},k}\,v\|_{L^2(\Gamma)} + \|h_\bullet^{1/2}\,\gamma_1^{\rm int}\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{L^2(\Gamma)},$$

We treat each term on the right-hand side separately. First, [AFF+17, Theorem 3.1] yields that

$$\|h_\bullet^{1/2}\,\mathfrak{W}_0\,v\|_{L^2(\Gamma)} \lesssim \|v\|_{\widetilde{H}^{1/2}(\Gamma)} + \|h_\bullet^{1/2}\,\nabla_\Gamma\,v\|_{L^2(\Gamma)}.$$

Recall from Step 4 that $\widetilde{\mathfrak{S}}_{\mathfrak{K},k}v, \widetilde{\mathfrak{A}}_{\mathfrak{K},k}v \in H^2(B_R)$ and hence $\gamma_1^{\rm int}\widetilde{\mathfrak{S}}_{\mathfrak{K},k}\,v,\ \gamma_1^{\rm int}\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}v \in H^{1/2}(\partial\Omega)$. As in Step 3, stability of $\gamma_1^{\rm int}$ gives

$$\begin{aligned}\|h_\bullet^{1/2}\,\gamma_1^{\rm int}\,\widetilde{\mathfrak{S}}_{\mathfrak{K},k}\,v\|_{L^2(\Gamma)} &\lesssim \|\widetilde{\mathfrak{S}}_{\mathfrak{K},k}\,v\|_{H^2(B_R)} \overset{(6.40)}{\lesssim} k\,\|v\|_{L^2(\Gamma)},\\ \|h_\bullet^{1/2}\,\gamma_1^{\rm int}\,\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{L^2(\Gamma)} &\lesssim \|\widetilde{\mathfrak{A}}_{\mathfrak{K},k}\,v\|_{H^2(B_R)} \overset{(6.41)}{\lesssim} (1+k^3)\,\|v\|_{L^2(\Gamma)}.\end{aligned}$$

Combining the latter four estimates, we conclude the proof of (6.26).

**Step 7: Proof of equations (6.27)–(6.28).** According to [GHS05, Geo08] or [AFF$^+$17, Lemma A.1], there hold the following inverse estimates

$$\|h_\bullet^{1/2}\,(p+1)^{-1}\,\Psi_\bullet\|_{L^2(\Gamma)} \lesssim \|\Psi_\bullet\|_{H^{-1/2}(\Gamma)} \quad \text{for all } \Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet), \tag{6.42}$$

$$\|h_\bullet^{1/2}\,(p+1)^{-1}\,\nabla_\Gamma V_\bullet\|_{L^2(\Gamma)} \lesssim \|V_\bullet\|_{\widetilde{H}^{1/2}(\Gamma)} \quad \text{for all } V_\bullet \in \widetilde{\mathcal{S}}^p(\mathcal{T}_\bullet), \tag{6.43}$$

where $p$ is the fixed polynomial degree. The hidden constant depends only on $\partial\Omega$, $\Gamma$, and the shape regularity of $\mathcal{T}_\bullet$. Applying (6.42)–(6.43) to the right-hand side of equation (6.23) and (6.26), we conclude (6.27). Using (6.42)–(6.43) to estimate the right-hand side of (6.24) and (6.26), we reveal (6.28). This concludes the proof.

□

## 6.5 Verification of the axioms

In this section, we prove the estimator axioms (E1)–(E4) for the weighted-residual error estimator defined in (6.21). Further, we show that Algorithm 6.2 ensures (E5). Therefore, the current setting fits in the abstract framework of Chapter 4.

For the validity of (E1)–(E4) for the Laplace equation, we refer to [FFK$^+$14, FKMP13, Gan13] as well as the overview in [CFPP14]. In case of the Helmholtz equation, most of the proofs are similar to the Laplace case. For the sake of completeness, we include the most important steps.

**Proposition 6.6** (stability on non-refined element domains)**.** *There exists $C_{\mathrm{stb}} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, the following implication holds: Provided that the discrete solutions $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ and $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ exist, it holds that*

$$\big|\eta_\circ(\mathcal{T}_\circ \cap \mathcal{T}_\bullet) - \eta_\bullet(\mathcal{T}_\circ \cap \mathcal{T}_\bullet)\big| \le C_{\mathrm{stb}}\,\|\Phi_\circ - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)}. \tag{6.44}$$

*In particular, there holds* (E1)*. The constant $C_{\mathrm{stb}} > 0$ depends only on $\Gamma$, $\gamma$-shape regularity, $p$, and $k$.*

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$ such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ and the corresponding discrete solutions $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ and $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ exist. For all non-refined elements $T \in \mathcal{T}_\bullet \cap \mathcal{T}_\circ$, it holds that $h_\bullet(T) = h_\circ(T)$. Together with the inverse triangle inequality and the inverse estimate (6.27), we obtain that

$$\begin{aligned}
|\eta_\bullet(\mathcal{T}_\bullet \cap \mathcal{T}_\circ) - \eta_\circ(\mathcal{T}_\bullet \cap \mathcal{T}_\circ)| &= \big|\|h_\bullet^{1/2}\,\nabla_\Gamma\,(\mathfrak{V}_k\,\Phi_\bullet - f)\|_{L^2(\bigcup(\mathcal{T}_\bullet\cap\mathcal{T}_\circ))} \\
&\qquad - \|h_\circ^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_k\,(\Phi_\circ - f)\|_{L^2(\bigcup(\mathcal{T}_\bullet\cap\mathcal{T}_\circ))}\big| \\
&\le \|h_\bullet^{1/2}\,\nabla_\Gamma\,\mathfrak{V}_k\,(\Phi_\bullet - \Phi_\circ)\|_{L^2(\Gamma)} \\
&\le \widetilde{C}_{\mathrm{inv}}\,(1+k^3)\,\|\Phi_\bullet - \Phi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}.
\end{aligned}$$

This concludes (6.44) with $C_{\mathrm{stb}} := (1+k^3)\,\widetilde{C}_{\mathrm{inv}}$. □

**Proposition 6.7** (reduction on refined element domains). *There exist* $C_{\rm red} > 0$ *and* $0 < q_{\rm red} < 1$ *such that for all* $\mathcal{T}_\bullet \in \mathbb{T}$ *and all* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$*, the following implication holds: Provided that the discrete solutions* $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ *and* $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ *exist, it holds that*

$$\eta_\circ(\mathcal{T}_\circ \backslash \mathcal{T}_\bullet)^2 \le q_{\rm red}\, \eta_\bullet(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ)^2 + C_{\rm red}^2 \, \|\Phi_\circ - \Phi_\bullet\|^2_{\widetilde{H}^{-1/2}(\Gamma)}. \tag{6.45}$$

*In particular, there holds* (E2). *The constants* $q_{\rm red}$ *and* $C_{\rm red}$ *depend only on* $q_{\rm mesh}$*,* $\Gamma$*,* $\gamma$*-shape regularity,* $p$*, and* $k$*.*

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$ such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ and the corresponding discrete solutions $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ and $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ exist. For all $T \in \mathcal{T}_\circ \setminus \mathcal{T}_\bullet$, reduction of the local mesh size (R1) implies that $h_\circ|_T \le q_{\rm mesh}\, h_\bullet|_T$. Using the Young inequality with arbitrary $\delta > 0$, we estimate

$$\begin{aligned}
\eta_\circ(\mathcal{T}_\circ \setminus \mathcal{T}_\bullet)^2 &= \sum_{T \in \mathcal{T}_\circ \setminus \mathcal{T}_\bullet} \|h_\circ^{1/2}\, \nabla_\Gamma\, (\mathfrak{V}_k \Phi_\circ - f)\|^2_{L^2(T)} \\
&\le \sum_{T \in \mathcal{T}_\circ \setminus \mathcal{T}_\bullet} \Big( \|h_\circ^{1/2}\, \nabla_\Gamma\, (\mathfrak{V}_k \Phi_\bullet - f)\|_{L^2(T)} + \|h_\circ^{1/2}\, \nabla_\Gamma\, \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\|_{L^2(T)} \Big)^2 \\
&\le \sum_{T \in \mathcal{T}_\circ \setminus \mathcal{T}_\bullet} \Big( (1+\delta)\, q_{\rm mesh}\, \|h_\bullet^{1/2}\, \nabla_\Gamma\, (\mathfrak{V}_k \Phi_\bullet - f)\|^2_{L^2(T)} \\
&\qquad\qquad + (1+\delta^{-1})\, \|h_\circ^{1/2}\, \nabla_\Gamma\, \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\|^2_{L^2(T)} \Big).
\end{aligned}$$

Next, the inverse inequality (6.27) yields that

$$\eta_\bullet(\mathcal{T}_\circ \setminus \mathcal{T}_\bullet)^2 \le (1+\delta)\, q_{\rm mesh}\, \eta_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)^2 + (1+\delta^{-1})\,(1+k^3)^2\, \widetilde{C}_{\rm inv}^2\, \|\Phi_\bullet - \Phi_\circ\|^2_{\widetilde{H}^{-1/2}(\Gamma)}.$$

Choosing $\delta > 0$ sufficiently small such that $q_{\rm red} := (1+\delta)\, q_{\rm mesh} < 1$, we conclude (6.45) with $C_{\rm red} = (1+\delta^{-1})\,(1+k^3)^2\, \widetilde{C}_{\rm inv}^2$. $\square$

**Proposition 6.8** (discrete reliability). *There exists* $C_{\rm rel} > 0$*, such that for all* $\mathcal{T}_\bullet \in \mathbb{T}$ *and all* $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$*, there exists a set* $\mathcal{R}_{\bullet,\circ} \subseteq \mathcal{T}_\bullet$ *with* $\mathcal{T}_\bullet \backslash \mathcal{T}_\circ \subseteq \mathcal{R}_{\bullet,\circ}$ *and* $\#\mathcal{R}_{\bullet,\circ} \le C_{\rm rel}\,\#(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ)$*, such that the following implication holds: Provided that the discrete solutions* $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ *and* $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ *exist, it holds that*

$$\|\Phi_\circ - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \le C_{\rm rel}\, \beta_\circ^{-1}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ}), \tag{6.46}$$

*where* $\beta_\circ > 0$ *is the inf–sup constant associated with* $\mathcal{P}^p(\mathcal{T}_\circ)$*. In particular, there holds* (E4)*. The constant* $C_{\rm rel} \ge 1$ *depends only on the given data, the polynomial degree* $p$*, the initial mesh* $\mathcal{T}_0$*, and* $\gamma$*-shape regularity.*

*Proof.* We follow the arguments from [FKMP13, Theorem 5.3] for the case of $k = 0$. Recall that notation of Proposition 4.1. Then, existence and uniqueness of $\Phi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ is equivalent to $\beta_\circ > 0$. The discrete inf–sup condition (4.8) for $\mathcal{X}_\circ := \mathcal{P}^p(\mathcal{T}_\circ)$ and $W_\circ := \Phi_\circ - \Phi_\bullet$ reads as

$$\beta_\circ\, \|\Phi_\circ - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \le \sup_{\Psi_\circ \in \mathcal{X}_\circ \backslash \{0\}} \frac{\langle \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\,,\, \Psi_\circ \rangle}{\|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}}. \tag{6.47}$$

Let $\mathcal{N}_\bullet$ denote the set of nodes corresponding to a triangulation $\mathcal{T}_\bullet$. Let $\rho_z \in \mathcal{S}^1(\mathcal{T}_\bullet)$ denote the hat function associated with a node $z \in \mathcal{N}_\bullet$. Further, let $\mathcal{N}_\bullet^{\mathcal{R}} := \mathcal{N}_\bullet \cap \big(\bigcup(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)\big)$ be the set of all nodes which belong to the refined elements. Define $\mathcal{R}_{\bullet,\circ} := \omega_\bullet(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)$ and $\mathcal{Q}_\bullet := \mathcal{R}_{\bullet,\circ} \setminus (\mathcal{T}_\bullet \setminus \mathcal{T}_\circ)$. These definitions give rise to disjoint decompositions

$$\mathcal{R}_{\bullet,\circ} := (\mathcal{T}_\bullet \setminus \mathcal{T}_\circ) \,\dot\cup\, \mathcal{Q}_\bullet \quad \text{and} \quad \mathcal{T}_\bullet = (\mathcal{T}_\bullet \setminus \mathcal{R}_{\bullet,\circ}) \,\dot\cup\, (\mathcal{T}_\bullet \setminus \mathcal{T}_\circ) \,\dot\cup\, \mathcal{Q}_\bullet.$$

Define $\chi := \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z$. Then, $\chi \in \mathcal{S}^1(\mathcal{T}_\bullet)$ satisfies $\operatorname{supp}(\chi) = \bigcup \mathcal{R}_{\bullet,\circ}$ and $\chi|_{\bigcup(\mathcal{T}_\bullet\setminus\mathcal{T}_\circ)} \equiv 1$. We define the operator $\pi_\bullet : \mathcal{P}^p(\mathcal{T}_\circ) \to \mathcal{P}^p(\mathcal{T}_\bullet)$ by

$$\pi_\bullet(\Psi_\circ) := \begin{cases} 0 & \text{on } \bigcup(\mathcal{T}_\bullet \setminus \mathcal{T}_\circ), \\ \Psi_\circ & \text{elsewhere.} \end{cases}$$

For any $\Psi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ and $\Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$, the Galerkin orthogonality yields that

$$\langle \mathfrak{V}_k\,(\Phi_\circ - \Phi_\bullet)\,,\, \Psi_\circ\rangle = \langle f - \mathfrak{V}_k\,\Phi_\bullet\,,\, \Psi_\circ\rangle = \langle f - \mathfrak{V}_k\,\Phi_\bullet\,,\, \Psi_\circ - \Psi_\bullet\rangle. \tag{6.48}$$

Choose $\Psi_\bullet := \pi_\bullet(\Psi_\circ) \in \mathcal{P}^p(\mathcal{T}_\bullet)$ and note that $\operatorname{supp}\big((1-\pi_\bullet)\,\Psi_\circ\big) \subseteq \bigcup(\mathcal{T}_\bullet\setminus\mathcal{T}_\circ)$. Using (6.48), we derive that

$$\begin{aligned}\langle \mathfrak{V}_k\,(\Phi_\circ - \Phi_\bullet)\,,\, \Psi_\circ\rangle &= \langle f - \mathfrak{V}_k\Phi_\bullet\,,\, (1-\pi_\bullet)\,\Psi_\circ\rangle \\ &= \Big\langle \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\,,\, (1-\pi_\bullet)\,\Psi_\circ\Big\rangle \\ &= \Big\langle \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\,,\, \Psi_\circ\Big\rangle - \Big\langle \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\,,\, \Psi_\circ|_{\bigcup\mathcal{Q}_\bullet}\Big\rangle.\end{aligned}$$

Since $\mathcal{Q}_\bullet \subset \mathcal{T}_\bullet \cap \mathcal{T}_\circ$, we obtain that $h_\bullet(T) = h_\circ(T)$ for all $T \in \mathcal{Q}_\bullet$. We estimate

$$\begin{aligned}|\langle \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\,,\, \Psi_\circ\rangle| &\le \Big\| \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\Phi_\bullet)\Big\|_{H^{1/2}(\Gamma)} \|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)} \\ &\qquad + \Big\|h_\bullet^{-1/2} \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\Big\|_{L^2(\Gamma)} \|h_\bullet^{1/2}\,\Psi_\circ\|_{L^2(\bigcup\mathcal{Q}_\bullet)} \\ &\le \Big\| \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\Big\|_{H^{1/2}(\Gamma)} \|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)} \\ &\qquad + \Big\|h_\bullet^{-1/2} \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z\,(f - \mathfrak{V}_k\,\Phi_\bullet)\Big\|_{L^2(\Gamma)} \|h_\circ^{1/2}\,\Psi_\circ\|_{L^2(\Gamma)}.\end{aligned}$$

Applying the inverse estimate (6.42) to the right-hand side, we see that

$$\begin{aligned}|\langle \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\,,\, \Psi_\circ\rangle| \lesssim \Big(&\Big\| \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z(f - \mathfrak{V}_k\Phi_\bullet)\Big\|_{H^{1/2}(\Gamma)} \\ &+ \Big\|h_\bullet^{-1/2} \sum_{z\in\mathcal{N}_\bullet^{\mathcal{R}}} \rho_z(f - \mathfrak{V}_k\Phi_\bullet)\Big\|_{L^2(\Gamma)}\Big) \|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}.\end{aligned}$$

The terms in the parentheses are estimated as in [CMS01]. The sole difference is that compared to [CMS01, Theorem 3.2] only hat functions associated with nodes $z \in \mathcal{N}_\bullet^{\mathcal{R}}$ are involved. Hence, the upper bound affects only $\bigcup \mathcal{R}_{\bullet,\circ} \subset \Gamma$ and reads

$$|\langle \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\,,\,\Psi_\circ\rangle| \lesssim \|h_\bullet^{1/2}\,\nabla_\Gamma\,(f - \mathfrak{V}_k\Phi_\bullet)\|_{L^2(\bigcup \mathcal{R}_{\bullet,\circ})}\,\|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}. \tag{6.49}$$

Altogether, the combination of (6.47)–(6.49) proves that

$$\|\Phi_\circ - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \leq \frac{1}{\beta_\circ} \sup_{\Psi_\circ \in \mathcal{X}_\circ} \frac{\langle \mathfrak{V}_k(\Phi_\circ - \Phi_\bullet)\,,\,\Psi_\circ\rangle}{\|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}} \lesssim \beta_\circ^{-1}\,\|h_\bullet^{1/2}\,\nabla_\Gamma\,(f - \mathfrak{V}_k\Phi_\bullet)\|_{L^2(\bigcup \mathcal{R}_{\bullet,\circ})}.$$

This concludes the proof. □

**Corollary 6.9** (reliability). *There exists $C'_{\mathrm{rel}} > 0$ such that for all $\mathcal{T}_\bullet \in \mathbb{T}$, the following implication holds: Provided that there exists a discrete solution $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$, it holds that*

$$\|\phi - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \leq C'_{\mathrm{rel}}\,\eta_\bullet.$$

*In particular, there holds* (E3)*.*

***Sketch of proof.*** Reliability can be shown analogously to discrete reliability. Using $\mathcal{N}_\bullet^{\mathcal{R}} := \mathcal{N}_\bullet$ as well as the techniques from [CMS01], we obtain analogously to Proposition 6.8 that

$$\|\phi - \Phi_\bullet\|_{\widetilde{H}^{-1/2}(\Gamma)} \leq \frac{1}{\beta} \sup_{\Psi_\circ \widetilde{H}^{-1/2}(\Gamma)} \frac{\langle \mathfrak{V}_k(\phi - \Phi_\bullet)\,,\,\Psi_\circ\rangle}{\|\Psi_\circ\|_{\widetilde{H}^{-1/2}(\Gamma)}} \lesssim \beta^{-1}\,\|h_\bullet^{1/2}\,\nabla_\Gamma\,(f - \mathfrak{V}_k\Phi_\bullet)\|_{L^2(\Gamma)},$$

where $\beta > 0$ is the continuous inf–sup constant associated with $\widetilde{H}^{-1/2}(\Gamma)$. This concludes the proof. □

### 6.5.1 Definiteness on the "discrete" limit space (E5)

It remains to prove validity of (E5). Recall that $\mathcal{H} = \widetilde{H}^{-1/2}(\Gamma)$ and $\mathcal{X}_\ell = \mathcal{P}^p(\mathcal{T}_\ell)$ in the sense of Section 4.5.1. We define the discrete limit space $\mathcal{X}_\infty := \overline{\bigcup_{\ell=0}^\infty \mathcal{P}^p(\mathcal{T}_\ell)}$ and obtain the following lemma. The lemma is an analogon to Lemma 5.6 for finite elements.

**Lemma 6.10.** *Let $p \geq 0$. The triangulations $\mathcal{T}_\ell$ generated by Algorithm 6.2 are uniformly $\gamma$-shape regular with $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$. Moreover, there holds $\mathcal{X}_\infty = \widetilde{H}^{-1/2}(\Gamma)$ and hence assumption* (E5) *is satisfied.*

*Proof.* Recall the notation of Algorithm 6.2. We emphasize that both, EB and NVB guarantee uniform $\gamma$-shape regularity (3.4). The expanded Dörfler marking in Step (iv) and Step (v) of Algorithm 6.2 enforces $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$. It remains to show that $\mathcal{X}_\infty = \widetilde{H}^{-1/2}(\Gamma)$. Let $\Pi_\ell : L^2(\Gamma) \to \mathcal{P}^p(\mathcal{T}_\ell)$ denote the $L^2(\Gamma)$-orthogonal projection onto $\mathcal{P}^0(\mathcal{T}_\ell)$. For all $\psi \in L^2(\Gamma)$, it holds that

$$\|(1 - \Pi_\ell)\,\psi\|_{\widetilde{H}^{-1/2}(\Gamma)} \lesssim \|h_\ell^{1/2}\,(1 - \Pi_\ell)\,\psi\|_{L^2(\Gamma)}; \tag{6.50}$$

see [AFF$^+$17, Corollary 3.3] or [SS11, Section 4.3.4]. For $\psi \in L^2(\Gamma)$, nestedness $\mathcal{P}^0(\mathcal{T}_\ell) \subset \mathcal{P}^p(\mathcal{T}_\ell)$ implies that

$$\inf_{\Psi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)} \|\psi - \Psi_\ell\|_{\widetilde{H}^{-1/2}(\Gamma)} \lesssim \|h_\ell^{1/2}\,(1-\Pi_\bullet)\,\psi\|_{L^2(\Gamma)} \lesssim \|h_\ell^{1/2}\|_{L^\infty(\Gamma)}\,\|\psi\|_{L^2(\Gamma)}.$$

Recall that the embedding $L^2(\Gamma) \subset \widetilde{H}^{-1/2}(\Gamma)$ is dense. Hence, with $\|h_\ell\|_{L^\infty(\Omega)} \to 0$ as $\ell \to \infty$, we conclude $\mathcal{X}_\infty = \widetilde{H}^{-1/2}(\Gamma)$. $\square$

## 6.6 Optimal Convergence

The next theorem is the main result of this chapter. It states that Algorithm 6.2 does not only lead to linear convergence, but also guarantees optimal algebraic convergence rates for the sequence of *a posteriori* error estimators. Recall that either NVB or EB guarantee (R3)–(R6) and the estimator satisfies (E1)–(E4). Then, the theorem is a direct consequence of Theorem 4.14 and Theorem 4.21.

**Theorem 6.11.** *Employ the notation of Algorithm 6.2. Suppose* $0 < \theta \le 1$. *Then, there exist* $\ell_{\rm lin} > 0$ *as well as constants* $0 < q_{\rm lin} < 1$ *and* $C_{\rm lin} > 0$ *such that Algorithm 6.2 guarantees that*

$$\eta_{\ell+n} \le C_{\rm lin}\, q_{\rm lin}^n\, \eta_\ell \quad \text{for all } \ell, n \in \mathbb{N} \text{ with } \ell \ge \ell_{\rm lin}. \tag{6.51}$$

*The constants* $\ell_{\rm lin}$, $q_{\rm lin}$, $C_{\rm lin}$ *depend only on* $q_{\rm est}$, $C_{\rm rel}$, *and* $\ell_3$ *from Lemma 4.13. Moreover, there exists* $\beta_0 > 0$, $\ell_{\rm opt} > 0$, *as well as* $\theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$ *such that for all* $0 < \theta < \theta_{\rm opt}$ *and all* $s > 0$, *it holds that*

$$\|\phi\|_{\mathbb{A}_s} < \infty \quad\Longleftrightarrow\quad \exists C_{\rm opt} > 0\ \forall \ell \ge \ell_{\rm opt} \quad \eta_\ell \le C_{\rm opt}\,(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}. \tag{6.52}$$

*The constants* $\ell_{\rm opt}$, $C_{\rm opt}$ *depend only on* $C_{\rm son}$, $\mathcal{T}_0$, $\theta$, $s$, $\|\phi\|_{\mathbb{A}_s}$, *and on the constants in* (E1)–(E4). $\square$

**Remark 6.12.** *For the presentation, we focused on the model problem* (6.13) *for some indirect boundary element method. In the case of a direct boundary element approach, the model problem reads*

$$\mathfrak{V}_k \phi = \big(\mathfrak{K}_k + \frac{1}{2}\,\mathrm{Id}\big)\, g \quad \text{on } \Gamma, \tag{6.53}$$

*where* $g \in H^{1/2}(\partial\Omega)$ *are the given Dirichlet data and* $\phi = \partial_n u \in H^{-1/2}(\partial\Omega)$ *is the sought normal derivative of the solution* $u \in H^1(\Omega)$ *of the (equivalent) boundary value problem*

$$-\Delta u - k^2\, u = 0 \quad \text{in } \Omega \qquad \text{subject to} \qquad u = g \quad \text{on } \Gamma.$$

*The implementation of the right-hand side requires to approximate* $g \approx G_\bullet \in \mathcal{S}^{p+1}(\mathcal{T}_\bullet)$. *Suitable approximations* $G_\bullet = I_\bullet\, g$ *together with some local data oscillations which control the additional approximation error* $\|g - G_\bullet\|_{H^{1/2}(\partial\Omega)}$ *are discussed and analyzed in [FFK$^+$14] for the Laplace problem. Provided that* $g \in H^1(\partial\Omega)$, *it is shown that the adaptive algorithm then still leads to optimal convergence behavior. Together with the present analysis, the results of [FFK$^+$14] transfer immediately to the direct boundary element approach* (6.53).

## 6.7 Hyper-singular integral equation

In this section, we briefly comment on the extension of our analysis to the hyper-singular integral equation. In case of the Laplace equation ($k = 0$), a proof of optimal algebraic convergence rates is found in [FFK+15, Gan13]. Throughout this section, we additionally suppose that $\partial\Omega$ is connected.

We define the spaces $H_\star^{\pm 1/2}(\partial\Omega)$ consisting of $H^{\pm 1/2}(\partial\Omega)$-functions with integral mean zero by

$$\begin{aligned} H_\star^{-1/2}(\partial\Omega) &:= \big\{\phi \in H^{-1/2}(\partial\Omega) \,:\, \langle \phi\,,\, 1\rangle = 0\big\} \qquad \text{and} \\ H_\star^{1/2}(\partial\Omega) &:= \big\{v \in H^{1/2}(\partial\Omega) \,:\, \langle 1\,,\, v\rangle = 0\big\}. \end{aligned}$$

First, we recap some important properties of the hyper-singular operator $\mathfrak{W}_k := -\gamma_1^{\mathrm{int}}\,\widetilde{\mathfrak{K}}_k$. For $k = 0$, the operator $\mathfrak{W}_0$ is symmetric and positive semi-definite on $H^{1/2}(\partial\Omega)$, i.e.,

$$\langle \mathfrak{W}_0\, v\,,\, w\rangle = \langle \mathfrak{W}_0\, w\,,\, v\rangle \quad \text{and} \quad \langle \mathfrak{W}_0\, v\,,\, v\rangle \geq 0 \quad \text{for all } v, w \in H^{1/2}(\partial\Omega).$$

Since $\partial\Omega$ is connected, the kernel of $\mathfrak{W}_0$ consists of the constant functions. Hence, the bilinear form $\langle \mathfrak{W}_0(\cdot)\,,\,\cdot\rangle$ provides a scalar product on $H_\star^{1/2}(\partial\Omega)$. This can be extended to

$$a(u\,,\, v) := \langle \mathfrak{W}_0 v\,,\, w\rangle + \langle 1\,,\, v\rangle\langle 1\,,\, w\rangle \quad \text{for all } v, w \in H^{1/2}(\partial\Omega), \tag{6.54}$$

which defines a scalar product on $H^{1/2}(\partial\Omega)$. According to the Rellich compactness theorem, there holds the norm equivalence $|\!|\!|\, v \,|\!|\!|^2 := a(v\,,\, v) \simeq \|v\|^2_{H^{1/2}(\partial\Omega)}$ for all $v \in H^{1/2}(\partial\Omega)$.

For $k \neq 0$, it is well-known that the hyper-singular integral operator $\mathfrak{W}_k$ is invertible, if and only if $k^2$ is not an eigenvalue of the interior Neumann eigenvalue problem (see [Ste13, Proposition 2.5]), i.e., it holds that

$$\forall u \in H^1(\Omega) \quad \Big(\Delta u = k^2 u \text{ with } \gamma_1^{\mathrm{int}} u = 0 \text{ and } \int_\Gamma u\,dx = 0 \quad \Longrightarrow \quad u = 0 \text{ in } \Omega\Big). \tag{INP}$$

To ensure solvability, we assume throughout this section that $k^2$ is not an eigenvalue of the (INP). Then, the model problem for the hyper-singular operator $\mathfrak{W}_k$ reads as follows: Given $f \in H_\star^{-1/2}(\partial\Omega)$, find $u \in H_\star^{1/2}(\partial\Omega)$ such that

$$\mathfrak{W}_k\, u = f \quad \text{on } \partial\Omega. \tag{6.55}$$

### 6.7.1 Framework

The proofs of the abstract properties of Section 4.2 and the estimator axioms (E1)–(E4) for model problem (6.55) are very similar to the weakly-singular case in Section 6.3 and Section 6.5. Therefore, we focus only on the occurring differences and highlight the necessary modifications. We define the operator

$$\widetilde{\mathfrak{C}}_{W_k} : H^{1/2}(\partial\Omega) \to H^{-1/2}(\partial\Omega) \qquad \widetilde{\mathfrak{C}}_{W_k} := \mathfrak{W}_k - \mathfrak{W}_0.$$

On Lipschitz boundaries, $\widetilde{\mathfrak{C}}_{W_k}$ is compact, (see [SS11, Lemma 3.9.8]). For all $v, w \in H^{1/2}(\partial\Omega)$, define $\mathfrak{C}_{W_k}$ by

$$\langle \mathfrak{C}_{W_k} v \,,\, w \rangle := \langle \widetilde{\mathfrak{C}}_{W_k} v \,,\, w \rangle - \langle 1 \,,\, v \rangle \langle 1 \,,\, w \rangle$$

Note that, $\mathfrak{C}_{W_k}$ is uniquely defined and compact; see, e.g., [Ste08a, Section 6.9]. Reformulation of (6.55) yields the following equivalent formulation: Given $f \in H_\star^{-1/2}(\partial\Omega)$, find $u \in H^{1/2}(\partial\Omega)$ such that

$$(\mathfrak{W}_0 + \mathfrak{C}_{W_k})\, u = f. \tag{6.56}$$

With the bilinear form $a(\cdot\,,\,\cdot)$, from (6.54), the corresponding discrete formulation of (6.56) reads as follows: Find $U_\star \in \mathcal{S}^p(\mathcal{T}_\star)$ such that

$$a(U_\star\,,\, V_\star) + \langle \mathfrak{C}_{W_k} U_\star\,,\, V_\star \rangle = \langle f\,,\, V_\star \rangle \quad \text{for all } V_\star \in \mathcal{S}^p(\mathcal{T}_\star). \tag{6.57}$$

Then, the discrete formulation (6.57) fits in the abstract framework of Section 4.2. It remains to show that uniform refinement leads to a dense sequence of subspaces. Since we use globally continuous and piecewise polynomials $\mathcal{S}^p(\mathcal{T}_\star)$, the proof follows the same lines as the proof of Lemma 5.6, where we additionally exploit the density of $H^1(\partial\Omega) \subseteq H^{1/2}(\partial\Omega)$. Then, Proposition 4.1 with $\mathcal{H} = H^{1/2}(\partial\Omega)$ and $\mathcal{X}_\bullet = \mathcal{S}^p(\mathcal{T}_\bullet)$ guarantees existence and uniqueness of discrete solutions of (6.57).

### 6.7.2 Weighted-residual error estimator

Analogously to $H_\star^{\pm 1/2}(\partial\Omega)$, we define $L_\star^2(\partial\Omega) := \big\{\phi \in L^2(\partial\Omega) \,:\, \int_\Gamma \phi\, ds = 0\big\}$. Let $\mathcal{T}_\bullet \in \mathbb{T} := \text{refine}(\mathcal{T}_0)$ such that the corresponding discrete solution $U_\bullet \in \mathcal{S}^p(\mathcal{T}_\bullet)$ of (6.57) exists. Suppose that $f \in L_\star^2(\partial\Omega)$. Then, the local contributions of the weighted-residual error estimator for hyper-singular integral equation are defined by

$$\eta_\bullet(T) := \|h_\bullet^{1/2}(f - \mathfrak{W}_k U_\bullet)\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}_\bullet. \tag{6.58}$$

The proofs of stability on non-refined domains (E1), reduction on refined element domains (E2), discrete reliability (E4) as well as reliability (E3) are similar to Section 6.5 and can be found in [FFK+15, Proposition 3.5]. The sole difference is in the use of inverse inequality (6.26) instead of (6.23).

### 6.7.3 Adaptive Algorithm and optimal convergence rates

For the hyper-singular integral equation (6.57), we seek a solution $U_\ell \in \mathcal{S}^p(\mathcal{T}_\ell)$ and use the *a posteriori* error estimator from (6.58). Then, Algorithm 6.2 transfers into the following formulation:

**Algorithm 6.13.** Input: *Parameters* $0 < \theta \le 1$ *and* $C_{\text{mark}} \ge 1$ *as well as initial triangulation* $\mathcal{T}_0$ *with* $U_{-1} := 0 \in \mathcal{S}^p(\mathcal{T}_0)$ *and* $\eta_{-1} := 1$.
Adaptive loop: *For all* $\ell = 0, 1, 2, \dots,$ *iterate the following Steps* (i)–(vi)*:*

(i) `If` (6.57) *does not admit a unique solution in* $\mathcal{S}^p(\mathcal{T}_\ell)$*:*

– *Define* $U_\ell := U_{\ell-1} \in \mathcal{S}^p(\mathcal{T}_\ell)$ *and* $\eta_\ell := \eta_{\ell-1}$,
– *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ *be the uniform refinement of* $\mathcal{T}_\ell$,
– *Increase* $\ell \to \ell + 1$ *and continue with Step* (i).

(ii) `Else` *compute the unique solution* $U_\ell \in \mathcal{S}^p(\mathcal{T}_\ell)$ *to* (6.57).

(iii) *Compute the corresponding indicators* $\eta_\ell(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iv) *Determine a set* $\mathcal{M}'_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\mathrm{mark}}$ *minimal cardinality such that* $\theta\eta_\ell^2 \le \eta_\ell(\mathcal{M}'_\ell)^2$.

(v) *Find* $\mathcal{M}''_\ell \subseteq \mathcal{T}_\ell$ *such that* $\#\mathcal{M}''_\ell = \#\mathcal{M}'_\ell$ *as well as* $h_\ell(T) \ge h_\ell(T')$ *for all* $T \in \mathcal{M}''_\ell$ *and* $T' \in \mathcal{T}_\ell \setminus \mathcal{M}''_\ell$. *Define* $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \mathcal{M}''_\ell$.

(vi) *Generate* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, *increase* $\ell \to \ell + 1$, *and continue with Step* (i).

OUTPUT: *Sequences of successively refined triangulations* $\mathcal{T}_\ell$, *discrete solutions* $U_\ell$, *and corresponding estimators* $\eta_\ell$.

Note that Steps (iii)–(vi) are verbatim to Algorithm 6.2. Recall that both refinement strategies, NVB and EB guarantee the refinement axioms (R3)–(R6). Therefore, the optimal convergence behaviour of Algorithm 6.13 relies only on the validity of the estimator axioms (E1)–(E4). Further, definiteness of the discrete limit space (E5) can be shown verbatim to Lemma 5.6. Hence, the following theorem for the hyper-singular integral equation is direct consequence of Theorem 4.14 and Theorem 4.21.

**Theorem 6.14.** *Employ the notation of Algorithm 6.13. Suppose* $0 < \theta \le 1$. *Then, there exist* $\ell_{\mathrm{lin}} > 0$ *as well as constants* $0 < q_{\mathrm{lin}} < 1$ *and* $C_{\mathrm{lin}} > 0$ *such that Algorithm 6.13 guarantees that*

$$\eta_{\ell+n} \le C_{\mathrm{lin}}\, q_{\mathrm{lin}}^n\, \eta_\ell \quad \text{for all } \ell, n \in \mathbb{N} \text{ with } \ell \ge \ell_{\mathrm{lin}}. \tag{6.59}$$

*The constants* $\ell_{\mathrm{lin}}$, $q_{\mathrm{lin}}$, $C_{\mathrm{lin}}$ *depend only on* $q_{\mathrm{est}}$, $C_{\mathrm{rel}}$, *and* $\ell_3$ *from Lemma 4.13. Moreover, there exists* $\widehat{\beta}_0 > 0$, $\ell_{\mathrm{opt}} > 0$, *as well as* $\theta_{\mathrm{opt}} := (1 + C_{\mathrm{stb}}^2 C_{\mathrm{rel}}^2/\widehat{\beta}_0^2)^{-1}$, *such that for all* $0 < \theta < \theta_{\mathrm{opt}}$ *and all* $s > 0$, *it holds that*

$$\|\phi\|_{\mathbb{A}_s} < \infty \quad \Longleftrightarrow \quad \exists C_{\mathrm{opt}} > 0\ \forall \ell \ge \ell_{\mathrm{opt}} \quad \eta_\ell \le C_{\mathrm{opt}}\,(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}. \tag{6.60}$$

*The constants* $\ell_{\mathrm{opt}}$, $C_{\mathrm{opt}}$ *depend only on* $C_{\mathrm{son}}$, $\mathcal{T}_0$, $\theta$, $s$, $\|\phi\|_{\mathbb{A}_s}$, *and on the constants in* (E1)–(E4). □

**Remark 6.15.** *Similar to Remark 6.12, one may consider a direct formulation for the Neumann boundary-value problem. In this case, the model problem reads as follows: Given Neumann data* $\phi \in H^{-1/2}(\partial\Omega)$, *find* $u \in H^{1/2}(\partial\Omega)$ *such that*

$$\mathfrak{W}_k u = (\frac{1}{2}\,\mathrm{Id} - \mathfrak{K}'_k)\,\phi \quad \text{on } \partial\Omega. \tag{6.61}$$

*In practice, the implementation of the right-hand side requires to approximate* $\phi \approx \Phi_\bullet \in \mathcal{P}^{p-1}(\mathcal{T}_\star)$. *Provided that* $\phi \in L^2(\partial\Omega)$, *a suitable approximation* $\Phi_\bullet := \Pi_\bullet\phi$ *is given by* *the $L^2$-orthogonal projection onto $\mathcal{P}^{p-1}(\mathcal{T}_\bullet)$. The local data oscillations which control the additional approximation error $\|\phi - \Phi_\bullet\|_{H^{-1/2}(\Gamma)}$ are discussed and analyzed in [FFK+15] for the Laplace problem. There, it is shown that the adaptive algorithm then still leads to optimal convergence behavior. Together with the present analysis, the results of [FFK+15] transfer immediately to the direct boundary element approach* (6.61). $\square$

## 6.8 Numerical experiments

In this section, we present some numerical experiments for the 3D Helmholtz equation that underpin the theoretical findings of this chapter. We use lowest order BEM and consider $\mathcal{X}_\bullet = \mathcal{P}^0(\mathcal{T}_\bullet)$ for the weakly-singular integral equation and $\mathcal{X}_\bullet = \mathcal{S}^1(\mathcal{T}_\bullet)$ for the hypersingular equation. The numerical computations were done with help of BEM++, which is an open-source Galerkin boundary element library. We refer to [SBA+15, GBB+15, vtWGBA15] for details on BEM++.

We consider sound-soft (exterior Dirichlet) and sound-hard (exterior Neumann) acoustic scattering problems in $\mathbb{R}^3 \setminus \Omega$ where $\Omega \subset \mathbb{R}^3$ denotes the scatterer. Let $a \in \mathbb{R}^3$ with $|a| = 1$ denote the directional vector of the incident wave. Then, the incident (plane-) wave is given by $u^{\text{inc}} = \exp(ika \cdot x)$. Let $u^{\text{scat}}$ be the scattered field and the resulting total field is defined by $u^{\text{tot}} = u^{\text{inc}} + u^{\text{scat}}$.

For sake of simplicity we restrict the numerical examples to an indirect approach, in which the solution is in the form of a layer potential with some unknown density. For the sound-soft scattering problem, we obtain: Find $u^{\text{scat}} = \widetilde{\mathfrak{V}}_k(\phi)$ such that

$$\mathfrak{V}_k\,\phi = g \quad \text{subject to} \quad g = -u^{\text{inc}} \quad \text{on } \Gamma. \tag{6.62}$$

The indirect approach for the sound-hard reads: Find $u^{\text{scat}} = \widetilde{\mathfrak{K}}_k(\phi)$ such that

$$\mathfrak{W}_k\,\phi = g \quad \text{subject to} \quad g = -\partial_n u^{\text{inc}} \quad \text{on } \Gamma. \tag{6.63}$$

### 6.8.1 Sound-soft scattering on an L-shaped domain

As first numerical example we consider the so called L-shaped domain in $x - y$-direction and expand it on the $z$-axis up to $[-1, 1]$, see Figure 6.1. We compare two directions of the incident wave. One with $a = (-1/\sqrt{2}, 1/\sqrt{2}, 0)^T$ (Figure 6.2, left) hitting the scatterer on the non-convex part vs. $a = (1/\sqrt{2}, -1/\sqrt{2}, 0)^T$ hitting the convex part of $\Omega$ (Figure 6.2, right).

First, we comment on the non-convex case. Figure 6.3 (top), shows the convergence rate of $\eta_\ell^2$ for $k = 1$, and different marking strategies. We compare uniform refinement to normal Dörfler marking as well as to the expanded Döfler marking from Proposition 4.7, both using $\theta = 0.4$. The experiments show that uniform mesh-refinement leads to a suboptimal rate of $\mathcal{O}(N^{-2/3})$ for $\eta_\ell^2$, while adaptive refinement with Algorithm 6.2 regains the improved rate $\mathcal{O}(N^{-\delta})$ with $\delta = 1.075$, independently of the actual marking. Empirically, the results generated by employing the standard Dörfler marking are of no difference compared to the results generated by employing the expanded Dörfler marking. The same observation is made for all computations (not desplayed), so that we only focus on the expanded marking in the remaining plots.

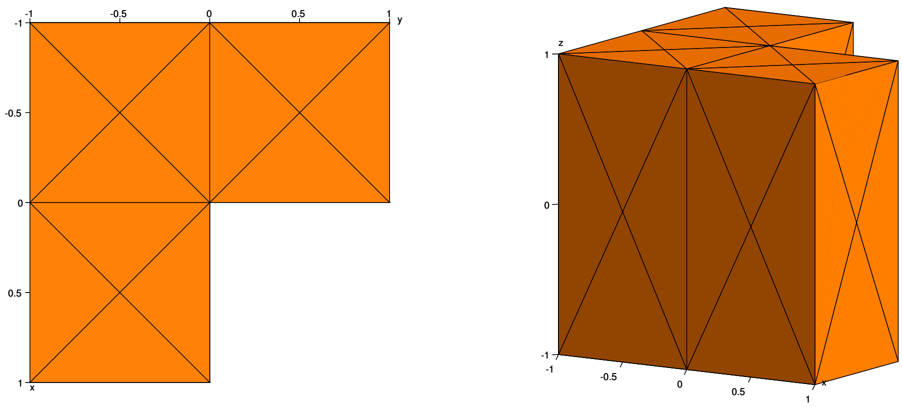



Figure 6.1: Geometry and initial mesh $\mathcal{T}_0$ with 56 elements (left: top view, right: 3D view). The reentrant edge has the coordinates $(0,0,t)$ with $t \in [-1,1]$.

Figure 6.3 (bottom) compares uniform vs. adaptive refinement for fixed $\theta = 0.4$ but various $k \in \{1,2,4,8,16\}$. As expected, the pre-asymptotic phase increases with $k$, but adaptive mesh-refinement asymptotically regains improved convergence rates for every $k$.

Figure 6.5 compares the convergence of the estimator for different values of the marking parameter $\theta \in \{0.2,0.4,0.6,0.8\}$ as well as uniform mesh-refinement. Again, uniform mesh-refinement leads to a suboptimal rate of convergence for the error estimator, while adaptive refinement with Algorithm 6.2 regains the improved rate of convergence, independently of the actual choice of the marking parameter. Although Theorem 6.11 predicts optimal convergence rates only for small marking parameters $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$, we observe that Algorithm 6.2 is stable in $\theta$, and any choice of $\theta \le 0.8$ leads to the improved convergence behavior. In Figure 6.4, one can see some of the obtained adaptive meshes $\mathcal{T}_\ell$ with $\ell = 4, 8, 12, 16$. The mesh-refinement is focused around the facets and edges hit by the incoming wave, while all facets in the shadow remain coarse.

In the second case, i.e., the scatterer is hit on the convex part of the domain (Figure 6.3, right), we compute very similar results as in the non-convex case. As shown in Figure 6.6 above, expanded as well as normal Dörfler marking lead to improved rates of $\mathcal{O}(N^{-1.06})$ for the squared error estimator, while uniform refinement leads only to $\mathcal{O}(N^{-2/3})$. The rate of convergence is independent of the wavelength $k > 0$, but increasing $k$ leads to a longer preasymtotic phase. Figure 6.7 shows the triangulations $\mathcal{T}_{12}$ and $\mathcal{T}_{16}$. Again, the mesh-refinement is focused around the facets and edges hit by the incoming wave, all facets in the shadow remain coarse.

Finally, we emphasize that Algorithm 6.2 never enforced uniform mesh-refinement in Step (i).

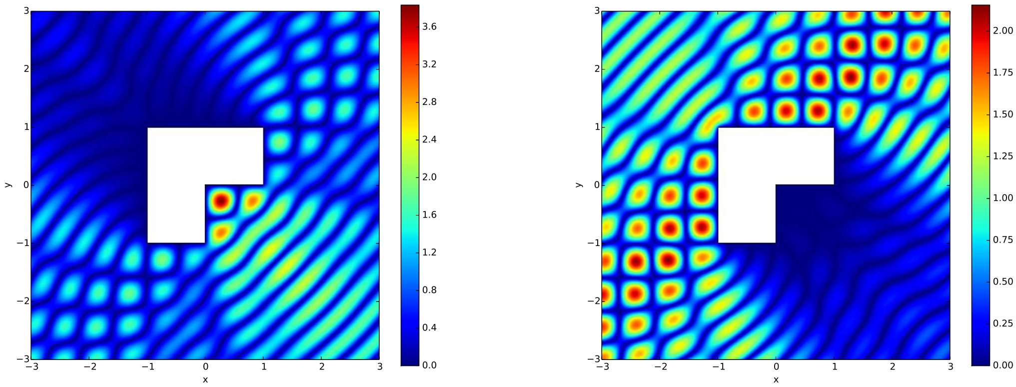


Figure 6.2: Ex. 6.8.1: Total field $u^{\text{tot}}$ at the plane $z = 0$ for different directions of $u^{\text{inc}}$ with $k = 8$. The incident wave hits the scatterer on the non-convex part (left), and on the convex part (right).

### 6.8.2 Sound-hard scattering on a L-shaped domain

For the second example we consider sound-hard scattering on an L-shaped domain given in Figure 6.1. The direction of the incident wave is given by $a = (-1/\sqrt{2}, 1/\sqrt{2}, 0)^T$, hitting the scatterer on a the non-convex part; see Figure 6.8.

Figure 6.9 compares uniform vs. adaptive mesh-refinement for fixed $k = 1$ and various $\theta = \{0.2, 0.4, 0.6, 0.8\}$. Again, Algorithm 6.2 realizes the improved rate $\mathcal{O}(N^{-1})$ for the squared estimator $\eta_\ell^2$, while the uniform strategy leads to a reduced rate of $\mathcal{O}(N^{-2/3})$. Figure 6.10 shows the adaptive rate for various $k \in \{1, 2, 4, 5\}$ and fixed $\theta = 0.2$ (above) as well as $\theta = 0.4$ (below). A higher wavenumber $k$ just influences the invoked constants and the length of the pre-asymptotic phase, but does not effect the rate of convergence. For $k = 16$, we admit that the computed range is not sufficient to observe a better rate of convergence for the adaptive scheme. Finally, we emphasize that Algorithm 6.2 never enforced uniform mesh-refinement in Step (i).

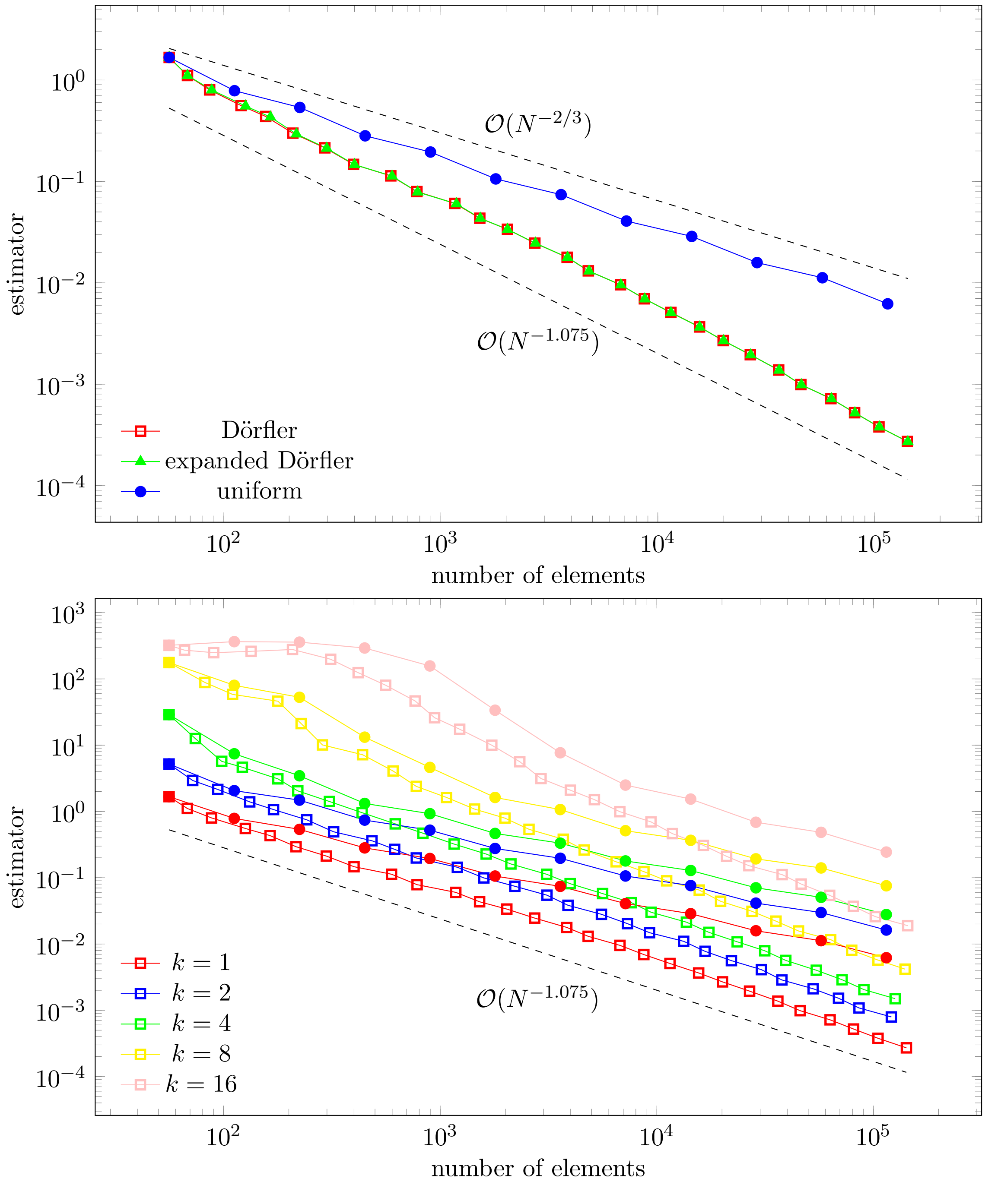



Figure 6.3: Ex. 6.8.1: Convergence of $\eta_\ell^2$ for standard Dörfler marking vs. expanded Dörfler and uniform refinement with $k = 1$ (above). Below, expanded Dörfler marking (squares) vs. uniform refinement (circles) for different values of $k > 0$. Both plots are computed in the non-convex setting with $\theta = 0.4$.

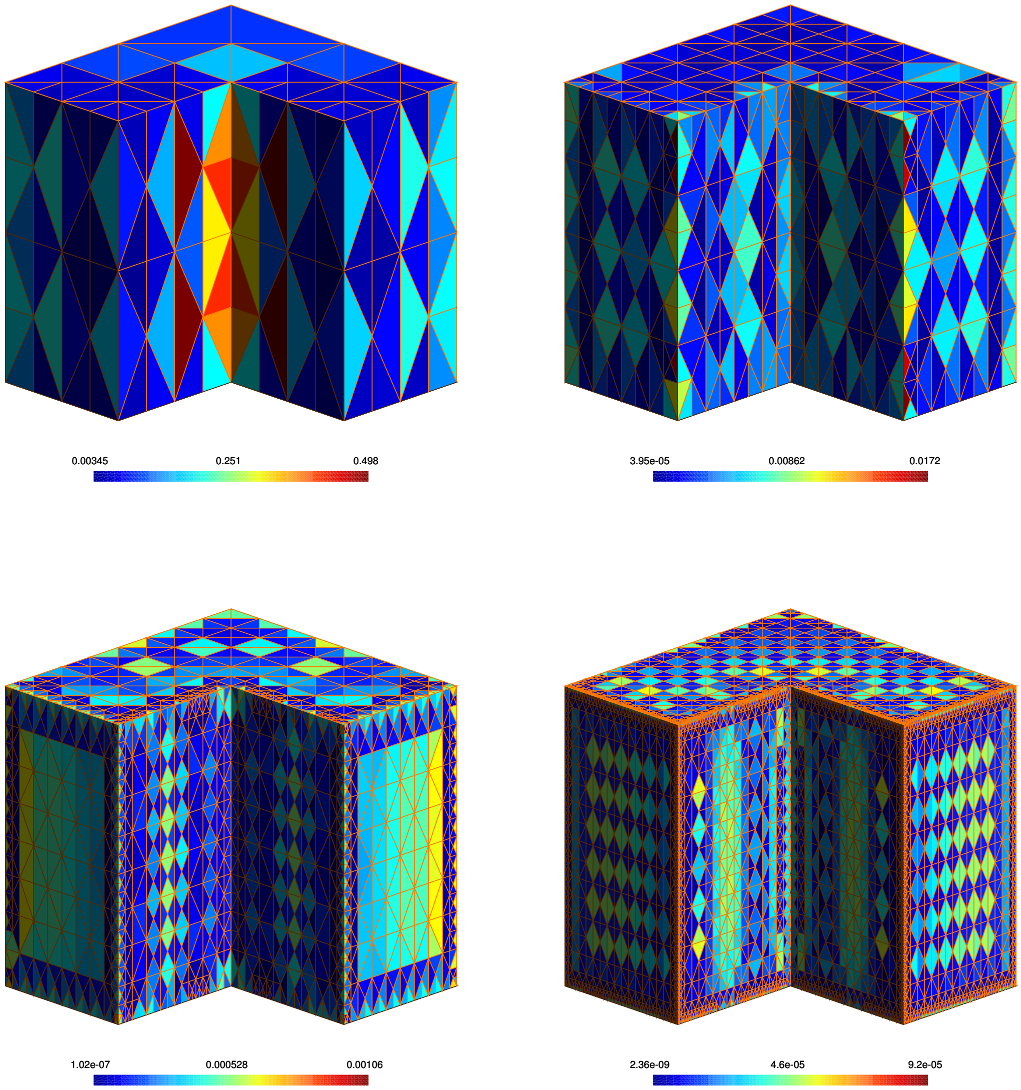



Figure 6.4: Ex. 6.8.1: Triangulations $\mathcal{T}_4$, $\mathcal{T}_8$, $\mathcal{T}_{12}$ and $\mathcal{T}_{16}$ with $208, 766, 2332$ and $6746$ elements. The color indicates the element contribution of the error estimator $\eta_\ell(T)^2$ for all $T \in \mathcal{T}_\ell$.

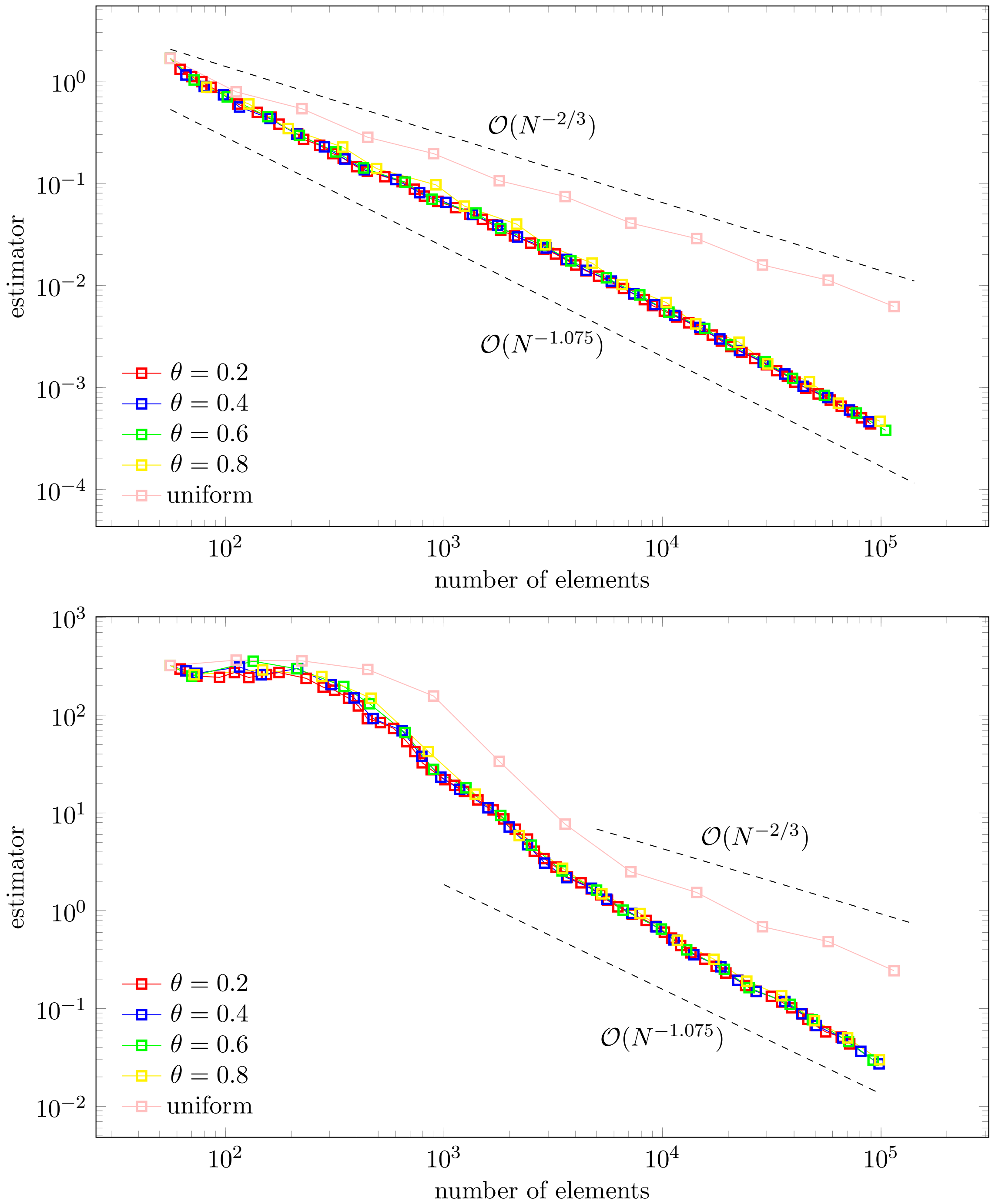



Figure 6.5: Ex. 6.8.1: Convergence of $\eta_\ell^2$ for different values of $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ as well as uniform refinement. Both plots use expanded Dörfler marking with $k = 1$ (above) and $k = 16$ below.

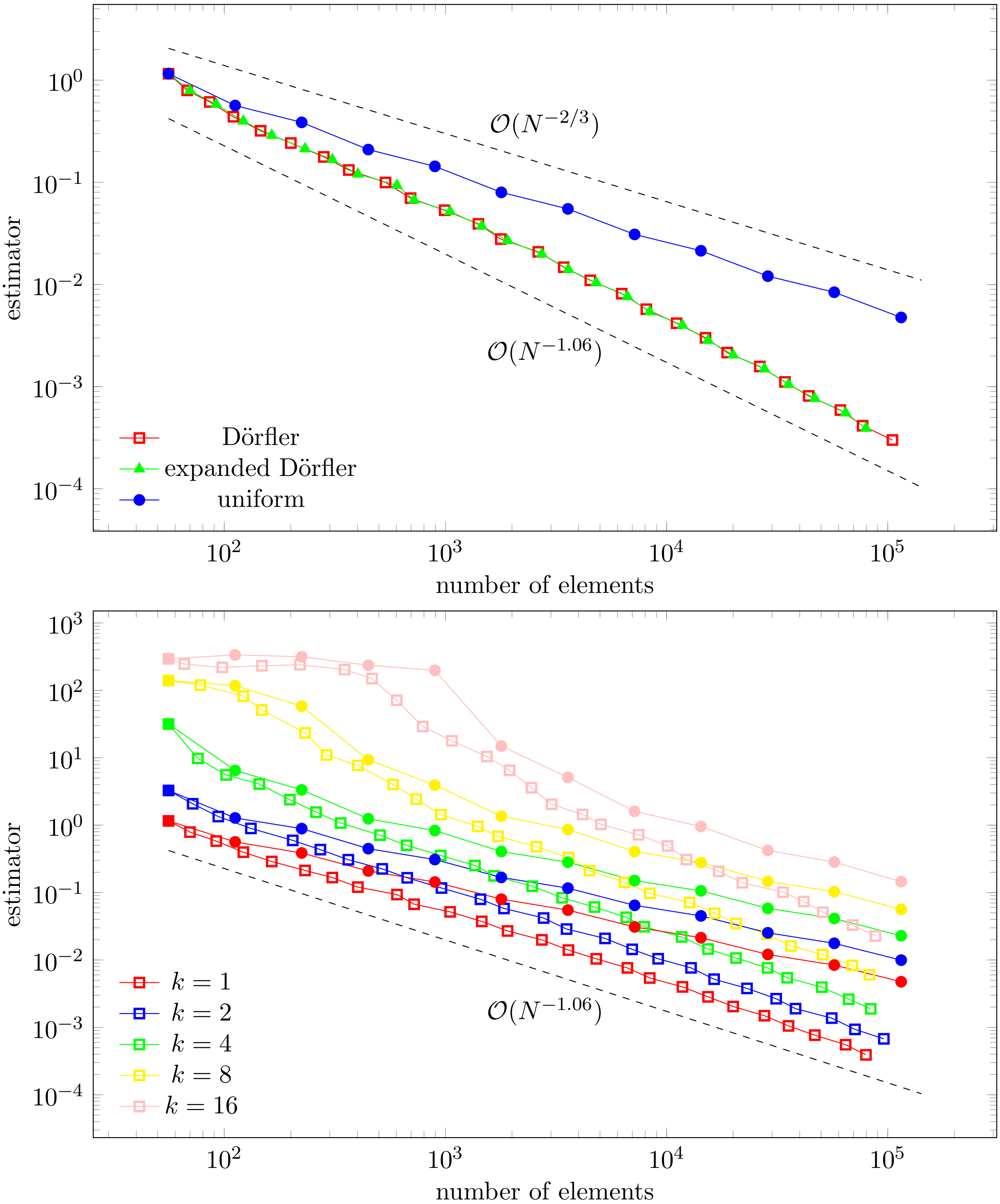



Figure 6.6: Ex. 6.8.1: Convergence of $\eta_\ell^2$ for standard Dörfler marking vs. expanded Dörfler with $\theta = 0.4$ and uniform refinement (above) in the convex case. Below, expanded Dörfler marking (squares) vs. uniform refinement (circles) for different values of $k > 0$.

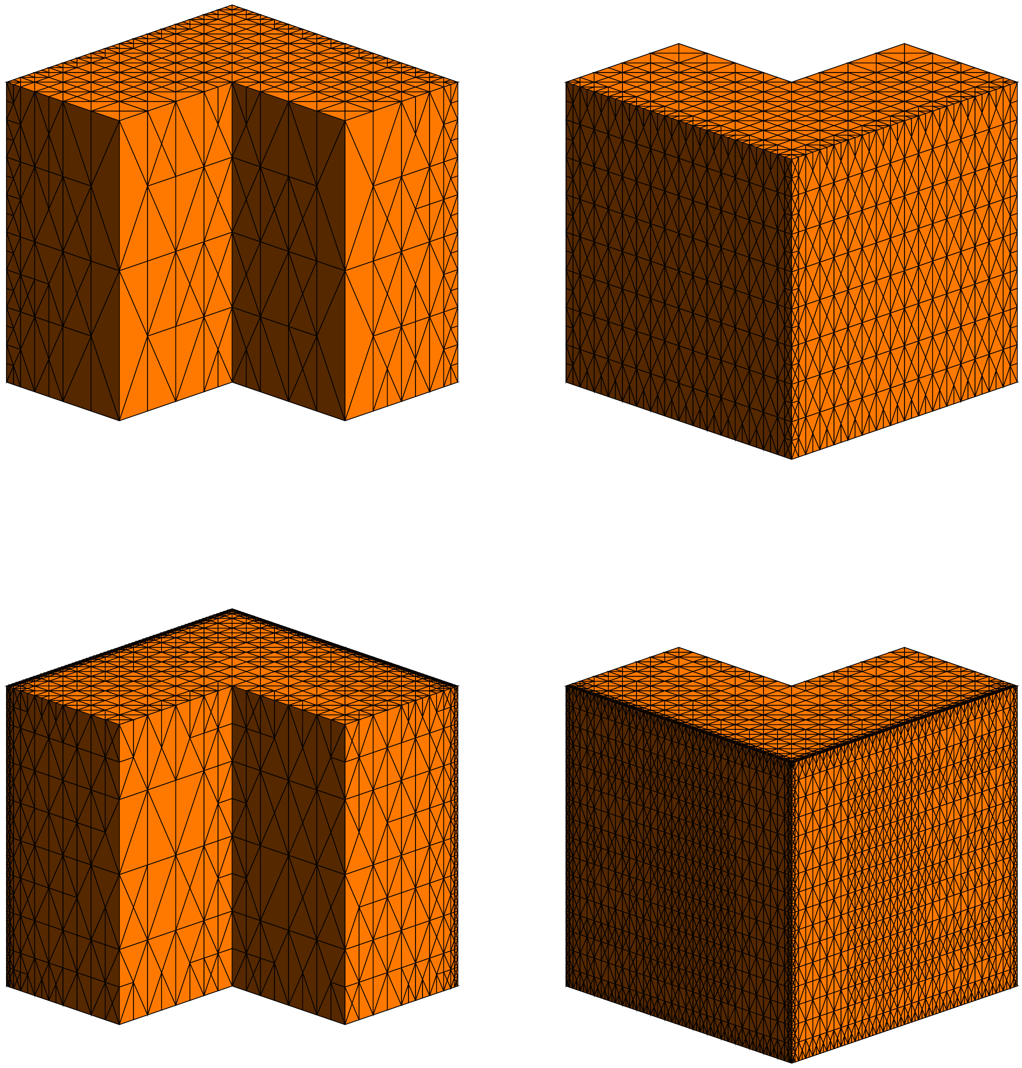


Figure 6.7: Ex. 6.8.1: Triangulations $\mathcal{T}_{12}$ and $\mathcal{T}_{16}$ with 2446 and 7472 elements in the convex case. The refinement focuses on the surface hit by the incoming wave (right), where instead all parts in the shadow remain relatively coarse (left).

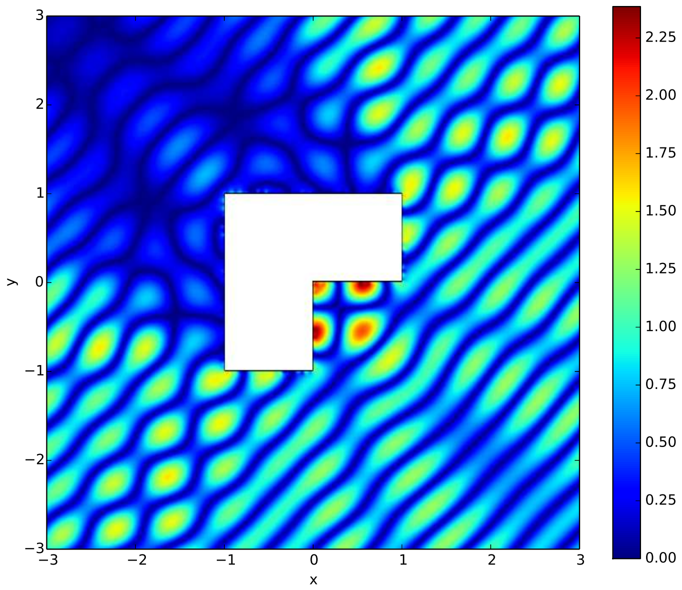



Figure 6.8: Ex. 6.8.2: Total field $u^{\text{tot}}$ for sound-hard scattering with wavenumber $k = 8$. The incident wave $u^{\text{inc}}$ hits the scatterer on the non-convex part of the domain.

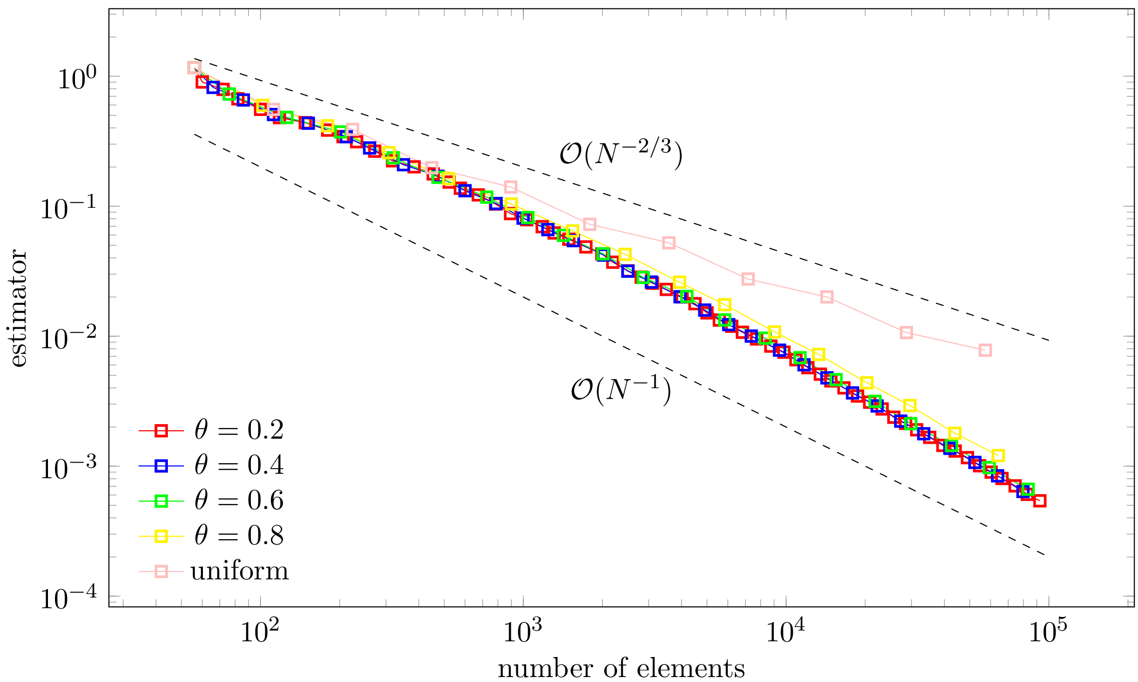



Figure 6.9: Ex. 6.8.2: Convergence of $\eta_\ell^2$ for different values of $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ as well as uniform refinement. The plot uses expanded Dörfler marking with $k = 1$.

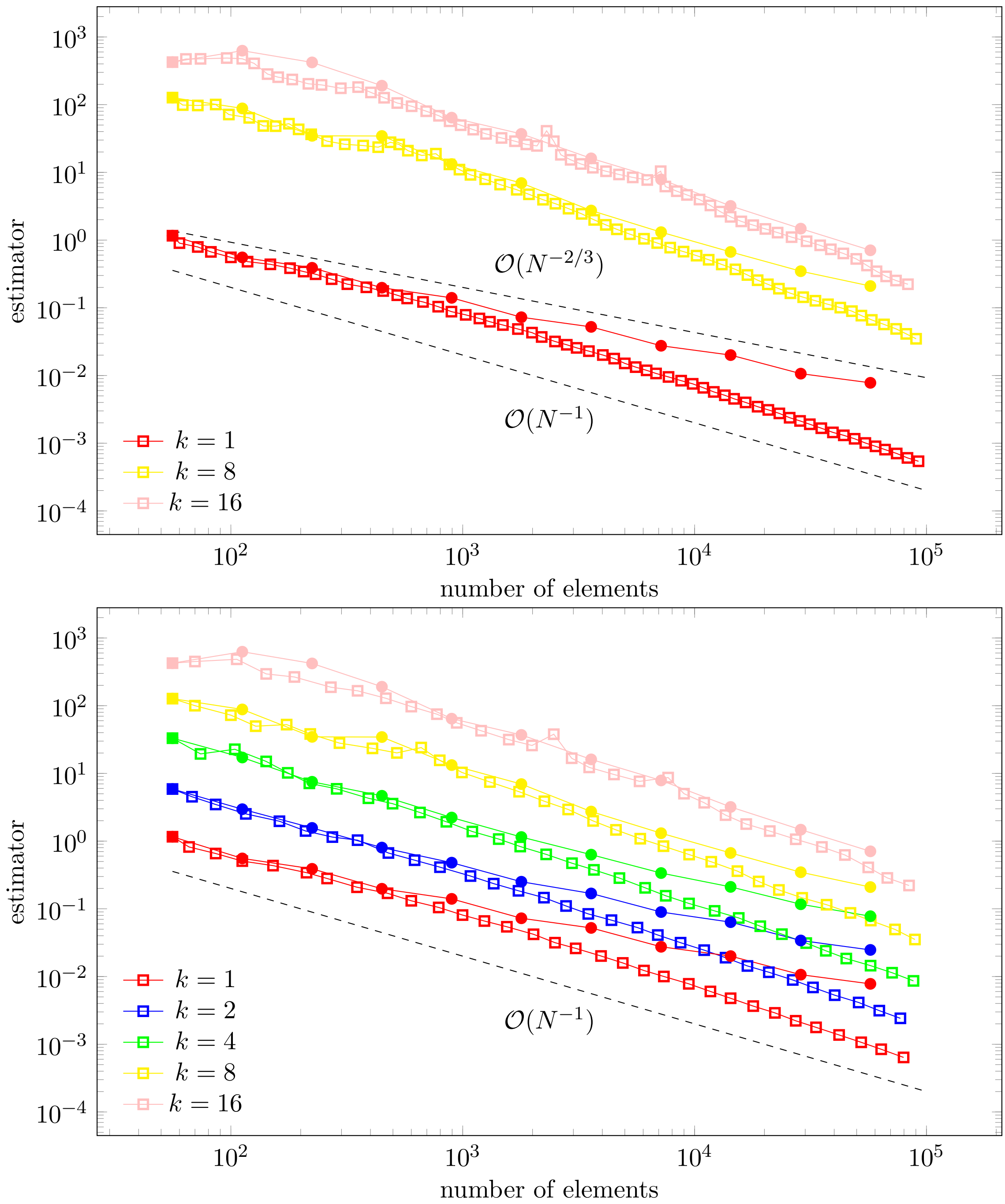



Figure 6.10: Ex. 6.8.2: Convergence of $\eta_\ell^2$ for expanded Dörfler (squares) vs. uniform refinement (circles) for different values of $k \in \{1, 2, 4, 8, 16\}$. The computations use $\theta = 0.2$ (above) as well as $\theta = 0.4$ (below).

# 7 Adaptive BEM for optimal convergence of point errors

One particular strength of the boundary element method is, that it allows for a higher-order point-wise approximation of the solution of the underlying PDE via the representation formula (6.4). As an extension of the previous Chapters 4 and 6, we propose two adaptive algorithms and prove quasi-optimal convergence behavior with respect an *a posteriori* computable bound for the point error of the Helmholtz equation.

For boundary elements based on piecewise polynomials of degree $p$ and a smooth solution $\phi$, it holds that $\|\phi - \Phi_\ell\|_{H^{-1/2}(\Gamma)} = \mathcal{O}(h_\ell^{p+3/2})$ for the energy-error, where $h_\ell$ denotes the mesh-size of $\mathcal{T}_\ell$. On the other, hand the point error decays with a higher rate $|u(\widetilde{x})-u_\ell(\widetilde{x})| = \mathcal{O}(h_\ell^{2p+3})$. However, these convergence rates are usually spoiled by singularities of the (unknown) solution $u$ and hence lack of regularity.

For the Laplace equation, earlier works [AFF$^+$13, FKMP13, FFK$^+$14, FFK$^+$15, Gan13] focused on $h$-adaptive strategies which aim to recover the optimal rate of convergence of the energy error $\|\phi - \Phi_\ell\|_{H^{-1/2}(\Gamma)}$. Using ideas for goal oriented adaptivity for finite elements, see e.g, [MS09, BET11, FPZ16], our own work [FGH$^+$16] proposes and analyzes optimal adaptive strategies for the point-wise error $|u(\widetilde{x}) - u_\ell(\widetilde{x})|$ for $\widetilde{x} \in \Omega$ in case of BEM for the Laplace equation. Using Chapters 4 and 6, we extend the analysis in [FGH$^+$16] from $k = 0$ to the Helmholtz equation for arbitrary wavenumber $k \geq 0$.

**Outline.** Section 7.1 introduces the model problem and some key ideas. Section 7.2 formulates two adaptive algorithms (Algorithm 7.1, Algorithm 7.3) and states the main result (Theorem 7.5) of this chapter which yields optimal convergence for both algorithms. While Algorithm 7.1, follows ideas from [MS09] and employs a separate Dörfler marking strategy, Algorithm 7.3 is inspired by [BET11] and uses a combined Dörfler marking instead. Besides the marking strategies from [MS09, BET11], both algorithms employ the extended Dörfler marking from [BHP17] or Chapter 4 to ensure (E5). The proof of the main result for Algorithms 7.1 and 7.3 is found in Section 7.4 and 7.5 respectively.

This chapter extends the work [FGH$^+$16], where we developed a similar results for the Laplace equation.

## 7.1 Model problem

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with polygonal (not necessarily connected) boundary $\Gamma = \partial\Omega$. In this chapter, we consider the interior Helmholtz–Dirichlet problem

$$-\Delta u - k^2\, u = 0 \quad \text{in } \Omega \qquad \text{subject to} \qquad u = g \text{ on } \Gamma, \tag{7.1}$$

with wavenumber $k \geq 0$ and given Dirichlet data $g \in H^{1/2}(\Gamma)$. Recall the notation of Section 6.2. To guarantee solvability of (7.1), we assume throughout this chapter that $k^2$ is not an eigenvalue of the (IDP), cf. Section 6.2.1.

We recall some properties of Section 6.2. The solution $u \in H^1(\Omega)$ of (7.1) is given by the representation formula (6.4) as

$$u(x) = \widetilde{\mathfrak{V}}_k \, \phi(x) - \widetilde{\mathfrak{K}}_k \, g(x) \quad \text{for all } x \in \Omega, \tag{7.2}$$

where $\phi = \partial_{\mathbf{n}} u$ is the normal derivative of the solution $u$. Further, $\phi \in H^{-1/2}(\Gamma)$ can be obtained by the weakly-singular integral equation

$$\mathfrak{V}_k \, \phi = \big(\mathfrak{K}_k + \frac{1}{2}\,\mathrm{Id}\big)\, g \tag{7.3}$$

where $\mathfrak{V}_k$ denotes the simple-layer and $\mathfrak{K}_k$ the double-layer integral operator corresponding to the wavenumber $k > 0$. For the definition and some fundamental properties of the integral operators we refer to Section 6.2.1. Further details can be found in, e.g., [McL00, SS11].

### 7.1.1 Weakly-singular integral equation

We recap some important properties of the weakly-singular integral equation $\mathfrak{V}_k \, \xi = f$ from Section 6.3. For any given $f \in H^{1/2}(\Gamma)$, the variational formulation reads as: Find $\xi \in H^{-1/2}(\Gamma)$ such that

$$\langle \mathfrak{V}_k \, \xi \, , \, \chi \rangle = \langle f \, , \, \chi \rangle_{L^2(\Gamma)} \quad \text{for all } \chi \in H^{-1/2}(\Gamma). \tag{7.4}$$

Since $k^2$ is not an eigenvalue of the (IDP), the variational formulation (7.4) is well-posed in the sense of (4.4). Let $\mathcal{T}_0$ be an admissible initial mesh and $\mathbb{T} := \mathrm{refine}(\mathcal{T}_0)$ be the set of possible refinements. As in Chapter 6, we use extended 1D bisection (Section 3.4) in case of $d = 2$ and newest vertex bisection (Section 3.5) in case of $d = 3$. For the discretization, we consider standard piecewise polynomial ansatz and test spaces based on regular triangulations of $\Gamma$; see Section 3.6. Then, the Galerkin discretization of (7.4) reads as follows: Find $\Xi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ such that

$$\langle \mathfrak{V}_k \, \Xi_\bullet \, , \, X_\bullet \rangle = \langle f \, , \, X_\bullet \rangle_{L^2(\Gamma)} \quad \text{for all } X_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet). \tag{7.5}$$

According to Section 6.3, the variational formulation (7.4) as well as the Galerkin formulation (7.5) can be recast in the following way

$$\langle \mathfrak{V}_0 \, \Xi_\bullet \, , \, X_\bullet \rangle + \langle \mathfrak{C}_k \, \Xi_\bullet \, , \, X_\bullet \rangle = \langle f \, , \, \Psi_\bullet \rangle_{L^2(\Gamma)} \quad \text{for all } X_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet), \tag{7.6}$$

where the operator $\mathfrak{V}_0$ is symmetric and elliptic and $\mathfrak{C}_k$ is a compact operator. Since we use a verbatim (analytic and discrete) setting as in Section 6.3, existence and uniqueness of the solutions of (7.6) and (7.5) are guaranteed by Proposition 4.1 with $\mathcal{H} := H^{-1/2}(\Gamma)$ and sufficiently fine $\mathcal{X}_\bullet = \mathcal{P}^p(\mathcal{T}_\bullet)$.

### 7.1.2 Weighted-residual error estimator

For any $f \in H^1(\Gamma)$, let $\xi := \mathfrak{V}_k^{-1} f$ be the solution of (7.4) and $\Xi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ denote the corresponding Galerkin solution. For any $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet$ and any $T \in \mathcal{T}_\bullet$, the weighted-residual error estimator is given by

$$\begin{aligned} \eta_{\xi,\bullet}(T)^2 &:= h_\bullet^{1/2} \, \|\nabla(f - \mathfrak{V}_k \, \Xi_\bullet)\|_{L^2(T)}^2, \\ \eta_{\xi,\bullet} &:= \eta_{\xi,\bullet}(\mathcal{T}_\star), \quad \text{where} \quad \eta_{\xi,\bullet}(\mathcal{U}_\bullet) := \Big( \sum_{T \in \mathcal{U}_\bullet} \eta_{\xi,\bullet}(T)^2 \Big)^{1/2}. \end{aligned} \tag{7.7}$$

According to Section 6.5 the error estimator $\eta_{\xi,\bullet}$ satisfies the following properties:

- Stability on non refined element domains (E1), see Proposition 6.6.
- Reduction on refined element domains (E2), see Proposition 6.7.
- Reliability (E3), see Corollary 6.9.
- Discrete reliability (E4), see Proposition 6.8.

### 7.1.3 Main idea and dual problem

The main idea of the following adaptive strategies reads as follows: Let $\phi \in H^{-1/2}(\Gamma)$ be the unique solution of the weakly-singular integral equation (7.3) with $f = (\mathfrak{K}_k + 1/2)\, g$. Let $\mathcal{T}_\bullet \in \mathbb{T}$ and suppose that the corresponding Galerkin approximation $\Phi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ of (7.5) exists. Having obtained $\Phi_\bullet$, the representation formula gives rise to an approximation $U_\bullet \in H^1(\Omega)$ of $u$ by

$$U_\bullet(x) := \widetilde{\mathfrak{V}}_k \, \Phi_\bullet(x) - \widetilde{\mathfrak{K}}_k \, g(x) \quad \text{for all } x \in \Omega. \tag{7.8}$$

Recall that $\mathfrak{V}_k$ is symmetric. Then, the Galerkin orthogonality directly implies that

$$\langle \mathfrak{V}_k \, X_\bullet \, , \, \phi - \Phi_\bullet \rangle = \langle \mathfrak{V}_k \, (\phi - \Phi_\bullet) \, , \, X_\bullet \rangle = 0 \quad \text{for all } X_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet). \tag{7.9}$$

Suppose that $\widetilde{x} \in \Omega$ is an arbitrary but fixed evaluation point. With the representation formula (7.2) and (7.8), it thus holds that

$$u(\widetilde{x}) - U_\bullet(\widetilde{x}) = \widetilde{\mathfrak{V}}_k \, (\phi - \Phi_\bullet)(\widetilde{x}). \tag{7.10}$$

Since $\widetilde{x} \in \Omega$, the fundamental solution $G_k(\widetilde{x}, \cdot)$ is a smooth function on $\Gamma$; see, e.g., [SS11, Chapter 3.1]. Further, $\mathfrak{V}_k$ is invertible if and only if $k^2$ is not an eigenvalue of the (IDP). Galerkin orthogonality (7.9) and the definition of $\mathfrak{V}_k$ yield for any $X_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ that

$$\widetilde{\mathfrak{V}}_k \, (\phi - \Phi_\bullet)(\widetilde{x}) \stackrel{(6.8)}{=} \langle G_k(\widetilde{x}, \cdot) \, , \, \phi - \Phi_\bullet \rangle \stackrel{(7.9)}{=} \langle G_k(\widetilde{x}, \cdot) - \mathfrak{V}_k \, X_\bullet \, , \, \phi - \Phi_\bullet \rangle. \tag{7.11}$$

Suppose that $\psi^{[\widetilde{x}]}$ solves the following *dual problem*: Given $\widetilde{x}$ and $G_k(\widetilde{x}, \cdot) \in H^{1/2}(\Gamma)$ find $\psi^{[\widetilde{x}]} \in H^{-1/2}(\Gamma)$ such that

$$\mathfrak{V}_k \, \psi^{[\widetilde{x}]} = G_k(\widetilde{x}, \cdot). \tag{7.12}$$

Note that, the dual problem is a weakly-singular integral equation in the sense of (7.5) with right-hand side $f = G_k(\widetilde{x}, \cdot)$. The corresponding Galerkin discretization of (7.12) is given by: Find $\Psi_\bullet^{[\widetilde{x}]} \in \mathcal{P}^p(\mathcal{T}_\bullet)$ such that

$$\langle \mathfrak{V}_k \, \Psi_\bullet^{[\widetilde{x}]} \, , \, X_\bullet \rangle = \langle G_k(\widetilde{x}, \cdot) \, , \, X_\bullet \rangle \quad \text{for all } X_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet). \tag{7.13}$$

We emphasize that the right hand side $G_k(\widetilde{x}, \cdot)$ and therefore, $\psi^{[\widetilde{x}]}$ as well as $\Psi_\bullet^{[\widetilde{x}]}$ depend on the arbitrary but fixed evaluation point $\widetilde{x} \in \Omega$. Since the only difference of the primal and dual problem is the right hand side $f$, existence and uniqueness of solutions $\Psi_\bullet^{[\widetilde{x}]} \in \mathcal{P}^p(\mathcal{T}_\bullet)$ of (7.13) is equivalent to solvability of the primal problem (7.6) and guaranteed by Proposition 4.1 even with the same index $\ell_0$.

Now, suppose that $\mathcal{T}_\bullet \in \mathbb{T}$ admits discrete solutions $\Phi_\bullet, \Psi_\bullet^{[\widetilde{x}]} \in \mathcal{P}^p(\mathcal{T}_\bullet)$ to the corresponding Galerkin formulation (7.6). With (7.10) and (7.11), we derive for $X_\bullet := \Psi_\bullet^{[\widetilde{x}]}$ that

$$\begin{aligned} |u(\widetilde{x}) - U_\bullet(\widetilde{x})| &= \big|\langle G_k(\widetilde{x}, \cdot) - \mathfrak{V}_k \Psi_\bullet^{[\widetilde{x}]} \, , \, \phi - \Phi_\bullet \rangle\big| \\ &\le \|G_k(\widetilde{x}, \cdot) - \mathfrak{V}_k \Psi_\bullet^{[\widetilde{x}]}\|_{H^{1/2}(\Gamma)} \, \|\phi - \Phi_\bullet\|_{H^{-1/2}(\Gamma)} \\ &\simeq \|\psi^{[\widetilde{x}]} - \Psi_\bullet^{[\widetilde{x}]}\|_{H^{-1/2}(\Gamma)} \, \|\phi - \Phi_\bullet\|_{H^{-1/2}(\Gamma)}, \end{aligned}$$

where the hidden constants depend only on $\Gamma$. Either of these Galerkin errors will be controlled by the respective weighted-residual error estimator which requires additional regularity $g \in H^1(\Gamma)$ for the Dirichlet data. To this end, let $\eta_{\phi,\bullet}$, $\eta_{\psi^{[\widetilde{x}]},\bullet}$ denote the corresponding error estimators for the primal and the dual problem. With reliability (E3), the latter estimate turns into

$$|u(\widetilde{x}) - U_\bullet(\widetilde{x})| \lesssim \|\psi^{[\widetilde{x}]} - \Psi_\bullet^{[\widetilde{x}]}\|_{H^{-1/2}(\Gamma)} \, \|\phi - \Phi_\ell\|_{H^{-1/2}(\Gamma)} \lesssim \eta_{\phi,\bullet} \, \eta_{\psi^{[\widetilde{x}]},\bullet}. \tag{7.14}$$

To abbreviate notation and if its clear from the context, we omit the dependence on $\widetilde{x}$ and use $\psi$ instead of $\psi^{[\widetilde{x}]}$ as well as $\Psi$ instead of $\Psi^{[\widetilde{x}]}$.

## 7.2 Adaptive algorithm

The previous section gives rise to the following two adaptive algorithms. These have been proposed and analyzed by [MS09, BET11] for goal-oriented adaptivity in the context of FEM for the Poisson problem and in [FGH+16] for ABEM for point-wise approximation of the solutions of the Laplace equation. The first algorithm goes back to [MS09].

**Algorithm 7.1.** INPUT: *Parameters* $0 < \theta \le 1$ *and* $C_{\rm mark} \ge 1$ *as well as initial triangulation* $\mathcal{T}_0$ *with* $\Psi_{-1} := \Phi_{-1} := 0 \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\phi,-1} = \eta_{\psi,-1} := 1$.
ADAPTIVE LOOP: *For all* $\ell = 0, 1, 2, \dots,$ *iterate the following Steps* (i)–(vii)*:*

(i) `If` (7.3) *does not admit a unique solution* $\Phi_\ell \in \mathcal{P}_0^p(\mathcal{T}_\ell)$*:*
  - *Define* $\Phi_\ell := \Phi_{\ell-1} \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\phi,\ell} := \eta_{\phi,\ell-1}$*,*
  - *Define* $\Psi_\ell := \Psi_{\ell-1} \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\psi,\ell} := \eta_{\psi,\ell-1}$*,*

– *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ *be the uniform refinement of* $\mathcal{T}_\ell$*,*

– *Increase* $\ell \to \ell + 1$*, and continue with Step* (i)*.*

(ii.a) `Else` *Compute Galerkin approximation* $\Phi_\ell$ *to* $\phi$*.*

(ii.b) *Compute Galerkin approximation* $\Psi_\ell$ *to* $\psi$*.*

(iii.a) *Compute refinement indicators* $\eta_{\phi,\ell}(T)$ *for all* $T \in \mathcal{T}_\ell$*.*

(iii.b) *Compute refinement indicators* $\eta_{\psi,\ell}(T)$ *for all* $T \in \mathcal{T}_\ell$*.*

(iv.a) *Determine a set* $\mathcal{M}_{\phi,\ell} \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\mathrm{mark}}$ *minimal cardinality such that*

$$\theta\, \eta_{\phi,\ell}^2 \le \sum_{T \in \mathcal{M}_{\phi,\ell}} \eta_{\phi,\ell}(T)^2. \tag{7.15}$$

(iv.b) *Determine a set* $\mathcal{M}_{\psi,\ell} \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\mathrm{mark}}$ *minimal cardinality such that*

$$\theta\, \eta_{\psi,\ell}^2 \le \sum_{T \in \mathcal{M}_{\psi,\ell}} \eta_{\psi,\ell}(T)^2. \tag{7.16}$$

(v) *Choose* $\mathcal{M}'_\ell \in \{\mathcal{M}_{\phi,\ell}, \mathcal{M}_{\psi,\ell}\}$ *as the set of minimal cardinality.*

(vi) *Find* $\mathcal{M}''_\ell \subseteq \mathcal{T}_\ell$ *such that* $\#\mathcal{M}''_\ell = \#\mathcal{M}'_\ell$ *as well as* $h_\ell(T) \ge h_\ell(T')$ *for all* $T \in \mathcal{M}''_\ell$ *and* $T' \in \mathcal{T}_\ell \setminus \mathcal{M}''_\ell$*. Define* $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \mathcal{M}''_\ell$*.*

(vii) *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ *be the coarsest refinement of* $\mathcal{T}_\ell$ *such that all marked elements* $T \in \mathcal{M}_\ell$ *have been refined. Increase* $\ell \to \ell + 1$ *and continue with Step* (i)*.*

Output: *Discrete approximations* $\Phi_\ell, \Psi_\ell$ *and corresponding error estimators* $\eta_{\phi,\ell}, \eta_{\psi,\ell}$ *for all* $\ell \in \mathbb{N}_0$*.* □

**Remark 7.2.** *i) Besides Step* (i) *and Step* (vi)*, Algorithm 7.1 coincides with the adaptive Algorithm proposed in [MS09, FGH+16].*

*ii) Recall that, the discrete problem* (7.5) *in general does not admit a unique solution for all* $\mathcal{T}_\bullet \in \mathbb{T}$*. Instead, Proposition 4.1 guarantees an index* $\ell_0$*, such that for all* $\ell \ge \ell_0$ *the mesh* $\mathcal{T}_\ell$ *admits unique solutions* $\Phi_\ell, \Psi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$*; see also Lemma 4.6 and Lemma 4.17. Therefore, Step* (i) *will be only carried out at most* $\ell_0$*-times.*

*iii) Step* (vi) *is found verbatim in Algorithm 6.2 and applies the expanded Dörfler marking strategy of Proposition 4.7. According to Section 4.5.1 and Section 6.5.1 this ensures definiteness on the "discrete" limit space* (E5)*.*

The second algorithm has been analyzed in [BET11] in the context of finite elements and in [FGH+16] for ABEM for the Laplacian. Note that both algorithms only differ in the used marking strategy. While Algorithm 7.1 employs a separate Dörfler marking in Step (v)–(vii), Algorithm 7.3 computes a combined refinement indicator $\rho_\ell$ in Step (iv) and employs the expanded Dörfler marking for $\rho_\ell$ in Steps (v)–(vi).

**Algorithm 7.3.** INPUT: *Parameters* $0 < \theta \leq 1$ *and* $C_{\mathrm{mark}} \geq 1$ *as well as initial triangulation* $\mathcal{T}_0$ *with* $\Psi_{-1} := \Phi_{-1} := 0 \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\phi,-1} = \eta_{\psi,-1} := 1$.
ADAPTIVE LOOP: *For all* $\ell = 0, 1, 2, \dots$, *iterate the following Steps* (i)–(vii)*:*

(i) `If` (7.3) *does not admit a unique solution* $\Phi_\ell \in \mathcal{P}^p_0(\mathcal{T}_\ell)$*:*
   - *Define* $\Phi_\ell := \Phi_{\ell-1} \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\phi,\ell} := \eta_{\phi,\ell-1}$,
   - *Define* $\Psi_\ell := \Psi_{\ell-1} \in \mathcal{P}^p(\mathcal{T}_0)$ *and* $\eta_{\psi,\ell} := \eta_{\psi,\ell-1}$,
   - *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{T}_\ell)$ *be the uniform refinement of* $\mathcal{T}_\ell$,
   - *Increase* $\ell \to \ell + 1$, *and continue with Step* (i).

(ii.a) `Else` *Compute Galerkin approximation* $\Phi_\ell$ *to* $\phi$.

(ii.b) *Compute Galerkin approximation* $\Psi_\ell$ *to* $\psi$.

(iii.a) *Compute indicators* $\eta_{\phi,\ell}(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iii.b) *Compute indicators* $\eta_{\psi,\ell}(T)$ *for all* $T \in \mathcal{T}_\ell$.

(iv) *Assemble refinement indicators* $\rho_\ell(T)^2 := \eta_{\phi,\ell}(T)^2\eta^2_{\psi,\ell} + \eta^2_{\phi,\ell}\eta_{\psi,\ell}(T)^2$ *for all* $T \in \mathcal{T}_\ell$.

(v) *Determine a set* $\mathcal{M}'_\ell \subseteq \mathcal{T}_\ell$ *of up to the multiplicative factor* $C_{\mathrm{mark}}$ *minimal cardinality such that*

$$\theta\, \rho_\ell^2 \leq \rho_\ell(\mathcal{M}'_\ell)^2, \tag{7.17}$$

(vi) *Find* $\mathcal{M}''_\ell \subseteq \mathcal{T}_\ell$ *such that* $\#\mathcal{M}''_\ell = \#\mathcal{M}'_\ell$ *as well as* $h_\ell(T) \geq h_\ell(T')$ *for all* $T \in \mathcal{M}''_\ell$ *and* $T' \in \mathcal{T}_\ell \setminus \mathcal{M}''_\ell$. *Define* $\mathcal{M}_\ell := \mathcal{M}'_\ell \cup \mathcal{M}''_\ell$.

(vii) *Let* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ *be the coarsest refinement of* $\mathcal{T}_\ell$ *such that all marked elements* $T \in \mathcal{M}_\ell$ *have been refined. Increase* $\ell \to \ell + 1$ *and continue with Step* (i).

Output: *Discrete approximations* $\Phi_\ell, \Psi_\ell$ *and corresponding error estimators* $\eta_{\phi,\ell}, \eta_{\psi,\ell}$ *for all* $\ell \in \mathbb{N}_0$. □

Since Step (i) and Step (vi) of Algorithm 7.1 coincide with Step (i) and Step (vi) of Algorithm 7.3, Remark 7.2 holds verbatim for Algorithm 7.3.

**Remark 7.4.** *We note that the Algorithms 7.1 and 7.3 as well as Theorem 7.5 are independent of whether we use direct BEM for the interior or exterior problem (based on the representation formula (6.4)), or indirect BEM, where we solve* $\mathfrak{V}_k\,\phi = f$ *for some given right-hand side* $f \in H^1(\Gamma)$ *and aim to approximate* $\widetilde{\mathfrak{V}}_k\,\phi(\widetilde{x}) \approx \widetilde{\mathfrak{V}}_k\,\Phi_\ell(\widetilde{x})$ *for some* $\widetilde{x} \in \mathbb{R}^d \backslash \Gamma$. *Moreover, all results hold accordingly if the Dirichlet data g are given, while the hyper-singular integral equation is employed to approximate the (unknown) Neumann data, cf. Section 6.7.* □

## 7.3 Optimal convergence

The main result of this chapter is that Algorithm 7.1 as well as Algorithm 7.3 lead to optimal convergence of the estimator product $\eta_{\phi,\ell}\,\eta_{\psi,\ell}$. Throughout, to indicate that a result holds for the primal as well as the dual problem, we use the abbreviate notation $\xi \in \{\phi, \psi\}$, with corresponding Galerkin solution $\Xi_\bullet \in \{\Phi_\bullet, \Psi_\bullet\}$. Further, we write $\eta_{\xi,\bullet} \in \{\eta_{\phi,\bullet}, \eta_{\psi,\bullet}\}$ for the associated error estimator.

For the statement of the main result, recall the abstract approximation class from Section 4.8.1. For $s > 0$ and $\mathcal{T} \in \mathbb{T}$, we write $\xi \in \mathbb{A}_s(\mathcal{T})$ if

$$\|\xi\|_{\mathbb{A}_s(\mathcal{T})} := \sup_{N \in \mathbb{N}_0} \Big( (N+1)^s \min_{\mathcal{T}_\bullet \in \mathbb{T}_N(\mathcal{T})} \eta_{\xi,\bullet} \Big) < \infty, \tag{7.18}$$

where $\eta_{\xi,\bullet}$ is the weighted-residual error estimator associated with the optimal mesh $\mathcal{T}_\bullet \in \mathbb{T}_N$. The following main theorem states that both adaptive algorithms do not only lead to linear convergence, but also that each possible algebraic rate $s > 0$ will asymptotically be realized.

**Theorem 7.5.** *Suppose* (E1)–(E5). *For all $0 < \theta \le 1$, there exists constants $0 < q_{\rm lin} < 1$ and $C_{\rm lin} > 0$ and an index $\ell_{\rm lin} > 0$ such that the sequences of estimators $(\eta_{\xi,\ell})_{\ell \in \mathbb{N}_0}$ generated by Algorithm 7.1 and Algorithm 7.3 guarantee*

$$\eta_{\phi,\ell+n}\,\eta_{\psi,\ell+n} \le C_{\rm lin}\, q_{\rm lin}^n\, \eta_{\phi,\ell}\,\eta_{\psi,\ell} \quad \text{for all } \ell, n \in \mathbb{N}_0 \text{ with } \ell \ge \ell_{\rm lin}. \tag{7.19}$$

*Let $\widehat{\beta}_0 > 0$ be the lower-bound of the* inf-sup *constant* (4.6) *for the uniform refinement $\widehat{\mathcal{T}}_0$ from Lemma 4.17. Moreover, let $0 < \theta < \theta_{\rm opt} := (1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$ for Algorithm 7.1 and $0 < \theta < \theta_{\rm opt}/2$ in case of Algorithm 7.3. Then, for all $s, t > 0$ it holds that*

$$\|\phi\|_{\mathbb{A}_s}\,\|\psi\|_{\mathbb{A}_t} < \infty \Longrightarrow \exists \ell_{\rm opt} \in \mathbb{N}_0\, \exists C_{\rm opt} > 0\, \forall \ell \ge \ell_{\rm opt} \quad \eta_{\phi,\ell}\,\eta_{\psi,\ell} \le C_{\rm opt}\,(\#\mathcal{T}_\ell - \#\mathcal{T}_0)^{-(s+t)}. \tag{7.20}$$

*i.e., Algorithm 7.1 as well as Algorithm 7.3 guarantee that the product of the error estimators decays with any possible algebraic rate. The constant $C_{\rm opt}$ depends only on $\#\mathcal{T}_{\ell_{\rm opt}}$, $\mathcal{T}_0$, $\theta$, $s, t$, and validity of* (E1)–(E5).

**Remark 7.6.** *In principle, the analysis covers goal-oriented adaptivity if the goal function $u(\widetilde{x})$ and its approximation $U_\ell(\widetilde{x})$ satisfy* (7.14), *where the error estimators satisfy the properties* (E1)–(E4). *For instance, this is the case for goal-oriented FEM for symmetric and elliptic PDEs with $L^2$-goal functional; see Chapter 5 resp. [CFPP14] for the verification of the properties* (E1)–(E4). *Thus, our analysis also extends the results of [MS09, BET11] beyond the Poisson problem to the problem class of [CKNS08]. An abstract approach and analysis of goal-oriented adaptivity for finite elements is found in [FPZ16], where also non-symmetric differential operators are considered. In addition, our analysis avoids any (discrete) efficiency estimates (which are open for BEM) and allows for simple newest vertex bisection, while [MS09, BET11] required local bisec5-refinement in the spirit of [Ste07].*

For goal-oriented FEM for the Poisson problem with polynomial data, [BET11] proves that Algorithm 7.3 leads to linear convergence $\mathtt{err}_{\ell+1} \le q\, \mathtt{err}_\ell$, where instead Algorithm 7.1

leads only to a weaker contraction $\mathtt{err}_{\ell+1} \le q^{1/2}\,\mathtt{err}_\ell$. Here, $0 < q < 1$, and $\mathtt{err}_\bullet$ is the product of the energy errors with respect to some mesh $\mathcal{T}_\bullet \in \mathbb{T}$; see [BET11, eq. (2.12)] and [BET11, eq. (2.20)].

Hence, at least in particular situations, the combined Dörfler marking (7.17) leads to a reduction of adaptive steps and therefore appears to be more effective overall. To our best knowledge, no such result for goal-oriented adaptivity with boundary elements is known. Although we did not succeed to prove such a statement in the present case, this aspect is empirically addressed by appropriate numerical experiments in [FGH+16]. Finally, we note that our proof of Theorem 7.5 provides an upper bound $0 < \theta_{\rm opt} < 1$ such that optimal convergence rates (7.20) are guaranteed for Algorithm 7.1 for all $0 < \theta < \theta_{\rm opt}$, but for Algorithm 7.3 only for all $0 < \theta < \theta_{\rm opt}/2$.

### 7.3.1 Separated linear convergence

In order to prove linear convergence (7.19), we first show that each of the involved estimators satisfies a slightly generalized form of linear convergence. To that end let $\xi \in \{\phi, \psi\}$. The following corollary recaps the generalized reduction property from Chapter 4 in the current notation and is an immediate consequence of Lemma 4.15.

**Corollary 7.7** (generalized contraction)**.** *Let $0 < \theta \le 1$. Suppose that the corresponding error estimator $\eta_\xi$ satisfies* (E1)–(E5)*. Let $\mathcal{T}_\ell \in \mathbb{T}$ and $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\ell)$ be given meshes such that the corresponding discrete solutions $\Xi_\ell, \Xi_\circ$ exist. Further, suppose that the set of refined elements satisfies the Dörfler marking criterion, i.e., $\theta\,\eta_{\xi,\ell}^2 \le \eta_{\xi,\ell}(\mathcal{T}_\ell \backslash \mathcal{T}_\circ)^2$. Then, there exist constants $0 < q_{\mathrm{ctr},\xi}, \lambda_\xi < 1$ such that*

$$\Delta_{\xi,\circ} \le q_{\mathrm{ctr},\xi}\,\Delta_{\xi,\ell}, \quad \text{for all } \ell \ge \ell_{3,\xi}, \quad \text{where} \quad \Delta_{\xi,\bullet} := |\!|\!|\, \xi - \Xi_\bullet \,|\!|\!|^2 + \lambda_\xi\,\eta_{\xi,\bullet}^2, \tag{7.21}$$

*and $\ell_{3,\xi} \in \mathbb{N}_0$ is the index from Lemma 4.13. The constants $q_{\mathrm{ctr},\xi}$ and $\lambda_\xi$ depend on $C_{\rm rel}$ and $q_{\rm est}$.* □

The next proposition generalizes the concept of linear convergence (cf. Theorem 4.14) in the way that in $n$ steps of the adaptive algorithm, Dörfler marking for an estimator $\eta_{\xi_\ell}$ is performed only $k \le n$ times.

**Proposition 7.8.** *Suppose* (E1)–(E4)*. Let $0 < \theta \le 1$ and $\ell_{3,\xi}, \ell_{2,\xi} \in \mathbb{N}_0$ be the indices from Lemma 4.13 and Lemma 4.8. Let $\mathcal{T}_\ell$ be a sequence of successively refined meshes such that* (E5) *is satisfied and the corresponding discrete solutions $\Xi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$ exist for all $\ell \ge \ell_{3,\xi}$. Then, there are constants $0 < q_{\mathrm{ctr},\xi} < 1$ and $C_{\mathrm{lin},\xi} > 0$ such that the following holds:*

*Let $\ell \ge \max\{\ell_{2,\xi}, \ell_{3,\xi}\}$, $n \in \mathbb{N}_0$ and suppose that there are at least $k \le n$ indices $\ell \le j_1 < j_2 < \cdots < j_k < \ell + n$ such that the Dörfler marking for $\eta_\xi$ is satisfied on the refined elements, i.e.,*

$$\theta\,\eta_{\xi,j_m}^2 \le \eta_{\xi,j_m}(\mathcal{T}_{j_m} \setminus \mathcal{T}_{j_m+1})^2 \quad \text{for all } m = 1, \dots, k. \tag{7.22}$$

*Then, the error estimator satisfies*

$$\eta_{\xi,\ell+n}^2 \le C_{\mathrm{lin},\xi}\, q_{\mathrm{ctr},\xi}^k\, \eta_{\xi,\ell}^2 \quad \text{for all } \ell \ge \max\{\ell_{2,\xi}, \ell_{3,\xi}\}. \tag{7.23}$$

*The constant $C_{\mathrm{lin},\xi}$ depends only on $\lambda_\xi$, $C_{\mathrm{mon},\xi}$ and $C_{\rm rel}$.*

*Proof.* Recall the definition of the quantities $\Delta_{\xi,\bullet}$ from Corollary 7.7. With reliability (E3) as well as $\|\!| \cdot |\!\| \simeq \| \cdot \|_{H^{-1/2}(\Gamma)}$, we obtain that

$$\lambda_\xi \, \eta_{\xi,\bullet}^2 \le \Delta_{\xi,\bullet} \quad \text{and} \quad \Delta_{\xi,\bullet} \le (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, \eta_{\xi,\bullet}^2.$$

Since $\ell \ge \ell_{2,\xi}$, quasi-monotonicity from Lemma 4.8 applies. With $C'_{\mathrm{mon},\xi} := (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, C_{\mathrm{mon},\xi} \, \lambda_\xi^{-1}$, we obtain quasi-monotonicity for all $\Delta_{\xi,\circ}$ corresponding to $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\ell)$,

$$\begin{aligned} \Delta_{\xi,\circ} \le (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, \eta_{\xi,\circ}^2 &\le (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, C_{\mathrm{mon},\xi} \, \eta_{\xi,\ell}^2 \\ &\le (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, C_{\mathrm{mon},\xi} \, \lambda_\xi^{-1} \, \Delta_{\xi,\ell} = C'_{\mathrm{mon},\xi} \, \Delta_{\xi,\ell}. \end{aligned} \tag{7.24}$$

Note that, Dörfler marking in Step $j_k$ for $\mathcal{T}_{j_k} \setminus \mathcal{T}_{j_k+1}$ implies Dörfler marking on $\mathcal{T}_{j_k} \setminus \mathcal{T}_{\ell+n} \subseteq \mathcal{T}_{j_k} \setminus \mathcal{T}_{j_k+1}$. Then, (7.24) with $\Delta_{\xi,\circ} = \Delta_{\xi,\ell+n}$ and Corollary 7.7 yield that

$$\lambda_\xi \, \eta_{\xi,\ell+n}^2 \le \Delta_{\xi,\ell+n} \overset{(7.21)}{\le} q_{\mathrm{ctr},\xi} \, \Delta_{\xi,j_k}.$$

Iterative application of (7.21) gives

$$\begin{aligned} \lambda_\xi \, \eta_{\xi,\ell+n}^2 \le q_{\mathrm{ctr},\xi} \, \Delta_{\xi,j_k} \overset{(7.21)}{\le} \cdots \overset{(7.21)}{\le} q_{\mathrm{ctr},\xi}^k \, \Delta_{\xi,j_1} &\overset{(7.24)}{\le} q_{\mathrm{ctr},\xi}^k \, C'_{\mathrm{mon},\xi} \, \Delta_{\xi,\ell} \\ &= \quad q_{\mathrm{ctr},\xi}^k \, C'_{\mathrm{mon},\xi} \, (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) \, \eta_{\xi,\ell}^2. \end{aligned}$$

Hence, we obtain (7.23) with $C_{\mathrm{lin},\xi} = \lambda_\xi^{-1} \, C'_{\mathrm{mon},\xi} \, (C_{\mathrm{rel},\xi}^2 + \lambda_\xi) = (C_{\mathrm{rel},\xi}^2 + \lambda_\xi)^2 \, C_{\mathrm{mon},\xi} \, \lambda_\xi^{-2}$. This concludes the proof. $\square$

## 7.4 Proof of Theorem 7.5 for Algorithm 7.1

Throughout this section and for the ease of presentation, we suppose without loss of generality that $\eta_{\phi,\ell}$ and $\eta_{\psi,\ell}$ satisfy (E1)–(E4) even with the same constants. Then, for $\xi \in \{\phi, \psi\}$ and $\mathcal{T}_\circ, \mathcal{T}_\bullet \in \mathbb{T}$ such that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, Lemma 4.8 implies quasi-monotonicity $\eta_{\xi,\bullet} \le C_{\mathrm{mon}} \, \eta_{\xi,\circ}$ even with the same constant. With the help of Proposition 7.8, we first prove linear convergence.

***Proof of linear convergence* (7.19) *of Algorithm 7.1.*** Recall the notation of Algorithm 7.1. In each adaptive step, the set $\mathcal{M}'_\ell$ and hence $\mathcal{M}_\ell$ satisfies either the Dörfler marking (7.15) for $\eta_{\phi,\ell}$ or (7.16) for $\eta_{\psi,\ell}$. With $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \backslash \mathcal{T}_{\ell+1}$, this guarantees that within $n$ successive steps $j = \ell, \dots, \ell+n$ of the adaptive algorithm, $\mathcal{T}_j \backslash \mathcal{T}_{j+1}$ satisfies

- $k$-times the Dörfler marking for $\eta_{\phi,\ell}$ and
- $(n-k)$-times the Dörfler marking for $\eta_{\psi,\ell}$.

Define $\ell_{\mathrm{lin}} := \max\{\ell_{2,\phi}, \ell_{2,\psi}, \ell_{3,\phi}, \ell_{3,\psi}\}$. Then, Proposition 7.8 implies that

$$\eta_{\phi,\ell+n}^2 \le C_{\mathrm{lin},\phi} \, q_{\mathrm{ctr},\phi}^k \, \eta_{\phi,\ell}^2 \quad \text{as well as} \quad \eta_{\psi,\ell+n}^2 \le C_{\mathrm{lin},\psi} \, q_{\mathrm{ctr},\psi}^{n-k} \, \eta_{\psi,\ell}^2 \quad \text{for all } \ell \ge \ell_{\mathrm{lin}}.$$

Altogether, with $C_{\mathrm{lin}} := \max\{C_{\mathrm{lin},\phi}, C_{\mathrm{lin},\psi}\}$ and $q_{\mathrm{lin}}^2 := \max\{q_{\mathrm{ctr},\phi}, q_{\mathrm{ctr},\psi}\}$ this proves that

$$\eta_{\phi,\ell+n}^2 \, \eta_{\psi,\ell+n}^2 \le C_{\mathrm{lin}}^2 \, q_{\mathrm{lin}}^{2n} \, \eta_{\phi,\ell}^2 \, \eta_{\psi,\ell}^2 \quad \text{for all } \ell \ge \ell_{\mathrm{lin}}$$

and concludes the proof of (7.19). $\square$

The proof of optimal convergence rates (7.20) in Theorem 7.5 is similar to Section 4.8, but we additionally have to deal with the product structure $\eta_{\phi,\bullet}\,\eta_{\psi,\bullet}$ of the underlying error indicator. We obtain the following two technical lemmas, which are an analogon of Lemma 4.23 in the current setting.

**Lemma 7.9.** *Under the assumptions of Theorem 7.5, let $\ell_5$ be the index from Lemma 4.17 and $0 < \kappa < 1$. There exist a refinement $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\ell)$ for all $\ell \geq \ell_5$ such that the following holds: For all $s, t > 0$ with $\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} < \infty$, it holds*

$$\#\mathcal{T}_\circ - \#\mathcal{T}_\ell \leq 2\,\big(C_{\rm mon}^2\,\kappa^{-1/2}\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}\big)^{1/(s+t)}\,(\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)} \tag{7.25}$$

*as well as*

$$\eta_{\phi,\circ}^2\,\eta_{\psi,\circ}^2 \leq \kappa\,\eta_{\phi,\ell}^2\,\eta_{\psi,\ell}^2. \tag{7.26}$$

*Proof.* Recall the notation of Lemma 4.17. There exists $m_{\rm unif} \in \mathbb{N}_0$ and an index $\ell_5 > 0$ such that $\mathcal{T}_\ell \in \mathrm{refine}(\widehat{\mathcal{T}}_0)$ for all $\ell \geq \ell_5$, where $\widehat{\mathcal{T}}_0$ denotes the $m_{\rm unif}$-times uniform refinement of $\mathcal{T}_0$. Further, any refinement $\mathcal{T}_\bullet \in \mathrm{refine}(\widehat{\mathcal{T}}_{\ell_5})$, admits unique solutions $\Phi_\bullet, \Psi_\bullet \in \mathcal{P}^p(\mathcal{T}_\bullet)$ of the primal and dual problem. Hence, there holds $\mathbb{T}_N(\mathcal{T}_{\ell_5}) \neq \emptyset$ for all $\mathbb{N} \in \mathbb{N}_0$. We split the remainder of the proof into three steps.

**Step 1: Construction of the mesh $\mathcal{T}_\circ$.** Let $\ell \geq \ell_5$, $0 < \kappa < 1$ and define $\varepsilon := C_{\rm mon}^{-2}\,\kappa^{1/2}\,\eta_{\phi,\ell}\,\eta_{\psi,\ell}$. Then quasi-monotonicity implies that

$$\varepsilon \leq \kappa^{1/2}\,\eta_{\phi,\ell_5}\,\eta_{\psi,\ell_5} \leq \|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} < \infty.$$

Choose the minimal $N \in \mathbb{N}_0$ with $\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} \leq \varepsilon\,(N+1)^{s+t}$. Since $\mathbb{T}_N(\mathcal{T}_{\ell_5}) \neq \emptyset$, choose $\mathcal{T}_{\varepsilon_1}, \mathcal{T}_{\varepsilon_2} \in \mathbb{T}_N(\mathcal{T}_{\ell_5})$ with

$$\eta_{\phi,\varepsilon_1} = \min_{\mathcal{T}_\star \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_{\phi,\star} \quad \text{and} \quad \eta_{\psi,\varepsilon_2} = \min_{\mathcal{T}_\star \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_{\psi,\star}. \tag{7.27}$$

Define $\mathcal{T}_\varepsilon := \mathcal{T}_{\varepsilon_1} \oplus \mathcal{T}_{\varepsilon_2}$ as well as $\mathcal{T}_\circ := \mathcal{T}_\varepsilon \oplus \mathcal{T}_\ell$. Note that $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_{\ell_5})$ and hence, the corresponding Galerkin solutions $\Xi_\circ \in \mathcal{P}^p(\mathcal{T}_\circ)$ exists.

**Step 2: Proof of (7.26).** Quasi-monotonicity, the definition of the approximation classes (7.18), and minimality of $N$ give

$$\begin{aligned} \eta_{\phi,\circ}\,\eta_{\psi,\circ} \leq C_{\rm mon}^2\,\eta_{\phi,\varepsilon_1}\,\eta_{\psi,\varepsilon_2} &\overset{(7.27)}{=} C_{\rm mon}^2 \min_{\mathcal{T}_\star \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_{\phi,\star} \min_{\mathcal{T}_\star \in \mathbb{T}_N(\mathcal{T}_{\ell_5})} \eta_{\psi,\star} \\ &\overset{(7.18)}{\leq} C_{\rm mon}^2\,(N+1)^{-(s+t)}\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} \\ &\leq C_{\rm mon}^2\,\varepsilon = \kappa^{1/2}\,\eta_{\phi,\ell}\,\eta_{\psi,\ell}. \end{aligned} \tag{7.28}$$

This concludes the proof of (7.26).

**Step 3: Proof of (7.25).** Application of the overlay estimate (R4) for $\mathcal{T}_\circ$ as well as $\mathcal{T}_\varepsilon$ yields that

$$\#\mathcal{T}_\circ - \#\mathcal{T}_\ell \overset{\rm (R4)}{\leq} \#\mathcal{T}_\varepsilon - \#\mathcal{T}_{\ell_5} \overset{\rm (R4)}{\leq} \#\mathcal{T}_{\varepsilon_1} + \#\mathcal{T}_{\varepsilon_2} - 2\,\#\mathcal{T}_{\ell_5} \leq 2N. \tag{7.29}$$

Since $N$ is minimal, it holds that

$$\begin{aligned} N &< \big(\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}\big)^{1/(s+t)}\varepsilon^{-1/(s+t)} \\ &= \big(C_{\mathrm{mon}}^2\,\kappa^{-1/2}\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}\big)^{1/(s+t)}\,\big(\eta_{\phi,\ell}\,\eta_{\psi,\ell}\big)^{-1/(s+t)}. \end{aligned} \tag{7.30}$$

Combining the estimates (7.29) and (7.30), we see that

$$\#\mathcal{T}_\circ - \#\mathcal{T}_\ell \le 2N \le 2\,\big(C_{\mathrm{mon}}^2\kappa^{-1/2}\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}\big)^{1/(s+t)}\,\big(\eta_{\phi,\ell}\eta_{\psi,\ell}\big)^{-1/(s+t)}.$$

This concludes the proof. □

**Lemma 7.10.** *Under the assumptions of Theorem 7.5. Let $\ell_5$ be the index from Lemma 4.17 and let $0 < \theta < \theta_\star := (1 + C_{\mathrm{stb}}^2 C_{\mathrm{rel}}^2/\widehat{\beta}_0^2)^{-1}$. There exist constants $C_1, C_2 > 0$ such that the following holds: For all $s, t > 0$ with $\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} < \infty$, the set of marked elements $\mathcal{M}_\ell$ generated by Algorithm 7.1 satisfies that*

$$\#\mathcal{M}_\ell \le C_1\,(C_2\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})})^{1/(s+t)}\,(\eta_{\phi,\ell}\eta_{\psi,\ell})^{-1/(s+t)}. \tag{7.31}$$

*The constant $C_1, C_2$ depend only on $\theta$, $\widehat{\beta}_0$, and* (E1)–(E4).

*Proof.* Adopt the notation of Lemma 4.22 and Lemma 7.9. Choose $\kappa := \kappa_{\mathrm{opt}}^4$ in Lemma 7.9. Then, equation (7.26) shows that the constructed mesh $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\ell)$ satisfies $\eta_{\phi,\circ}^2\,\eta_{\psi,\circ}^2 \le \kappa_{\mathrm{opt}}^4\,\eta_{\phi,\ell}^2\,\eta_{\psi,\ell}^2$. This implies that

$$\eta_{\phi,\circ}^2 \le \kappa_{\mathrm{opt}}^2\,\eta_{\phi,\ell}^2 \quad \text{or} \quad \eta_{\psi,\circ}^2 \le \kappa_{\mathrm{opt}}^2\,\eta_{\psi,\ell}^2.$$

According to Lemma 4.22, this already implies Dörfler marking for either $\eta_{\phi,\ell}$ and $\mathcal{R}_{\phi,\ell,\circ}$ or $\eta_{\psi,\ell}$ and $\mathcal{R}_{\psi,\ell,\circ}$, where $\mathcal{R}_{\xi,\ell,\circ} \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_\circ$ denotes the extended set of refined elements from discrete reliability (E4). Recall, that the expanded Dörfler marking strategy guarantees $\#\mathcal{M}_\ell \le C_{\mathrm{mark}}\,\#\mathcal{M}_\ell'$. Then, minimality of $\mathcal{M}_\ell'$ and the splitting property (R3) imply that

$$\begin{aligned} \#\mathcal{M}_\ell \le C_{\mathrm{mark}}\,\#\mathcal{M}_\ell' \le 2C_{\mathrm{mark}}\,\min\{\#\mathcal{M}_{\phi,\ell}\,,\,\#\mathcal{M}_{\psi,\ell}\} &\le 2C_{\mathrm{mark}}\,\max\{\#\mathcal{R}_{\phi,\ell,\circ}\,,\,\#\mathcal{R}_{\psi,\ell,\circ}\} \\ &\le 2C_{\mathrm{mark}}\,C_{\mathrm{rel}}\,\#(\mathcal{T}_\ell \setminus \mathcal{T}_\circ) \\ &\overset{\text{(R3)}}{\le} 2C_{\mathrm{mark}}\,C_{\mathrm{rel}}\,(\#\mathcal{T}_\ell - \#\mathcal{T}_\circ). \end{aligned}$$

In combination with Lemma 7.9, we obtain that

$$\#\mathcal{M}_\ell \le 2C_{\mathrm{mark}}\,C_{\mathrm{rel}}\,(\#\mathcal{T}_\ell - \#\mathcal{T}_\circ) \le C_1\,(C_2\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}\,\big(\eta_{\phi,\ell}\,\eta_{\psi,\ell}\big)^{-1/(s+t)},$$

where $C_1 = 4C_{\mathrm{mark}}C_{\mathrm{rel}}$ and $C_2 = C_{\mathrm{mon}}^2\kappa_{\mathrm{opt}}^{-2}$. This concludes the proof. □

With the help of Lemma 7.10, the proof of optimal convergence follows analogously to the proof of Theorem 4.21. For the sake of completeness, we recap the important steps.

***Proof of optimal convergence* (7.20) *of Algorithm 7.1.*** Let $\ell_{\rm opt} := \max\{\ell_{\rm lin}, \ell_5\}$. Analoguosly to the proof of Theorem 4.21, for all $\ell \geq \ell_{\rm opt}$ the mesh-closure estimate (R5) and Lemma 7.10 imply that

$$\begin{aligned} \#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 &\overset{\rm (R5)}{\leq} C_{\rm mesh}\,(\ell_6\,C + 1) \sum_{j=\ell_{\rm opt}}^{\ell-1} \#\mathcal{M}_j \\ &\overset{(7.25)}{\leq} C_{\rm mesh}\,(\ell_6\,C + 1) C_1\,(C_2 \|\phi\|_{\mathbb{A}_s} \|\psi\|_{\mathbb{A}_t})^{1/(s+t)} \sum_{j=\ell_{\rm opt}}^{\ell-1} (\eta_{\phi,j}\,\eta_{\psi,j})^{-1/(s+t)}, \end{aligned} \tag{7.32}$$

where $C := \max_{j=0,\dots,\ell_{\rm opt}} \frac{\#\mathcal{M}_j}{\#\mathcal{M}_{\ell_{\rm opt}}}$. Then, linear convergence (7.19) yields that

$$\eta_{\phi,\ell}\,\eta_{\psi,\ell} \leq C_{\rm lin}\, q_{\rm lin}^{\ell-j}\,\eta_{\phi,j}\,\eta_{\psi,j} \quad \text{for all } \ell_{\rm lin} \leq j \leq \ell.$$

Hence,

$$(\eta_{\phi,j}\,\eta_{\psi,j})^{-1/(s+t)} \leq C_{\rm lin}^{1/(s+t)} q_{\rm lin}^{(\ell-j)/(s+t)} (\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)} \quad \text{for all } \ell_{\rm opt} \leq j \leq \ell.$$

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

$$\begin{aligned} \sum_{j=\ell_{\rm opt}}^{\ell-1} (\eta_{\phi,j}\,\eta_{\psi,j})^{-1/(s+t)} &\leq C_{\rm lin}^{1/(s+t)} (\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)} \sum_{j=\ell_{\rm opt}}^{\ell-1} q^{\ell-j} \\ &\leq \frac{C_{\rm lin}^{1/(s+t)}}{1 - q_{\rm lin}^{1/(s+t)}}\,(\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)}. \end{aligned}$$

Combining the latter estimate with (7.32), we prove that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \leq \frac{C_{\rm mesh}(\ell_6\,C + 1) C_1}{1 - q_{\rm lin}^{1/(s+t)}}\,(C_{\rm lin} C_2\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} \|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})})^{1/(s+t)}\,(\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)}.$$

Rearranging the terms in the estimate above, we see $\eta_{\phi,\ell}\,\eta_{\psi,\ell} \lesssim (\#\mathcal{T}_\ell - \#\mathcal{T}_{\ell_0} + 1)^{-(s+t)}$. Using the definitions of $C_1, C_2 > 0$, this implies (7.20) with

$$C_{\rm opt} := \left( \frac{2\#\mathcal{T}_{\ell_{\rm opt}}\,C_{\rm mesh}\,(\ell_6\,C + 1)\,C_{\rm mark}^2 C_{\rm rel}}{1 - q_{\rm lin}^{1/(s+t)}} \right)^{(s+t)} C_{\rm lin}\,C_{\rm mon}^2\,\kappa_{\rm opt}^{-2}\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})}.$$

Finally, recall from Lemma 4.18 that $\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})} \|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} < \infty$ if and only if $\|\phi\|_{\mathbb{A}_s} \|\psi\|_{\mathbb{A}_t} < \infty$. This concludes the proof. $\square$

## 7.5 Proof of Theorem 7.5 for Algorithm 7.3

As in the previous section, we suppose that $\eta_{\phi,\ell}$ and $\eta_{\psi,\ell}$ satisfy the properties (E1)–(E4) even with the same constants. We start with the proof of linear convergence.

***Proof of linear convergence* (7.19) *of Algorithm 7.3.*** Suppose that $\ell \geq \ell_{\mathrm{lin}} := \max\{\ell_{\phi,3}, \ell_{\psi,3}\}$ and hence the discrete solutions $\xi_\ell \in \mathcal{P}^p(\mathcal{T}_\ell)$ and corresponding error estimators $\eta_{\xi,\ell}$ exist. Recall the definition of the combined error indicator $\rho_\ell(T)^2 := \eta_{\phi,\ell}(T)^2\eta_{\psi,\ell}^2 + \eta_{\phi,\ell}^2\eta_{\psi,\ell}(T)^2$. Summing over all elements $T \in \mathcal{T}_\ell$, we see that

$$\rho_\ell^2 = \eta_{\psi,\ell}^2 \sum_{T\in\mathcal{T}_\ell} \eta_{\phi,\ell}(T)^2 + \eta_{\phi,\ell}^2 \sum_{T\in\mathcal{T}_\ell} \eta_{\psi,\ell}(T)^2 = 2\,\eta_{\phi,\ell}^2\,\eta_{\psi,\ell}^2.$$

Therefore, the Dörfler marking criterion (7.17) of Algorithm 7.3 reads as

$$2\,\theta\,\eta_{\phi,\ell}^2\eta_{\psi,\ell}^2 = \theta\,\rho_\ell^2 \leq \rho_\ell(\mathcal{M}_\ell)^2 \leq \eta_{\phi,\ell}(\mathcal{M}_\ell)^2\,\eta_{\psi,\ell}^2 + \eta_{\phi,\ell}^2\,\eta_{\psi,\ell}(\mathcal{M}_\ell)^2.$$

In particular, this shows that (7.17) implies that

$$\theta\,\eta_{\phi,\ell}^2 \leq \eta_{\phi,\ell}(\mathcal{M}_\ell)^2 \quad \text{or} \quad \theta\,\eta_{\psi,\ell}^2 \leq \eta_{\psi,\ell}(\mathcal{M}_\ell)^2.$$

Hence, in each adaptive step, either Dörfler marking for the primal or the dual problem is satisfied. As we have seen in the proof of linear convergence for Algorithm 7.1 with separate marking, this already implies linear convergence

$$\eta_{\phi,\ell+n}^2\,\eta_{\psi,\ell+n}^2 \leq C_{\mathrm{lin}}^2\,q_{\mathrm{lin}}^n\,\eta_{\phi,\ell}^2\,\eta_{\psi,\ell}^2 \quad \text{for all } \ell \geq \ell_{\mathrm{lin}},$$

and concludes (7.19) for Algorithm 7.3. ☐

The proof of optimal convergence rates (7.20) for Algorithm 7.3 is essentially a consequence of the following elementary observation.

**Lemma 7.11.** *Let $0 < \theta \leq 1/2$ and $\ell \in \mathbb{N}_0$ such that $\mathcal{T}_\ell$ admits unique Galerkin solutions $\Xi_\ell$. Suppose $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$ such that*

$$2\theta\,\eta_{\phi,\ell}^2 \leq \eta_{\phi,\ell}(\mathcal{R}_\ell)^2 \quad \textit{or} \quad 2\theta\,\eta_{\psi,\ell}^2 \leq \eta_{\psi,\ell}(\mathcal{R}_\ell)^2.$$

*Then, the combined indicator $\rho_\ell$ satisfies that,*

$$\theta\,\rho_\ell^2 \leq \rho_\ell(\mathcal{R}_\ell)^2, \tag{7.33}$$

*i.e., if $\mathcal{R}_\ell$ satisfies Dörfler marking for the primal or dual problem with parameter $2\theta$, then $\mathcal{R}_\ell$ satisfies the combined Dörfler marking with parameter $\theta$.*

*Proof.* Using the definition of $\rho_\ell$ in Step (iv) of Algorithm 7.3, an elementary calculation yields that

$$\theta\,\rho_\ell^2 = 2\theta\,\eta_{\phi,\ell}^2\,\eta_{\psi,\ell}^2 \leq \eta_{\phi,\ell}(\mathcal{R}_\ell)^2\,\eta_{\psi,\ell}^2 + \eta_{\phi,\ell}^2\,\eta_{\psi,\ell}(\mathcal{R}_\ell)^2 = \rho_\ell(\mathcal{R}_\ell)^2. \tag{7.34}$$

This concludes the proof. ☐

Note that, Lemma 7.9 does not rely on the specific marking strategy and holds for the present setting as well. Then, the key idea of Lemma 7.11 yields that the upper bound $\theta_{\mathrm{opt}}$ for optimal marking parameters in Algorithm 7.3 is half that for optimal marking parameter in Algorithm 7.1.

**Lemma 7.12.** *Under the assumptions of Theorem 7.5. Let $\ell_5$ be the index from Lemma 4.17 and let $0 < \theta < \theta_{\rm opt}/2 := 1/2\,(1 + C_{\rm stb}^2 C_{\rm rel}^2/\widehat{\beta}_0^2)^{-1}$. There exist constants $C_1, C_2 > 0$ such that the following holds: For all $s, t > 0$ with $\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})} < \infty$, the set of marked elements $\mathcal{M}_\ell$ generated by Algorithm 7.3 satisfies*

$$\#\mathcal{M}_\ell \le C_1\,(C_2\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})})^{1/(s+t)}\,(\eta_{\phi,\ell}\eta_{\psi,\ell})^{-1/(s+t)}. \tag{7.35}$$

*The constant $C_1, C_2$ depends only on $\theta$, $\widehat{\beta}_0$, and* (E1)–(E4).

*Proof.* Adopt the notation of Lemma 4.22 and Lemma 7.9. Apply Lemma 4.22 with $0 < 2\,\theta < \theta_{\rm opt}$ to obtain $\kappa_{\rm opt}$ and choose $\kappa := \kappa_{\rm opt}^4$ in Lemma 7.9. Then, equation (7.28) shows that the constructed mesh $\mathcal{T}_\circ \in \text{refine}(\mathcal{T}_{\ell_5})$ satisfies $\eta_{\phi,\circ}^2\eta_{\psi,\circ}^2 \le \kappa_{\rm opt}^4\,\eta_{\phi,\ell}^2\eta_{\psi,\ell}^2$.

This implies that $\eta_{\phi,\circ}^2 \le \kappa_{\rm opt}^2\,\eta_{\phi,\ell}^2$ or $\eta_{\psi,\circ}^2 \le \kappa_{\rm opt}^2\,\eta_{\psi,\ell}^2$. Hence, according to Lemma 4.22, there holds Dörfler marking

$$2\,\theta\,\eta_{\phi,\ell}^2 \le \eta_{\phi,\ell}(\mathcal{R}_{\phi,\ell,\circ})^2 \quad \text{or} \quad 2\,\theta\,\eta_{\psi,\ell}^2 \le \eta_{\psi,\ell}(\mathcal{R}_{\psi,\ell,\circ})^2.$$

Therefore, Lemma 7.11 implies Dörfler marking (7.33), i.e.,

$$\theta\,\rho_\ell^2 \le \rho_\ell(\mathcal{R}_{\phi,\ell,\circ})^2 \quad \text{or} \quad \theta\,\rho_\ell^2 \le \rho_\ell(\mathcal{R}_{\psi,\ell,\circ})^2$$

Similarly to the proof of Lemma 7.10, this implies that

$$\#\mathcal{M}_\ell \le C_{\rm mark}\,\#\mathcal{M}_\ell' \le 2C_{\rm mark}\,\max\{\#\mathcal{R}_{\phi,\ell,\circ}\,,\,\#\mathcal{R}_{\psi,\ell,\circ}\} \le C_{\rm mark}\,C_{\rm rel}\,\#(\mathcal{T}_\ell \setminus \mathcal{T}_\circ).$$

Note that Lemma 7.9 holds independently of the marking strategy. With the splitting property (R3), we conclude that

$$\#(\mathcal{T}_\ell \setminus \mathcal{T}_\circ) \overset{\rm (R3)}{\le} \#\mathcal{T}_\circ - \#\mathcal{T}_\ell \le C_1\,(C_2\,\|\phi\|_{\mathbb{A}_s(\mathcal{T}_{\ell_5})}\,\|\psi\|_{\mathbb{A}_t(\mathcal{T}_{\ell_5})})^{1/(s+t)}\,(\eta_{\phi,\ell}\,\eta_{\psi,\ell})^{-1/(s+t)},$$

with constants $C_1 = 2\,C_{\rm mark}^2 C_{\rm rel}$ and $C_2 = C_{\rm mon}^2\kappa_{\rm opt}^{-2}$. □

***Proof of optimal convergence* (7.20) *of Algorithm 7.3.*** Since that proof of (7.20) in case of Algorithm 7.1 essentially relies only on the validity of Lemma 7.10, the proof of (7.20) for Algorithm 7.3 follows verbatim, where Lemma 7.10 is replaced by Lemma 7.12. □

# 8 Abstract theory on strongly monotone nonlinear operators

## 8.1 State of the art and outline

As for linear problems, the analysis of convergence and optimal convergence behavior of AFEM for nonlinear problems has been a fertile field for new publications. We refer to [Vee02, DK08, BDK12, GMZ12] for some major contributions and to [CFPP14] for some general abstract framework. While the interplay of adaptive algorithms, optimal convergence rates, and inexact solvers is already well understood and analyzed, e.g., in [Ste07, AGL13, ALMS13, CFPP14] for linear PDEs and in [CG12] for eigenvalue problems, the influence of inexact solvers for nonlinear equations has not been analyzed yet. The work [GMZ11] considers adaptive mesh-refinement in combination with a Kačanov-type iterative solver for strongly monotone operators. Following [MSV08, Sie11], the work focuses on plain convergence, whereas the proof of optimal convergence rates remains open.

On the other hand, there is a rich body on *a posteriori* error estimation which also includes the iterative and inexact solution for nonlinear problems; see, e.g., [EV13]. Based on our own work [GHPS17], we aim to close this gap between numerical analysis (e.g., [CFPP14]) and empirical evidence of optimal convergence rates (e.g., [GMZ11, EV13]) by analyzing an adaptive algorithm from [CW17].

We consider nonlinear elliptic equations in their variational formulation: Given $F \in \mathcal{H}^*$, find $u^\star \in \mathcal{H}$ such that

$$\langle \mathfrak{A}\, u^\star \,,\, v\rangle = \langle F \,,\, v\rangle \quad \text{for all } v \in \mathcal{H}, \tag{8.1}$$

where $\mathfrak{A}$ is a strongly monotone (A1) and Lipschitz continuous (A2) operator. In view of applications, we admit that (A1)–(A2) exclude the $p$-Laplacian [Vee02, DK08, BDK12], but cover the same problem class as, e.g., [CW17, GMZ11, GMZ12]. We refer also to [BSF$^+$14] for strongly monotone nonlinearities arising in magnetostatics.

The discrete formulation of (8.1) reads: Find $u_\ell^\star \in \mathcal{X}_\ell$ such that

$$\langle \mathfrak{A} u_\ell^\star \,,\, v_\ell\rangle = \langle F \,,\, v_\ell\rangle \quad \text{for all } v_\ell \in \mathcal{X}_\ell. \tag{8.2}$$

We emphasize that the discrete nonlinear system (8.2) cannot be solved exactly in practice. Instead, [CW17, GHPS17] use a Picard approximate $u_\ell^n := \Phi_\ell(u_\ell^{n-1}) \approx u_\ell^\star$, where $u_\ell^\star \in \mathcal{X}_\ell$ is the exact solution of (8.2) and the involved nonlinear mapping $\Phi : \mathcal{H} \to \mathcal{H}$ is a contraction (see Section 8.2 for details). Unlike [GMZ12, BDK12], there holds $u_\ell^n \neq u_\ell^\star$ in general.

The computation of each Picard step requires to solve a discrete Laplace problem. In contrast to [GHPS17, CW17], we do not assume that these arising linear systems are solved exactly and use an inexact iterative PCG-solver instead. Therefore, the proposed algorithm steers not only the local mesh-refinement and the Picard iteration, but also the inexact solver for the invoked linear system, where we employ *nested iteration* to lower the number of iterative steps.

The algorithm generates a sequence of conforming nested subspaces $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{H}$, corresponding discrete solutions $u_\ell^{n,k} \in \mathcal{X}_\ell$, and estimators $\eta_\ell(u_\ell^{n,k})$ such that $\|u^\star - u_\ell^{n,k}\|_\mathcal{H} \leq C_{\rm rel}^\star \, \eta_\ell(u_\ell^{n,k}) \to 0$ as $\ell \to \infty$ with optimal algebraic rate in the sense of certain approximation classes [CKNS08, FFP14, CFPP14]. Moreover, under an additional assumption, we prove optimal convergence rates with respect to the cumulative computational costs, which implies optimal computational complexity and improves the existing result of [GHPS17]. While the plain convergence from [GMZ11] applies to various marking strategies, the convergence analysis in this chapter is tailored to the Dörfler marking strategy. We emphasize that the whole analysis is done in a complete abstract setting in the spirit of [CFPP14].

**Outline of tis chapter.** Section 8.2 recalls the well-known proof that (8.1) admits a unique solution. Section 8.3 comments on the discrete problem (8.2) and introduces the corresponding Picard mapping $\Phi_\ell$. Section 8.3.2 introduces the PCG-solver for the arising linear system. The adaptive strategy (Algorithm 8.7) is given in Section 8.5. There, we also formulate the estimator axioms in the current setting. Linear convergence of the proposed algorithm is proved in Section 8.8 (Theorem 8.20). Optimal algebraic convergence behavior, is proved in Section 8.9 (Theorem 8.21). As a consequence of the preceding results and with an additional assumption, we also obtain optimal computational complexity (Theorem 8.32), i.e., the error estimator converges optimal with respect to the cumulative computational effort.

## 8.2 Abstract setting

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with $d \geq 2$. Further, let $\mathcal{H}$ be a Hilbert space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with dual space $\mathcal{H}^*$. The $\mathcal{H}$-scalar product is given by $(\cdot\,,\,\cdot)_\mathcal{H}$ . Let $\langle \cdot\,,\,\cdot \rangle$ denote the corresponding duality bracket. We consider nonlinear elliptic equations in the following abstract setting with variational formulation: Given $F \in \mathcal{H}^*$, find $u^\star \in \mathcal{H}$ such that

$$\langle \mathfrak{A}\, u^\star\,,\, v\rangle = \langle F\,,\, v\rangle \quad \text{for all } v \in \mathcal{H}. \tag{8.3}$$

In order to guarantee solvability, we additionally suppose that the operator $\mathfrak{A} : \mathcal{H} \to \mathcal{H}^*$ satisfies the following conditions:

**A1) $\mathfrak{A}$ is strongly monotone:** There exists $\alpha > 0$ such that

$$\alpha\, \|w - v\|_\mathcal{H}^2 \leq \operatorname{Re} \langle \mathfrak{A}w - \mathfrak{A}v\,,\, w - v\rangle \quad \text{for all } v, w \in \mathcal{H}.$$

**A2) $\mathfrak{A}$ is Lipschitz continuous:** There exists $L > 0$ such that

$$\|\mathfrak{A}w - \mathfrak{A}v\|_{\mathcal{H}^*} \leq L\, \|w - v\|_\mathcal{H} \quad \text{for all } v, w \in \mathcal{H}.$$

**A3) $\mathfrak{A}$ has a potential:** There exists a Gâteaux differentiable function $P : \mathcal{H} \to \mathbb{K}$ such that its derivative $\mathrm{d}P : \mathcal{H} \to \mathcal{H}^*$ coincides with $\mathfrak{A}$, i.e., for all $v, w \in \mathcal{H}$, it holds that

$$\langle \mathfrak{A} w \,,\, v \rangle = \langle \mathrm{d}P(w) \,,\, v \rangle = \lim_{\substack{r \to 0 \\ r \in \mathbb{R}}} \frac{P(w + rv) - P(w)}{r}. \tag{8.4}$$

We note that (A1)–(A2) are sufficient to guarantee existence and uniqueness of the solution $u^\star \in \mathcal{H}$ of (8.3) (see Section 8.2.2). On the other hand, Assumption (A3) is used to prove linear convergence in Section 8.8.

### 8.2.1 Nonlinear discrete problem

Let $\mathcal{T}_0$ be a given regular initial mesh. Suppose that refine$(\cdot)$ is a fixed refinement strategy satisfying the axioms (R1)–(R6); see Chapter 3. For each $\mathcal{T}_\bullet \in \mathbb{T} = \mathrm{refine}(\mathcal{T}_0)$, let $\mathcal{X}_\bullet \subset \mathcal{H}$ denote the related conforming finite-dimensional subspace of $\mathcal{H}$. Further, suppose that refinement $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$ leads to nestedness $\mathcal{X}_\bullet \subset \mathcal{X}_\circ$ of the corresponding subspaces. Then, the discrete formulation of (8.3) reads as: Find $u_\bullet^\star \in \mathcal{X}_\bullet$ such that

$$\langle \mathfrak{A} u_\bullet^\star \,,\, v_\bullet \rangle = \langle F \,,\, v_\bullet \rangle \quad \text{for all } v_\bullet \in \mathcal{X}_\bullet. \tag{8.5}$$

Note that, if (A1)–(A2) are satisfied, the restriction $\mathfrak{A}_\bullet : \mathcal{X}_\bullet \to \mathcal{X}_\bullet^*$ of $\mathfrak{A}$ is also strongly monotone and Lipschitz continuous, even with the same constants $\alpha, L > 0$ as in the continuous case.

### 8.2.2 Existence of solutions

In this section, we prove that the model problem (8.3) as well as its discrete version (8.5) admit unique solutions $u^\star \in \mathcal{H}$ and $u_\bullet^\star \in \mathcal{X}_\bullet$. The proof follows essentially from the Banach fixpoint theorem and relies only on the validity of (A1)–(A2). To that end, recall the Riesz mapping

$$I_{\mathcal{H}} : \mathcal{H} \to \mathcal{H}^* \quad \text{with} \quad I_{\mathcal{H}} w := (\cdot \,,\, w)_{\mathcal{H}}.$$

We emphasize that $I_{\mathcal{H}}$ is (up to complex conjugation) an isometric isomorphism; see e.g, [Yos80, Chapter III.6]. Further, the Banach–Picard iteration $\Phi : \mathcal{H} \to \mathcal{H}$ is given by

$$\Phi(v) := v - (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - F). \tag{8.6}$$

We obtain the following proposition.

**Proposition 8.1.** *Let $\mathfrak{A}$ satisfy (A1)–(A2) with constants $0 < \alpha, L \le \infty$. Then, $\Phi : \mathcal{H} \to \mathcal{H}$ defined in (8.6) is a contraction with Lipschitz constant*

$$0 \le q_{\mathrm{pic}} := (1 - \alpha^2/L^2)^{1/2} < 1. \tag{8.7}$$

*There exist unique solutions $u^\star \in \mathcal{H}$ to (8.3) as well as $u_\bullet^\star \in \mathcal{X}_\bullet$ to (8.5). Moreover, for any initial guess $u^0 \in \mathcal{H}$, the Picard iteration $u^{n,\star} := \Phi(u^{n-1,\star})$ converges to $u^\star$ as $n \to \infty$ with*

$$\|u^\star - u^{n,\star}\|_{\mathcal{H}} \le \frac{q_{\rm pic}}{1-q_{\rm pic}} \|u^{n,\star} - u^{n-1,\star}\|_{\mathcal{H}} \le \frac{q_{\rm pic}^n}{1-q_{\rm pic}} \|u^{1,\star} - u^0\|_{\mathcal{H}} \tag{8.8a}$$

*as well as*

$$\|u^{n,\star} - u^{n-1,\star}\|_{\mathcal{H}} \le (1+q_{\rm pic}) \|u^\star - u^{n-1,\star}\|_{\mathcal{H}} \tag{8.8b}$$

*Proof.* We split the proof into two steps.

**Step 1: Existence of solutions.** Let $v, w \in \mathcal{H}$. Note that (A1)–(A2) immediately imply that

$$\alpha \|v-w\|_{\mathcal{H}}^2 \overset{\text{(A1)}}{\le} \operatorname{Re} \langle \mathfrak{A}v - \mathfrak{A}w \,,\, v-w \rangle \le \|\mathfrak{A}v - \mathfrak{A}w\|_{\mathcal{H}^*} \, \|v-w\|_{\mathcal{H}} \overset{\text{(A2)}}{\le} L \, \|v-w\|_{\mathcal{H}}^2,$$

and hence, $\alpha \le L$ as well as $0 \le q_{\rm pic} < 1$. Using the definition of $\Phi(\cdot)$ and $\|v\|_{\mathcal{H}}^2 := (v \,,\, v)_{\mathcal{H}}$, we obtain that

$$\begin{aligned}
\|\Phi v - \Phi w\|_{\mathcal{H}}^2 &= \big\| v - (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - F) - \big(w - (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}w - F)\big)\big\|_{\mathcal{H}}^2 \\
&= \big(v - w - (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w) \,,\, v - w - (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w)\big)_{\mathcal{H}} \\
&= \|v-w\|_{\mathcal{H}}^2 - 2\,\frac{\alpha}{L^2} \operatorname{Re}\, \big(v-w \,,\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w)\big)_{\mathcal{H}} + \frac{\alpha^2}{L^4} \big\| I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w)\big\|_{\mathcal{H}}^2.
\end{aligned}$$

We treat each term on the right hand side separately. First, by definition of $I_{\mathcal{H}}$, it holds that

$$\operatorname{Re}\, \big(v-w \,,\, I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w)\big)_{\mathcal{H}} = \operatorname{Re} \langle \mathfrak{A}v - \mathfrak{A}w \,,\, v-w \rangle \overset{\text{(A1)}}{\ge} \alpha \, \|v-w\|_{\mathcal{H}}^2.$$

Second, since $I_{\mathcal{H}}$ is an isometric isomorphism, we obtain that

$$\|I_{\mathcal{H}}^{-1}(\mathfrak{A}v - \mathfrak{A}w)\|_{\mathcal{H}}^2 = \|\mathfrak{A}v - \mathfrak{A}w\|_{\mathcal{H}^*}^2 \overset{\text{(A2)}}{\le} L^2 \, \|v-w\|_{\mathcal{H}}^2.$$

Combining these observations, we see that

$$\begin{aligned}
\|\Phi v - \Phi w\|_{\mathcal{H}}^2 \; &\le \; \Big(1 + \frac{\alpha^2}{L^2}\Big) \, \|v-w\|_{\mathcal{H}}^2 - 2\,\frac{\alpha}{L^2} \operatorname{Re} \langle \mathfrak{A}v - \mathfrak{A}w \,,\, v-w \rangle \\
&\le \; \Big(1 - \frac{\alpha^2}{L^2}\Big) \, \|v-w\|_{\mathcal{H}}^2 \\
&\overset{(8.7)}{=} \; q_{\rm pic}^2 \, \|v-w\|_{\mathcal{H}}^2.
\end{aligned} \tag{8.9}$$

Hence, $\Phi : \mathcal{H} \to \mathcal{H}$ is a contraction with Lipschitz constant $0 \le q_{\rm pic} < 1$. According to the Banach fixpoint theorem, $\Phi$ has a unique fixpoint $u^\star \in \mathcal{H}$, i.e., $u^\star = \Phi(u^\star)$. By definition of $\Phi$, the fixpoint $u^\star$ satisfies

$$0 = \big(v \,,\, (\alpha/L^2)\, I_{\mathcal{H}}^{-1}(\mathfrak{A}u_\star - F)\big)_{\mathcal{H}} = \frac{\alpha}{L^2} \langle \mathfrak{A}u_\star - F \,,\, v \rangle \quad \text{for all } v \in \mathcal{H}, \tag{8.10}$$

and hence, $u^\star = \Phi(u^\star)$ is equivalent to the strong form (8.3). Overall, we conclude that (8.3) admits a unique solution. Recall that the restriction $\mathfrak{A}_\bullet : \mathcal{X}_\bullet \to \mathcal{X}_\bullet^*$ of $\mathfrak{A}$ satisfies (A1)– (A2) even with the same constants. Hence, the proof also applies to the discrete setting and yields existence and uniqueness of the solution $u_\bullet^\star \in \mathcal{X}_\bullet$ to (8.5).

**Step 2: Proof of (8.8a) and (8.8b) .** The Banach fixpoint theorem guarantees for each initial guess $u^0 \in \mathcal{H}$, that the Picard iteration $u^{n,\star} := \Phi(u^{n-1,\star})$ converges to $u^\star$ as $n \to \infty$. For all $n \in \mathbb{N}$, we obtain that

$$\begin{aligned}\|u^\star - u^{n,\star}\|_{\mathcal{H}} = \|\Phi(u^\star) - \Phi(u^{n-1,\star})\|_{\mathcal{H}} &\overset{(8.9)}{\leq} q_{\mathrm{pic}}\,\|u^\star - u^{n-1,\star}\|_{\mathcal{H}} \\ &\leq q_{\mathrm{pic}}\,\|u^\star - u^{n,\star}\|_{\mathcal{H}} + q_{\mathrm{pic}}\,\|u^{n,\star} - u^{n-1,\star}\|_{\mathcal{H}}.\end{aligned}$$

Rearranging this estimate and aguing by induction on $n$ with (8.9), we derive the following well-known *a posteriori* and *a priori* estimate for the Picard iterates,

$$\|u^\star - u^{n,\star}\|_{\mathcal{H}} \leq \frac{q_{\mathrm{pic}}}{1-q_{\mathrm{pic}}}\,\|u^{n,\star} - u^{n-1,\star}\|_{\mathcal{H}} \overset{(8.9)}{\leq} \frac{q_{\mathrm{pic}}^n}{1-q_{\mathrm{pic}}}\,\|u^{1,\star} - u^0\|_{\mathcal{H}}. \tag{8.11}$$

This concludes (8.8a). Moreover, it holds that

$$\|u^{n,\star} - u^{n-1,\star}\|_{\mathcal{H}} \leq \|u^\star - u^{n,\star}\|_{\mathcal{H}} + \|u^\star - u^{n-1,\star}\|_{\mathcal{H}} \overset{(8.9)}{\leq} (1+q_{\mathrm{pic}})\,\|u^\star - u^{n-1,\star}\|_{\mathcal{H}}. \tag{8.12}$$

This concludes (8.8b). Thus, the *a posteriori* computable term $\|u^n - u^{n-1}\|_{\mathcal{H}}$ provides an upper bound for $\|u^\star - u^n\|_{\mathcal{H}}$ as well as a lower bound for $\|u^\star - u^{n-1}\|_{\mathcal{H}}$. □

Recall the following well-known Céa-type estimate for strongly monotone operators. For the sake of completeness we include its proof.

**Lemma 8.2.** *Suppose that the operator* $\mathfrak{A}$ *satisfies* (A1)–(A2)*. Then, it holds that*

$$\|u^\star - u_\bullet^\star\|_{\mathcal{H}} \leq \frac{L}{\alpha} \min_{w_\bullet \in \mathcal{X}_\bullet} \|u^\star - w_\bullet\|_{\mathcal{H}}. \tag{8.13}$$

*Proof.* Note the Galerkin orthogonality $\langle \mathfrak{A}u^\star - \mathfrak{A}u_\bullet^\star\,,\, v_\bullet\rangle = 0$ for all $v_\bullet \in \mathcal{X}_\bullet$. For $w_\bullet \in \mathcal{X}_\bullet$ and $u^\star \neq u_\bullet^\star$, this results in

$$\begin{aligned}\alpha\,\|u^\star - u_\bullet^\star\|_{\mathcal{H}} &\overset{(A1)}{\leq} \frac{\mathrm{Re}\,\langle \mathfrak{A}u^\star - \mathfrak{A}u_\bullet^\star\,,\, u^\star - u_\bullet^\star\rangle}{\|u^\star - u_\bullet^\star\|_{\mathcal{H}}} \\ &= \frac{\mathrm{Re}\,\langle \mathfrak{A}u^\star - \mathfrak{A}u_\bullet^\star\,,\, u^\star - w_\bullet\rangle}{\|u^\star - u_\bullet^\star\|_{\mathcal{H}}} \overset{(A2)}{\leq} L\,\|u^\star - w_\bullet\|_{\mathcal{H}}.\end{aligned}$$

Finite dimension concludes that the infimum over all $w_\bullet \in \mathcal{X}_\bullet$ is, in fact, attained. □

## 8.3 Discretization and a priori error estimation

We emphasize that the nonlinear system (8.5) can hardly be solved exactly even on the discrete level. Instead, we introduce the exact discrete Picard iterates $u_\bullet^{n,\star} := \Phi(u_\bullet^{n-1,\star})$, with $u_\bullet^{n,\star} \to u_\bullet^\star$ as $n \to \infty$.

### 8.3.1 Linearized discrete problem

Let $I_\bullet : \mathcal{X}_\bullet \to \mathcal{X}_\bullet^*$ denote the discrete Riesz mapping. Define the restriction $F_\bullet \in \mathcal{X}_\bullet^*$ of $F$ to $\mathcal{X}_\bullet$. Then, the discrete Picard function is given by

$$\Phi_\bullet : \mathcal{X}_\bullet \to \mathcal{X}_\bullet \quad \text{ with } \Phi_\bullet(v_\bullet) := v_\bullet - (\alpha/L^2) I_\bullet^{-1}(\mathfrak{A}_\bullet v_\bullet - F_\bullet).$$

Given $u_\bullet^{n-1,\star} \in \mathcal{X}_\bullet$, the discrete Picard iterate $u_\bullet^{n,\star} := \Phi_\bullet(u_\bullet^{n-1,\star})$ can be computed as follows:

(P.i) Solve the linear system

$$(v_\bullet, w_\bullet^{n,\star})_\mathcal{H} = \langle \mathfrak{A} u_\bullet^{n-1,\star} - F \,,\, v_\bullet \rangle \quad \text{for all } v_\bullet \in \mathcal{X}_\bullet. \tag{8.14}$$

(P.ii) Define $u_\bullet^{n,\star} := u_\bullet^{n-1,\star} - \alpha/L^2 \, w_\bullet^{n,\star}$.

Then, Proposition 8.1 holds verbatim for $\Phi_\bullet$ instead of $\Phi$ and implies that $\Phi_\bullet$ is a contraction on $\mathcal{X}_\bullet$. For each discrete initial guess $u_\bullet^0 \in \mathcal{X}_\bullet$, the discrete Picard iteration $u_\bullet^{n+1,\star} = \Phi_\bullet(u_\bullet^{n,\star})$ converges to $u_\bullet^\star$ as $n \to \infty$. Moreover, the error estimates (8.8a)–(8.8b) also hold for the discrete Picard iteration, i.e., for all $n \in \mathbb{N}$, it holds that

$$\begin{aligned} \|u_\bullet^\star - u_\bullet^{n,\star}\|_\mathcal{H} &\le \frac{q_{\rm pic}}{1-q_{\rm pic}} \|u_\bullet^{n,\star} - u_\bullet^{n-1,\star}\|_\mathcal{H} \\ &\le \min\Big\{ \frac{q_{\rm pic}^n}{1-q_{\rm pic}} \|u_\bullet^{1,\star} - u_\bullet^0\|_\mathcal{H}, \, \frac{q_{\rm pic}(1+q_{\rm pic})}{1-q_{\rm pic}} \|u_\bullet^\star - u_\bullet^{n-1,\star}\|_\mathcal{H} \Big\}. \end{aligned} \tag{8.15}$$

To simplify the notation, and if it is clear from the context, we write $\Phi(\cdot)$ instead of $\Phi_\bullet(\cdot)$. Finally, we recall the following *a priori* estimate for the discrete Picard iteration from [CW17, Proposition 2.1]. We also include its simple proof for the sake of completeness.

**Lemma 8.3.** *Suppose that the operator* $\mathfrak{A}$ *satisfies* (A1)–(A2). *Then, it holds that*

$$\|u^\star - u_\bullet^{n,\star}\|_\mathcal{H} \le \frac{L}{\alpha} \min_{w_\bullet \in \mathcal{X}_\bullet} \|u^\star - w_\bullet\|_\mathcal{H} + \frac{q_{\rm pic}^n}{1-q_{\rm pic}} \|u_\bullet^{1,\star} - u_\bullet^0\|_\mathcal{H} \quad \textit{for all } n \in \mathbb{N}. \tag{8.16}$$

*Proof.* With (8.15), we estimate that

$$\|u^\star - u_\bullet^{n,\star}\|_\mathcal{H} \le \|u^\star - u_\bullet^\star\|_\mathcal{H} + \|u_\bullet^\star - u_\bullet^{n,\star}\|_\mathcal{H} \overset{(8.15)}{\le} \|u^\star - u_\bullet^\star\|_\mathcal{H} + \frac{q_{\rm pic}^n}{1-q_{\rm pic}} \|u_\bullet^{1,\star} - u_\bullet^0\|_\mathcal{H}.$$

Then, (8.16) follows from the Céa-type estimate of Lemma 8.2. □

**Remark 8.4.** *Note that, for any* $u_\bullet^0 \in \mathcal{X}_\bullet$ *and* $u_\bullet^{1,\star} = \Phi_\bullet(u_\bullet^0)$, *there holds that*

$$(u_\bullet^{1,\star}\,,\, v_\bullet)_\mathcal{H} = (u_\bullet^0\,,\, v_\bullet)_\mathcal{H} - \frac{\alpha}{L^2} \langle \mathfrak{A} u_\bullet^0 - F \,,\, v_\bullet \rangle \quad \textit{for all } v_\bullet \in \mathcal{X}_\bullet.$$

*For* $v_\bullet = u_\bullet^{1,\star} - u_\bullet^0$, *this reveals that*

$$\|u_\bullet^{1,\star} - u_\bullet^0\|_\mathcal{H}^2 = -\frac{\alpha}{L^2} \langle \mathfrak{A} u_\bullet^0 - F \,,\, u_\bullet^{1,\star} - u_\bullet^0 \rangle \le \frac{\alpha}{L^2} \|\mathfrak{A} u_\bullet^0 - F\|_{\mathcal{H}^*} \|u_\bullet^{1,\star} - u_\bullet^0\|_\mathcal{H}.$$

*Consequently, we get*

$$\|u_\bullet^{1,\star} - u_\bullet^0\|_{\mathcal{H}} \leq \frac{\alpha}{L^2}\,\|\mathfrak{A}u_\bullet^0 - F\|_{\mathcal{H}^*} \overset{(8.3)}{=} \frac{\alpha}{L^2}\,\|\mathfrak{A}u_\bullet^0 - \mathfrak{A}u^\star\|_{\mathcal{H}^*} \overset{\text{(A2)}}{\leq} \frac{\alpha}{L}\,\|u_\bullet^0 - u^\star\|_{\mathcal{H}}. \tag{8.17}$$

*Therefore, boundedness of* $\|u_\bullet^{1,\star} - u_\bullet^0\|_{\mathcal{H}}$ *in the a priori estimate of Lemma 8.3 can be guaranteed independently of the space* $\mathcal{X}_\bullet \subset \mathcal{H}$ *by choosing, e.g.,* $u_\bullet^0 := 0$. *If* $\min_{w_\bullet \in \mathcal{X}_\bullet} \|u^\star - w_\bullet\|_{\mathcal{H}} = \mathcal{O}(N^{-s})$ *for some* $s > 0$ *and with* $N > 0$ *being the degrees of freedom associated with* $\mathcal{X}_\bullet$*, this suggests the choice* $n = \mathcal{O}(\log N)$ *in Lemma 8.3; see the discussion in [CW17, Remark 3.7]. Moreover, we shall see below that the choice of* $u_\bullet^0$ *by nested iteration leads to optimal computational complexity; see Section 8.10.* ☐

### 8.3.2 Inexact PCG solver for the Picard system

To solve the linearized discrete system (8.14) in Step (P.i) of each discrete Picard iteration, we use a preconditioned conjugate gradient method (PCG); see, e.g., [FP17] where PCG is used for an Uzawa-type solver for transmission problems.

To this end, let $\{\xi_\bullet^1, \dots, \xi_\bullet^N\} \subseteq \mathcal{X}_\bullet$ denote a basis of $\mathcal{X}_\bullet$. Given initial guess $u_\bullet^{n-1,\circ} \approx u_\bullet^{n-1,\star}$, we define the stiffness matrix $S_\bullet$ and right hand side as

$$S_\bullet := \Big( (\xi_\bullet^k\,,\,\xi_\bullet^j)_{\mathcal{H}} \Big)_{j,k=1,\dots,N} \in \mathbb{R}^{N\times N} \quad \text{and} \quad b_\bullet^n = \Big( \langle \mathfrak{A}u_\bullet^{n-1,\circ} - F\,,\,\xi_j\rangle \Big)_{j=1,\dots,N} \in \mathbb{R}^N.$$

Let $w_\bullet^{n,\star}$ denotes the exact solution of the linear system (8.14) with representation $w_\bullet^{n,\star} = \sum_{j=1}^N x_j \xi_\bullet^j$ and coefficient vector $x = (x_1, \dots, x_n)$. Then, the linear system (8.14) is equivalent to solve $S_\bullet x = b_\bullet^n$. Note that $S_\bullet$ is symmetric and positive definite. This allows to use PCG as inexact solver to approximate the exact Picard iterate $w_\bullet^{n,j} \approx w_\bullet^{n,\star}$; see, e.g., [GVL13, Saa03, FP17].

Instead of solving $S_\bullet x = b_\bullet$, the PCG iteration considers the preconditioned system

$$P_\bullet^{-1/2}\, S_\bullet\, P_\bullet^{-1/2} \widetilde{x} = P_\bullet^{-1/2}\, b_\bullet^n, \tag{8.18}$$

and formally applies the conjugate gradient method to (8.18); see e.g., [GVL13, Algorithm 11.3.2]. Further, we note that $x$ and $\widetilde{x}$ are connected through $x = P_\bullet^{-1/2}\widetilde{x}$. We suppose that the matrix $P_\bullet \in \mathbb{R}^{N\times N}$ in (8.18) is symmetric and positive definite. Additionally, $P_\bullet$ is called an optimal preconditioner for $S_\bullet$, if there exists constants $c_{\mathrm{P}}, C_{\mathrm{P}} > 0$ which are independent of the space $\mathcal{X}_\bullet$, such that

$$c_{\mathrm{P}}\, y^T\, P_\bullet\, y \leq y^T\, S_\bullet\, y \leq C_{\mathrm{P}}\, y^T\, P_\bullet\, y \quad \text{for all } y \in \mathbb{R}^N,$$

i.e., $P_\bullet$ is spectrally equivalent to $S_\bullet$. The latter assumption implies that

$$\mathrm{cond}_2(P_\bullet^{-1/2}\, S_\bullet\, P_\bullet^{-1/2}) \leq C_{\mathrm{PCG}}, \tag{8.19}$$

where $C_{\mathrm{PCG}} > 0$ depends only on $c_{\mathrm{P}}, C_{\mathrm{P}}$ and is independent of the discrete subspace $\mathcal{X}_\bullet \subset \mathcal{H}$. For details on optimal preconditioners for finite and boundary elements we refer to [FFPS17a, FFPS17b, WC06, XCH10]. To solve the preconditioned system (8.18), we use the PCG algorithm proposed in [GVL13, Algorithm 11.5.1].

To that end, let $w_\bullet^{n,0} \in \mathcal{X}_\bullet$ be an initial guess with corresponding coefficient vector $x^{n,0} \in \mathbb{R}^N$ and representation $w_\bullet^{n,0} = \sum_{j=0}^N x_j^{n,0} \xi_\bullet^j$. For all $k = 1, \ldots, N$, let $x^{n,k} \in \mathbb{R}^N$ denote the approximate solution of (8.18) and $S_\bullet x = b_\bullet^n$ after $k$ iterations of the PCG algorithm. This gives rise to approximate discrete solutions

$$w_\bullet^{n,k} \in \mathcal{X}_\bullet \quad \text{and } w_\bullet^{n,k} := \sum_{j=1}^N x_j^{n,k} \xi_\bullet^j.$$

Recall, that $w_\bullet^{n,\star}$ denote the exact solution of the linear system (8.14). The next lemma summarizes some important properties of the PCG iteration; see [GVL13, Theorem 11.3.3].

**Lemma 8.5.** *Let $w_\bullet^{n,0}$ be given. For all $k = 1, \ldots, N$, the approximate solution $w_\bullet^{n,k} \in \mathcal{X}_\bullet$ in the $k$-th step of the PCG algorithm satisfies*

$$\big\| w_\bullet^{n,\star} - w_\bullet^{n,k} \big\|_{\mathcal{H}} \le q_{\mathrm{pcg}} \big\| w_\bullet^{n,\star} - w_\bullet^{n,k-1} \big\|_{\mathcal{H}}, \quad \text{with } q_{\mathrm{pcg}} := \big(1 - C_{\mathrm{PCG}}^{-1}\big)^{1/2} < 1 \tag{8.20}$$

*as well as*

$$\big\| w_\bullet^{n,\star} - w_\bullet^{n,k} \big\|_{\mathcal{H}} \le 2\kappa_{\mathrm{pcg}}^k \big\| w_\bullet^{n,\star} - w_\bullet^{n,0} \big\|_{\mathcal{H}}, \quad \text{with } \kappa_{\mathrm{pcg}} := \left( \frac{\sqrt{C_{\mathrm{PCG}}} - 1}{\sqrt{C_{\mathrm{PCG}}} + 1} \right) < 1, \tag{8.21}$$

*where $w_\bullet^{n,\star} \in \mathcal{X}_\bullet$ denotes the exact solution of the linear system* (8.14). *In particular, given any tolerance $\varepsilon > 0$, there exists a constant $C_{\mathrm{CG}} \in \mathbb{N}$ such that*

$$\big\| w_\bullet^{n,\star} - w_\bullet^{n,k} \big\|_{\mathcal{H}} \le \varepsilon \big\| w_\bullet^{n,\star} - w_\bullet^{n,0} \big\|_{\mathcal{H}} \quad \text{for all } C_{\mathrm{CG}} \le k \le N.$$

*The constant $C_{\mathrm{CG}}$ is independent of the discrete space $\mathcal{X}_\bullet$ and depends only on $c_{\mathrm{P}}, C_{\mathrm{P}}$ as well as $\varepsilon$.* □

We approximate one step of the discrete Picard iteration as follows. Given an initial guess $u_\bullet^{n-1,\circ} \approx u_\bullet^{n-1,\star}$ and $w_\bullet^{n,0} := 0$, let $w_\bullet^{n,k}$ be the solution after the $k$-th step of the PCG iteration to (8.14). Then, $w_\bullet^{n,k}$ gives rise to an approximation $u_\bullet^{n,k} \approx u_\bullet^{n,\star}$ where

$$u_\bullet^{n,\star} := \Phi_\bullet(u_\bullet^{n-1,\circ}) = u_\bullet^{n-1,\circ} - \frac{\alpha}{L^2} w_\bullet^{n,\star} \quad \text{and} \quad u_\bullet^{n,k} := u_\bullet^{n-1,\circ} - \frac{\alpha}{L^2} w_\bullet^{n,k}, \tag{8.22}$$

for all $k \ge 0$. Note that, (8.22) with $w_\bullet^{n,0} = 0$ leads to nested iteration $u_\bullet^{n,0} = u_\bullet^{n-1,\circ}$. Further, we emphasize the preconditioner $P_\bullet \in \mathbb{R}^{N \times N}$ has only to be computed once for each adaptive step and is independent of the Picard iteration.

The contraction property for the PCG iterations $w_\bullet^{n,k}$ in Lemma 8.5 directly transfers to approximations $u_\bullet^{n,k}$ and gives the following corollary.

**Corollary 8.6.** *Given an initial guess $u_\bullet^{n-1,\circ}$, the approximative solutions $u_\bullet^{n,k} \in \mathcal{X}_\bullet$ defined in* (8.22) *satisfy*

$$\| u_\bullet^{n,\star} - u_\bullet^{n,k} \|_{\mathcal{H}} \le \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}} \| u_\bullet^{n,k} - u_\bullet^{n,k-1} \|_{\mathcal{H}}, \tag{8.23}$$

*as well as*

$$\| u_\bullet^{n,\star} - u_\bullet^{n,k} \|_{\mathcal{H}} \le \min \big\{ q_{\mathrm{pcg}} \, \| u_\bullet^{n,\star} - u_\bullet^{n,k-1} \|_{\mathcal{H}} \, , \, 2\kappa_{\mathrm{pcg}}^k \, \| u_\bullet^{n,\star} - u_\bullet^{n,0} \|_{\mathcal{H}} \big\}. \tag{8.24}$$

*Proof.* Recall the definition of $u_\bullet^{n,\star} := \Phi_\bullet(u_\bullet^{n,\circ}) = u_\bullet^{n,\circ} - \frac{\alpha}{L^2} w_\bullet^{n,\star}$. For all $k \geq 1$, Lemma 8.5 implies that

$$\begin{aligned}
\|u_\bullet^{n,\star} - u_\bullet^{n,k}\|_{\mathcal{H}} &\overset{(8.22)}{=} \|u_\bullet^{n,\circ} - \frac{\alpha}{L^2} w_\bullet^{n,\star} - u_\bullet^{n,\circ} + \frac{\alpha}{L^2} w_\bullet^{n,k}\|_{\mathcal{H}} \\
&= \frac{\alpha}{L^2} \|w_\bullet^{n,\star} - w_\bullet^{n,k}\|_{\mathcal{H}} \\
&\overset{(8.20)}{\leq} q_{\mathrm{pcg}} \frac{\alpha}{L^2} \|w_\bullet^{n,\star} - w_\bullet^{n,k-1}\|_{\mathcal{H}} \\
&\overset{(8.22)}{=} q_{\mathrm{pcg}} \|u_\bullet^{n,\star} - u_\bullet^{n,k-1}\|_{\mathcal{H}} \\
&\leq q_{\mathrm{pcg}} \|u_\bullet^{n,\star} - u_\bullet^{n,k}\|_{\mathcal{H}} + q_{\mathrm{pcg}} \|u_\bullet^{n,k} - u_\bullet^{n,k-1}\|_{\mathcal{H}}.
\end{aligned}$$

This directly gives (8.23) and the first estimate in (8.24). Analogous argumentation with (8.21) yields that

$$\|u_\bullet^{n,\star} - u_\bullet^{n,k}\|_{\mathcal{H}} = \frac{\alpha}{L^2} \|w_\bullet^{n,\star} - w_\bullet^{n,k}\|_{\mathcal{H}} \leq 2\kappa_{\mathrm{pcg}}^k \frac{\alpha}{L^2} \|w_\bullet^{n,\star} - w_\bullet^{n,0}\|_{\mathcal{H}} = 2\kappa_{\mathrm{pcg}}^k \|u_\bullet^{n,\star} - u_\bullet^{n,0}\|_{\mathcal{H}}.$$

This concludes the proof. □

## 8.4 A posteriori error estimator

Suppose that for each $T \in \mathcal{T}_\bullet \in \mathbb{T}$ and each discrete function $v_\bullet \in \mathcal{X}_\bullet$, one can compute an associated *error estimator* $\eta_\bullet(T, v_\bullet) \geq 0$. To abbreviate notation, we define

$$\begin{aligned}
\eta_\bullet(v_\bullet) &:= \eta_\bullet(\mathcal{T}_\bullet, v_\bullet), \qquad \text{where} \\
\eta_\bullet(\mathcal{U}_\bullet, v_\bullet) &:= \Big( \sum_{T \in \mathcal{U}_\bullet} \eta_\bullet(T, v_\bullet)^2 \Big)^{1/2} \quad \text{for all } \mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet.
\end{aligned} \tag{8.25}$$

We emphasize, that in contrast to the error estimator for indefinite problems in Chapter 4 (see e.g, (4.13)), the estimator (8.25) is defined for arbitrary discrete functions $v_\bullet \in \mathcal{X}_\bullet$.

## 8.5 Adaptive algorithm

We analyze the following adaptive algorithm which is based on the works [CW17] and [GHPS17]. In contrast to [CW17], where the algorithm is considered with a different *a posteriori* error estimation based on *elliptic reconstruction*, the following algorithm works for general error estimators satisfying the estimator axioms in Section 8.6. Algorithm 8.7 also considers an iterative PCG-solver to compute the discrete Picard iteration and hence expands the adaptive scheme in [GHPS17].

**Algorithm 8.7.** INPUT: *Initial triangulation* $\mathcal{T}_0$, *parameters* $0 < \theta \leq 1$, $\lambda_{\mathrm{PCG}}, \lambda_{\mathrm{Pic}} > 0$, $C_{\mathrm{mark}} \geq 1$, *counters* $\ell = 0$ *as well as* $n = k = 1$ *and arbitrary initial guess, e.g.* $u_0^{1,0} := 0$.

ADAPTIVE LOOP: *Iterate the following Steps (i)–(v).*

(i) *Compute iterative solution* $u_\ell^{n,k}$ *as in* (8.22) *via one step of the PCG algorithm.*

(ii) *Compute the error estimators* $\eta_\ell(T, u_\ell^{n,k})$ *for all* $T \in \mathcal{T}_\ell$*.*

(iii) `If` $\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} > \lambda_{\mathrm{PCG}}\, \eta_\ell(u_\ell^{n,k})$*:*

(iii.a) *Update* $(\ell, n, k) \to (\ell, n, k+1)$ *and go to* (i).

(iv) `Elseif` $\|u_\ell^{n,k} - u_\ell^{n,0}\|_{\mathcal{H}} > \lambda_{\mathrm{Pic}}\, \eta_\ell(u_\ell^{n,k})$*:*

(iv.a) *Define* $u_\ell^{n+1,0} := u_\ell^{n,k}$*.*

(iv.b) *Update* $(\ell, n, k) \to (\ell, n+1, 1)$ *and go to* (i).

(v) `Else`

(v.a) *Determine a set* $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ *of marked elements which has minimal cardinality up to the multiplicative constant* $C_{\mathrm{mark}}$ *and which satisfies the Dörfler marking criterion*

$$\theta\, \eta_\ell(u_\ell^{n,k}) \le \eta_\ell(\mathcal{M}_\ell, u_\ell^{n,k}).$$

(v.b) *Generate the new triangulation* $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ *by refinement of (at least) all marked elements* $T \in \mathcal{M}_\ell$*.*

(v.c) *Define* $u_{\ell+1}^{1,0} := u_\ell^{n,k} \in \mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$*.*

(v.d) *Update* $(\ell, n, k) \to (\ell+1, 1, 1)$ *and go to (i).*

OUTPUT: *Sequence of discrete solutions* $u_\ell^{n,k}$ *and corresponding error estimators* $\eta_\ell(u_\ell^{n,k})$*.*

**Remark 8.8.** *Recall that* $u_\bullet^\star \in \mathcal{X}_\bullet$ *denote the exact solution of* (8.5) *and* $u_\bullet^{n,\star} \in \mathcal{X}_\bullet$ *denotes the exact solution of one discrete Picard iteration; see* (8.22)*. The PCG algorithm computes the exact solution* $w_\ell^{n,\star}$ *to* (8.14) *in at most* $\dim(\mathcal{X}_\ell)$ *steps. This directly implies* $u_\ell^{n,k+1} = u_\ell^{n,k} = u_\ell^{n,\star}$ *for all* $k \ge \dim(\mathcal{X}_\ell)$*. Hence, after at most* $\dim(\mathcal{X}_\ell) + 1$ *steps it holds that*

$$\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} \le \lambda_{\mathrm{PCG}}\, \eta_\ell(u_\ell^{n,k}), \tag{8.26}$$

*and Step* (iv) *in Algorithm 8.7 is executed.*

To abbreviate notation, we make the following notational convention. We emphasize that in an actual finite element implementation of Algorithm 8.7 the triple index $(\ell, n, k)$ will be replaced by one single index $j$ which will be increased in Step (iii.a), Step (iv.b) and Step (v.d). However, the above statement of Algorithm 8.7 is more intuitive and leads to an easier access for the analysis.

**Definition 8.9.** *Let* $\mathcal{I} := \big\{(\ell, n, k) \,:\, u_\ell^{n,k} \text{ is defined by Algorithm 8.7}\big\}$ *be the set off all triple indices. Then, we make the following definitions:*

1. $\underline{k}(\ell, n) := \max\big\{k \in \mathbb{N} \,:\, (\ell, n, k) \in \mathcal{I}\big\}$*, i.e,* $\underline{k}(\ell, n)$ *is the smallest index such that the PCG approximation* $u_\ell^{n,\underline{k}} \approx u_\ell^{n,\star}$ *is sufficiently accurate.*

2. *$\underline{n}(\ell) := \max\{n \in \mathbb{N}_0 \,:\, (\ell, n, 0) \in \mathcal{I}\}$, i.e, $\underline{n}(\ell)$ is the smallest index such that Picard iterate $u_\ell^{n,\underline{k}} \approx u_\ell^\star$ is sufficiently accurate.*

To shorten the notation and if it is clear from the context, we omit the dependencies and write $\underline{k} := \underline{k}(\ell, n)$ and $\underline{n} := \underline{n}(\ell)$.

**Remark 8.10.** *With the definition of $\underline{n}$ and $\underline{k}$, there holds $u_\ell^{n+1,0} := u_\ell^{n,\underline{k}}$ and $u_{\ell+1}^{1,0} := u_\ell^{\underline{n},\underline{k}}$ in Algorithm 8.7. Then, Step* (iv) *reads for all $n \geq 2$ as*

$$\texttt{Elseif}\ \|u_\ell^{n,k} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}} > \lambda_{\mathrm{Pic}}\, \eta_\ell(u_\ell^{n,k}).$$

*Further, by definition of $u_\ell^{n,k} = u_\ell^{n-1,\underline{k}} - \frac{\alpha}{L^2} w_\ell^{n,k}$, we obtain that*

$$\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} = \frac{\alpha}{L^2} \|w_\ell^{n,k} - w_\ell^{n,k-1}\|_{\mathcal{H}}.$$

*Hence with $\lambda'_{\mathrm{PCG}} = L^2/\alpha\, \lambda_{\mathrm{PCG}}$, Step* (iii) *can be equivalently written as*

$$\texttt{If}\ \|w_\ell^{n,k} - w_\ell^{n,k-1}\|_{\mathcal{H}} > \lambda'_{\mathrm{PCG}}\, \eta_\ell(u_\ell^{n,k}).$$

## 8.6 Axioms of adaptivity

Since the error estimator $\eta_\bullet(T, v_\bullet)$ is defined for arbitrary functions $v_\bullet \in \mathcal{X}_\bullet$, the following axioms (E1)–(E4) are a slight generalization of Chapter 4 resp. [CFPP14]. We suppose that the estimator satisfies the properties (E1)–(E4) with fixed constants $C_{\mathrm{stb}} \geq 1$, $C_{\mathrm{red}} \geq 1$, $C^\star_{\mathrm{rel}} \geq 1$, $C^\star_{\mathrm{drel}} \geq 1$ and $0 < q_{\mathrm{red}} < 1$.

**E1) stability on non-refined element domains:** For all triangulations $\mathcal{T}_\bullet \in \mathbb{T}$ and $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, arbitrary discrete functions $v_\bullet \in \mathcal{X}_\bullet$ and $v_\circ \in \mathcal{X}_\circ$, and an arbitrary set $\mathcal{U} \subseteq \mathcal{T}_\bullet \cap \mathcal{T}_\circ$ of non-refined elements, it holds that

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

**E2) reduction on refined element domains:** For all triangulations $\mathcal{T}_\bullet \in \mathbb{T}$ and $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, and arbitrary $v_\bullet \in \mathcal{X}_\bullet$ and $v_\circ \in \mathcal{X}_\circ$, it holds that

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

**E3) reliability:** For all triangulations $\mathcal{T}_\bullet \in \mathbb{T}$, the error of the exact discrete solution $u^\star_\bullet \in \mathcal{X}_\bullet$ to (8.5) is controlled by

$$\|u^\star - u^\star_\bullet\|_{\mathcal{H}} \leq C^\star_{\mathrm{rel}}\, \eta_\bullet(u^\star_\bullet).$$

**E4) discrete reliability:** For all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\mathcal{T}_\circ \in \mathrm{refine}(\mathcal{T}_\bullet)$, 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 $\#\mathcal{R}_{\bullet,\circ} \leq C^\star_{\mathrm{drel}}\, \#(\mathcal{T}_\bullet \backslash \mathcal{T}_\circ)$ such that the difference of the exact discrete solutions $u^\star_\bullet \in \mathcal{X}_\bullet$ and $u^\star_\circ \in \mathcal{X}_\circ$ is controlled by

$$\|u^\star_\circ - u^\star_\bullet\|_{\mathcal{H}} \leq C^\star_{\mathrm{drel}}\, \eta_\bullet(\mathcal{R}_{\bullet,\circ}, u^\star_\bullet).$$

For convenience of the reader, we used the same notation for (E1)–(E4) as for the axioms in Chapter 4. But for the rest of this thesis, all references to axioms will use the latter definition.

**Remark 8.11.** *Suppose the following approximation property of $u^\star \in \mathcal{H}$: For all $\mathcal{T}_\bullet \in \mathbb{T}$ and all $\varepsilon > 0$, there exists a refinement $\mathcal{T}_\circ \in \operatorname{refine}(\mathcal{T}_\bullet)$ such that $\|u^\star - u_\circ^\star\|_{\mathcal{H}} \le \varepsilon$. Then, discrete reliability* (E4) *already implies reliability* (E3)*, see (d) of Lemma 4.8 or [CFPP14, Lemma 3.4].*

We note that (E3)–(E4) are formulated for the non-computable exact Galerkin solution $u_\bullet^\star \in \mathcal{X}_\bullet$ to (8.5), while Algorithm 8.7 generates approximations $u_\bullet^{n,k} \approx u_\bullet^\star \in \mathcal{X}_\bullet$. The following lemma proves that reliability (E3) transfers to certain approximations.

**Lemma 8.12.** *Suppose* (A1)–(A2) *for the operator $\mathfrak{A}$ as well as stability* (E1) *and reliability* (E3) *for the a posteriori error estimator. Let $\lambda_{\rm PCG}, \lambda_{\rm Pic} > 0$. Then, there exists $C_{\rm rel} > 0$ such that for all $(\ell, \underline{n}, \underline{k}) \in \mathcal{I}$, it holds that*

$$\big\|u^\star - u_\bullet^{\underline{n},\underline{k}}\big\|_{\mathcal{H}} \le C_{\rm rel}\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}). \tag{8.27}$$

*The constant $C_{\rm rel}$ depends only on $C_{\rm rel}^\star$, $q_{\rm pic}$, $q_{\rm pcg}$, as well as $\lambda_{\rm PCG}, \lambda_{\rm Pic}$.*

*Proof.* Recall that the approximation $u_\ell^{\underline{n},\underline{k}}$ generated by Algorithm 8.7 satisfies

$$\big\|u_\bullet^{n,\underline{k}} - u_\bullet^{n,\underline{k}-1}\big\|_{\mathcal{H}} \le \lambda_{\rm PCG}\, \eta_\bullet(u_\bullet^{n,\underline{k}}) \quad \text{and} \quad \big\|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\big\|_{\mathcal{H}} \le \lambda_{\rm Pic}\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}). \tag{8.28}$$

The triangle inequality implies that

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \quad \le \quad \|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} + \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}}. \tag{8.29}$$

Using the definition $u_\bullet^{n,\star} = \Phi(u_\bullet^{n,0}) = \Phi(u_\bullet^{n-1,\underline{k}})$, we estimate the first term on the right hand side of (8.29) by

$$\begin{aligned}\|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} = \|\Phi(u_\bullet^\star) - \Phi(u_\bullet^{\underline{n},0})\|_{\mathcal{H}} &\le q_{\rm pic}\, \|u_\bullet^\star - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \\ &\le q_{\rm pic}\, \|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} + q_{\rm pic}\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}}.\end{aligned}$$

Rearranging the terms gives

$$\begin{aligned}\|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} &\le \frac{q_{\rm pic}}{1-q_{\rm pic}}\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \\ &\le \frac{q_{\rm pic}}{1-q_{\rm pic}}\Big(\|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} + \|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}}\Big).\end{aligned}$$

The latter estimate in combination with (8.29) gives

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \le \frac{q_{\rm pic}}{1-q_{\rm pic}}\, \|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} + \Big(1 + \frac{q_{\rm pic}}{1-q_{\rm pic}}\Big)\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}}. \tag{8.30}$$

We treat each term in (8.30) separately. Then, (8.28) implies for the first term that

$$\|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \le \lambda_{\rm Pic}\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).$$

The last term in the right hand side of (8.30), we see that

$$\|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.23)}{\leq} \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},\underline{k}-1}\|_{\mathcal{H}} \overset{(8.28)}{\leq} \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).$$

Combining the latter estimates, we obtain that

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \leq \Big(\frac{q_{\rm pic}}{1-q_{\rm pic}}\lambda_{\rm Pic} + \Big(1+\frac{q_{\rm pic}}{1-q_{\rm pic}}\Big)\lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\Big)\,\eta_\bullet(u_\bullet^{\underline{n},\underline{k}}). \tag{8.31}$$

With reliability (E3) and stability (E1), we estimate that

$$\begin{aligned}
\|u^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} &\leq \|u^\star - u_\bullet^\star\|_{\mathcal{H}} + \|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \\
&\overset{(E3)}{\leq} C_{\rm rel}^\star\,\eta_\bullet(u_\bullet^\star) + \|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \\
&\overset{(E1)}{\leq} C_{\rm rel}^\star\,\eta_\bullet(u_\bullet^{\underline{n},\underline{k}}) + (1 + C_{\rm rel}^\star C_{\rm stb})\,\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \\
&\overset{(8.31)}{\lesssim} \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).
\end{aligned}$$

This concludes the proof. □

## 8.7 Convergence

### 8.7.1 Lucky breakdown

The following two results analyze the possible (lucky) breakdown of Algorithm 8.7. The first proposition shows that, if $\max\big\{\ell' \in \mathbb{N}_0 : (\ell', 1, 0) \in \mathcal{I}\big\} < \infty$, i.e., there exists an index $\ell \in \mathbb{N}$ such that Algorithm 8.7 does not reach Step (v) in the $\ell$-th adaptive step, then the exact solution $u^\star = u_\ell^\star$ belongs to the discrete space $\mathcal{X}_\ell$.

**Proposition 8.13.** *Suppose* (A1)–(A2) *for the nonlinear operator $\mathfrak{A}$ as well as stability* (E1) *and reliability* (E3) *for the a posteriori error estimator. Let $\lambda_{\rm Pic} > 0$ and assume that Step* (v) *in Algorithm 8.7 is never reached for some $\ell \in \mathbb{N}_0$, i.e., $\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} > \lambda_{\rm Pic}\,\eta_\ell(u_\ell^{n,\underline{k}})$ for all $n \in \mathbb{N}$. Further, suppose that $\lambda_{\rm PCG}$ is sufficiently small (in particular, with respect to $\lambda_{\rm Pic}$) such that*

$$q_1 := \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\lambda_{\rm Pic}^{-1} < 1 \quad \textit{and} \quad q_{\rm lucky} := \frac{(q_1 + q_{\rm pcg})}{(1-q_1)} < 1. \tag{8.32}$$

*Then, it holds that $\lim_{n\to\infty} u_\ell^{n,\star} = u^\star = u_\ell^\star \in \mathcal{X}_\ell$ and*

$$\eta_\ell(u_\ell^{n,\underline{k}}) \leq q_{\rm lucky}^{n-1}\,\|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal{H}} \xrightarrow{n\to\infty} 0 = \eta_\ell(u^\star). \tag{8.33}$$

*Proof.* We prove the assertion in two steps. First note, that Step (iii) of Algorithm 8.7 implies that, for $(\ell, n, \underline{k}) \in \mathcal{I}$,

$$\|u_\ell^{n,\star} - u_\ell^{n,\underline{k}}\|_{\mathcal{H}} \overset{(8.23)}{\leq} \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\|u_\ell^{n,\underline{k}} - u_\ell^{n,\underline{k}-1}\|_{\mathcal{H}} \leq \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\lambda_{\rm PCG}\,\eta_\ell(u_\ell^{n,\underline{k}}). \tag{8.34}$$

**Step 1: Proof of $\boldsymbol{\eta_\ell(u_\ell^{n,\underline{k}}) \le q_{\rm lucky}^{n-1}\,\|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal{H}}}$.** Let $n \ge 2$. The definition of $u_\ell^{n,0} := u_\ell^{n-1,\underline{k}}$, the triangle inequality, and Step (iv) of Algorithm 8.7, (i.e., $\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} > \lambda_{\rm Pic}\,\eta_\ell(u_\ell^{n,\underline{k}})$), yield that

$$
\begin{aligned}
\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \;&\le\; \|u_\ell^{n,\star} - u_\ell^{n,\underline{k}}\|_{\mathcal{H}} + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}} + \|u_\ell^{n-1,\star} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}} \\
&\overset{(8.34)}{\le}\; \lambda_{\rm PCG}\,\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\eta_\ell(u_\ell^{n,\underline{k}}) + \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}} \\
&\overset{(iv)}{<}\; \lambda_{\rm PCG}\,\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm Pic}^{-1}\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} + \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}}.
\end{aligned}
$$

For sufficiently small $\lambda_{\rm PCG}$ such that $\lambda_{\rm PCG}\,\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm Pic}^{-1} \le q_1 < 1$, we obtain that

$$
(1-q_1)\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \le \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}}.
$$

Recall the definition of $u_\ell^{n,\star} := \Phi(u_\ell^{n,0}) = \Phi(u_\ell^{n-1,\underline{k}})$. We estimate the last term on the right hand side by

$$
\|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}} = \|\Phi(u_\ell^{n-1,\underline{k}}) - \Phi(u_\ell^{n-1,0})\|_{\mathcal{H}} \le q_{\rm pic}\,\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}.
$$

Step (iv) of Algorithm 8.7 implies that $\eta_\ell(u_\ell^{n-1,\underline{k}}) < \lambda_{\rm Pic}^{-1}\,\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}$. With (8.32), the latter estimate yields that

$$
\begin{aligned}
(1-q_1)\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} &\le \lambda_{\rm PCG}\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\eta(u_\ell^{n-1,\underline{k}}) + q_{\rm pic}\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}} \\
&\le (q_1 + q_{\rm pcg})\,\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}.
\end{aligned} \tag{8.35}
$$

Rearranging the terms, we obtain that

$$
\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \le \frac{(q_1 + q_{\rm pcg})}{(1-q_1)}\,\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}. \tag{8.36}
$$

By assumption (8.32), there holds that $q_{\rm lucky} = \frac{(q_1+q)}{(1-q_1)} < 1$. Recall that $u_\ell^{1,\star} = \Phi(u_\ell^{1,0}) = \Phi(u_{\ell-1}^{\underline{n},\underline{k}})$. Now, for any $n \ge 2$, inductive application of (8.36) finally reveals that

$$
\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \le q_{\rm lucky}^{n-1}\,\|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal{H}}. \tag{8.37}
$$

With Step (iv) of Algorithm 8.7, this implies that

$$
\eta(u_\ell^{n,\underline{k}}) \overset{(iv)}{<} \lambda_{\rm Pic}^{-1}\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \le q_{\rm lucky}^{n-1}\,\|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal{H}} \to 0 \quad \text{as } n \to \infty. \tag{8.38}
$$

**Step 2: Proof of $\boldsymbol{u_\ell^{n,\star} \to u_\ell^\star = u^\star \in \mathcal{X}_\ell}$.** Note that $u_\ell^\star$ is a fixpoint of $\Phi(\cdot)$. By definition of $u_\ell^{n,\star}$, we obtain that

$$
\begin{aligned}
\|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}} = \|\Phi(u_\ell^\star) - \Phi(u_\ell^{n-1,\underline{k}})\|_{\mathcal{H}} &\le q_{\rm pic}\|u_\ell^\star - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}} \\
&\le q_{\rm pic}\|u_\ell^\star - u_\ell^{n-1,\star}\|_{\mathcal{H}} + q_{\rm pic}\|u_\ell^{n-1,\star} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}}.
\end{aligned}
$$

With (8.34), this implies that

$$\|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}} \overset{(8.34)}{\leq} q_{\mathrm{pic}}\,\|u_\ell^\star - u_\ell^{n-1,\star}\|_{\mathcal{H}} + q_{\mathrm{pic}}\,\frac{q_{\mathrm{pcg}}}{1-q_{\mathrm{pcg}}}\,\lambda_{\mathrm{PCG}}\,\eta(u_\ell^{n,\underline{k}}).$$

The sequence $\|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}}$ is contractive up to a non-negative perturbation which tends to zero. Basic calculus (e.g., [AFLP12, Lemma 2.3]) proves that $\|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}} \to 0$ as $n \to \infty$. The triangle inequality and Step 1 further imply that

$$\begin{aligned}\|u_\ell^\star - u_\ell^{n,\underline{k}}\|_{\mathcal{H}} &\leq \|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}} + \|u_\ell^{n,\star} - u_\ell^{n,\underline{k}}\|_{\mathcal{H}} \\ &\overset{(8.34)}{\leq} \|u_\ell^\star - u_\ell^{n,\star}\|_{\mathcal{H}} + \frac{q_{\mathrm{pcg}}}{1-q_{\mathrm{pcg}}}\,\lambda_{\mathrm{PCG}}\,\eta_\ell(u_\ell^{n,\underline{k}}) \xrightarrow{n\to\infty} 0.\end{aligned} \tag{8.39}$$

It remains to show that $u_\ell^\star = u^\star$. According to (E1), $\eta_\ell(v_\ell)$ depends Lipschitz continuously on $v_\ell \in \mathcal{X}_\ell$. Reliability (E3) and stability (E1) imply that

$$\|u^\star - u_\ell^\star\|_{\mathcal{H}} \overset{(E3)}{\lesssim} \eta_\ell(u_\ell^\star) \overset{(8.39)}{\leq} \lim_{n\to\infty} \eta_\ell(u_\ell^{n,\underline{k}}) \overset{\text{Step 1}}{=} 0. \tag{8.40}$$

This concludes that $u^\star = u_\ell^\star \in \mathcal{X}_\ell$ and $\eta_\ell(u^\star) = \eta_\ell(u_\ell^\star) = 0$. □

The second proposition shows that, if $\max\big\{\ell' \in \mathbb{N}_0 : (\ell',1,0) \in \mathcal{I}\big\} = \infty$ and additionally $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) = 0$ for some $\ell$ in Step (v) of Algorithm 8.7, then it holds that $u_\ell^{\underline{n},\underline{k}} = u^\star$ as well as $\eta_j(u_j^{\underline{n},\underline{k}}) = 0$ and $u_j^{\underline{n},\underline{k}} = u_j^{1,0} = u^\star$ for all $j > \ell$.

**Proposition 8.14.** *Suppose* (A1)–(A2) *for the nonlinear operator $\mathfrak{A}$ as well as stability* (E1) *and reliability* (E3) *for the error estimator. Suppose that* $\max\big\{\ell' \in \mathbb{N}_0 : (\ell',1,0) \in \mathcal{I}\big\} = \infty$ *and that $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) = 0$ for some $(\ell,\underline{n},\underline{k}) \in \mathcal{I}$ (or equivalently $\mathcal{M}_\ell = \emptyset$ in Step* (v.a) *for some $\ell \in \mathbb{N}_0$). Then, $u_j^{\underline{n},\underline{k}} = u_j^{1,0} = u^\star$ as well as $\mathcal{M}_j = \emptyset$ for all $j \geq \ell$.*

*Proof.* Clearly, $\mathcal{M}_\ell = \emptyset$ implies that $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) = 0$. Conversely, $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) = 0$ also implies $\mathcal{M}_\ell = \emptyset$. In this case, Lemma 8.12 yields that $\|u^\star - u_\ell^{\underline{n},\underline{k}}\|_{\mathcal{H}} \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) = 0$ and hence, $u_\ell^{\underline{n},\underline{k}} = u^\star$. Moreover, $\mathcal{M}_\ell = \emptyset$ implies that $\mathcal{T}_{\ell+1} = \mathcal{T}_\ell$. Nested iteration guarantees that $u_{\ell+1}^{1,0} = u_\ell^{\underline{n},\underline{k}} = u^\star$ and hence concludes the proof. □

### 8.7.2 Estimator convergence

In this section, we show that, if $\max\{\ell' : (\ell',1,0) \in \mathcal{I}\} = \infty$, i.e., Step (v) of Algorithm 8.7 is executed for every $\ell \in \mathbb{N}$, then Algorithm 8.7 yields convergence $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \to 0$ as $\ell \to \infty$.

We first show that the iterates $u_\bullet^{n,k}$ of Algorithm 8.7 are close to the non-computable exact Galerkin approximation $u_\bullet^\star \in \mathcal{X}_\bullet$ to (8.5). Then, the corresponding error estimators are equivalent.

**Lemma 8.15.** *Suppose* (A1)–(A2) *for the nonlinear operator $\mathfrak{A}$ and stability* (E1) *for the a posteriori error estimator. There exists a constant $C_\lambda > 0$ given by*

$$C_\lambda := \Big(\lambda_{\mathrm{Pic}}\,\frac{q_{\mathrm{pic}}}{1-q_{\mathrm{pic}}} + \lambda_{\mathrm{PCG}}\Big(1 + \frac{q_{\mathrm{pic}}}{1-q_{\mathrm{pic}}}\Big)\frac{q_{\mathrm{pcg}}}{1-q_{\mathrm{pcg}}}\Big), \tag{8.41}$$

*such that for all sufficiently small* $\lambda_{\mathrm{Pic}}, \lambda_{\mathrm{PCG}} < 1$ *with* $C_\lambda\, C_{\mathrm{stb}} < 1$*, it holds that*

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \le C_\lambda\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}), \tag{8.42}$$

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \le \frac{C_\lambda}{1 - C_\lambda C_{\mathrm{stb}}}\, \eta(u_\bullet^\star). \tag{8.43}$$

*Moreover, there holds equivalence*

$$(1 - C_\lambda C_{\mathrm{stb}})\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}) \le \eta_\ell(u_\bullet^\star) \le (1 + C_\lambda C_{\mathrm{stb}})\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}). \tag{8.44}$$

*Proof.* For convenience of the reader, we split the proof into two steps.

**Step 1: Proof of (8.42).** The proof is similar to the proof of Lemma 8.12. Recall that the approximation $u_\ell^{\underline{n},\underline{k}}$ generated by Algorithm 8.7 satisfies that

$$\big\|u_\bullet^{n,\underline{k}} - u_\bullet^{n,\underline{k}-1}\big\|_{\mathcal{H}} \le \lambda_{\mathrm{PCG}}\, \eta_\bullet(u_\bullet^{n,\underline{k}}) \quad \text{and} \quad \big\|u_\bullet^{\underline{n},k} - u_\bullet^{\underline{n},0}\big\|_{\mathcal{H}} \le \lambda_{\mathrm{Pic}}\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}). \tag{8.45}$$

The triangle inequality implies that

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \quad \le \quad \|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} + \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}}. \tag{8.46}$$

Using the definition $u_\ell^{n,\star} = \Phi(u_\ell^{n,0}) = \Phi(u_\ell^{n-1,\underline{k}})$, we estimate the first term on the right-hand side of (8.46) by

$$\begin{aligned}\|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} = \|\Phi(u_\bullet^\star) - \Phi(u_\bullet^{\underline{n},0})\|_{\mathcal{H}} &\le q_{\mathrm{pic}}\, \|u_\bullet^\star - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \\ &\le q_{\mathrm{pic}}\, \|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} + q_{\mathrm{pic}}\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}}.\end{aligned}$$

Rearranging the terms, we see that

$$\begin{aligned}\|u_\bullet^\star - u_\bullet^{\underline{n},\star}\|_{\mathcal{H}} &\le \frac{q_{\mathrm{pic}}}{1 - q_{\mathrm{pic}}}\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \\ &\le \frac{q_{\mathrm{pic}}}{1 - q_{\mathrm{pic}}} \Big(\|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} + \|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}}\Big).\end{aligned}$$

The latter estimate in combination with (8.46) gives

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \le \frac{q_{\mathrm{pic}}}{1 - q_{\mathrm{pic}}}\, \|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} + \Big(1 + \frac{q_{\mathrm{pic}}}{1 - q_{\mathrm{pic}}}\Big)\, \|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}}. \tag{8.47}$$

We treat each term in (8.47) separately. For the first term, (8.28) implies that

$$\|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},0}\|_{\mathcal{H}} \le \lambda_{\mathrm{Pic}}\, \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).$$

The last term in the right hand side of (8.47) is estimated by

$$\|u_\bullet^{\underline{n},\star} - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.23)}{\le} \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}} \|u_\bullet^{\underline{n},\underline{k}} - u_\bullet^{\underline{n},\underline{k}-1}\|_{\mathcal{H}} \overset{(8.45)}{\le} \lambda_{\mathrm{PCG}}\, \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}} \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).$$

Combining the latter estimates, we conclude (8.42).

**Step 2: Proof of (8.43) and (8.44).** Choose $\lambda_{\rm PCG}, \lambda_{\rm Pic}$ sufficiently small such that $C_\lambda C_{\rm stb} < 1$. This implies that

$$\eta(u_\bullet^{\underline{n},\underline{k}}) \leq \eta(u_\bullet^\star) + C_{\rm stb}\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.42)}{\leq} \eta(u_\bullet^\star) + C_\lambda C_{\rm stb}\,\eta(u_\bullet^{\underline{n},\underline{k}})$$

and hence

$$(1 - C_\lambda C_{\rm stb})\,\eta(u_\bullet^{\underline{n},\underline{k}}) \leq \eta(u_\bullet^\star). \tag{8.48}$$

This concludes the first estimate in (8.44). Combining (8.48) and (8.42), we see that

$$\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \leq C_\lambda\,\eta(u_\bullet^{\underline{n},\underline{k}}) \leq \frac{C_\lambda}{1 - C_\lambda C_{\rm stb}}\,\eta(u_\bullet^\star).$$

This proves (8.43). With (8.42), we see that

$$\eta_\bullet(u_\bullet^\star) \overset{\rm (E1)}{\leq} \eta_\bullet(u_\bullet^{\underline{n},\underline{k}}) + C_{\rm stb}\|u_\bullet^\star - u_\bullet^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.42)}{\leq} (1 + C_\lambda C_{\rm stb})\,\eta_\bullet(u_\bullet^{\underline{n},\underline{k}}).$$

This verifies (8.44) and concludes the proof. ☐

The following proposition gives a first convergence result for Algorithm 8.7. Unlike the stronger convergence result of linear convergence in Theorem 8.20, plain convergence only relies on (A1)–(A2), but avoids the use of (A3).

**Proposition 8.16.** *Suppose* (A1)–(A2) *for the nonlinear operator* $\mathfrak{A}$ *and* (E1)–(E2) *for the error estimator. Let* $0 < \theta \leq 1$ *and let* $C_\lambda$ *be the constant from Lemma 8.15. Choose* $\lambda_{\rm PCG}, \lambda_{\rm Pic}$ *sufficiently small such that* $0 < C_\lambda C_{\rm stb} < \theta \leq 1$*. Then, there exist constants* $0 < q_{\rm est} < 1$ *and* $C_{\rm est} > 0$ *which depend only on* (E1)–(E2) *as well as* $C_\lambda, \lambda_{\rm PCG}, \lambda_{\rm Pic}$*, and* $\theta$*, such that the following implication holds: If* $\max\{\ell' : (\ell', 1, 0) \in \mathcal{I}\} = \infty$*, then*

$$\eta_{\ell+1}(u_{\ell+1}^\star)^2 \leq q_{\rm est}\,\eta_\ell(u_\ell^\star)^2 + C_{\rm est}\,\|u_{\ell+1}^\star - u_\ell^\star\|_{\mathcal{H}}^2 \quad \textit{for all } \ell \in \mathbb{N}_0, \tag{8.49}$$

*where* $u_\bullet^\star \in \mathcal{X}_\bullet$ *in the (non-computable) Galerkin solution to* (8.5)*. Moreover, there holds estimator convergence* $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \to 0$ *as* $\ell \to \infty$*.*

*Proof.* We prove the assertion in three steps.

**Step 1: Proof of (8.49).** Arguing as in the proof of Lemma 8.15, stability (E1) proves that

$$\eta_\ell(\mathcal{M}_\ell, u_\ell^{\underline{n},\underline{k}}) \overset{\rm (E1)}{\leq} \eta_\ell(\mathcal{M}_\ell, u_\ell^\star) + C_{\rm stb}\,\|u_\ell^\star - u_\ell^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.42)}{\leq} \eta_\ell(\mathcal{M}_\ell, u_\ell^\star) + C_\lambda C_{\rm stb}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}).$$

Together with the Dörfler marking strategy in Step (v.a) of Algorithm 8.7, this proves that

$$\begin{aligned}\theta'\,\eta_\ell(u_\ell^\star) := \frac{\theta - C_\lambda C_{\rm stb}}{1 + C_\lambda C_{\rm stb}}\,\eta_\ell(u_\ell^\star) &\overset{(8.44)}{\leq} (\theta - C_\lambda C_{\rm stb})\,\eta_\ell(u_\ell^{\underline{n},\underline{k}})\\ &\overset{\rm (v.a)}{\leq} \eta_\ell(\mathcal{M}_\ell, u_\ell^{\underline{n},\underline{k}}) - C_\lambda C_{\rm stb}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq \eta_\ell(\mathcal{M}_\ell, u_\ell^\star).\end{aligned}$$

Note that $0 < C_\lambda C_{\mathrm{stb}} < \theta \le 1$ implies that $\theta' > 0$. Hence, the letter estimate shows the Dörfler marking for $u_\ell^\star$ with parameter $0 < \theta' < \theta$. Therefore, [CFPP14, Lemma 4.7] proves (8.49).

**Step 2:** Next, we adopt an argument from [BV84, MSV08] to prove *a priori* convergence of the sequence $(u_\ell^\star)_{\ell\in\mathbb{N}_0}$: Since the discrete subspaces are nested, $\mathcal{X}_\infty := \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ is a closed subspace of $\mathcal{H}$ and hence a Hilbert space. Arguing as above (for the continuous and discrete problem), there exists a unique solution $u_\infty^\star \in \mathcal{X}_\infty$ of

$$\langle \mathfrak{A}\, u_\infty^\star\,,\, v_\infty\rangle = \langle F\,,\, v_\infty\rangle \quad \text{for all } v_\infty \in \mathcal{X}_\infty.$$

Note that $\mathcal{X}_\ell \subseteq \mathcal{X}_\infty$ implies that $u_\ell^\star$ is a Galerkin approximation to $u_\infty^\star$. Hence, the Céa lemma (Lemma 8.2) is valid with $u^\star \in \mathcal{H}$ replaced by $u_\infty^\star \in \mathcal{X}_\infty$. Together with the definition of $\mathcal{X}_\infty$, this proves that

$$\|u_\infty^\star - u_\ell^\star\|_{\mathcal{H}} \le \frac{L}{\alpha} \min_{w_\ell\in\mathcal{X}_\ell} \|u_\infty^\star - w_\ell\|_{\mathcal{H}} \xrightarrow{\ell\to\infty} 0.$$

In particular, we infer that $\|u_{\ell+1}^\star - u_\ell^\star\|_{\mathcal{H}}^2 \to 0$ as $\ell \to \infty$.

**Step 3:** According to (8.49) and Step 2, the sequence $\big(\eta_\ell(u_\ell^\star)\big)_{\ell\in\mathbb{N}_0}$ is contractive up to a non-negative perturbation which tends to zero. Basic calculus (e.g., [AFLP12, Lemma 2.3]) proves that $\eta_\ell(u_\ell^\star) \to 0$ as $\ell \to \infty$. Lemma 8.15 guarantees the equivalence $\eta_\ell(u_\ell^\star) \simeq \eta_\ell(u_\ell^{\underline{n},\underline{k}})$. This concludes the proof. $\square$

Note that, $\lambda_{\mathrm{PCG}}, \lambda_{\mathrm{Pic}}$ in Lemma 8.15 and Proposition 8.16 can be chosen independently of each other. To see this, we pick $\lambda_{\mathrm{PCG}}, \lambda_{\mathrm{Pic}} > 0$ sufficiently small such that

$$\lambda_{\mathrm{Pic}} < \theta\, \frac{1-q_{\mathrm{pic}}}{2\,C_{\mathrm{stb}}\,q_{\mathrm{pic}}} \qquad \text{and} \qquad \lambda_{\mathrm{PCG}} < \theta\, \frac{(1-q_{\mathrm{pcg}})\,(1-q_{\mathrm{pic}})}{2\,C_{\mathrm{stb}}\,q_{\mathrm{pcg}}}. \tag{8.50}$$

In combination with (8.41), this implies that $C_\lambda\, C_{\mathrm{stb}} < \theta \le 1$.

**Remark 8.17.** *As in Proposition 8.16, the linear convergence result of Theorem 8.20 below allows arbitrary $0 < \theta \le 1$, but requires sufficiently small parameters $\lambda_{\mathrm{PCG}}, \lambda_{\mathrm{Pic}} \ll 1$ such that $0 < C_\lambda C_{\mathrm{stb}} < \theta$ with $C_\lambda > 0$ being the constant from Lemma 8.15. In many situations, the weaker constraint $0 < C_\lambda C_{\mathrm{stb}} < 1$ which avoids any coupling of $\theta$ to $\lambda_{\mathrm{PCG}}, \lambda_{\mathrm{Pic}}$, appears to be sufficient to guarantee* plain *convergence. To see this, note that usually the error estimator is equivalent to error plus data oscillations*

$$\eta_\ell(u_\ell^\star) \simeq \|u^\star - u_\ell^\star\|_{\mathcal{H}} + \mathrm{osc}_\ell(u_\ell^\star).$$

*If the "discrete limit space" $\mathcal{X}_\infty := \overline{\bigcup_{\ell=0}^\infty \mathcal{X}_\ell}$ satisfies $\mathcal{X}_\infty = \mathcal{H}$, possible smoothness of $\mathfrak{A}$ guarantees $\|u^\star - u_\ell^\star\|_{\mathcal{H}} + \mathrm{osc}_\ell(u_\ell^\star) \to 0$ as $\ell \to \infty$; see e.g., Section 4.5.1 or the argumentation in [EP16, Proof of Theorem 3]. Moreover, $\mathcal{X}_\infty = \mathcal{H}$ follows either implicitly if $u^\star$ is "nowhere discrete", or can explicitly be ensured by the marking strategy without deteriorating optimal convergence rates; see Section 4.5.1 or Section 5.2.1. Since $0 < C_\lambda C_{\mathrm{stb}} < 1$, Lemma 8.15 guarantees estimator equivalence $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \simeq \eta_\ell(u_\ell^\star)$. Overall, such a situation leads to $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \to 0$ as $\ell \to \infty$.*

## 8.8 Linear convergence

Suppose that $\mathfrak{A}$ additionally satisfies (A3). For $v \in \mathcal{H}$, we define the energy functional $E$ by

$$E : \mathcal{H} \to \mathbb{R} \quad \text{with} \quad E(v) := \operatorname{Re}(P - F)v,$$

where $P(\cdot)$ denotes the potential associated with $\mathfrak{A}$ from (8.4) and $F \in \mathcal{H}^*$ is the right-hand side of (8.3). The next lemma generalizes [DK08, Lemma 16] and [GMZ12, Theorem 4.1] and states equivalence of the energy difference and the difference in norm.

**Lemma 8.18.** *Suppose* (A1)–(A3). *Let $\mathcal{X}_\bullet$ be a closed subspace of $\mathcal{H}$ (which also allows $\mathcal{X}_\bullet = \mathcal{H}$). If $u_\bullet^\star \in \mathcal{X}_\bullet$ denotes the corresponding Galerkin approximation* (8.2)*, it holds that*

$$\frac{\alpha}{2}\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2 \le E(v_\bullet) - E(u_\bullet^\star) \le \frac{L}{2}\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2 \quad \text{for all } v_\bullet \in \mathcal{X}_\bullet. \tag{8.51}$$

*Proof.* Since $\mathcal{H}$ is also a Hilbert space over $\mathbb{R}$, we interpret $E$ as an $\mathbb{R}$-functional. Since $F$ is linear with Gâteaux derivative $\langle \mathrm{d}F(v)\,,\, w\rangle = \langle F\,,\, w\rangle$ for all $v, w \in \mathcal{H}$, the energy $E$ is also Gâteaux differentiable with

$$\langle \mathrm{d}E(v)\,,\, w\rangle = \operatorname{Re}\langle \mathrm{d}P(v) - F\,,\, w\rangle = \operatorname{Re}\langle \mathfrak{A}v - F\,,\, w\rangle.$$

Define $\psi(t) := E(u_\bullet^\star + t(v_\bullet - u_\bullet))$ for $t \in [0,1]$. We first prove that $\psi$ is differentiable. For $t \in [0,1]$, it holds that

$$\begin{aligned}
\psi'(t) &= \lim_{\substack{r\to 0\\ r\in\mathbb{R}}} \frac{E\big(u_\bullet^\star + t(v_\bullet - u_\bullet^\star) + r(v_\bullet - u_\bullet^\star)\big) - E\big(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\big)}{r} \\
&= \langle \mathrm{d}E\big(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\big)\,,\, v_\bullet - u_\bullet^\star\rangle \\
&= \operatorname{Re}\langle \mathfrak{A}\left(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\right) - F\,,\, v_\bullet - u_\bullet^\star\rangle.
\end{aligned} \tag{8.52}$$

Hence, $\psi$ is differentiable. For $s, t \in [0,1]$, Lipschitz continuity (A2) of $\mathfrak{A}$ proves that

$$\begin{aligned}
|\psi'(s) - \psi'(t)| &= \big|\operatorname{Re}\left\langle \mathfrak{A}\big(u_\bullet^\star + s(v_\bullet - u_\bullet^\star)\big) - \mathfrak{A}\big(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\big)\,,\, v_\bullet - u_\bullet^\star\right\rangle\big| \\
&\le L\|(s-t)(v_\bullet - u_\bullet^\star)\|_{\mathcal{H}}\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}} = L\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2|s-t|,
\end{aligned}$$

i.e., $\psi'$ is Lipschitz continuous with constant $L\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2$. By Rademacher's theorem, $\psi'$ is almost everywhere differentiable and there additionally holds $|\psi''| \le L\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2$ almost everywhere. Moreover, the fundamental theorem of calculus applies and integration by parts yields that

$$E(v_\bullet) - E(u_\bullet^\star) = \psi(1) - \psi(0) = \psi'(0) + \int_0^1 \psi''(t)(1-t)\,\mathrm{d}t\,.$$

Since $\mathcal{X}_\bullet \subset \mathcal{H}$ is a closed subspace, there also holds $\mathrm{d}P_\bullet = \mathfrak{A}_\bullet$ with the restriction $P_\bullet := P|_{\mathcal{X}_\bullet}$. Hence, we may also define the restricted energy $E_\bullet := E|_{\mathcal{X}_\bullet}$. With (8.52) and $\mathrm{d}E_\bullet\, u_\bullet^\star = \mathfrak{A}_\bullet u_\bullet^\star - F_\bullet = 0$, we see that $\psi'(0) = 0$. Hence, we obtain that

$$E(v_\bullet) - E(u_\bullet^\star) = \int_0^1 \psi''(t)(1-t)\,dt. \tag{8.53}$$

Since $|\psi''| \le L\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2$ almost everywhere, we get the upper bound in (8.51). To see the lower bound, we compute for almost every $t \in [0,1]$ that

$$\begin{aligned}\psi''(t) &\overset{(8.52)}{=} \lim_{\substack{r\to 0\\ r\in\mathbb{R}}} \frac{1}{r}\Big( \operatorname{Re}\Big\langle \mathfrak{A}\big(u_\bullet^\star + (t+r)(v_\bullet - u_\bullet^\star)\big) - F\,,\, v_\bullet - u_\bullet^\star \Big\rangle \\ &\qquad - \operatorname{Re}\Big\langle \mathfrak{A}\big(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\big) - F\,,\, v_\bullet - u_\bullet^\star \Big\rangle\Big) \\ &= \lim_{\substack{r\to 0\\ r\in\mathbb{R}}} \frac{1}{r^2}\operatorname{Re}\Big\langle \mathfrak{A}\left(u_\bullet^\star + (t+r)(v_\bullet - u_\bullet^\star)\right) - \mathfrak{A}\left(u_\bullet^\star + t(v_\bullet - u_\bullet^\star)\right),\, r(v_\bullet - u_\bullet^\star)\Big\rangle \\ &\overset{(A1)}{\ge} \lim_{\substack{r\to 0\\ r\in\mathbb{R}}} \frac{\alpha}{r^2}\,\|r(v_\bullet - u_\bullet^\star)\|_{\mathcal{H}}^2 = \alpha\|v_\bullet - u_\bullet^\star\|_{\mathcal{H}}^2.\end{aligned}$$

Together with (8.53), we conclude the proof. □

**Remark 8.19.** *Lemma 8.18 immediately implies that the Galerkin solution $u_\bullet^\star \in \mathcal{X}_\bullet$ to* (8.5) *minimizes the energy $E$ in $\mathcal{X}_\bullet$, i.e., $E(u_\bullet^\star) \le E(v_\bullet)$ for all $v_\bullet \in \mathcal{X}_\bullet$. On the other hand, if $w_\bullet \in \mathcal{X}_\bullet$ is a minimizer of the energy in $\mathcal{X}_\bullet$, we deduce $E(w_\bullet) = E(u_\bullet^\star)$. Lemma 8.18 thus implies $w_\bullet = u_\bullet^\star$. Therefore, solving the Galerkin formulation* (8.5) *is equivalent to the minimization of the energy $E$ in $\mathcal{X}_\bullet$.* □

Next, we prove a contraction property as in [DK08, Theorem 20], [BDK12, Theorem 4.7], and [GMZ12, Theorem 4.2] and, in particular, obtain linear convergence of Algorithm 8.7 in the sense of [CFPP14]. The proof is similar to the proof of linear convergence for compactly perturbed problems ( see Theorem 4.14 in Chapter 4).

**Theorem 8.20.** *Suppose* (A1)–(A3) *for the nonlinear operator $\mathfrak{A}$ and* (E1)–(E3) *for the error estimator. Let $C_\lambda$ be the constant from Lemma 8.15 and $0 \le \theta \le 1$. Let $0 < \lambda_{\rm PCG}, \lambda_{\rm Pic} \le 1$ be sufficiently small such that $0 < C_\lambda C_{\rm stb} \le \theta$. Then, there exist constants $0 < q_{\rm lin} < 1$ and $\rho > 0$ which depend only on* (A1)–(A2) *and* (E1)–(E3) *as well as on $C_\lambda$, $\lambda_{\rm PCG}, \lambda_{\rm Pic}$, and $\theta$, such that the following implication holds: If $\max\{\ell' : (\ell',1,0) \in \mathcal{I}\} = \infty$, then there holds contraction*

$$\Delta_{\ell+1} \le q_{\rm lin}\,\Delta_\ell \text{ for all } \ell \in \mathbb{N}_0, \quad \text{where} \quad \Delta_\bullet := E(u_\bullet^\star) - E(u^\star) + \rho\,\eta_\bullet(u_\bullet^\star)^2. \tag{8.54}$$

*Moreover, there exists a constant $C_{\rm lin} > 0$ such that*

$$\eta_{\ell+j}(u_{\ell+j}^{\underline{n},\underline{k}})^2 \le C_{\rm lin}\, q_{\rm lin}^j\, \eta_\ell(u_\ell^{\underline{n},\underline{k}})^2 \quad \text{for all } j,\ell \in \mathbb{N}_0. \tag{8.55}$$

*Proof.* Recall that refinement of meshes $\mathcal{T}_{\ell+1} \in \operatorname{refine}(\mathcal{T}_\ell)$ leads to nestedness of the corresponding discrete spaces $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \subset \mathcal{H}$. Then, Lemma 8.18 proves that

$$\frac{\alpha}{2}\|u_j^\star - u_\ell^\star\|_{\mathcal{H}}^2 \le E(u_j^\star) - E(u_\ell^\star) \le \frac{L}{2}\|u_j^\star - u_\ell^\star\|_{\mathcal{H}}^2 \quad \text{for all } j,\ell,\in \mathbb{N}_0 \text{ with } j \le \ell. \tag{8.56}$$

Verbatim argumentation yields that

$$\frac{\alpha}{2}\|u_\ell^\star - u^\star\|_{\mathcal{H}}^2 \le E(u_\ell^\star) - E(u^\star) \le \frac{L}{2}\|u_\ell^\star - u^\star\|_{\mathcal{H}}^2 \quad \text{for all } \ell \in \mathbb{N}_0. \tag{8.57}$$

We set $\rho := \alpha/(2C_{\rm est})$. Together with (8.56), estimator reduction (8.49) gives

$$\begin{aligned}
\Delta_{\ell+1} &= E(u_{\ell+1}^\star) - E(u^\star) + \rho\,\eta_{\ell+1}(u_{\ell+1}^\star)^2 \\
&\overset{(8.49)}{\le} \big(E(u_\ell^\star) - E(u^\star)\big) - \big(E(u_\ell^\star) - E(u_{\ell+1}^\star)\big) + \rho\,\big(q_{\rm est}\eta_\ell(u_\ell^\star)^2 + C_{\rm est}\|u_\ell^\star - u_{\ell+1}^\star\|_{\mathcal{H}}^2\big) \\
&\overset{(8.56)}{\le} E(u_\ell^\star) - E(u^\star) + \rho\,q_{\rm est}\,\eta_\ell(u_\ell^\star)^2.
\end{aligned}$$

Let $\varepsilon > 0$. Combining this estimate with reliability (E3) and (8.57), we see that

$$\begin{aligned}
\Delta_{\ell+1} &\le E(u_\ell^\star) - E(u^\star) + \rho\,(q_{\rm est}+\varepsilon)\,\eta_\ell(u_\ell^\star)^2 - \rho\,\varepsilon\,\eta_\ell(u_\ell^\star)^2 \\
&\overset{(E3)}{\le} E(u_\ell^\star) - E(u^\star) + \rho\,(q_{\rm est}+\varepsilon)\,\eta_\ell(u_\ell^\star)^2 - \rho\,\varepsilon\,(C_{\rm rel}^\star)^{-2}\,\|u_\ell^\star - u^\star\|_{\mathcal{H}}^2 \\
&\overset{(8.57)}{\le} E(u_\ell^\star) - E(u^\star) + \rho\,(q_{\rm est}+\varepsilon)\,\eta_\ell(u_\ell^\star)^2 - \rho\,\frac{2\varepsilon}{L(C_{\rm rel}^\star)^2}\big(E(u_\ell^\star) - E(u^\star)\big) \\
&= (1-\rho\,\frac{2\varepsilon}{L(C_{\rm rel}^\star)^2})\big(E(u_\ell^\star) - E(u^\star)\big) + \rho\,(q_{\rm est}+\varepsilon)\,\eta_\ell(u_\ell^\star)^2 \\
&\le \max\Big\{(1-\rho\,\frac{2\varepsilon}{L(C_{\rm rel}^\star)^2}), (q_{\rm est}+\varepsilon)\Big\}\,\Delta_\ell.
\end{aligned}$$

This proves (8.54) with

$$0 < q_{\rm lin} := \inf_{\varepsilon>0} \max\Big\{(1-\rho\,\frac{2\varepsilon}{L(C_{\rm rel}^\star)^2})\,,\,(q_{\rm est}+\varepsilon)\Big\} < 1$$

Moreover, induction on $j$ proves that $\Delta_{\ell+j} \le q_{\rm lin}^j\,\Delta_\ell$ for all $j,\ell \in \mathbb{N}_0$. In combination with (8.57) and reliability (E3), the estimator equivalence (8.44) of Lemma 8.15 proves that, for all $j,\ell \in \mathbb{N}_0$,

$$\eta_{\ell+j}(u_{\ell+j}^{n,k})^2 \overset{(8.44)}{\simeq} \eta_{\ell+j}(u_{\ell+j}^\star)^2 \simeq \Delta_{\ell+j} \le q_{\rm lin}^j\,\Delta_\ell \simeq q_{\rm lin}^j\,\eta_\ell(u_\ell^\star)^2 \overset{(8.44)}{\simeq} q_{\rm lin}^j\,\eta_{\ell+j}(u_{\ell+j}^{n,k})^2.$$

This concludes the proof. $\square$

## 8.9 Optimal convergence rates

### 8.9.1 Approximation class

Similar to Section 4.8.1, we define the following approximation class in the sense of [CFPP14]. For $N \in \mathbb{N}_0$, we define the set

$$\mathbb{T}_N := \big\{\mathcal{T}_\bullet \in \mathrm{refine}(\mathcal{T}_0)\,:\,\#\mathcal{T}_\bullet - \#\mathcal{T}_0 \le N\big\}, \tag{8.58}$$

of all refinements of $\mathcal{T}_0$ which have at most $N$ elements more than $\mathcal{T}_0$. For $s > 0$, we define the approximation norm $\|\cdot\|_{\mathbb{A}_s}$ by

$$\|u^\star\|_{\mathbb{A}_s} := \sup_{N\in\mathbb{N}_0}\Big((N+1)^s \min_{\mathcal{T}_\bullet\in\mathbb{T}_N} \eta_\bullet(u_\bullet^\star)\Big), \tag{8.59}$$

where $\eta_\bullet(u_\bullet^\star)$ is the error estimator corresponding to the optimal triangulation $\mathcal{T}_\bullet \in \mathbb{T}_N$. Note that $\|u^\star\|_{\mathbb{A}_s} < \infty$ implies the existence of a (not necessarily nested) sequence of triangulations, such that the error estimator $\eta_\bullet(u_\bullet^\star)$ corresponding to the (non-computable) Galerkin approximation $u_\bullet^\star$ decays at least with algebraic rate $s > 0$.

### 8.9.2 Main result

The following theorem is the main result of this section. It proves that Algorithm 8.7 does not only lead to linear convergence, but also guarantees the best possible algebraic convergence rate for the error estimator $\eta_\ell(u_\ell^{n,k})$.

**Theorem 8.21.** *Suppose* (A1)–(A3) *for the nonlinear operator* $\mathfrak{A}$ *and* (E1)–(E4) *for the error estimator. Let $C_\lambda > 0$ be the constant from Lemma 8.15. Suppose that* $0 < \theta \le 1$, *$\lambda_{\rm PCG}$ and $\lambda_{\rm Pic}$ are chosen sufficiently small, such that*

$$0 < C_\lambda C_{\rm stb} < \theta \quad \textit{as well as} \quad \theta'' := \frac{\theta + C_\lambda C_{\rm stb}}{1 - C_\lambda C_{\rm stb}} < \theta_{\rm opt} := \frac{1}{1 + C_{\rm stb}^2 (C_{\rm drel}^\star)^2} \tag{8.60}$$

*(which is satisfied, e.g., for $0 < \theta < \theta_{\rm opt}$ and sufficiently small $\lambda_{\rm PCG}, \lambda_{\rm Pic}$). Suppose that* $\max\{\ell' : (\ell', n, 0) \in \mathcal{I}\} = \infty$, *i.e, both iterations (over $n$ and $k$) terminate for all $\ell \in \mathbb{N}_0$. Then, for all $s > 0$, there holds the equivalence*

$$\|u^\star\|_{\mathbb{A}_s} < \infty \quad \Longleftrightarrow \quad \exists C_{\rm opt} > 0 \, \forall \ell \in \mathbb{N}_0 \quad \eta_\ell(u_\ell^{n,k}) \le C_{\rm opt} \big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\big)^{-s}. \tag{8.61}$$

*Moreover, there holds $C_{\rm opt} = C'_{\rm opt} \|u^\star\|_{\mathbb{A}_s}$, where $C'_{\rm opt} > 0$ depends only on $\mathcal{T}_0$, $\theta$, $C_\lambda$, $\lambda_{\rm PCG}$, $\lambda_{\rm Pic}$, $s$,* (E1)–(E4), (A1)–(A2), *and on the refinement axioms* (R1)–(R6).

The following comparison lemma is very similar to Lemma 4.23 and is found in [CFPP14]. For sake of completeness we include its formulation in the current setting. The proof is verbatim to Lemma 4.23 where $\ell_5$ and $\theta$ are replaced by $\ell_0$ and $\theta''$.

**Lemma 8.22.** *Suppose* (E1), (E2), *and* (E4). *Let $0 < \theta'' < \theta_{\rm opt}$. Then, there exist constants $C_1, C_2 > 0$, such that for all $s > 0$ with $\|u^\star\|_{\mathbb{A}_s} < \infty$ and all $\ell \in \mathbb{N}_0$, there exists $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$ which satisfies*

$$\#\mathcal{R}_\ell \le C_1 \big(C_2 \|u^\star\|_{\mathbb{A}_s}\big)^{1/s} \eta_\ell(u_\ell^\star)^{-1/s}, \tag{8.62}$$

*as well as the Dörfler marking criterion*

$$\theta'' \eta_\ell(u_\ell^\star) \le \eta_\ell(\mathcal{R}_\ell, u_\ell^\star). \tag{8.63}$$

*The constants $C_1, C_2$ depend only on $\theta'', s$, and the constants in* (E1), (E2), *and* (E4). ☐

With Lemma 8.22 at hand, we can proof the main result of this section. The proof of Theorem 8.21 is similar to that of Theorem 4.21 and follows ideas from [CFPP14, Theorem 4.1].

***Proof of Theorem 8.21.*** We prove the assertion in three steps.

**Step 1:** The implication "$\Longleftarrow$" follows by definition of the approximation class, the equivalence $\eta_\ell(u_\ell^\star) \simeq \eta_\ell(u_\ell)$ from Lemma 8.15, and the upper bound of (R3); see Step 1 of the proof of Theorem 4.21 or [CFPP14, Proposition 4.15]). We thus focus on the converse, more important implication "$\Longrightarrow$".

**Step 2:** Suppose $\|u^\star\|_{\mathbb{A}_s} < \infty$. By Assumption (8.60), Lemma 8.22 provides a set $\mathcal{R}_\ell \subseteq \mathcal{T}_\ell$ with (8.62)–(8.63). Arguing as in the proof of Proposition 8.16, stability on non refined elements (E1) proves that

$$\eta_\ell(\mathcal{R}_\ell, u_\ell^\star) \overset{\text{(E1)}}{\le} \eta_\ell(\mathcal{R}_\ell, u_\ell^{\underline{n},\underline{k}}) + C_{\text{stb}}\,\|u_\ell^\star - u_\ell^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.42)}{\le} \eta_\ell(\mathcal{R}_\ell, u_\ell^{\underline{n},\underline{k}}) + C_\lambda C_{\text{stb}}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}),$$

where $C_\lambda > 0$ with $C_\lambda C_{\text{stb}} < 1$ denotes the constant from Lemma 8.15. Together with $\theta''\eta_\ell(u_\ell^\star) \le \eta_\ell(\mathcal{R}_\ell, u_\ell^\star)$, this proves that

$$(1 - C_\lambda C_{\text{stb}})\theta''\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \overset{(8.44)}{\le} \theta''\,\eta_\ell(u_\ell^\star) \le \eta_\ell(\mathcal{R}_\ell, u_\ell^\star) \le \eta_\ell(\mathcal{R}_\ell, u_\ell^{\underline{n},\underline{k}}) + C_\lambda C_{\text{stb}}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}).$$

Using the definition in $\theta''$ in (8.60), the latter estimate results in

$$\theta\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \overset{(8.60)}{=} \Big((1 - C_\lambda C_{\text{stb}})\theta'' - C_\lambda C_{\text{stb}}\Big)\,\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \le \eta_\ell(\mathcal{R}_\ell, u_\ell^{\underline{n},\underline{k}}). \tag{8.64}$$

Hence, $\mathcal{R}_\ell$ satisfies the Dörfler marking for the computed solution $u_\ell^{\underline{n},\underline{k}}$ with parameter $\theta$. By minimality (up to constant $C_{\text{mark}} > 0$) of $\mathcal{M}_\ell$ in Step (v) of Algorithm 8.7, we thus infer that, for all $\ell \in \mathbb{N}_0$,

$$\begin{aligned} \#\mathcal{M}_\ell &\overset{(8.64)}{\le} C_{\text{mark}}\,\#\mathcal{R}_\ell \overset{(8.62)}{\le} C_{\text{mark}}\,C_1\,\big(C_2\,\|u^\star\|_{\mathbb{A}_s}\big)^{1/s}\eta_\ell(u_\ell^\star)^{-1/s} \\ &\overset{(8.44)}{\simeq} \|u^\star\|_{\mathbb{A}_s}^{1/s}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}})^{-1/s}. \end{aligned} \tag{8.65}$$

The mesh-closure estimate (R5) guarantees that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \lesssim \sum_{j=0}^{\ell-1} \#\mathcal{M}_j \overset{(8.65)}{\lesssim} \|u^\star\|_{\mathbb{A}_s}^{1/s} \sum_{j=0}^{\ell-1} \eta_j(u_j^{\underline{n},\underline{k}})^{-1/s} \quad \text{for all } \ell > 0. \tag{8.66}$$

**Step 3:** Recall linear convergence of Theorem 8.20. This implies that

$$\eta_\ell(u_\ell^{\underline{n},\underline{k}})^2 \le C_{\text{lin}}\,q_{\text{lin}}^{\ell-i}\,\eta_i(u_i^{\underline{n},\underline{k}})^2 \quad \text{for all } 0 \le i \le \ell.$$

In particular, this leads to

$$\eta_i(u_i^{\underline{n},\underline{k}})^{-1/s} \le C_{\text{lin}}^{1/(2s)}\,q_{\text{lin}}^{(\ell-i)/(2s)}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}})^{-1/s} \quad \text{for all } 0 \le i \le \ell.$$

By use of the geometric series with $0 < q_{\text{lin}}^{1/(2s)} < 1$, we obtain that

$$\sum_{j=0}^{\ell-1} \eta_j(u_j^{\underline{n},\underline{k}})^{-1/s} \le C_{\text{lin}}^{1/(2s)}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}})^{-1/s} \sum_{j=0}^{\ell-1} \big(q_{\text{lin}}^{1/(2s)}\big)^{\ell-j} \le C_{\text{lin}}^{1/(2s)}\,\frac{q_{\text{lin}}^{1/(2s)}}{1 - q_{\text{lin}}^{1/(2s)}}\,\eta_\ell(u_\ell)^{-1/s}.$$

Combining the latter estimate with (8.66), we derive that

$$\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \lesssim \|u^\star\|_{\mathbb{A}_s}^{1/s}\,\eta_\ell(u_\ell^{\underline{n},\underline{k}})^{-1/s} \quad \text{for all } \ell > 0.$$

Since $\eta_0(u_0^{1,0}) \simeq \eta_0(u_0^\star) \lesssim \|u^\star\|_{\mathbb{A}_s}$, the latter inequality holds, in fact, for all $\ell \ge 0$. Rearranging this estimate, we conclude the proof of (8.61). □

**Remark 8.23.** *We emphasize that linear convergence in Theorem 8.20 and optimal convergence in Theorem 8.21 hold for sufficiently small, but independent parameters $\lambda_{\rm PCG}, \lambda_{\rm Pic} > 0$; see* (8.50)*. On the other hand, a more subtle choice of $\lambda_{\rm PCG} > 0$, i.e., $\lambda_{\rm PCG}$ depending on $\lambda_{\rm Pic}$, implies even linear convergence (Theorem 8.30) and optimal convergence of the full estimator sequence $\eta_\ell(u_\ell^{n,k})$ (Proposition 8.27).*

## 8.10 Optimal complexity

Throughout this section, we suppose that there holds $\max\big\{\ell' \in \mathbb{N}_0 \,:\, (\ell',1,0) \in \mathcal{I}\big\} = \infty$, i.e., the Picard iteration in Step (iv) of Algorithm 8.7 terminates with $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) > 0$ for all $\ell \in \mathbb{N}_0$. The main result of this section is that Algorithm 8.7 does not only lead to linear convergence with optimal convergence rates with respect to the degrees of freedom (Theorem 8.20, Theorem 8.21), but also guarantees linear convergence of the full estimator sequence $\big(\eta_\ell(u_\ell^{n,k})\big)_{(\ell,n,k)\in\mathcal{I}}$ (Theorem 8.30) and optimal convergence rates with respect to the overall computational effort (Theorem 8.32).

### 8.10.1 Optimal convergence of the full estimator sequence

In order to prove the main results (Theorem 8.30, Theorem 8.32), we first prove optimal convergence of the full estimator sequence $\eta_\ell(u_\ell^{n,k})$ for all $(\ell,n,k) \in \mathcal{I}$ (Proposition 8.27).

With (8.23) and Step (iii) of Algorithm 8.7, there holds the following observation for all $n \geq 1$

$$\|u_\ell^{n,\star} - u_\ell^{n,\underline{k}}\|_{\mathcal{H}} \overset{(8.23)}{\leq} \frac{q_{\rm pcg}}{1-q_{\rm pcg}} \|u_\ell^{n,\underline{k}} - u_\ell^{n,\underline{k}-1}\|_{\mathcal{H}} \leq \frac{q_{\rm pcg}}{1-q_{\rm pcg}} \lambda_{\rm PCG}\, \eta_\ell(u_\ell^{n,\underline{k}}). \tag{8.67}$$

To prove Proposition 8.27, we need the following two technical lemmas. We emphasize that both lemmas hold for all $\lambda_{\rm Pic} > 0$, but sufficiently small $\lambda_{\rm PCG} > 0$ depending on $\lambda_{\rm Pic}$.

**Lemma 8.24.** *Suppose* (A1)–(A2) *for the operator $\mathfrak{A}$ as well as stability* (E1) *and reliability* (E3) *for the a posteriori error estimator. Let $0 < \theta \leq 1$ and $\lambda_{\rm Pic} > 0$ be arbitrary. Suppose that $\lambda_{\rm PCG}$ is sufficiently small with*

$$q_2 := \lambda_{\rm PCG} \frac{q_{\rm pcg}}{1-q_{\rm pcg}} \max\{\lambda_{\rm Pic}^{-1}, C_{\rm stb}\} < 1 \quad \textit{and} \quad q_{\rm ctr} := \frac{(q_2 + q_{\rm pic})}{(1-q_2)} < 1, \tag{8.68}$$

*where $q_{\rm pic} < 1$ is the contraction constant for the Picard iteration $\Phi(\cdot)$. Then, there exists $C_1 > 0$ such that for all $(\ell, n, \underline{k}) \in \mathcal{I}$ with $n < \underline{n}$, the discrete solutions of Algorithm 8.7 satisfy that*

$$\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \leq q_{\rm ctr}\, \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}} \quad \textit{for all } 2 \leq n < \underline{n}. \tag{8.69}$$

*Moreover, it holds that*

$$\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \leq C_1\, q_{\rm ctr}^{n-1}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \textit{for all } n = 1, \dots, \underline{n}-1, \tag{8.70}$$

*as well as*

$$\eta_\ell(u_\ell^{n,\underline{k}}) \le \lambda_{\rm Pic}^{-1}\, C_1\, q_{\rm ctr}^{n-1}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \textit{for all } n = 1, \dots, \underline{n} - 1. \tag{8.71}$$

*Note that in* (8.70), *there holds that* $\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} = \|u_\ell^{n,\underline{k}} - u_\ell^{n-1,\underline{k}}\|_{\mathcal H}$ *for all* $n \ge 2$. *The constant* $C_1$ *depends on* $q_{\rm pcg}$, $\lambda_{\rm PCG}$, $\alpha$, $L$, $C_{\rm stb}$, *as well as* $C_{\rm rel}$.

*Proof.* Recall that by definition, it holds that $u_\ell^{n,0} = u_\ell^{n-1,\underline{k}}$ and consequently $\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} = \|u_\ell^{n,\underline{k}} - u_\ell^{n-1,\underline{k}}\|_{\mathcal H}$ for all $n \ge 2$. We split the remainder of the proof into three steps.

**Step 1: Proof of (8.69).** Let $2 \le n < \underline{n}$. The proof is similar to Step 1 in the proof of Proposition 8.13. The triangle inequality and Step (iv) of Algorithm 8.7, i.e. $\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} > \lambda_{\rm Pic}\, \eta_\ell(u_\ell^{n,\underline{k}})$, yield that

$$\begin{aligned}
\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} &\le \|u_\ell^{n,\star} - u_\ell^{n,\underline{k}}\|_{\mathcal H} + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal H} + \|u_\ell^{n-1,\star} - u_\ell^{n-1,\underline{k}}\|_{\mathcal H} \\
&\overset{(8.67)}{\le} \lambda_{\rm PCG}\, \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \eta_\ell(u_\ell^{n,\underline{k}}) + \lambda_{\rm PCG} \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal H} \\
&\overset{(iv)}{\le} \lambda_{\rm PCG}\, \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \lambda_{\rm Pic}^{-1} \|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} + \lambda_{\rm PCG} \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal H}.
\end{aligned}$$

For sufficiently small $\lambda_{\rm PCG}$ such that $\lambda_{\rm PCG}\, \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\, \lambda_{\rm Pic}^{-1} \le q_2 < 1$, we obtain that

$$(1 - q_2)\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} \le \lambda_{\rm PCG}\, \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \eta_\ell(u_\ell^{n-1,\underline{k}}) + \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal H}.$$

Recall the definition of $u_\ell^{n,\star} := \Phi(u_\ell^{n,0}) = \Phi(u_\ell^{n-1,\underline{k}})$. We estimate the last term on the right hand side as

$$\|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal H} = \|\Phi(u_\ell^{n-1,\underline{k}}) - \Phi(u_\ell^{n-1,0})\|_{\mathcal H} \le q_{\rm pic}\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H}$$

Since $n-1 < \underline{n}$, Step (iv) of Algorithm 8.7 yields that $\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H} > \lambda_{\rm Pic}\eta_\ell(u_\ell^{n-1,\underline{k}})$. With (8.68), the latter estimate implies that

$$\begin{aligned}
(1 - q_2)\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} &\le \lambda_{\rm PCG}\, \frac{q_{\rm pcg}}{1 - q_{\rm pcg}}\, \eta(u_\ell^{n-1,\underline{k}}) + q_{\rm pic}\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H} \\
&\le (q_2 + q_{\rm pic})\, \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H}.
\end{aligned} \tag{8.72}$$

Hence,

$$\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} \le \frac{(q_2 + q_{\rm pic})}{(1 - q_2)}\, \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H} = q_{\rm ctr}\, \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal H}. \tag{8.73}$$

This proves (8.69).

**Step 2: Proof of (8.70).** Recall that $u_\ell^{1,\star} = \Phi(u_\ell^{1,0}) = \Phi(u_{\ell-1}^{\underline{n},\underline{k}})$. For any $1 \le n < \underline{n}$, inductive application of (8.73) reveals that

$$\begin{aligned}
\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal H} &\le q_{\rm ctr}^{n-1}\, \|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal H} \\
&\le q_{\rm ctr}^{n-1}\, \big(\|u_\ell^{1,\star} - u_\ell^{1,\underline{k}}\|_{\mathcal H} + \|u_\ell^{1,\star} - u_\ell^{1,0}\|_{\mathcal H}\big).
\end{aligned} \tag{8.74}$$

Estimate (8.67) and stability (E1) yield for the first term on the right hand side

$$\begin{aligned}\|u_\ell^{1,\star} - u_\ell^{1,\underline{k}}\|_{\mathcal{H}} &\overset{(8.67)}{\leq} \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,\eta_\ell(u_\ell^{1,\underline{k}}).\\ &\overset{(E1)}{\leq} \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,\Big(\eta_\ell(u_{\ell-1}^{\underline{n},\underline{k}}) + C_{\rm stb}\|u_\ell^{1,\underline{k}} - u_{\ell-1}^{\underline{n},\underline{k}}\|_{\mathcal{H}}\Big).\end{aligned}\tag{8.75}$$

Reduction on refined elements (E2) and stability yield that $\eta_\ell(v_{\ell-1}) \leq \eta_{\ell-1}(v_{\ell-1})$ for all $v_{\ell-1} \in \mathcal{X}_{\ell-1}$. Recall that $u_{\ell-1}^{\underline{n},\underline{k}} = u_\ell^{1,0}$. Using (8.74) for $n=1$, we obtain that

$$\Big(1 - \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,C_{\rm stb}\Big)\|u_\ell^{1,\underline{k}} - u_\ell^{1,0}\|_{\mathcal{H}} \leq \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) + \|u_\ell^{1,\star} - u_\ell^{1,0}\|_{\mathcal{H}}.$$

Note that by assumption (8.68) there holds $\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,C_{\rm stb} \leq q_2 < 1$ For the second term on the right-hand side, Remark 8.4 applies. With reliability of Lemma 8.12, we see that

$$\|u_\ell^{1,\star} - u_\ell^{1,0}\|_{\mathcal{H}} \overset{(8.17)}{\leq} \frac{\alpha}{L}\|u^\star - u_\ell^{1,0}\|_{\mathcal{H}} = \frac{\alpha}{L}\,\|u^\star - u_{\ell-1}^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.27)}{\leq} \frac{\alpha}{L}\,C_{\rm rel}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

Combining the latter estimates with the first inequality in (8.74), we obtain that

$$\|u_\ell^{n,\underline{k}} - u_\ell^{n,0}\|_{\mathcal{H}} \leq q_{\rm ctr}^{n-1}\Big(1 - \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG}\,C_{\rm stb}\Big)^{-1}\Big(\frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm PCG} + \frac{\alpha}{L}\,C_{\rm rel}\Big)\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

This proves (8.70).

**Step 3: Proof of (8.71).** To see (8.71), recall that Step (iv) of Algorithm 8.7 implies that $\|u_\ell^{n,\underline{k}} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}} > \lambda_{\rm Pic}\eta_\ell(u_\ell^{n,\underline{k}})$ for all $n < \underline{n}$. With (8.70), this yields that

$$\eta_\ell(u_\ell^{n,\underline{k}}) < \lambda_{\rm Pic}^{-1}\|u_\ell^{n,\underline{k}} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}} \lesssim q_{\rm ctr}^{n-1}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

This concludes the proof. □

**Remark 8.25.** *Note that assumption* (8.68) *is similar to* (8.32)*, but* $q_2$ *additionally relies on the constant* $C_{\rm stb}$*. For* $\lambda_{\rm Pic}^{-1} \geq C_{\rm stb}$*, it holds that* $q_2 = q_1$ *and* $q_{\rm ctr} = q_{\rm lucky}$*. Further,* (8.69) *implies that for sufficiently small* $\lambda_{\rm PCG} > 0$*, we even get a contraction in each step of the perturbed Picard iteration.*

The next lemma shows a similar result including also the PCG iteration.

**Lemma 8.26.** *Suppose the same assumptions as in Lemma 8.24. Then, there exist constants* $C_2, C_2' > 0$ *such that for all* $(\ell, n, k) \in \mathcal{I}$ *with* $\ell \geq 1$*, it holds that*

$$\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} \leq C_2'\,q_{\rm pcg}^{k-1}\,q_{\rm ctr}^{n-1}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \textit{if } n,k \geq 1. \tag{8.76}$$

*Moreover, it holds that*

$$\eta_\ell(u_\ell^{n,k}) \leq C_2\,q_{\rm pcg}^{k-1}\,q_{\rm ctr}^{n-1}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \textit{for all } k = 0, \dots, \underline{k}-1. \tag{8.77}$$

*The constants* $C_2, C_2'$ *depend on* $\lambda_{\rm PCG}$*,*$\lambda_{\rm Pic}$*,* $q_{\rm pic}$*,*$q_{\rm pic}$*,* $C_1$*,* $\alpha$*, and* $L$*.*

*Proof.* For $k \geq 1$, the triangle inequality and Corollary 8.6 imply that

$$\begin{aligned} \|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} &\leq \|u_\ell^{n,\star} - u_\ell^{n,k-1}\|_{\mathcal{H}} + \|u_\ell^{n,\star} - u_\ell^{n,k}\|_{\mathcal{H}} \\ &\overset{(8.24)}{\leq} (1 + q_{\mathrm{pcg}})\, \|u_\ell^{n,\star} - u_\ell^{n,k-1}\|_{\mathcal{H}} \\ &\overset{(8.24)}{\leq} (1 + q_{\mathrm{pcg}})\, q_{\mathrm{pcg}}^{k-1} \|u_\ell^{n,\star} - u_\ell^{n,0}\|_{\mathcal{H}}. \end{aligned} \tag{8.78}$$

**Step 1: Proof of (8.76) for $n \geq 2$.** Let $n \geq 2$. The triangle inequality implies that

$$\|u_\ell^{n,\star} - u_\ell^{n,0}\|_{\mathcal{H}} \leq \|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}} + \|u_\ell^{n-1,\star} - u_\ell^{n,0}\|_{\mathcal{H}}.$$

The definition of $u_\ell^{n,\star} = \Phi(u_\ell^{n,0}) = \Phi(u_\ell^{n-1,\underline{k}})$ and Step (iv) of Algorithm 8.7 yield that

$$\|u_\ell^{n,\star} - u_\ell^{n-1,\star}\|_{\mathcal{H}} = \|\Phi(u_\ell^{n-1,\underline{k}}) - \Phi(u_\ell^{n-1,0})\|_{\mathcal{H}} \leq q_{\mathrm{pic}}\, \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}.$$

Together with (8.78), (8.67), and $u_\ell^{n,0} = u_\ell^{n-1,\underline{k}}$, this proves that

$$\begin{aligned} \|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} &\overset{(8.67)}{\leq} (1 + q_{\mathrm{pcg}})\, q_{\mathrm{pcg}}^{k-1} \Big( q_{\mathrm{pic}} \|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}} \\ &\qquad + \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}}\, \lambda_{\mathrm{PCG}}\, \eta_\ell(u_\ell^{n-1,\underline{k}}) \Big) \end{aligned} \tag{8.79}$$

Let $C_1 > 0$ be the constant from Lemma 8.24. Since $n - 1 < \underline{n}$, Lemma 8.24 yields that

$$\|u_\ell^{n-1,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}} \leq C_1\, q_{\mathrm{ctr}}^{n-2}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \textit{and} \quad \eta_\ell(u_\ell^{n-1,\underline{k}}) \leq C_1\, \lambda_{\mathrm{Pic}}^{-1}\, q_{\mathrm{ctr}}^{n-2}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

In combination with (8.79) we obtain that

$$\begin{aligned} &\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} \\ &\leq q_{\mathrm{pcg}}^{k-1} (1 + q_{\mathrm{pcg}}) \Big( q_{\mathrm{pic}}\, C_1\, q_{\mathrm{ctr}}^{n-2}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) + \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}}\, \lambda_{\mathrm{PCG}}\, \lambda_{\mathrm{Pic}}^{-1}\, C_1\, q_{\mathrm{ctr}}^{n-2}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \Big) \\ &= q_{\mathrm{pcg}}^{k-1}\, q_{\mathrm{ctr}}^{n-2}\, (1 + q_{\mathrm{pcg}})\, C_1 \Big( q_{\mathrm{pic}} + \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}}\, \lambda_{\mathrm{PCG}}\, \lambda_{\mathrm{Pic}}^{-1} \Big)\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \\ &= q_{\mathrm{pcg}}^{k-1}\, q_{\mathrm{ctr}}^{n-1}\, (1 + q_{\mathrm{pcg}})\, C_1 \Big( q_{\mathrm{ctr}}^{-1}\, q_{\mathrm{pic}} + q_{\mathrm{ctr}}^{-1}\, \frac{q_{\mathrm{pcg}}}{1 - q_{\mathrm{pcg}}}\, \lambda_{\mathrm{PCG}}\, \lambda_{\mathrm{Pic}}^{-1} \Big)\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}), \end{aligned}$$

and concludes Step 1.

**Step 2: Proof of (8.76) for $n = 1$.** Analogously to the proof of Lemma 8.6, the definition $u_\ell^{1,0} = u_{\ell-1}^{\underline{n},\underline{k}}$ and reliability of Lemma 8.12 imply that

$$\|u_\ell^{1,\star} - u_\ell^{1,0}\|_{\mathcal{H}} \overset{(8.17)}{\leq} \frac{\alpha}{L} \|u^\star - u_\ell^{1,0}\|_{\mathcal{H}} = \frac{\alpha}{L}\, \|u^\star - u_{\ell-1}^{\underline{n},\underline{k}}\|_{\mathcal{H}} \overset{(8.27)}{\leq} \frac{\alpha}{L}\, C_{\mathrm{rel}}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

With (8.78) we obtain that

$$\|u_\ell^{1,k} - u_\ell^{1,k-1}\|_{\mathcal{H}} \leq (1 + q_{\mathrm{pcg}})\, q_{\mathrm{pcg}}^{k-1}\, \|u_\ell^{1,\star} - u_\ell^{1,0}\|_{\mathcal{H}} \leq q_{\mathrm{pcg}}^{k-1}\, (1 + q_{\mathrm{pcg}})\, \frac{\alpha}{L}\, C_{\mathrm{rel}}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

This concludes (8.76) with

$$C_2' := (1+q_{\rm pcg})\,\max\Big\{q_{\rm ctr}^{-1}\Big(q_{\rm pic}\,C_1 + \frac{q_{\rm pcg}}{1-q_{\rm pcg}}\,\lambda_{\rm Pic}^{-1}\,\lambda_{\rm PCG}\,C_1\Big)\,,\,\frac{\alpha}{L}\,C_{\rm rel}\Big\}$$

**Step 3: Proof of (8.77).** According to Step (iii) of Algorithm 8.7, it holds that

$$\eta_\ell(u_\ell^{n,k}) < \lambda_{\rm PCG}^{-1}\,\|u_\ell^{n,k} - u_\ell^{n,k-1}\|_{\mathcal{H}} \quad \text{for all } 1 \le k < \underline{k}.$$

This yields (8.77) for all $k \geq 1$. For $k = 0$, it holds by definition that $\eta_\ell(u_\ell^{n,0}) = \eta_\ell(u_\ell^{n-1,\underline{k}})$ if $n-1 < \underline{n}$. Hence, analogously to Step 1, Lemma 8.24 implies that,

$$\eta_\ell(u_\ell^{n,0}) = \eta_\ell(u_\ell^{n-1,\underline{k}}) \overset{(8.71)}{\le} C_1\,\lambda_{\rm Pic}^{-1}\,q_{\rm ctr}^{n-2}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

Recall $k = 0$. Multiplying the right-hand side with $q_{\rm pcg}^{k-1} = q_{\rm pcg}^{-1} > 1$ concludes the proof. □

Next, we show a generalization of Theorem 8.21, which proves optimal convergence even for the full estimator sequence $\eta_\ell(u_\ell^{n,k})$. The proposition is a consequence of Theorem 8.21 combined with Lemma 8.5 and Lemma 8.24.

**Proposition 8.27.** *Suppose the assumptions of Theorem 8.21. Let $\lambda_{\rm PCG} > 0$ is sufficiently small such that* (8.68) *be satisfied. Then, it holds that*

$$\|u^\star\|_{\mathbb{A}_s} < \infty \iff \exists C_{\rm opt,f} > 0\,\forall(\ell,n,k)\in\mathcal{I}\quad \eta_\ell(u_\ell^{n,k}) \le C_{\rm opt,f}\,\Big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\Big)^{-s} \tag{8.80}$$

*There holds $C_{\rm opt,f} = C'_{\rm opt,f}\,\|u^\star\|_{\mathbb{A}_s}$, where $C'_{\rm opt,f} > 0$ depends only on $\theta$, $\lambda_{\rm PCG}$, $\lambda_{\rm Pic}$, $s$, $q_{\rm pcg}$,* (E1)–(E4)*,* (A1)–(A2)*, and* (R1)–(R5)*, as well as on $\mathcal{T}_0$.*

*Proof.* With stability (E1) and the triangle inequality, we make the following observation

$$\begin{aligned}
\eta_\ell(u_\ell^{n,k}) &\overset{\rm (E1)}{\le} \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + C_{\rm stb}\,\|u_\ell^{\underline{n},\underline{k}} - u_\ell^{n,k}\|_{\mathcal{H}}\\
&\le\; \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + C_{\rm stb}\,\|u_\ell^{n,\underline{k}} - u_\ell^{n,k}\|_{\mathcal{H}} + C_{\rm stb}\sum_{m=n}^{\underline{n}(\ell)-1}\|u_\ell^{m+1,\underline{k}} - u_\ell^{m,\underline{k}}\|_{\mathcal{H}}\\
&\le\; \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + C_{\rm stb}\sum_{r=k}^{\underline{k}(\ell,n)-1}\|u_\ell^{n,r+1} - u_\ell^{n,r}\|_{\mathcal{H}} + C_{\rm stb}\sum_{m=n}^{\underline{n}(\ell)-1}\|u_\ell^{m+1,\underline{k}} - u_\ell^{m,\underline{k}}\|_{\mathcal{H}}.
\end{aligned} \tag{8.81}$$

For the last sum on the right-hand side, Lemma 8.24 yields that

$$\|u_\ell^{m+1,\underline{k}} - u_\ell^{m,\underline{k}}\|_{\mathcal{H}} = \|u_\ell^{m+1,\underline{k}} - u_\ell^{m+1,0}\|_{\mathcal{H}} \overset{(8.70)}{\lesssim} q_{\rm ctr}^m\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \quad \text{for all } m+1 < \underline{n}(\ell).$$

Hence, the geometric series proves that

$$\sum_{m=n}^{\underline{n}(\ell)-2}\|u_\ell^{m+1,\underline{k}} - u_\ell^{m,\underline{k}}\|_{\mathcal{H}} \lesssim \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}})\sum_{m=n}^{\underline{n}(\ell)-2} q_{\rm ctr}^m \le \frac{1}{1-q_{\rm ctr}}\,\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}). \tag{8.82}$$

Further, for $m+1 = \underline{n}(\ell)$, Step (iv) of Algorithm 8.7 implies that $\|u_\ell^{\underline{n},\underline{k}} - u_\ell^{\underline{n}-1,\underline{k}}\|_{\mathcal{H}} \le \lambda_{\mathrm{Pic}}\eta_\ell(u_\ell^{\underline{n},\underline{k}})$. Combining this estimate with (8.81), we obtain that

$$\eta_\ell(u_\ell^{n,k}) \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + \sum_{r=k}^{\underline{k}(\ell,n)-1} \|u_\ell^{n,r+1} - u_\ell^{n,r}\|_{\mathcal{H}} + \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}). \tag{8.83}$$

For the remaining sum, Lemma 8.26 reveals that

$$\|u_\ell^{n,r+1} - u_\ell^{n,r}\|_{\mathcal{H}} \overset{(8.76)}{\lesssim} q_{\mathrm{pcg}}^r\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

With $\|u_\ell^{n,\underline{k}} - u_\ell^{n,\underline{k}-1}\|_{\mathcal{H}} \le \lambda_{\mathrm{PCG}}\, \eta_\ell(u_\ell^{n,\underline{k}})$, this further implies that

$$\begin{aligned}\sum_{r=k}^{\underline{k}(\ell,n)-1} \|u_\ell^{n,r+1} - u_\ell^{n,r}\|_{\mathcal{H}} &\lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \sum_{r=k}^{\underline{k}(\ell,n)-2} q_{\mathrm{pcg}}^r \\ &\lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + \eta_{j-1}(u_{j-1}^{\underline{n},\underline{k}}).\end{aligned} \tag{8.84}$$

Combining the latter estimates with (8.81), Theorem 8.20 proves that

$$\eta_\ell(u_\ell^{n,k}) \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) + \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \overset{(8.55)}{\lesssim} \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}).$$

The splitting property (R3) directly implies that $\#\mathcal{T}_{\ell-1} \le \#\mathcal{T}_\ell \le C_{\mathrm{son}}\#\mathcal{T}_{\ell-1}$ for all $\ell \ge 1$. With Lemma 4.19 or [BHP17, Lemma 22], this translates to

$$\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1 \le \#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1 \le (C_{\mathrm{son}}\#\mathcal{T}_0)(\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1). \tag{8.85}$$

Then, optimal convergence rates of Theorem 8.21 for $\eta_\ell(u_\ell^{\underline{n},\underline{k}})$ imply that

$$\eta_\ell(u_\ell^{n,k}) \lesssim \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \overset{(8.61)}{\lesssim} (\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1)^{-s} \simeq (\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1)^{-s}.$$

This concludes the proof. $\square$

### 8.10.2 Linear convergence of the full estimator sequence

In this section, we show linear convergence of the full estimator sequence (see Theorem 8.30). We first prove the following technical lemma.

**Lemma 8.28.** *Let $\lambda_{\mathrm{Pic}}, \lambda_{\mathrm{PCG}} < C_{\mathrm{stb}}$ such that (8.68) is satisfied. For $C_3 > \max\{(1 - C_{\mathrm{stb}}\,\lambda_{\mathrm{Pic}})^{-1}, (1 - C_{\mathrm{stb}}\,\lambda_{\mathrm{PCG}})^{-1}\} > 1$, it holds that*

$$\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \le C_3^2\, \eta_\ell(u_\ell^{n,k}) \quad \textit{for all } (\ell,n,k) \in \mathcal{I}. \tag{8.86}$$

*Proof.* For $(\ell, n, k) = (\ell, \underline{n}, \underline{k})$ the statement is trivial. Hence we focus on $(\ell, n, k) \neq (\ell, \underline{n}, \underline{k})$. Stability (E1) and the Step (iv) of Algorithm 8.7 yield

$$\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq \eta_\ell(u_\ell^{\underline{n}-1,\underline{k}}) + C_{\mathrm{stb}}\, \|u_\ell^{\underline{n},\underline{k}} - u_\ell^{\underline{n}-1,\underline{k}}\|_{\mathcal{H}} \leq \eta_\ell(u_\ell^{\underline{n}-1,\underline{k}}) + C_{\mathrm{stb}}\, \lambda_{\mathrm{Pic}}\, \eta_\ell(u_\ell^{\underline{n},\underline{k}}).$$

This directly implies

$$\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq \frac{1}{1 - C_{\mathrm{stb}}\, \lambda_{\mathrm{Pic}}}\, \eta_\ell(u_\ell^{\underline{n}-1,\underline{k}}). \tag{8.87}$$

Verbatim argumentation for with Step (iii) of Algorithm 8.7 yields for $(\ell, n, \underline{k})$ that

$$\eta_\ell(u_\ell^{n,\underline{k}}) \leq \frac{1}{1 - C_{\mathrm{stb}}\, \lambda_{\mathrm{PCG}}}\, \eta_\ell(u_\ell^{n,\underline{k}-1}). \tag{8.88}$$

**Step 1:** Let $n < \underline{n}$. We first prove $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq C_3\, \eta_\ell(u_\ell^{n,\underline{k}})$ by contradiction. Therefore, suppose $C_3\, \eta_\ell(u_\ell^{n,\underline{k}}) < \eta_\ell(u_\ell^{\underline{n},\underline{k}})$. With Step (iv) of Algorithm 8.7 and the definition of $C_3$, we obtain that

$$\begin{aligned} C_3\, \eta_\ell(u_\ell^{n,\underline{k}}) < \eta_\ell(u_\ell^{\underline{n},\underline{k}}) \overset{(8.87)}{\leq} C_3\, \eta_\ell(u_\ell^{\underline{n}-1,\underline{k}}) &\overset{\text{(iv)}}{<} C_3\, \lambda_{\mathrm{Pic}}^{-1}\, \|u_\ell^{\underline{n}-1,\underline{k}} - u_\ell^{\underline{n}-2,\underline{k}}\|_{\mathcal{H}} \\ &\overset{(8.69)}{\leq} C_3\, \lambda_{\mathrm{Pic}}^{-1}\, \|u_\ell^{n,\underline{k}} - u_\ell^{n-1,\underline{k}}\|_{\mathcal{H}}. \end{aligned}$$

The latter estimate directly implies that

$$\lambda_{\mathrm{Pic}}\, \eta_\ell(u_\ell^{n,\underline{k}}) \leq \|u_\ell^{n,\underline{k}} - u_\ell^{n-1,0}\|_{\mathcal{H}}.$$

This implies $n = \underline{n}$ and contradicts $n < \underline{n}$. Hence, the contradiction proves $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq C_3\, \eta_\ell(u_\ell^{n,\underline{k}})$.

**Step 2:** Analogous argumentation as in Step 1 in combination with Step (iii) of Algorithm 8.7 and (8.88) yields that $\eta_\ell(u_\ell^{n,k}) \leq C_3\, \eta_\ell(u_\ell^{n,\underline{k}})$. Combining this with Step 1, we obtain that

$$\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \leq C_3\, \eta_\ell(u_\ell^{n,\underline{k}}) \leq C_3^2\, \eta_\ell(u_\ell^{n,k}) \quad \text{for all } (\ell, n, k) \in \mathcal{I}.$$

This concludes the proof. $\square$

In order to prove linear convergence of the full estimator sequence, we need the following assumption.

**E5) finite improvement:** There exists $C_4 > 0$ such that for all $(\ell, n, k) \in \mathcal{I}$ with $(\ell, n, k) \neq (\ell, \underline{n}, \underline{k})$, it holds that

$$\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \leq C_4\, \eta_\ell(u_\ell^{n,k}).$$

**Remark 8.29.** *Even though, we cannot thoroughly prove this fact, such a behavior is observed in practice. In particular, assumption* (E5) *is satisfied in the following cases:*

- *Linear convergence proves, in particular, that* $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \le C_{\rm lin}\, q_{\rm lin}\, \eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}})$ *for all* $(\ell,\underline{n},\underline{k}) \in \mathcal{I}$*. In practice, we observe that* $\eta_\ell(u_\ell^{\underline{n},\underline{k}})$ *is only improved by a fixed factor, i.e., there also holds the converse estimate* $\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \le \widetilde{C}\, \eta_\ell(u_\ell^{\underline{n},\underline{k}})$ *for all* $(\ell,\underline{n},\underline{k}) \in \mathcal{I}$*. In this case, Lemma 8.28 implies that*

$$\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}) \overset{(8.86)}{\lesssim} \eta_\ell(u_\ell^{n,k}) \quad \textit{for all } (\ell,n,k) \in \mathcal{I}.$$

- *Optimal convergence (see Theorem 8.27) proves* $\eta_\ell(u_\ell^{\underline{n},\underline{k}}) \lesssim \Big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\Big)^{-s}$ *for all* $(\ell,\underline{n},\underline{k}) \in \mathcal{I}$*. In numeric experiments, we observe that there also holds the converse estimate* $\Big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\Big)^{-s} \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}})$ *for all* $\ell \in \mathbb{N}$*. This implies that*

$$\eta_{\ell-1}(u_{\ell-1}^{\underline{n},\underline{k}}) \overset{(8.80)}{\lesssim} \Big(\#\mathcal{T}_{\ell-1} - \#\mathcal{T}_0 + 1\Big)^{-s} \overset{(8.85)}{\lesssim} \Big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\Big)^{-s} \lesssim \eta_\ell(u_\ell^{\underline{n},\underline{k}}),$$

  *and guarantees* (E5).

To simplify notation in the upcoming theorem, we introduce the following notation. For all $(\ell',n',k'), (\ell,n,k) \in \mathcal{I}$, we define the ordering

$$(\ell',n',k') < (\ell,n,k) \quad \overset{\text{def}}{\Longleftrightarrow} \quad \left\{ \begin{array}{ll} \text{either:} & \ell' < \ell \\ \text{or:} & \ell' = \ell \text{ and } n' < n \\ \text{or:} & \ell' = \ell \text{ and } n' = n \text{ and } k' < k \end{array} \right\}.$$

Let $|(\ell,n,k)| := 0$ for the initial index $(0,1,0)$. For $\ell \ge 0$ or $n \ge 1$ or $k \ge 0$, we define

$$|(\ell,n,k)| := \#\big\{(\ell',n',k') \in \mathcal{Q} \,:\, (\ell',n',k') < (\ell,n,k)\big\}. \tag{8.89}$$

The next theorem shows, that each step of Algorithm 8.7 (i.e., every time the counter is increased in any way), leads to a contraction of the corresponding error estimator. This linear convergence thus applies to the full estimator sequence and hence improves the result of Theorem 8.20.

**Theorem 8.30.** *Suppose* (A1)–(A2) *for the operator* $\mathfrak{A}$ *as well as stability* (E1)*, reliability* (E3) *and* (E5) *for the error estimator. Suppose* $\lambda_{\rm Pic}, \lambda_{\rm PCG}$ *are sufficiently small such that the assumptions of Proposition 8.27 are satisfied. Then, for all* $0 < \theta \le 1$ *and all* $s > 0$*, the estimator satisfies the inverse summability*

$$\sum_{(\ell,n,k)=(1,1,0)}^{(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^{-1/s} \le C\, \eta_{\ell'}(u_{\ell'}^{n',k'})^{-1/s} \quad \textit{for all } (\ell',n',k') \in \mathcal{I}, \tag{8.90}$$

*where* $C > 0$ *is independent of* $(\ell',n',k') \in \mathcal{I}$*. Moreover, there exist* $0 < q_{\rm slin} < 1$ *and* $C_{\rm slin} > 0$*, such that for all* $(1,1,0) \le (\ell',n',k') \le (\ell,n,k)$

$$\eta_\ell(u_\ell^{n,k}) \le C_{\rm slin}\, q_{\rm slin}^{|(\ell,n,k)|-|(\ell',n',k')|}\, \eta_{\ell'}(u_{\ell'}^{n',k'}). \tag{8.91}$$

*Proof.* We split the proof into two steps.

**Step 1:** According to [CFPP14, Lemma 4.9], inverse summability (8.90) for all $s > 0$ is equivalent to linear convergence (8.91) and also equivalent to the following summability condition

$$\sum_{(\ell,n,k)>(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^2 \lesssim \eta_{\ell'}(u_{\ell'}^{n',k'})^2 \quad \text{for all } (1,1,0) \le (\ell',n',k') \in \mathcal{I}, \tag{8.92}$$

where the involved constant is independent of $(\ell',n',k') \in \mathcal{I}$.

**Step 2:** It thus remains to prove (8.92). Let $(\ell',n',k') \in \mathcal{I}$ with $\ell' > 0$. Recall that nested iteration in Algorithm 8.7 yields that $u_\ell^{\underline{n},\underline{k}} = u_{\ell+1}^{1,0}$ as well as $u_\ell^{n,\underline{k}} = u_\ell^{n+1,0}$ for all $(\ell,n,0) \in \mathcal{I}$. This directly implies that $\eta_\ell(u_\ell^{n,\underline{k}}) = \eta_\ell(u_\ell^{n+1,0})$ for all $n < \underline{n}$. Combining this observation with Lemma 8.26, we obtain that

$$\begin{aligned}
\sum_{(\ell,n,k)>(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^2 \quad &\le \sum_{(\ell,n,k)\ge(\ell',1,0)} \eta_\ell(u_\ell^{n,k})^2 \\
&= \sum_{j=\ell'}^{\infty} \sum_{m=1}^{\underline{n}(j)} \sum_{r=0}^{\underline{k}(j,m)} \eta_j(u_j^{m,r})^2 \\
&\le \sum_{j=\ell'}^{\infty} \Big( \eta_j(u_j^{\underline{n},\underline{k}})^2 + 2 \sum_{m=1}^{\underline{n}(j)} \sum_{r=0}^{\underline{k}(j,m)-1} \eta_j(u_j^{n,k})^2 \Big) \\
&\overset{(8.77)}{\le} 2 \sum_{j=\ell'}^{\infty} \Big( \eta_j(u_j^{\underline{n},\underline{k}})^2 + C_2 \sum_{m=1}^{\underline{n}(j)} \sum_{r=0}^{\underline{k}(j,m)-1} q_{\mathrm{pcg}}^{2(r-1)}\, q_{\mathrm{ctr}}^{2(m-1)}\, \eta_{j-1}(u_{j-1}^{\underline{n},\underline{k}})^2 \Big) \\
&= 2 \sum_{j=\ell'}^{\infty} \Big( \eta_j(u_j^{\underline{n},\underline{k}})^2 + \eta_{j-1}(u_{j-1}^{\underline{n},\underline{k}})^2 \big( C_2 \sum_{m=1}^{\underline{n}(j)} q_{\mathrm{ctr}}^{2(m-1)} \sum_{r=0}^{\underline{k}(j,m)-1} q_{\mathrm{pcg}}^{2(r-1)} \big) \Big).
\end{aligned}$$

By use of the geometric series, we estimate the inner sums by

$$\sum_{r=0}^{\underline{k}(j,m)-1} q_{\mathrm{pcg}}^{2(r-1)} \le q_{\mathrm{pcg}}^2 \sum_{r=0}^{\infty} q_{\mathrm{pcg}}^{2r} = \frac{q_{\mathrm{pcg}}^2}{1-q_{\mathrm{pcg}}^2} \quad \text{and} \quad \sum_{m=1}^{\underline{n}(j)} q_{\mathrm{ctr}}^{2(m-1)} \le \frac{q_{\mathrm{ctr}}^2}{1-q_{\mathrm{ctr}}^2}.$$

With this, the previous estimate turns into

$$\sum_{(\ell,n,k)>(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^2 \lesssim \sum_{j=\ell'}^{\infty} \Big( \eta_j(u_j^{\underline{n},\underline{k}})^2 + \eta_{j-1}(u_{j-1}^{\underline{n},\underline{k}})^2 \Big) \lesssim \sum_{j=\ell'-1}^{\infty} \eta_j(u_j^{\underline{n},\underline{k}})^2.$$

With linear convergence (Theorem 8.20), we obtain that

$$\begin{aligned}
\sum_{(\ell,n,k)>(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^2 \quad &\le \sum_{j=\ell'-1}^{\infty} \eta_j(u_j^{\underline{n},\underline{k}})^2 \\
&\overset{(8.55)}{\le} C_{\mathrm{lin}} \eta_{\ell'-1}(u_{\ell'-1}^{\underline{n},\underline{k}})^2 \sum_{j=\ell'-1}^{\infty} q_{\mathrm{lin}}^{j-(\ell'-1)} \lesssim \eta_{\ell'-1}(u_{\ell'-1}^{\underline{n},\underline{k}})^2.
\end{aligned}$$

Finally, assumption (E5) guarantees that

$$\sum_{(\ell,n,k)>(\ell',n',k')} \eta_\ell(u_\ell^{n,k})^2 \lesssim \eta_{\ell'-1}(u_{\ell'-1}^{\underline{n},\underline{k}})^2 \overset{\text{(E5)}}{\lesssim} \eta_{\ell'}(u_{\ell'}^{n,k})^2.$$

This concludes the proof.

□

### 8.10.3 Main result

We show that Algorithm 8.7 does not only lead to optimal algebraic convergence rates for the error estimator $\eta_\ell(u_\ell^{n,k})$, but also guarantees optimal convergence behavior with respect to the computational complexity.

Optimal convergence behavior of Algorithm 8.7 means that, given $\|u^\star\|_{\mathbb{A}_s} < \infty$, the error estimator $\eta_\ell(u_\ell^{n,k})$ for $(\ell, n, k) \in \mathcal{I}$ decays with rate $s > 0$ with respect to the degrees of freedom $\mathcal{O}(\#\mathcal{T}_\ell)$.

*Optimal computational complexity* means that, given $\|u^\star\|_{\mathbb{A}_s} < \infty$, the error estimator $\eta_\ell(u_\ell^{n,k})$ for $(\ell, n, k) \in \mathcal{I}$ decays with rate $s > 0$ with respect to the computational cost; see, e.g, [Fei15] for linear problems.

Given a mesh $\mathcal{T}_\bullet \in \mathbb{T}$, we define *single-step* complexity by

$$\text{work}(\mathcal{T}_\bullet) := \text{computational effort to compute exact solution } u_\bullet^\star \text{ of (8.5) and } \eta_\bullet(u_\bullet^\star).$$

Then, a single-step complexity rate of $s > 0$ is possible, if and only if, there exists a sequence of successively refined meshes $\big(\mathcal{T}_j^{\text{opt}}\big)_{j\in\mathbb{N}} \subset \mathbb{T}$ with $\mathcal{T}_{j+1}^{\text{opt}} = \text{refine}(\mathcal{T}_j^{\text{opt}}, \mathcal{M}_j)$ for some $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell^{\text{opt}}$ such that

$$\|u^\star\|_{\mathbb{W}_s} := \sup_{\ell\in\mathbb{N}_0} \Big(\text{work}(\mathcal{T}_j^{\text{opt}})^s\, \eta_j(u_j^{\text{opt},\star})\Big) < \infty. \tag{8.93}$$

Here for $\mathcal{T}_j^{\text{opt}}$, the corresponding exact solution of (8.5) and the estimator are given by $u_j^{\text{opt},\star} \in \mathcal{X}_\ell^{\text{opt}}$ and $\eta_j(u_j^{\text{opt},\star})$.

Then, there holds the following correlation between the approximation class $\|u^\star\|_{\mathbb{A}_s}$ and the complexity class $\|u^\star\|_{\mathbb{W}_s}$.

**Lemma 8.31.** *Suppose linear single-step complexity* $\text{work}(\mathcal{T}_\bullet) \simeq \mathcal{T}_\bullet$ *for all* $\mathcal{T}_\bullet \in \mathbb{T}$*, i.e., given* $\mathcal{T}_\bullet \in \mathbb{T}$*, the exact solution* $u_\bullet^\star$ *of* (8.5) *and corresponding estimator* $\eta_\bullet(u_\bullet^\star)$ *can be computed in linear complexity. Then, it holds that*

$$\|u^\star\|_{\mathbb{W}_s} < \infty \quad \Longleftrightarrow \quad \|u^\star\|_{\mathbb{A}_s} < \infty.$$

*Proof.* The implication "$\Longleftarrow$" follows directly from Theorem 8.32 or Proposition 8.27. Hence, we focus on the implication "$\Longrightarrow$". To that end, let $\|u^\star\|_{\mathbb{W}_s} < \infty$ and $\big(\mathcal{T}_j^{\text{opt}}\big)_{j\in\mathbb{N}} \subset \mathbb{T}$ denote the corresponding sequence. For all $N \in \mathbb{N}$, there exist $\mathcal{T}_{j(N)}^{\text{opt}}, \mathcal{T}_{j(N+1)}^{\text{opt}}$ such that

$\#\mathcal{T}_{j(N)}^{\rm opt} \le N \le N+1 \le \#\mathcal{T}_{j(N)+1}^{\rm opt}$. With work$(\mathcal{T}) = \mathcal{O}(\mathcal{T})$ as well as the splitting property (R3), it follows that

$$\begin{aligned}
\|u^\star\|_{\mathbb{A}_s} &= \sup_{N\in\mathbb{N}} (N+1)^s \min_{\mathcal{T}_\bullet\in\mathbb{T}_N} \eta_\bullet(u_\bullet^\star) \\
&\le \sup_{N\in\mathbb{N}} (\#\mathcal{T}_{j(N)+1}^{\rm opt})^s\, \eta_{j(N)}(u_{j(N)}^{{\rm opt},\star}) \\
&\overset{(3.5)}{\le} C_{\rm son}^s \sup_{N\in\mathbb{N}} (\#\mathcal{T}_{j(N)}^{\rm opt})^s\, \eta_{j(N)}(u_{j(N)}^{{\rm opt},\star}) \\
&\lesssim \sup_{j\in\mathbb{N}_0} \Big(\mathrm{work}(\mathcal{T}_j^{\rm opt})^s\, \eta_j(u_j^{{\rm opt},\star})\Big) < \infty
\end{aligned}$$

This concludes the proof. □

On the other hand, the computational complexity of Algorithm 8.7 to compute an approximation $u_\ell^{\underline{n},\underline{k}}$ depends on the number of preceding adaptive steps as well as on the number of PCG and Picard iterations in each step.

- From now on, we assume that the preconditioner as well as each step of the PCG iteration can be computed in linear complexity $\mathcal{O}(\#\mathcal{T}_\ell)$. This can be guaranteed by using a multilevel additive-Schwarz preconditioner; see, e.g., [FFPS17a, Füh14]. Moreover, we suppose that the evaluation of $\langle \mathfrak{A} u_\ell^{n-1,\underline{k}} - F\,,\, v_\ell\rangle$ and $\eta_\ell(T, v_\ell)$ for one fixed $v_\ell \in \mathcal{X}_\ell$ and $T \in \mathcal{T}_\ell$ is of order $\mathcal{O}(1)$. Recall that $\underline{k}(\ell,n) > 0$ and $\underline{n}(\ell) > 0$ denote the number of PCG resp. Picard steps in the $\ell$-th adaptive step of Algorithm 8.7. In total, we thus require
$$\mathcal{O}\Big(\sum_{m=1}^{\underline{n}(\ell)} \underline{k}(m,\ell)\,\#\mathcal{T}_\ell\Big) = \mathcal{O}\Big(\sum_{m=1}^{\underline{n}(\ell)} \sum_{r=1}^{\underline{k}(\ell,m)} \#\mathcal{T}_\ell\Big)$$
operations to compute the discrete solution $u_\ell^{\underline{n},\underline{k}} \in \mathcal{X}_\ell$.
- We suppose that the construction of the set $\mathcal{M}_\ell$ in Step (v.a) as well as the local mesh-refinement $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$ in Step (v.b) of Algorithm 8.7 are performed in linear complexity $\mathcal{O}(\#\mathcal{T}_\ell)$; see, e.g., [Ste07] with $C_{\rm mark} = 2$ for Step (v.a).

Since one step of the adaptive algorithm depends on the full history of the adaptive meshes, the overall computational cost for $u_\ell^{n,k}$ in the $\ell$-th step of Algorithm 8.7 thus amounts to

$$\mathcal{O}\Big(\sum_{(\ell',n',k')\le(\ell,n,k)} \#\mathcal{T}_\ell\Big). \tag{8.94}$$

The next theorem shows that Algorithm 8.7 realizes every possible single-step complexity rate with respect to the cumulative effort (8.94).

**Theorem 8.32.** *Suppose* (A1)–(A2) *for the operator* $\mathfrak{A}$ *as well as stability* (E1)*, reliability* (E3) *and* (E5) *for the error estimator. Suppose* $\lambda_{\rm Pic}, \lambda_{\rm PCG}$ *are sufficiently small such* *that* (8.68) *and the assumptions of Theorem 8.21 are satisfied. Then, for all* $s > 0$*, it holds that*

$$\|u^\star\|_{\mathbb{A}_s} < \infty \quad \Longrightarrow \quad \exists C_{\rm work} > 0 \;\; \forall (\ell, n, k) \in \mathcal{I} \quad \eta_\ell(u_\ell^{n,k}) \le C_{\rm work} \Big( \sum_{(\ell',n',k') \le (\ell,n,k)} \#\mathcal{T}_{\ell'} \Big)^{-s} \tag{8.95}$$

*There holds* $C_{\rm work} = C'_{\rm work} \, \|u^\star\|_{\mathbb{A}_s}$*, where* $C'_{\rm work} > 0$ *depends only on* $\theta$*,* $\lambda_{\rm PCG}$*,* $\lambda_{\rm Pic}$*,* $s$*,* $q_{\rm pcg}$*,* (E1)–(E4)*,* (A1)–(A2)*, and* (R3)–(R5)*, as well as on* $\mathcal{T}_0$*,* $\underline{n}(0)$ *and* $\underline{k}(1,0)$*.*

*Proof.* Let $\|u^\star\|_{\mathbb{A}_s} < \infty$. First, we note that is holds that $\underline{n}(0)\,\underline{k}(1,0)\,\#\mathcal{T}_0 \lesssim \#\mathcal{T}_1$ for some constant depending on $\underline{n}(0)$, $\underline{k}(1,0)$ and $\#\mathcal{T}_0$. Using $\#\mathcal{T}_j \le \#\mathcal{T}_0(\#\mathcal{T}_j - \#\mathcal{T}_0 + 1)$ (see, e.g., Lemma 4.19), we obtain that

$$\sum_{(\ell',n',k') \le (\ell,n,k)} \#\mathcal{T}_{\ell'} \lesssim \sum_{(\ell',n',k')=(1,1,0)}^{(\ell,n,k)} \#\mathcal{T}_{\ell'} \overset{(4.26)}{\lesssim} \sum_{(\ell',n',k')=(1,1,0)}^{(\ell,n,k)} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1). \tag{8.96}$$

Optimal convergence in Proposition 8.27 implies further

$$\eta_{\ell'}(u_{\ell'}^{n',k'}) \overset{(8.80)}{\lesssim} (\#\mathcal{T}_{\ell'} - \#\mathcal{T}_0 + 1)^{-s} \quad \text{for all } (\ell', n', k') \in \mathcal{I}.$$

In combination with (8.96) and inverse summability from Theorem 8.30, we obtain that

$$\sum_{(\ell',n',k') \le (\ell,n,k)} \#\mathcal{T}_{\ell'} \lesssim \sum_{(\ell',n',k')=(1,1,0)}^{(\ell,n,k)} \eta_{\ell'}(u_{\ell'}^{n',k'})^{-1/s} \overset{(8.90)}{\lesssim} \eta_\ell(u_\ell^{n,k})^{-1/s}.$$

This concludes the proof. □

**Remark 8.33.** *In order to interpret the result of Theorem 8.32, we want to consider the following heuristic comparison. Suppose we have an oracle, which can do the following:*

- *Given* $N$*, the oracle produces the optimal mesh* $\mathcal{T}_{{\rm opt},N} \in \mathbb{T}_N$ *with* $\eta_{N,{\rm opt}} := \eta(\mathcal{T}_{\rm opt}, u^\star_{\rm opt}) = \min_{\mathcal{T}_\circ \in \mathbb{T}_N} \eta(\mathcal{T}_\circ, u^\star_\circ)$ *without any computational cost.*
- *The oracle can compute the exact solution of the nonlinear equation and the corresponding estimator in linear complexity* $\mathcal{O}(\mathcal{T}_{\rm opt})$*.*

*Theorem 8.32 shows that every single-step complexity rate* $s > 0$ *which can be achieved by using the oracle, is also realized by Algorithm 8.7, even with respect to the cumulative effort* (8.94)*.*

# 9 Adaptive FEM with fixpoint iteration

In this chapter, we focus once again on AFEM. As an application of the preceding analysis, we show that certain types of nonlinear PDEs fit in the abstract framework of Chapter 8. As model problems serve nonlinear boundary value problems similar to those of [GMZ11, GMZ12, BSF+14, CW17]; see (9.1). Then, applying Algorithm 8.7 leads to optimal algebraic convergence rates and optimal complexity for the underlying error estimator.

**Outline of this chapter.** Section 9.1 introduces the model problem and Section 9.1.2 the corresponding *a posteriori* error indicator. In Section 9.2, we prove that all operator axioms (A1)–(A3) and estimator axioms (E1)–(E4) of Chapter 8 are met. Section 9.3 recaps the main results on linear and optimal algebraic convergence. Finally, Section 9.4 underpins our theoretical findings with numerical experiments for AFEM in $\mathbb{R}^2$.

## 9.1 Model problem

Let $\Omega \subset \mathbb{R}^d$ with $d \in \{2,3\}$ be a bounded Lipschitz domain with polyhedral boundary $\Gamma = \partial\Omega$. To include also mixed boundary conditions, suppose that $\Gamma := \overline{\Gamma}_D \cup \overline{\Gamma}_N$ is split into relatively open and disjoint Dirichlet and Neumann boundaries $\Gamma_D, \Gamma_N \subseteq \Gamma$ with $|\Gamma_D| > 0$. For given $f \in L^2(\Omega)$, we consider second-order nonlinear problems of the following type:

$$\begin{aligned} -\operatorname{div}(\mu(x, |\nabla u^\star(x)|^2)\, \nabla u^\star(x)) &= f(x) && \text{in } \Omega, \\ u^\star(x) &= 0 && \text{on } \Gamma_D, \\ \mu(x, |\nabla u^\star(x)|^2)\partial_{\mathbf{n}} u^\star(x) &= g(x) && \text{on } \Gamma_N. \end{aligned} \tag{9.1}$$

As in [GMZ12], we suppose that the scalar nonlinearity $\mu : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies the following properties (M1)–(M4).

**M1)** There exist constants $0 < \gamma_{\mathrm{low}} < \gamma_{\mathrm{up}} < \infty$ such that

$$\gamma_{\mathrm{low}} \leq \mu(x,t) \leq \gamma_{\mathrm{up}} \quad \forall x \in \Omega \text{ and } \forall t \geq 0. \tag{9.2}$$

**M2)** There holds $\mu(x,\cdot) \in C^1(\mathbb{R}_{\geq 0}, \mathbb{R})$ for all $x \in \Omega$, and there exist $0 < \widetilde{\gamma}_{\mathrm{low}} < \widetilde{\gamma}_{\mathrm{up}} < \infty$ such that

$$\widetilde{\gamma}_{\mathrm{low}} \leq \mu(x,t) + 2t\frac{\mathrm{d}}{\mathrm{d}t}\mu(x,t) \leq \widetilde{\gamma}_{\mathrm{up}} \quad \text{for all } x \in \Omega \text{ and all } t \geq 0. \tag{9.3}$$

**M3)** Lipschitz-continuity of $\mu(x,t)$ in $x$, i.e., there exists $L_\mu > 0$ such that

$$|\mu(x,t) - \mu(y,t)| \leq L_\mu |x-y| \quad \forall x, y \in \Omega \text{ and } \forall t \geq 0. \tag{9.4}$$

**M4)** Lipschitz-continuity of $t\frac{\mathrm{d}}{\mathrm{d}t}\mu(x,t)$ in $x$, i.e., there exists $\widetilde{L}_\mu > 0$ such that

$$\Big|t\frac{\mathrm{d}}{\mathrm{d}t}\mu(x,t) - t\frac{\mathrm{d}}{\mathrm{d}t}\mu(y,t)\Big| \le \widetilde{L}_\mu|x-y| \quad \forall x,y \in \Omega \text{ and } \forall t \ge 0. \tag{9.5}$$

### 9.1.1 Weak formulation

To obtain the weak formulation of (9.1), we introduce the space of $H^1(\Omega)$-functions with homogeneous Dirichlet data

$$H^1_D(\Omega) := \{w \in H^1 : (\gamma_0^{\mathrm{int}} w)|_{\Gamma_D} = 0\},$$

where $\gamma_0^{\mathrm{int}} : H^1(\Omega) \to H^{1/2}(\Omega)$ denotes the interior trace operator, cf. Section 2.2. The weak formulation of (9.1) reads as follows: Given $f \in L^2(\Omega)$ and $g \in L^2(\Gamma)$, find $u \in H^1_D(\Omega)$ such that

$$\int_\Omega \mu(x, |\nabla u^\star(x)|^2)\,\nabla u^\star \cdot \nabla v\,\mathrm{d}x = \int_\Omega fv\,\mathrm{d}x + \int_{\Gamma_N} gv\,\mathrm{d}s \quad \text{for all } v \in H^1_D(\Omega). \tag{9.6}$$

Recall the abstract framework from Chapter 8. We define $\mathcal{H} := H^1_D(\Omega)$ with corresponding norm $\|v\|_{\mathcal{H}} := \|\nabla v\|_{L^2(\Omega)}$ and let $\langle \cdot\,,\,\cdot\rangle$ denote the extended $L^2(\Omega)$ scalar product. This gives rise to the operators

$$\langle \mathfrak{A}w\,,\,v\rangle := \int_\Omega \mu(x, |\nabla w(x)|^2)\,\nabla w(x) \cdot \nabla v(x)\,\mathrm{d}x, \tag{9.7a}$$

$$\langle F\,,\,v\rangle := \int_\Omega fv\,\mathrm{d}x + \int_\Omega gv\,\mathrm{d}s\,. \tag{9.7b}$$

In the framework of the abstract setting of Chapter 8, the weak formulation (9.6) can equivalently be stated as

$$\langle \mathfrak{A}\,u^\star\,,\,v\rangle = \langle F\,,\,v\rangle \quad \text{for all } v \in H^1_D(\Omega). \tag{9.8}$$

Let $\mathcal{T}_0$ be an initial conforming triangulation of $\Omega$. Analogously to Chapter 5, we use NVB for mesh-refinement. Then, Section 3.5 guarantees the validity of the refinement axioms (R1)–(R6). Let $\mathcal{T}_\bullet \in \mathbb{T} := \mathrm{refine}(\mathcal{T}_0)$ be a conforming triangulation of $\Omega$. We employ lowest-order finite element spaces

$$\mathcal{X}_\bullet := \mathcal{S}^1_D(\mathcal{T}_\bullet) := \mathcal{P}^1(\mathcal{T}_\bullet) \cap H^1_D(\Omega).$$

The discrete formulation corresponding to (9.8) reads as follows: Given $f$ and $g$, find $u^\star_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet)$ such that

$$\langle \mathfrak{A}\,u^\star_\bullet\,,\,v_\bullet\rangle = \langle F\,,\,v_\bullet\rangle \quad \text{for all } v_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet).$$

In order to apply the abstract analysis of Chapter 8, it remains to prove the operator properties (A1)–(A3) of Section 9.2.

### 9.1.2 Weighted-residual error estimator

Once again we consider the weighted-residual error estimator. To abbreviate notation, let $\mu_v(x) := \mu(x, |\nabla v(x)|^2)$ for all $v \in H^1_D(\Omega)$. The element contributions for an arbitrary discrete function $v_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet)$ are given by

$$\begin{aligned}\eta_\bullet(T, v_\bullet)^2 := h_T^2 \|f + \operatorname{div}(\mu_{v_\bullet} \nabla v_\bullet)\|^2_{L^2(T)} + h_T \|[\mu_{v_\bullet} v_\bullet \cdot \mathbf{n}]\|^2_{L^2(\partial T \cap \Omega)} \\ + h_T \|g - \mu_{v_\bullet} \nabla v_\bullet \cdot \mathbf{n}\|^2_{L^2(\partial T \cap \Gamma_N)}.\end{aligned} \tag{9.9}$$

where $[(\cdot) \cdot \mathbf{n}]$ denotes the normal jump over interior facets and $h_T := |T|^{1/d} \simeq \operatorname{diam}(T)$ denotes the local mesh-size (see Chapter 2–3).

## 9.2 Verification of the axioms

In this section we prove that the model problem (9.1) as well as the *a posteriori* error estimator of (9.9) fit in the abstract framework of Chapter 8. To that end, we first show the operator axioms (A1)–(A3) for the nonlinear operator $\mathfrak{A}$. Afterwards, we verify the axioms (E1)–(E4).

### 9.2.1 Verification of (A1)–(A3)

To prove (A1)–(A3), we first recall an auxiliary lemma which is just a simplified version of [LB96, Lemma 2.1] with $p := 2$ and $\delta := 0$.

**Lemma 9.1.** *Let $C_1 > 0$ as well as $0 < C_2 \le C_3 < \infty$. Further, suppose that $\kappa(x, \cdot) \in C^1(\mathbb{R}_{\ge 0}, \mathbb{R}_{\ge 0})$ satisfies that $\kappa(x,t) \le C_1$ for all $x \in \Omega$ and $t \ge 0$ as well as*

$$C_2 \le \frac{\mathrm{d}}{\mathrm{d}t}(t\kappa(x,t)) \le C_3 \quad \forall x \in \Omega \text{ and } \forall t \ge 0. \tag{9.10}$$

*Then, it holds that*

$$\big(\kappa(x,|y|)y - \kappa(x,|z|)z\big) \cdot \big(y - z\big) \ge C_2\, |y-z|^2 \quad \forall x \in \Omega \text{ and } \forall y, z \in \mathbb{R}^d, \tag{9.11}$$

*as well as*

$$\big|\kappa(x,|y|)y - \kappa(x,|z|)z\big| \le C_1\, |y-z| \quad \forall x \in \Omega \text{ and } \forall y, z \in \mathbb{R}^d. \tag{9.12}$$

□

Lemma 9.1 gives rise to the following proposition, which proves (A1)–(A3).

**Proposition 9.2.** *Suppose that $\mu : \Omega \times \mathbb{R}_{\ge 0} \to \mathbb{R}$ satisfies* (M1)–(M4)*. Then, the corresponding (nonlinear) operator $\mathfrak{A}$ satisfies the operator axioms* (A1)–(A3) *with constants $\alpha := \gamma_{\mathrm{low}}$ and $L := \widetilde{\gamma}_{\mathrm{up}}$.*

*Proof.* We split the proof into two steps.

**Step 1: Proof of (A1)–(A2).** Define $\kappa(x,t) := \mu(x,t^2)$. Note that (M1)–(M2) with $\frac{\mathrm{d}}{\mathrm{d}t}(t\kappa(x,t)) = \mu(x,t^2) + 2t^2\partial_2\mu(x,t^2)$ yield that

$$\kappa(x,t) \leq \gamma_{\mathrm{up}} \quad \text{and} \quad \widetilde{\gamma}_{\mathrm{low}} \leq \frac{\mathrm{d}}{\mathrm{d}t}(t\kappa(x,t)) \leq \widetilde{\gamma}_{\mathrm{up}}.$$

Hence, the assumptions of Lemma 9.1 are satisfied. The lemma implies for all $v, w \in H^1_D(\Omega)$ that

$$\alpha|\nabla v - \nabla w|^2 \overset{(9.11)}{\leq} \big(\mu(\cdot,|\nabla v|^2)\nabla v - \mu(\cdot,|\nabla w|^2)\nabla w\big) \cdot \big(\nabla v - \nabla w\big) \quad \text{a.e. in } \Omega, \tag{9.13}$$

as well as

$$\big|\mu(\cdot,|\nabla v|^2)\nabla v - \mu(\cdot,|\nabla w|^2)\nabla w\big|^2 \overset{(9.12)}{\leq} L^2\big|\nabla v - \nabla w\big|^2 \quad \text{a.e. in } \Omega. \tag{9.14}$$

Integration over $\Omega$ proves strong monotonicity (A1) and Lipschitz continuity (A2).

**Step 2: Proof of (A3).** Analogously to [Has10], we define

$$P : H^1_D(\Omega) \to \mathbb{R}_{\geq 0} : \quad w \mapsto \frac{1}{2}\int_\Omega \int_0^{|\nabla w|^2} \mu(x,\zeta)\,\mathrm{d}\zeta\,\mathrm{d}x\,. \tag{9.15}$$

Note that boundedness (M1) implies well-posedness of $P$. Next, we show that $\mathfrak{A}$ is the Gâteaux-derivative $dP$ of $P$. To that end, let $r > 0$ and $v, w \in H^1_D(\Omega)$. Define

$$H(r) := P(w + rv) \overset{(9.15)}{=} \frac{1}{2}\int_\Omega \int_0^{|\nabla w + r\nabla v|^2} \mu(x,\zeta)\,\mathrm{d}\zeta\,\mathrm{d}x\,.$$

With the Leibniz rule, we get

$$\begin{aligned} H'(r) &= \frac{1}{2}\int_\Omega \mu(x,|\nabla w + r\nabla v|^2)\frac{\mathrm{d}}{\mathrm{d}r}\big(|\nabla w + r\nabla v|^2\big)\,\mathrm{d}x \\ &= \int_\Omega \mu(x,|\nabla w + r\nabla v|^2)\,(\nabla w + r\nabla v)\cdot\nabla v\,\mathrm{d}x\,. \end{aligned}$$

With this, we obtain that

$$\langle dP(w)\,,\,v\rangle = H'(0) = \int_\Omega \mu(x,|\nabla w|^2)\,\nabla w\cdot\nabla v\,\mathrm{d}x \overset{(9.7a)}{=} \langle \mathfrak{A}\,w\,,\,v\rangle.$$

This concludes the proof. $\square$

### 9.2.2 Verification of (E1)–(E4)

The verification of stability (E1) and reduction (E2) requires the validity of a certain inverse estimate. We recall the following result from [GMZ12, Lemma 3.7].

**Lemma 9.3.** *Let $\mathcal{T}_\bullet \in \mathbb{T}$. Define $\rho_\bullet : \mathcal{T}_\bullet \times \mathcal{S}^1_D(\mathcal{T}_\bullet) \times \mathcal{S}^1_D(\mathcal{T}_\bullet)$ for all $T \in \mathcal{T}_\bullet$ and all $v_\bullet, w_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet)$ by*

$$\begin{aligned}\rho_\bullet(T, v_\bullet, w_\bullet) := h_T \| \operatorname{div}(\mu_{v_\bullet} \nabla v_\bullet) - \operatorname{div}(\mu_{w_\bullet} \nabla w_\bullet)\|_{L^2(T)} \\ + h_T^{1/2} \|[\mu_{v_\bullet} v_\bullet \cdot \mathbf{n}] - [\mu_{w_\bullet} w_\bullet \cdot \mathbf{n}]\|_{L^2(\partial T)}.\end{aligned}$$

*Then, there exists $C_{\mathrm{inv}} > 0$ such that for all $T \in \mathcal{T}_\bullet$ and $v_\bullet, w_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet)$ it holds that*

$$\rho_\bullet(T, v_\bullet, w_\bullet) \le C_{\mathrm{inv}} \|\nabla v_\bullet - \nabla w_\bullet\|_{L^2(\omega(T))}. \tag{9.16}$$

*Moreover, there exists $C_E > 1$, which depends only on $d$ and $\gamma$-shape regularity such that*

$$\sum_{T \in \mathcal{T}_\bullet} \rho_\bullet(T, v_\bullet, w_\bullet) \le C_E \|\nabla v_\bullet - \nabla w_\bullet\|_{L^2(\Omega)}. \tag{9.17}$$

The inverse estimate gives rise to the estimator axioms (E1)–(E4). We emphasize that the proofs are similar to the linear case; see, e.g., Section 5.2 or [CFPP14, CKNS08]. A proof for scalar nonlinearities can be found in [GMZ12, Lemma 3.7]. For sake of the completeness, we include the proof of stability (E1).

**Proposition 9.4** (stability on non-refined element domains). *Suppose* (M1)–(M4)*. Then the error estimator defined in* (9.9) *satisfies axiom* (E1).

*Proof.* Let $\mathcal{T}_\bullet, \mathcal{T}_\circ \in \mathbb{T}$ with $\mathcal{T}_\circ \in \operatorname{refine}(\mathcal{T}_\bullet)$. Let $\mathcal{U}_\bullet \subseteq \mathcal{T}_\bullet \cap \mathcal{T}_\circ$, and $v_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet)$ as well as $v_\circ \in \mathcal{S}^1_D(\mathcal{T}_\circ)$. Analogously to the proof of Proposition 5.3 we obtain

$$|\eta_\circ(\mathcal{U}_\bullet, v_\circ) - \eta_\bullet(\mathcal{U}_\bullet, v_\bullet)| \le \Big(\sum_{T \in \mathcal{U}_\bullet} R_\bullet(v_\circ, v_\bullet, T)^2\Big)^{1/2}.$$

Using $\rho_\bullet$ from Lemma 9.3, the function $R_\bullet(\cdot,\cdot,\cdot)$ is given by

$$R_\bullet(T, v_\circ, v_\bullet) := \rho_\bullet(T, v_\circ, v_\bullet) + h_T^{1/2} \|\mu_{v_\circ} v_\circ \cdot \nabla \mathbf{n} - \mu_{v_\bullet} \nabla v_\bullet \cdot \mathbf{n}\|_{L^2(\partial T \cap \Gamma_N)}.$$

Recall that each element patch contains at most finitely many elements. Note that nestedness of the discrete spaces implies that $v_\bullet \in \mathcal{S}^1_D(\mathcal{T}_\bullet) \subseteq \mathcal{S}^1_D(\mathcal{T}_\circ)$. The inverse estimate (9.16) applied to the first term on the right-hand side yields that

$$\sum_{T \in \mathcal{U}_\bullet} \rho_\bullet(T, v_\circ, v_\bullet)^2 \overset{(9.16)}{\lesssim} \sum_{T \in \mathcal{U}_\bullet} \|\nabla(v_\circ - v_\bullet)\|^2_{L^2(\omega(T))} \lesssim \|\nabla(v_\circ - v_\bullet)\|^2_{L^2(\Omega)}$$

The remaining term right-hand side is estimated analogously to (5.11). Using Lemma 9.1 and the trace inequality, we obtain for an edge $E \subset \partial T_E \cap \Gamma_N$ that

$$\begin{aligned}\|(\mu_{v_\circ} \nabla v_\circ - \mu_{v_\bullet} \nabla v_\bullet) \cdot \mathbf{n}\|^2_{L^2(E)} \le \|\mu_{v_\circ} \nabla v_\circ - \mu_{v_\bullet} \nabla v_\bullet\|^2_{L^2(E)} &\overset{(9.12)}{\le} L^2 \|\nabla(v_\circ - v_\bullet)\|^2_{L^2(E)} \\ &\lesssim h_{T_E}^{-1} \|v_\circ - v_\bullet\|^2_{L^2(T_E)}.\end{aligned}$$

Combining this with the latter estimate, we end up with

$$\sum_{T \in \mathcal{U}_\bullet} R_\bullet(T, v_\circ, v_\bullet)^2 \lesssim \|v_\circ - v_\bullet\|^2_{H^1(\Omega)},$$

which concludes the proof. ∎

With the inverse inequality (9.16), the proof of (E2) follows analogously to the linear case in Proposition 5.4. The Neumann term in the estimator (9.9) is estimated similarly to the proof of Proposition 9.4, where we additionally use the reduction of the local mesh size (3.3) on refined elements.

The well-posedness of the error estimator requires that the nonlinearity $\mu(x,t)$ is Lipschitz continuous in $x$, i.e. (M3). Then, reliability (E3) and discrete reliability (E4) are proved as in the linear case. We refer to, e.g., [CKNS08] for the linear case or [GMZ12, Theorem 3.3] and [GMZ12, Theorem 3.4] for strongly monotone nonlinearities. We emphasize that the arising constants in (E1)–(E4) depend also on the uniform $\gamma$-shape regularity of the triangulations generated by NVB.

## 9.3 Optimal convergence

Recall that NVB bisection guarantees the refinement axioms (R1)–(R6). With the operator axioms (A1)–(A3) as well as the estimator axioms (E1)–(E4) at hand, we can apply the abstract framework and analysis of Chapter 8. Let $u_\ell^{n,k} \in \mathcal{S}^1_D(\mathcal{T}_\ell)$ denote the sequence generated by Algorithm 8.7 applied to model problem (9.1). The following theorem recaps the main results of the previous chapter and is a direct consequence of Proposition 8.27, Theorem 8.30, and Theorem 8.32.

**Theorem 9.5.** *Suppose* (M1)–(M4) *for* $\mu(\cdot,\cdot)$. *Let* $0 < \lambda_{\rm Pic}, \lambda_{\rm PCG}, \theta < 1$ *be sufficiently small such that* (8.68) *and the assumptions of Theorem 8.21 are satisfied. Then, the sequence* $u_\ell^{n,k} \in \mathcal{S}^1_D(\mathcal{T}_\ell)$ *produced by Algorithm 8.7 satisfies linear convergence*

$$\eta_{\ell+j}(u_{\ell+j}^{\underline{n},\underline{k}}) \le C_{\rm lin}\, q_{\rm lin}^j\, \eta_\ell(u_\ell^{\underline{n},\underline{k}}) \quad \text{for all } (\ell,\underline{n},\underline{k}) \in \mathcal{I}. \tag{9.18}$$

*Moreover, for all* $s > 0$ *with* $\|u^\star\|_{\mathbb{A}_s} < \infty$, *there holds optimal algebraic convergence w.r.t. the degrees of freedom*

$$\eta_\ell(u_\ell^{n,k}) \lesssim \Big(\#\mathcal{T}_\ell - \#\mathcal{T}_0 + 1\Big)^{-s} \quad \text{for all } (\ell,n,k) \in \mathcal{I}. \tag{9.19}$$

*If we additionally assume* (E5), *we obtain optimal complexity*

$$\eta_\ell(u_\ell^{n,k}) \lesssim \Big(\sum_{(\ell,n,k)\ge(0,1,0)} \#\mathcal{T}_\ell\Big)^{-s} \text{for all } (\ell,n,k) \in \mathcal{I}. \tag{9.20}$$

*The involved constants depends only on* $\theta$, $\lambda_{\rm PCG}$, $\lambda_{\rm Pic}$, $s$, $q_{\rm pic}$, $q_{\rm pcg}$, *the constants in* (E1)–(E4), (A1)–(A2), (R1)–(R5). *The constant in* (9.20) *additionally depends on* $\mathcal{T}_0$, $\underline{n}(0)$, *and* $\underline{k}(1,0)$.

## 9.4 Numerical experiments

In this section, we present two numerical experiments in 2D to underpin our theoretical findings. In the experiments, we compare the performance of Algorithm 8.7 for

- different values of $\lambda_{\rm Pic}^2 \in \{1, 0.1, 0.01, 0.001\}$,

- different values of $\lambda_{\rm PCG}^2 \in \{\lambda_{\rm Pic}^2, 0.1\,\lambda_{\rm Pic}^2, 0.01\,\lambda_{\rm Pic}^2, 0.001\,\lambda_{\rm Pic}^2\}$,
- different values of $\theta \in \{0.2, 0.4, \dots, 1\}$,
- PCG with multilevel additive-Schwarz preconditioning vs. non preconditioned CG,
- nested iteration $u_\ell^{1,0} := u_{\ell-1}^{\underline{n},\underline{k}}$ compared to a naive initial guess $u_\ell^{1,0} := 0$.

As model problems serve nonlinear boundary value problems from [GHPS17], which are similar to those of [GMZ11, GMZ12, BSF$^+$14, CW17]. If it is not stated otherwise, we employ the multi-level additive-Schwarz (MLAS) preconditioner for the PCG-iteration; see, e.g., [Füh14].

### 9.4.1 Experiment with known solution (Ex. 1)

We consider the Z-shaped domain $\Omega \subset \mathbb{R}^2$ from Figure 9.1 (above) with mixed Dirichlet–Neumann boundary and the nonlinear problem (9.1), where the function $\mu(\cdot,\cdot)$ is given by

$$\mu(x, |\nabla u^\star(x)|^2) := 2 + \frac{1}{\sqrt{1 + |\nabla u^\star(x)|^2}}.$$

This choice of $\mu$ leads to $\alpha = 2$ and $L = 3$ in (O1)–(O2). We prescribe the solution $u^\star$ in polar coordinates by

$$u^\star(x, y) = r^\beta \cos\Big(\beta\,\phi\Big), \tag{9.21}$$

where $\beta = 4/7$ and $f$ as well as $g$ in (9.1) are chosen accordingly. We note that $u^\star$ has a generic singularity at the reentrant corner $(x, y) = (0, 0)$. Figure 9.1 (below) shows that in order to produce optimal rates, Algorithm 8.7 heavily refines towards the singularity.

Our empirical observations are the following: Due to the singular behavior of $u^\star$, uniform refinement leads to a reduced convergence rate $\mathcal{O}(N^{-\beta})$ for both, the energy error $\texttt{err}(u_\ell^{\underline{n},\underline{k}})^2 := \|\nabla u^\star - \nabla u_\ell^{\underline{n},\underline{k}}\|_{L^2(\Omega)}^2$ as well as the error estimator $\eta_\ell(u_\ell^{\underline{n},\underline{k}})^2$; see Figure 9.2. On the other hand, the adaptive refinement of Algorithm 8.7 regains the optimal convergence rate $\mathcal{O}(N^{-1})$, independently of the actual choice of $\theta \in \{0.2, 0.4, 0.6, 0.8\}$ and $\lambda_{\rm Pic}^2 \in \{1, 0.1, 0.01, 0.001\}$ if $\lambda_{\rm PCG}$ is chosen accordingly; see Figure 9.4–9.5. As it turns out, only $\lambda_{\rm Pic}^2 \in \{1, 0.1\}$ and $\lambda_{\rm PCG}^2 \geq 0.1\,\lambda_{\rm Pic}^2$ lead to reduced orders of convergence. All other pairings recover the optimal convergence rate of $\mathcal{O}(N^{-1})$.

Further, Figure 9.3 shows that Algorithm 8.7 leads to optimal rate with respect to the cumulative complexity (8.94), for $u_\ell^{\underline{n},\underline{k}}$, where instead uniform refinement leads to reduced order even for linear single step complexity.

Throughout all adaptive steps, Algorithm 8.7 guarantees that the number of Picard and PCG iterations remains bounded for all tested choices of $\lambda_{\rm PCG}, \lambda_{\rm Pic}$; see Figure 9.6. Moreover, we compare the performance of the PCG iteration with MLAS preconditioning and without any preconditioning. According to (8.50), $\lambda_{\rm PCG}$ has to be chosen sufficiently small in order to guarantee optimal rates. Note that $\lambda_{\rm PCG}$ depends on $q_{\rm pcg}$ and hence, on the condition number of the linear system. Figure 9.8 (top) shows that the condition number of the non-preconditioned system increases with $\mathcal{O}(N)$, where instead the MLAS preconditioner leads to a bounded condition number. This leads to $q_{\mathrm{pcg}} \to 1$ as $\ell \to \infty$ for the non-preconditioned CG iteration. As shown in Figure 9.7 and in contrast to PCG, non-preconditioned CG leads to stagnation of the energy error $\|\nabla u^\star - \nabla u_\ell^{n,k}\|_{L^2(\Omega)}^2$ for $\ell \geq \ell_\lambda$ depending on $\lambda_{\mathrm{PCG}}$ and $\lambda_{\mathrm{Pic}}$. Further, Figure 9.8 (bottom) shows that up to $\ell = \ell_\lambda$, the number of CG-iterations increases with $\ell$.

As expected from Remark 8.4, for the naive initial guess $u_{\ell+1}^{1,0} := 0$, the number of Picard iterations grows logarithmically with the number of elements $\#\mathcal{T}_\ell$, while we observe a bounded number of Picard iterations for nested iteration $u_{\ell+1}^{1,0} := u_{\ell-1}^{\underline{n},\underline{k}}$; see Figure 9.9.

### 9.4.2 Experiment with unknown solution (Ex. 2)

We consider the Z-shaped domain $\Omega \subset \mathbb{R}^2$ from Figure 9.1 (above). Moreover, we consider the nonlinear Dirichlet problem (9.1) with $\Gamma = \Gamma_D$ and constant right-hand side $f \equiv 1$, where $\mu(\cdot,\cdot)$ is given by

$$\mu(x, |\nabla u^\star|^2) = 1 + \arctan(|\nabla u^\star|^2).$$

According to [CW17, Example 1], there holds (O1)–(O2) with $\alpha = 1$ and $L = 1 + \sqrt{3}/2 + \pi/3$.

The exact solution is unknown. Therefore, our empirical observations are concerned with the error estimator only; see Figure 9.10–9.12. We derive similar results as in the previous example with known solution. Since we use the same geometry containing a reentrant corner, uniform mesh-refinement leads to a suboptimal rate of convergence $\mathcal{O}(N^{-\beta})$ for $\eta_\ell(u_\ell^{n,k})^2$, while the use of Algorithm 8.7 regains the optimal rate of convergence $\mathcal{O}(N^{-1})$. Figure 9.10 shows the convergence of the estimator sequence $\eta_\ell(u_\ell^{n,k})^2$ with respect to the degrees of freedom for different values of $\theta \in \{0.2, 0.4, 0.6, 0.8, 1\}$. Moreover, the estimator even realizes the optimal rate with respect to the cumulative effort (8.94); see Figure 9.12 (top). According to Figure 9.11, the optimal rate is achieved for all choices of $\lambda_{\mathrm{Pic}}$. A naive initial guess $u_\ell^{1,0} := 0$ for the iterative solver leads to a logarithmic growth of the number of Picard iterations, while the proposed use of nested iteration $u_\ell^{1,0} := u_{\ell-1}^{\underline{n},\underline{k}}$ again leads to bounded iteration numbers for all tested choices of $\lambda_{\mathrm{Pic}}$, see Figure 9.12 (below).

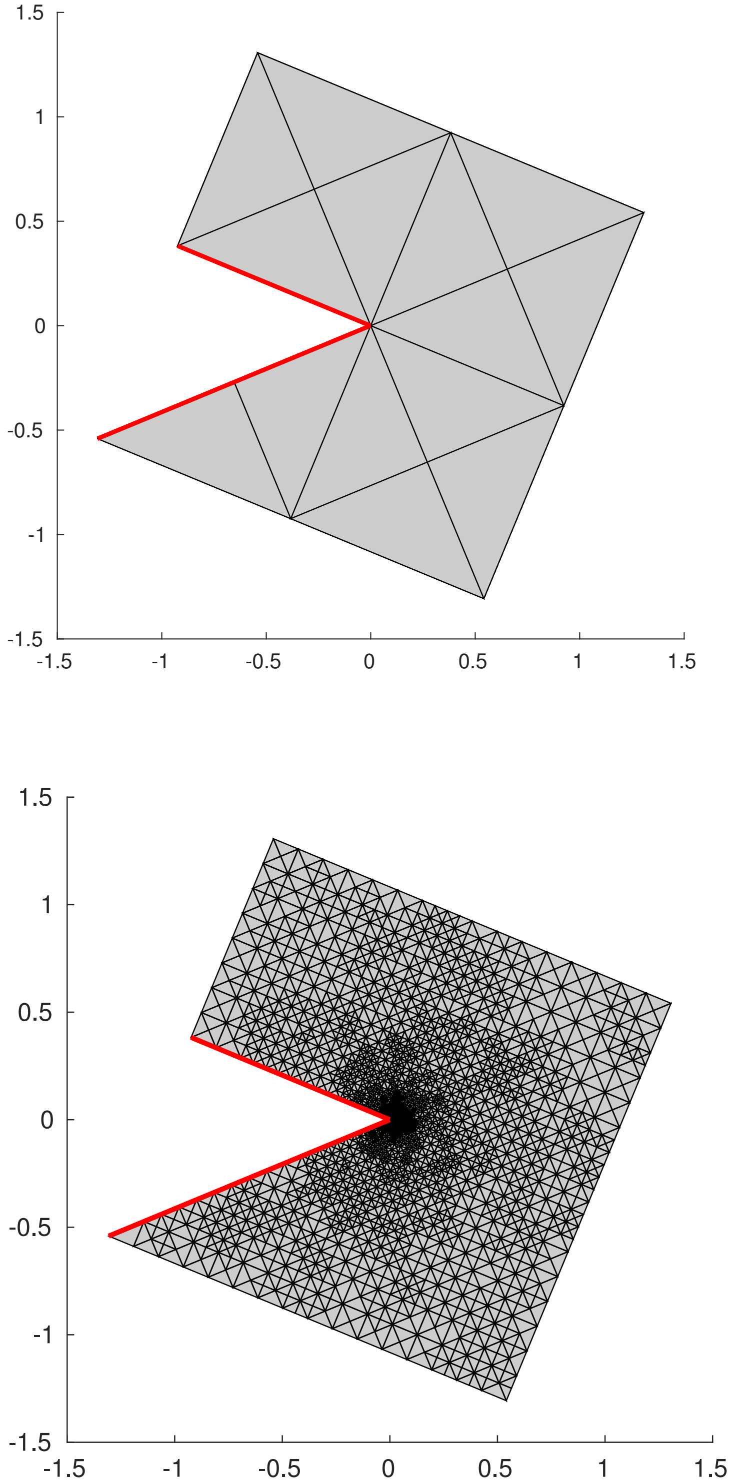



Figure 9.1: Geometry and initial partition $\mathcal{T}_0$ in the experiment from Section 5.4.1 (top). The Dirichlet boundary $\Gamma_D \subset \Gamma$ is marked by a thick red line. In addition we plot the mesh $\mathcal{T}_{18}$ with $\#\mathcal{T}_{18} = 4543$ generated by Algorithm 8.7 (bottom), where we used $\lambda^2_{\mathrm{Pic}} = 10^{-2}$, $\lambda^2_{\mathrm{PCG}} = 10^{-4}$, and $\theta = 0.2$.

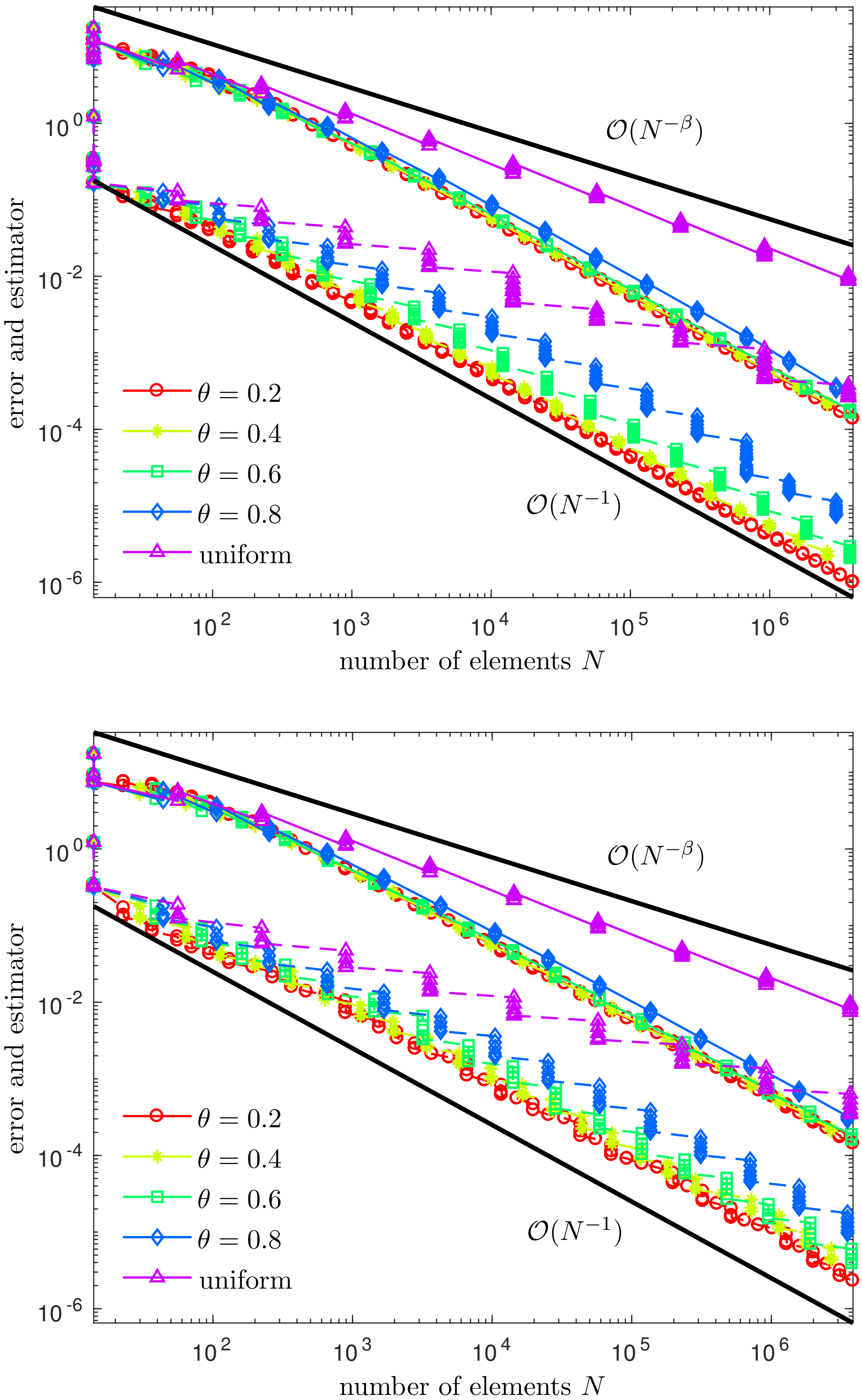
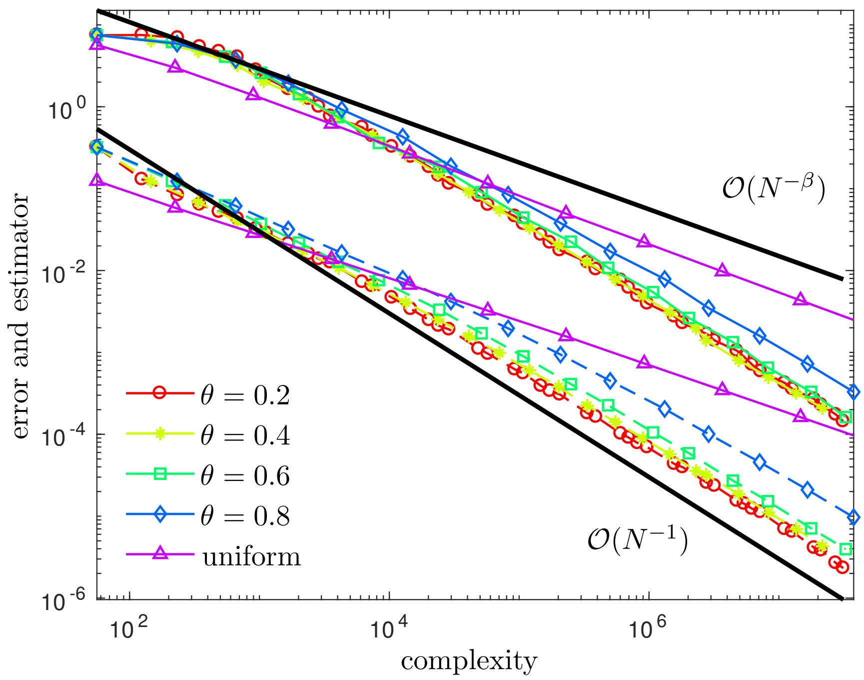



Figure 9.2: Ex. 1: Convergence rates of the full sequence $\eta_\ell(u_\ell^{n,k})^2$ (solid lines) and $\mathtt{err}(u_\ell^{n,k})^2$ (dashed lines) for different values of $\theta$ and $\lambda_{\mathrm{Pic}}^2 = 10^{-2}$ as well as $\lambda_{\mathrm{PCG}}^2 = 10^{-4}$ (top) and $\lambda_{\mathrm{Pic}}^2 = 1$ and $\lambda_{\mathrm{PCG}}^2 = 10^{-3}$ (bottom).



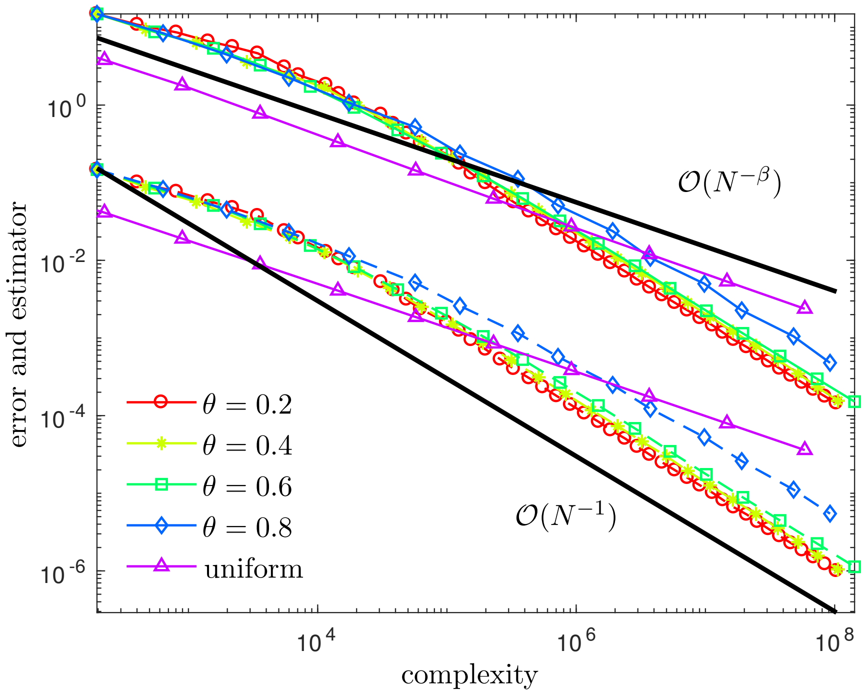
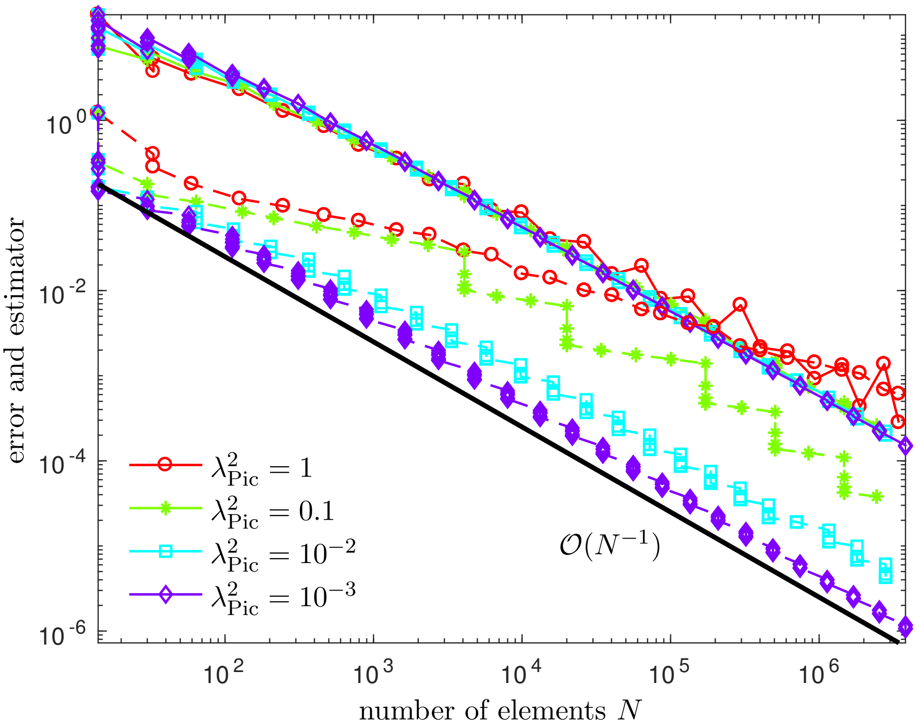




Figure 9.3: Ex. 1: Convergence rates of $\eta_\ell(u_\ell^{\underline{n},\underline{k}})^2$ (solid lines) and $\texttt{err}(u_\ell^{\underline{n},\underline{k}})^2$ (dashed lines) for different values of $\theta$ with respect to the cumulative complexity (8.94). For uniform refinement, we just plot the single step complexity, i.e., $\eta_\ell(u_\ell^{\underline{k},\underline{k}})^2$ w.r.t. $\#\mathcal{T}_\ell$. We used $\lambda_{\rm Pic}^2 = 0.1$, $\lambda_{\rm PCG} = 10^{-2}\,\lambda_{\rm Pic}^2$ (top) as well as $\lambda_{\rm Pic}^2 = 10^{-3}$, $\lambda_{\rm PCG} = 10^{-3}\,\lambda_{\rm Pic}^2$ (bottom).



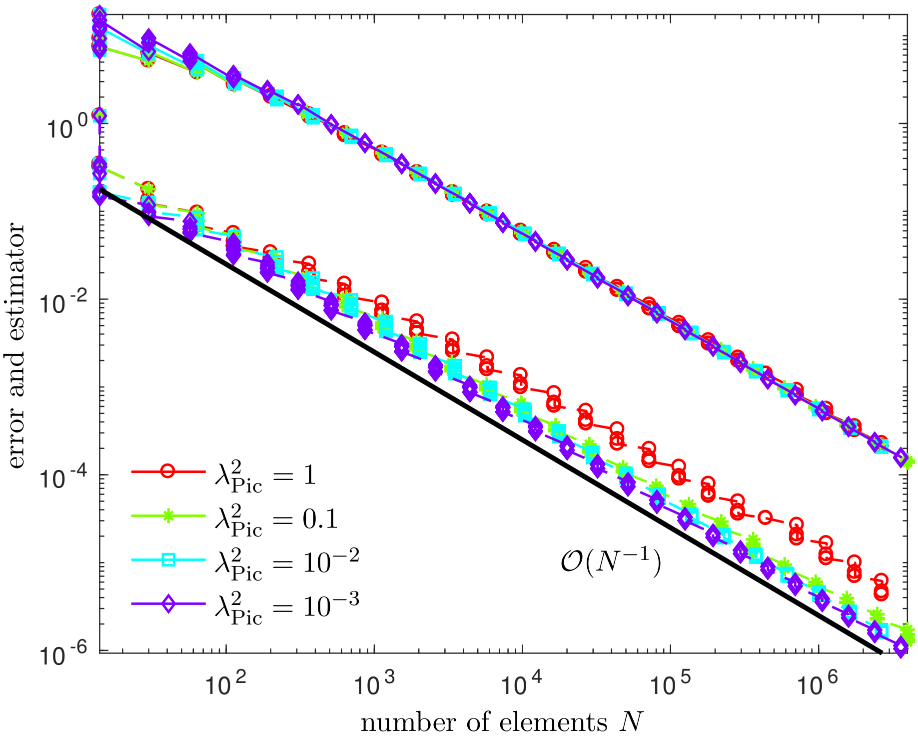




Figure 9.4: Ex. 1: Convergence rates of $\eta_\ell(u_\ell^{n,k})^2$ (solid lines) and $\mathtt{err}(u_\ell^{n,k})^2$ (dashed lines) for different values of $\lambda_{\mathrm{Pic}}$. We used $\theta = 0.4$ as well as $\lambda^2_{\mathrm{PCG}} = 0.1\,\lambda^2_{\mathrm{Pic}}$ (top) and $\lambda^2_{\mathrm{PCG}} = 10^{-3}\,\lambda^2_{\mathrm{Pic}}$ (bottom). For $\lambda_{\mathrm{Pic}} = 1$, the parameter $\lambda_{\mathrm{PCG}}$ has to be sufficiently small in order to get optimal rates.

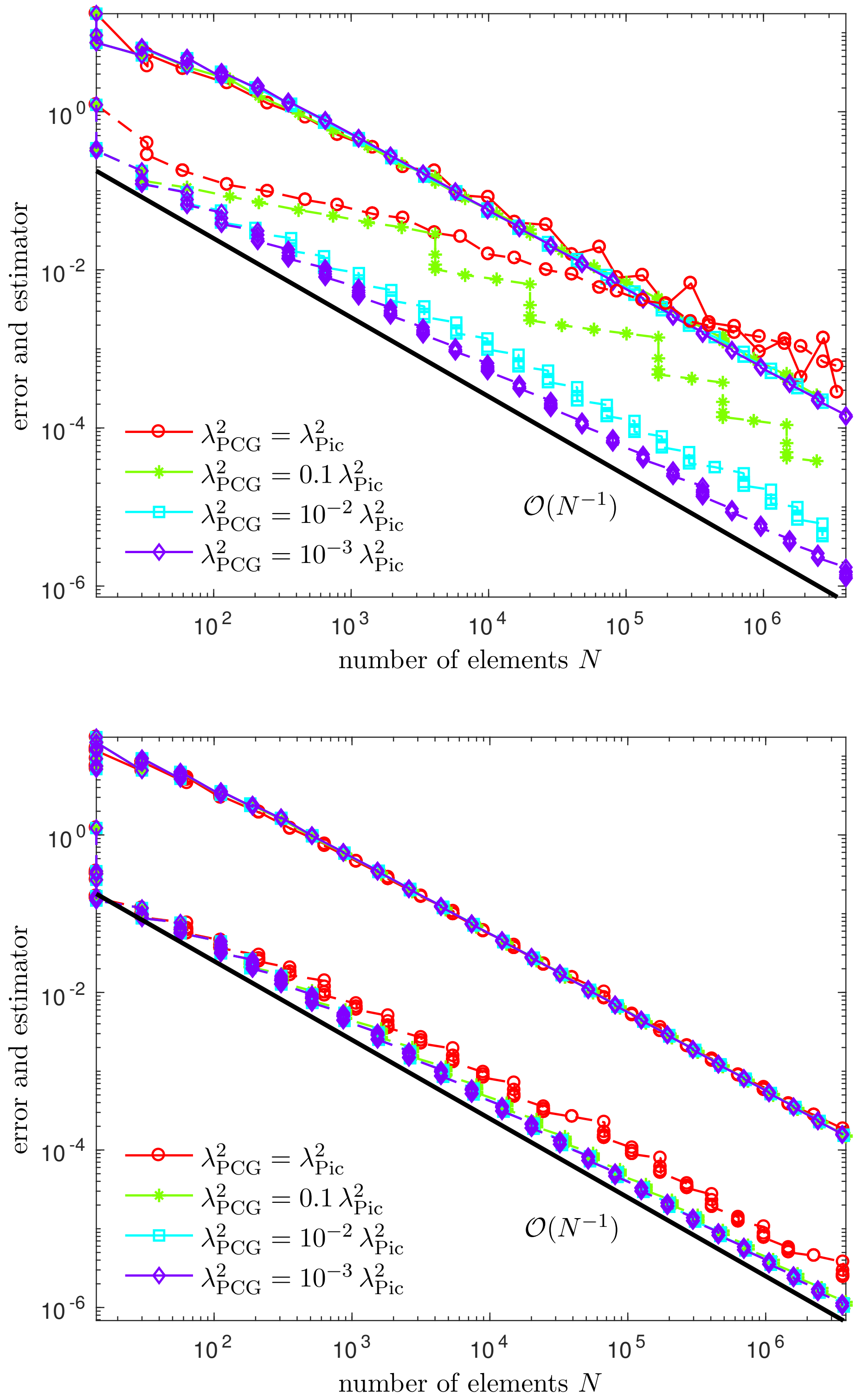



Figure 9.5: Ex. 1: Convergence rates of $\eta_\ell(u_\ell^{n,k})^2$ (solid lines) and $\mathtt{err}(u_\ell^{n,k})^2$ (dashed lines) for $\theta = 0.4$ as well as different values of $\lambda_{\mathrm{PCG}}$ depending on $\lambda_{\mathrm{Pic}} = 0.1$ (top) and $\lambda_{\mathrm{Pic}} = 10^{-3}$ (bottom).

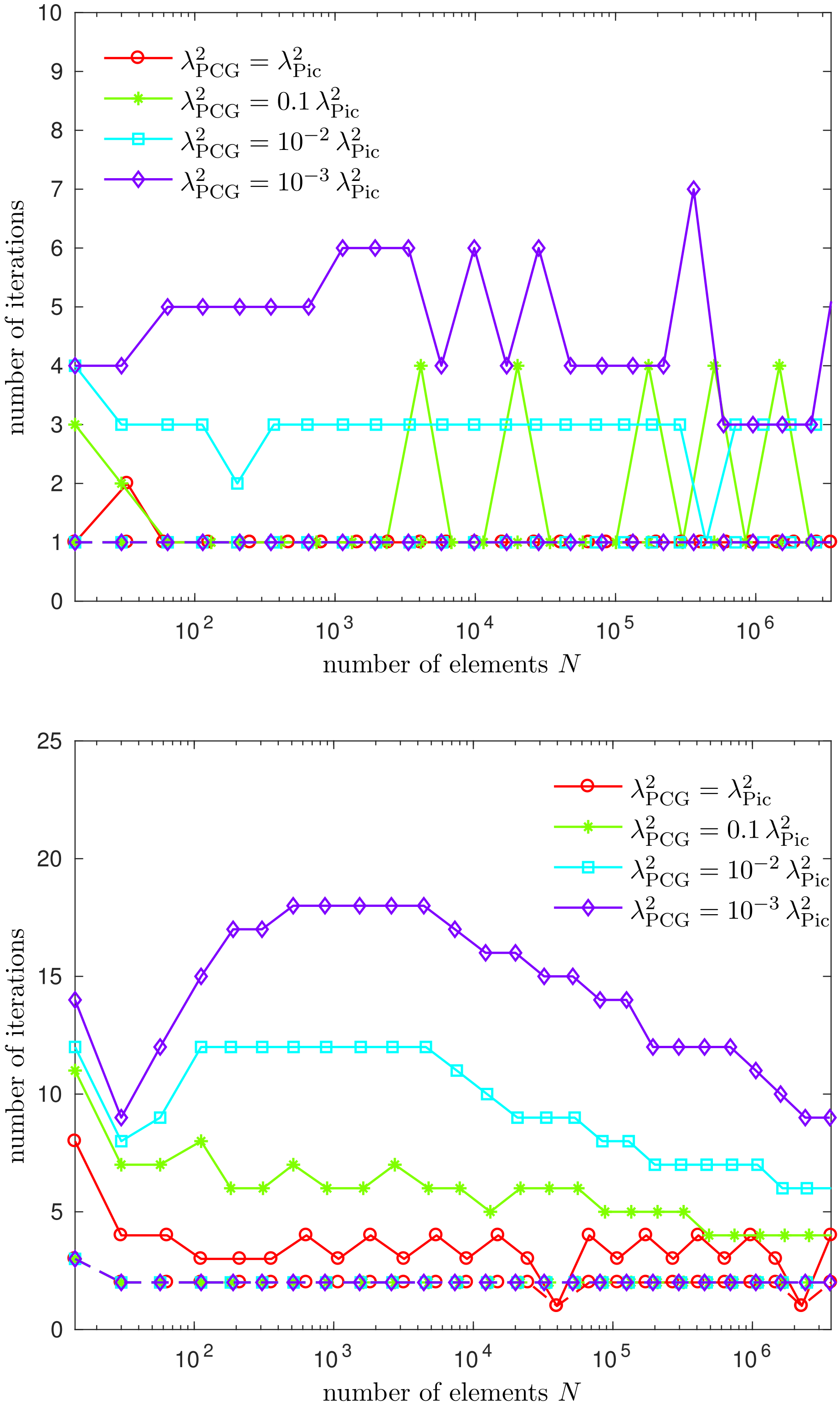
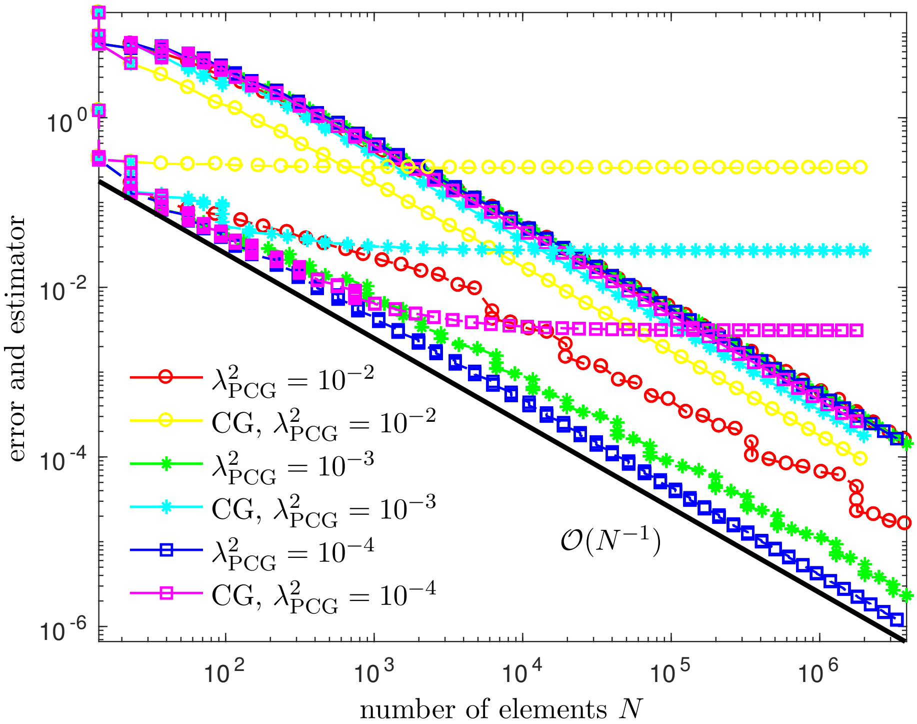



Figure 9.6: Ex. 1: Number of total PCG iterations $\sum_{m=1}^{\underline{n}(\ell)} \underline{k}$ (solid lines) and number total Picard iterations $\underline{n}(\ell)$ (dashed lines) in each adaptive step for $\theta = 0.4$, $\lambda^2_{\mathrm{Pic}} = 0.1$ (top) resp. $\lambda^2_{\mathrm{Pic}} = 10^{-3}$ (bottom) and different values of $\lambda^2_{\mathrm{PCG}}$.



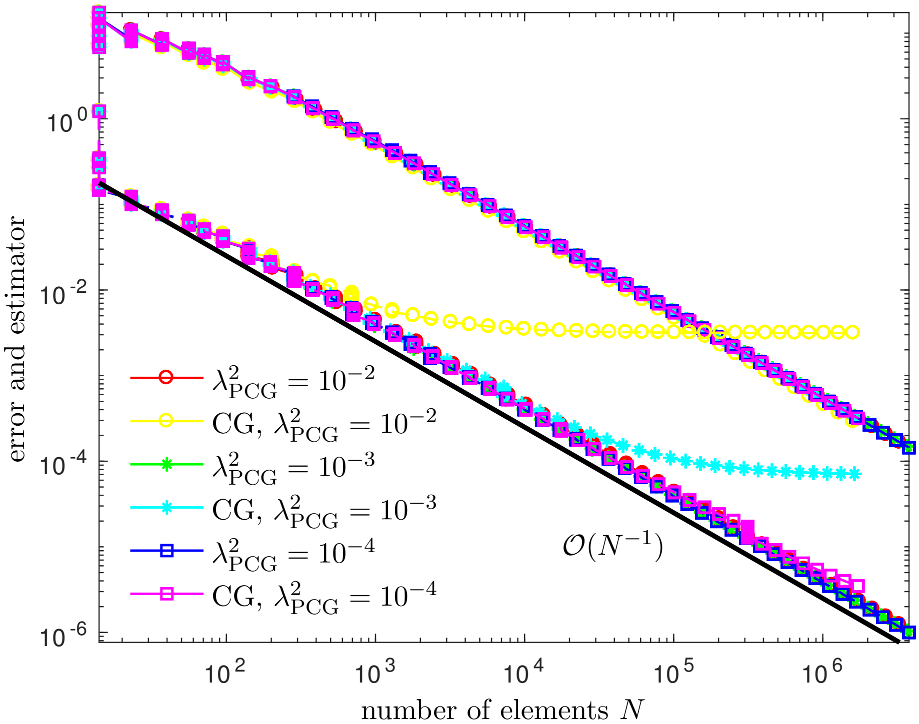




Figure 9.7: Ex. 1: Convergence rates of $\eta_\ell(u_\ell^{n,k})^2$ (solid lines) and $\mathtt{err}(u_\ell^{n,k})^2$ (dashed lines) for different values of $\lambda_{\mathrm{PCG}}$ depending on $\lambda_{\mathrm{Pic}} = 0.1$ (top) and $\lambda_{\mathrm{Pic}} = 10^{-3}$ (bottom) and $\theta = 0.2$. Additionally, we compare PCG to a non-preconditioned CG iteration to solve the linear system in (8.14).

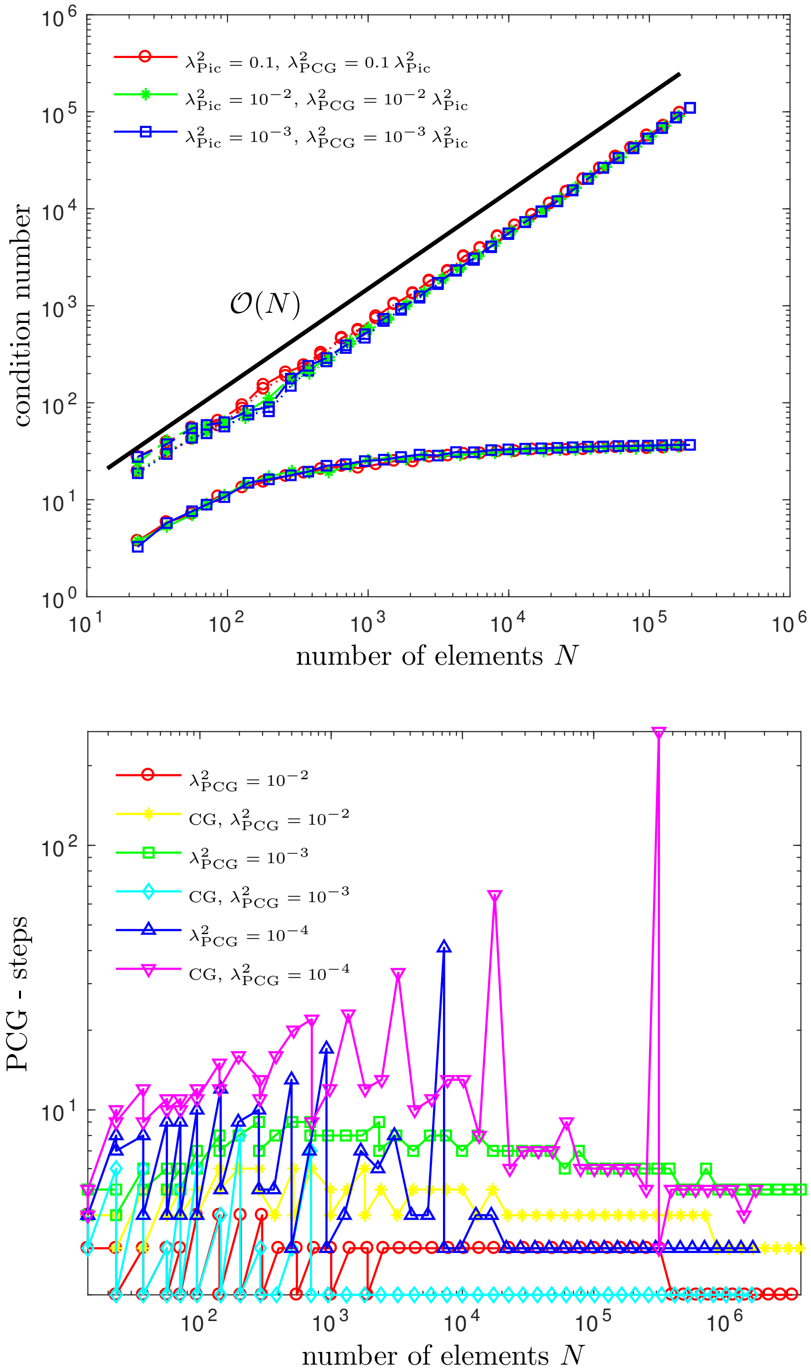



Figure 9.8: Ex. 1: Condition number estimate (top) for the linear system in (8.14) for MLAS-preconditioner (solid lines) to diagonal-preconditioner (dashed lines), and no preconditioning (dotted lines). Number of steps of the PCG-iteration (bottom) compared to non-preconditioned CG in (8.14).

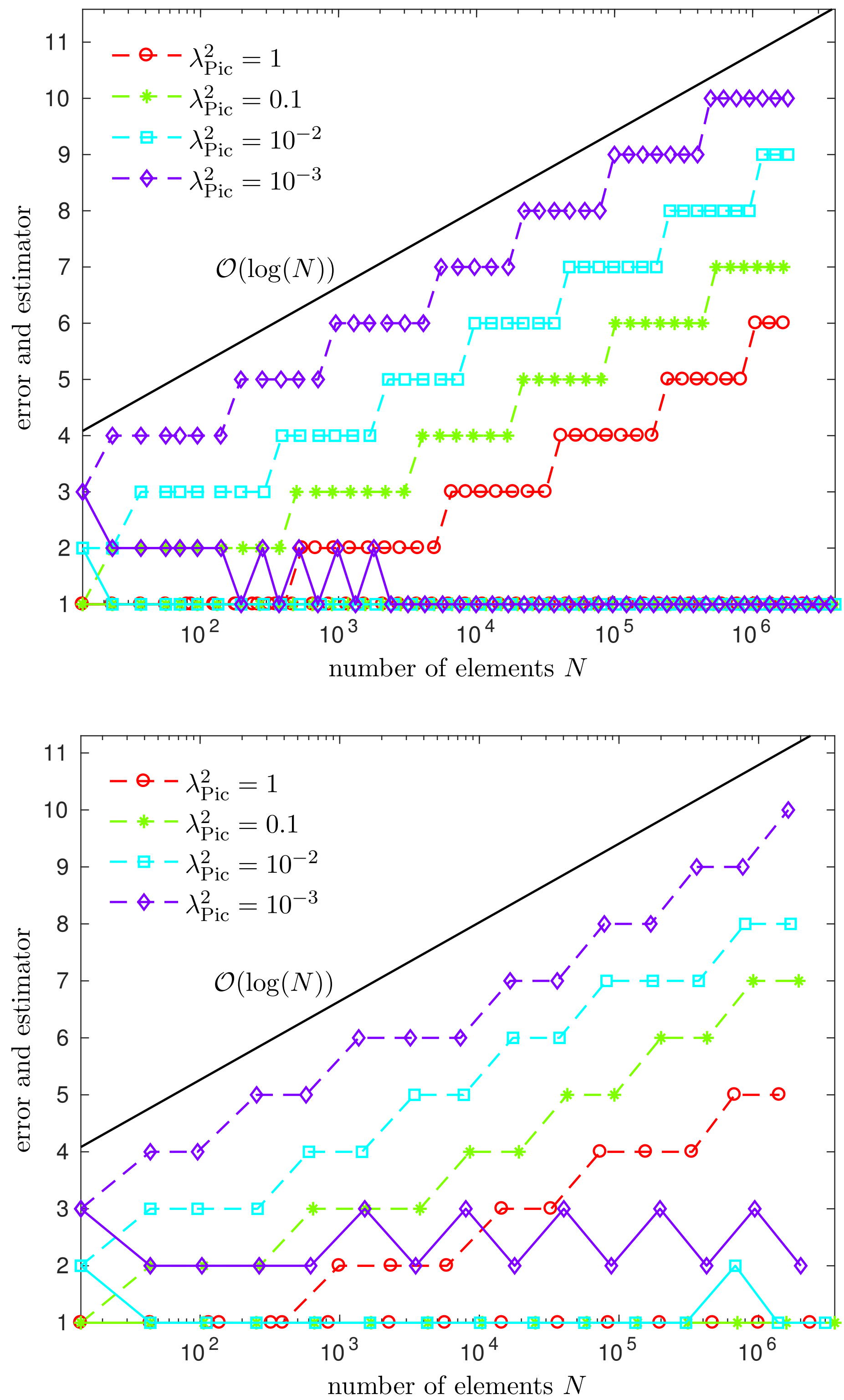



Figure 9.9: Ex. 1: Comparison of the number of Picard steps $\underline{n}(\ell)$ for nested iteration $u_{\ell+1}^{1,0} = u_\ell^{\underline{n},\underline{k}}$ (solid lines) to $u_{\ell+1}^{1,0} = 0$ (dashed lines). We used $\lambda_{\mathrm{PCG}}^2 = 10^{-3}\,\lambda_{\mathrm{Pic}}$ as well as $\theta = 0.2$ (top) and $\theta = 0.8$ (bottom).

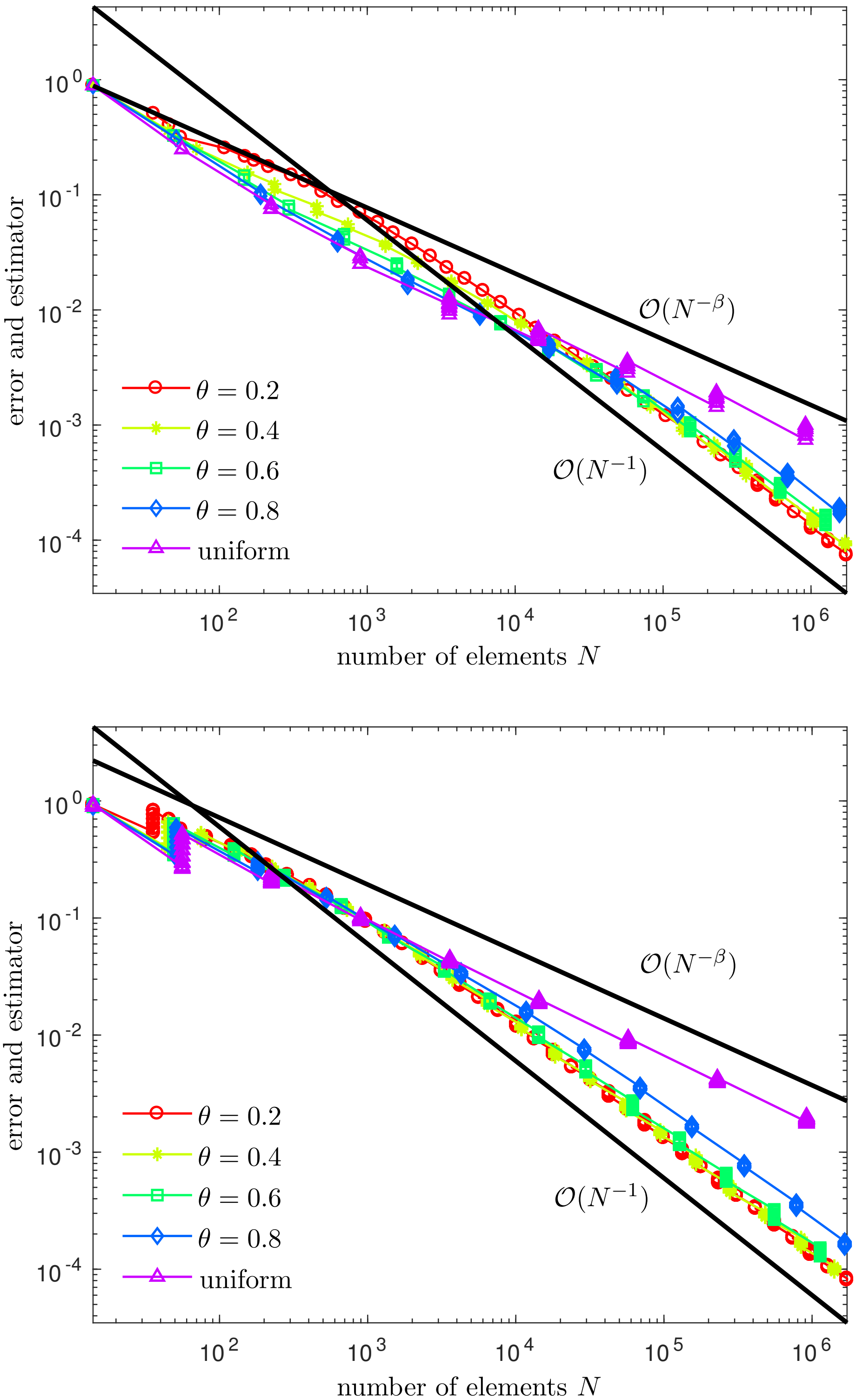



Figure 9.10: Ex. 2: We plot the convergence of the full estimator sequence $\eta_\ell(u_\ell^{n,k})^2$ for different values of $\theta$ with $\lambda_{\mathrm{Pic}}^2 = 10^{-2}$ and $\lambda_{\mathrm{PCG}}^2 = 10^{-4}$ (top) as well as $\lambda_{\mathrm{Pic}}^2 = 10^{-4}$ and $\lambda_{\mathrm{PCG}} = 10^{-2}\lambda_{\mathrm{Pic}}$ (bottom).

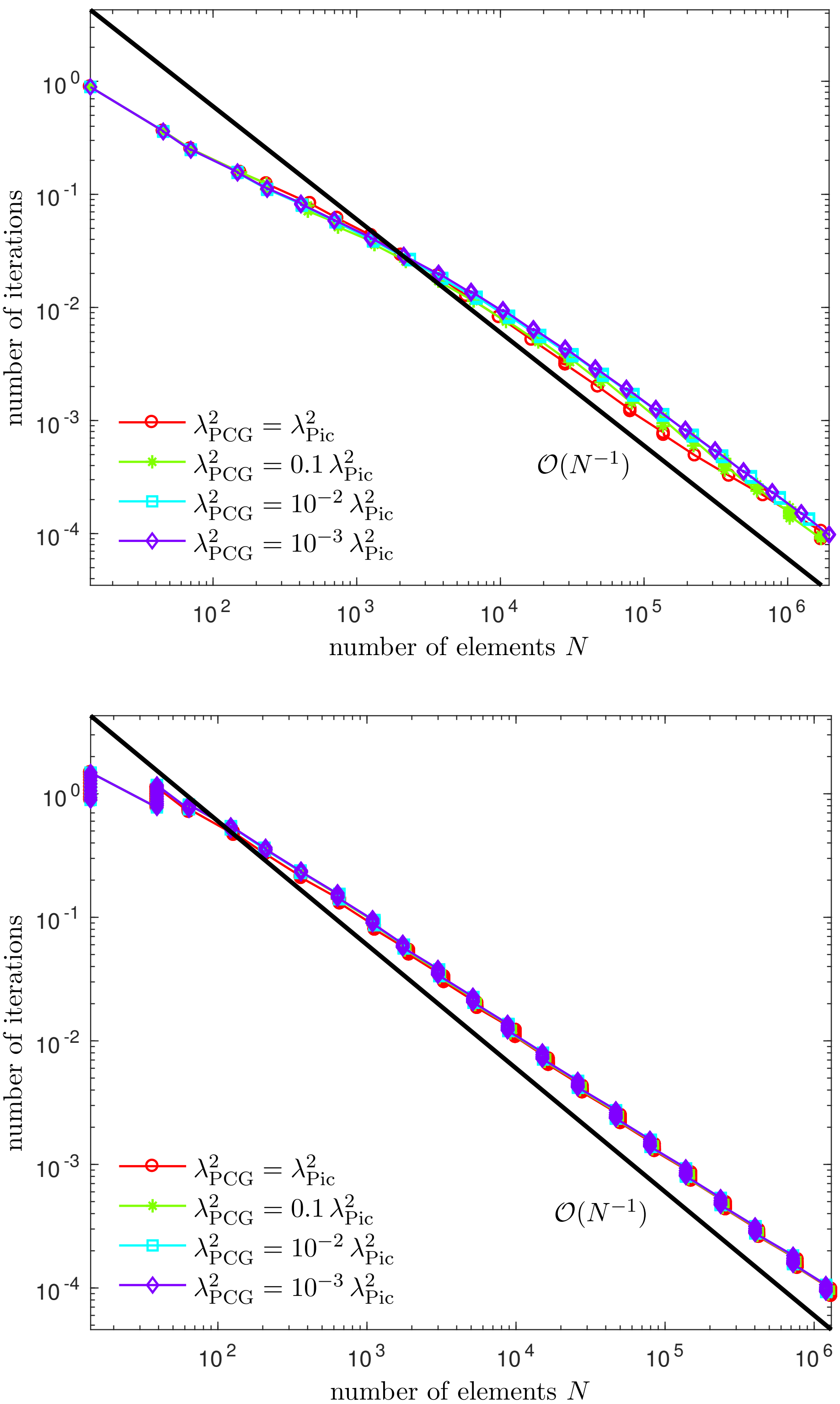



Figure 9.11: Ex. 2: Convergence rates of $\eta_\ell(u_\ell^{n,k})^2$ for different values of $\lambda_{\text{PCG}}$ depending on $\lambda_{\text{Pic}} = 0.1$ (top) and $\lambda_{\text{Pic}} = 10^{-3}$ (bottom) for $\theta = 0.4$

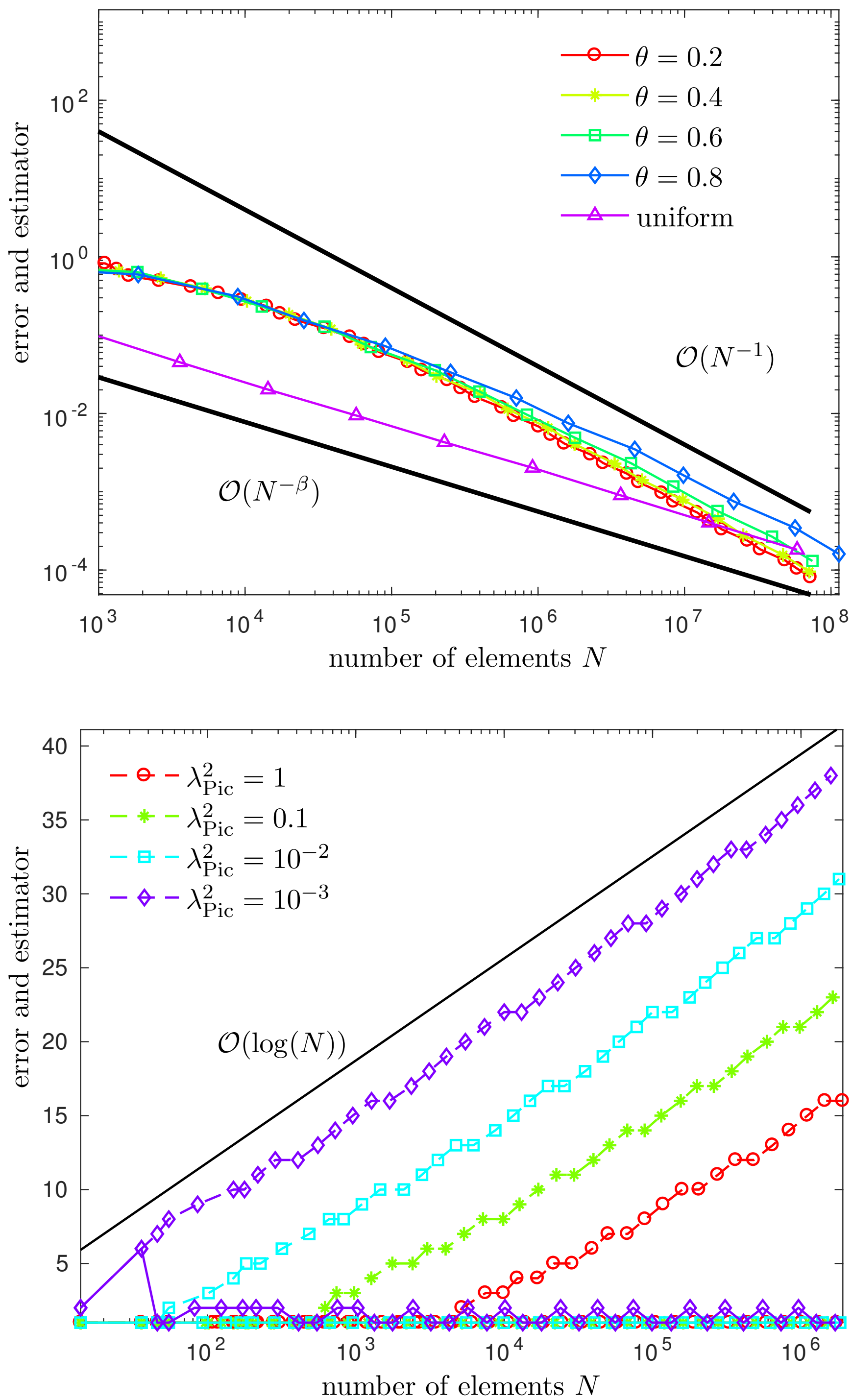



Figure 9.12: Ex. 2: Convergence rates of $\eta_\ell(u_\ell^{\underline{n},\underline{k}})^2$ for different values of $\theta$ with respect to the cumulative complexity (8.94), where $\lambda_{\mathrm{Pic}}^2 = 10^{-2}$ and $\lambda_{\mathrm{PCG}} = 10^{-4}$ (top). For uniform refinement, we plot the single step complexity, i.e., $\eta_\ell(u_\ell^{\underline{k},\underline{k}})^2$ with respect to $\#\mathcal{T}_\ell$. In addition, we compare the number of Picard steps $\underline{n}(\ell)$ for nested iteration $u_{\ell+1}^{1,0} = u_\ell^{\underline{n},\underline{k}}$ (solid lines) to $u_{\ell+1}^{1,0} = 0$ (dashed lines) for $\theta = 0.2$ and $\lambda_{\mathrm{Pic}}^2 = 10^{-2}\,\lambda_{\mathrm{PCG}}^2$ (bottom).

# Bibliography


[AFF+13] Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, and Dirk Praetorius. Efficiency and optimality of some weighted-residual error estimator for adaptive 2D boundary element methods. *Comput. Methods Appl. Math.*, 13(3):305–332, 2013.

[AFF+17] Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, J. Markus Melenk, and Dirk Praetorius. Local inverse estimates for non-local boundary integral operators. *Math. Comp.*, 86(308):2651–2686, 2017.

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

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

[AGL13] Mario Arioli, Emmanuil H. Georgoulis, and Daniel Loghin. Stopping criteria for adaptive finite element solvers. *SIAM J. Sci. Comput.*, 35(3):A1537–A1559, 2013.

[ALMS13] Mario Arioli, Jörg Liesen, Agnieszka Miedlar, and Zdeněk Strakoš. Interplay between discretization and algebraic computation in adaptive numerical solution of elliptic PDE problems. *GAMM-Mitt.*, 36(1):102–129, 2013.

[BC01] Annalisa Buffa and Patrick Ciarlet, Jr. On traces for functional spaces related to Maxwell's equations. II. Hodge decompositions on the boundary of Lipschitz polyhedra and applications. *Math. Methods Appl. Sci.*, 24(1):31–48, 2001.

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

[BDK12] Liudmila Belenki, Lars Diening, and Christian Kreuzer. Optimality of an adaptive finite element method for the $p$-Laplacian equation. *IMA J. Numer. Anal.*, 32(2):484–510, 2012.

[BET11] Roland Becker, Elodie Estecahandy, and David Trujillo. Weighted marking for goal-oriented adaptive finite element methods. *SIAM J. Numer. Anal.*, 49(6):2451–2469, 2011.

[BGMP16] Annalisa Buffa, Carlotta Giannelli, Philipp Morgenstern, and Daniel Peterseim. Complexity of hierarchical refinement for a class of admissible mesh configurations. *Comput. Aided Geom. Design*, 47:83 – 92, 2016.

[BHP17] Alex Bespalov, Alexander Haberl, and Dirk Praetorius. Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 317:318–340, 2017.

[BI98] Philippe Bouillard and Frank Ihlenburg. Error estimation and adaptivity for the finite element method in acoustics. In *Advances in adaptive computational methods in mechanics (Cachan, 1997)*, volume 47 of *Stud. Appl. Mech.*, pages 477–492. Elsevier, Amsterdam, 1998.

[BI99] Philippe Bouillard and Frank Ihlenburg. Error estimation and adaptivity for the finite element method in acoustics: 2D and 3D applications. *Comput. Methods Appl. Mech. Engrg.*, 176:147–163, 1999.

[BISG97] Ivo Babuška, Frank Ihlenburg, Theofanis Strouboulis, and Srihari K. Gangaraj. A posteriori error estimation for finite element solutions of Helmholtz' equation. I. The quality of local indicators and estimators. *Internat. J. Numer. Methods Engrg.*, 40(18):3443–3462, 1997.

[BN10] Andrea Bonito and Ricardo H. Nochetto. Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 48(2):734–771, 2010.

[BS08] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[BSF+14] Florian Bruckner, Dieter Suess, Michael Feischl, Thomas Führer, Petra Goldenits, Marcus Page, Dirk Praetorius, and Michele Ruggeri. Multiscale modeling in micromagnetics: Existence of solutions and numerical integration. *Math. Models Methods Appl. Sci.*, 24(13):2627–2662, 2014.

[BV84] Ivo Babuska and Michael Vogelius. Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, 44(1):75–102, 1984.

[Car96] Carsten Carstensen. Efficiency of a posteriori BEM-error estimates for first-kind integral equations on quasi-uniform meshes. *Math. Comp.*, 65(213):69–84, 1996.

[CFPP14] Carsten Carstensen, Michael Feischl, Marcus Page, and Dirk Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[CG12] Carsten Carstensen and Joscha Gedicke. An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. *SIAM J. Numer. Anal.*, 50(3):1029–1057, 2012.

[CK83] David L. Colton and Rainer Kress. *Integral equation methods in scattering theory*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1983. A Wiley-Interscience Publication.

[CKNS08] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[CMS01] Carsten Carstensen, Matthias Maischak, and Ernst P. Stephan. A posteriori error estimate and $h$-adaptive algorithm on surfaces for Symm's integral equation. *Numer. Math.*, 90(2):197–213, 2001.

[CN12] J. Manuel Cascon and Ricardo H. Nochetto. Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA J. Numer. Anal.*, 32(1):1–29, 2012.

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

[CW17] Scott Congreve and Thomas P. Wihler. Iterative Galerkin discretizations for strongly monotone problems. *J. Comput. Appl. Math.*, 311:457–472, 2017.

[CWGLS12] Simon N. Chandler-Wilde, Ivan G. Graham, Stephen Langdon, and Euan A. Spence. Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering. *Acta Numer.*, 21:89–305, 2012.

[Dau88] Monique Dauge. *Elliptic boundary value problems on corner domains*, volume 1341 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988. Smoothness and asymptotics of solutions.

[Dem06] Leszek Demkowicz. Babuška $\Leftrightarrow$ Brezzi?? *ICES Report*, 06-08, 2006.

[DK08] Lars Diening and Christian Kreuzer. Linear convergence of an adaptive finite element method for the $p$-Laplacian equation. *SIAM J. Numer. Anal.*, 46(2):614–638, 2008.

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

[EFGP13] Christoph Erath, Stefan Funken, Petra Goldenits, and Dirk Praetorius. Simple error estimators for the Galerkin BEM for some hypersingular integral equation in 2D. *Appl. Anal.*, 92(6):1194–1216, 2013.

[EFLFP09] Christoph Erath, Samuel Ferraz-Leite, Stefan Funken, and Dirk Praetorius. Energy norm based a posteriori error estimation for boundary element methods in two dimensions. *Appl. Numer. Math.*, 59(11):2713–2734, 2009.

[EP16] Christoph Erath and Dirk Praetorius. Adaptive finite volume methods with convergence rates. *SIAM J. Numer. Anal.*, 54:2228–2255, 2016.

[EV13] Alexandre Ern and Martin Vohralík. Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.*, 35(4):A1761–A1791, 2013.

[Fei15] Michael Feischl. *Rate optimality of adaptive algorithms.* PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, Wien, 2015.

[FFK+14] Michael Feischl, Thomas Führer, Michael Karkulik, J. Markus Melenk, and Dirk Praetorius. Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, part I: weakly-singular integral equation. *Calcolo*, 51(4):531–562, 2014.

[FFK+15] Michael Feischl, Thomas Führer, Michael Karkulik, J. Markus Melenk, and Dirk Praetorius. Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, part II: Hypersingular integral equation. *Electron. Trans. Numer. Anal.*, 44:153–176, 2015.

[FFP14] Michael Feischl, Thomas Führer, and Dirk Praetorius. Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. *SIAM J. Numer. Anal.*, 52(2):601–625, 2014.

[FFPS17a] Michael Feischl, Thomas Führer, Dirk Praetorius, and Ernst P. Stephan. Optimal additive Schwarz preconditioning for hypersingular integral equations on locally refined triangulations. *Calcolo*, 54(1):367–399, 2017.

[FFPS17b] Michael Feischl, Thomas Führer, Dirk Praetorius, and Ernst P. Stephan. Optimal preconditioning for the symmetric and nonsymmetric coupling of adaptive finite elements and boundary elements. *Numer. Methods Partial Differential Equations*, 33(3):603–632, 2017.

[FGH+16] Michael Feischl, Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Thomas Führer. Adaptive boundary element methods for optimal convergence of point errors. *Numer. Math.*, 132(3):541–567, 2016.

[FGHP16] Michael Feischl, Gregor Gantner, Alexander Haberl, and Dirk Praetorius. Adaptive 2D IGA boundary element methods. *Eng. Anal. Bound. Elem.*, 62:141–153, 2016.

[FGHP17] Michael Feischl, Gregor Gantner, Alexander Haberl, and Dirk Praetorius. Optimal convergence for adaptive IGA boundary element methods for weakly-singular integral equations. *Numer. Math.*, 136(1):147–182, 2017.

[FGP15] Michael Feischl, Gregor Gantner, and Dirk Praetorius. Reliable and efficient a posteriori error estimation for adaptive IGA boundary element methods for weakly-singular integral equations. *Comput. Methods Appl. Mech. Engrg.*, 290:362–386, 2015.

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

[FLP08] Samuel Ferraz-Leite and Dirk Praetorius. Simple a posteriori error estimators for the h-version of the boundary element method. *Computing*, 83(4):135–162, 2008.

[FP17] Thomas Führer and Dirk Praetorius. A linear Uzawa-type solver for nonlinear transmission problems. *Comput. Math. Appl., accepted for publication*, 2017.

[FPP14] Michael Feischl, Marcus Page, and Dirk Praetorius. Convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data. *J. Comput. Appl. Math.*, 255:481–501, 2014.

[FPZ16] Michael Feischl, Dirk Praetorius, and Kristoffer G. van der Zee. An abstract analysis of optimal goal-oriented adaptivity. *SIAM J. Numer. Anal.*, 54:1423–1448, 2016.

[Füh14] Thomas Führer. *Zur Kopplung von finiten Elementen und Randelementen (in German)*. PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, Wien, 2014.

[Gan13] Tsogtgerel Gantumur. Adaptive boundary element methods with convergence rates. *Numer. Math.*, 124(3):471–516, 2013.

[Gan17a] Gregor Gantner. *Optimal adaptivity for splines in finite and boundary element mehtods*. PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, Wien, 2017.

[Gan17b] Tsogtgerel Gantumur. Convergence rates of adaptive methods, Besov spaces, and multilevel approximation. *Found. Comput. Math.*, 17(4):917–956, 2017.

[GBB+15] S.P. Groth, A.J. Baran, T. Betcke, S. Havemann, and W. Śmigaj. The boundary element method for light scattering by ice crystals and its implementation in bem++. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 167(Supplement C):40 – 52, 2015.

[Geo08] Emmanuil H. Georgoulis. Inverse-type estimates on *hp*-finite element spaces and applications. *Math. Comp.*, 77(261):201–219, 2008.

[GHP17] Gregor Gantner, Daniel Haberlik, and Dirk Praetorius. Adaptive IGAFEM with optimal convergence rates: Hierarchical B-splines. *Math. Models Methods Appl. Sci.*, 27(14):2631–2674, 2017.

[GHPS17] Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Bernhard Stiftner. Rate optimal adaptive fem with inexact solver for nonlinear operators. *IMA J. Numer. Anal.*, 2017. DOI: 10.1093/imanum/drx050.

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

[GMZ11] Eduardo M. Garau, Pedro Morin, and Carlos Zuppa. Convergence of an adaptive Kačanov FEM for quasi-linear problems. *Appl. Numer. Math.*, 61(4):512–529, 2011.

[GMZ12] Eduardo M. Garau, Pedro Morin, and Carlos Zuppa. Quasi-optimal convergence rate of an AFEM for quasi-linear problems of monotone type. *Numer. Math. Theory Methods Appl.*, 5(2):131–156, 2012.

[GSS14] Dietmar Gallistl, Mira Schedensack, and Rob P. Stevenson. A remark on newest vertex bisection in any space dimension. *Comput. Methods Appl. Math.*, 14(3):317–320, 2014.

[GVL13] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[Has10] Alemdar Hasanov. Nonlinear monotone potential operators: From nonlinear ODE and PDE to computational material sciences. *Adv. Dyn. Syst. Appl.*, 5(2):173–190, 2010.

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

[KPP13] Michael Karkulik, David Pavlicek, and Dirk Praetorius. On 2D newest vertex bisection: optimality of mesh-closure and $H^1$-stability of $L_2$-projection. *Constr. Approx.*, 38(2):213–234, 2013.

[LB96] Wen B. Liu and John W. Barrett. Finite element approximation of some degenerate monotone quasilinear elliptic systems. *SIAM J. Numer. Anal.*, 33(1):88–106, 1996.

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

[Mel12] J. Markus Melenk. Mapping properties of combined field Helmholtz boundary integral operators. *SIAM J. Math. Anal.*, 44(4):2599–2636, 2012.

[MN05] Khamron Mekchay and Ricardo H. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.*, 43(5):1803–1827, 2005.

[MNS00] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.

[MP15] Philipp Morgenstern and Daniel Peterseim. Analysis-suitable adaptive T-mesh refinement with linear complexity. *Comput. Aided Geom. Design*, 34:50–66, 2015.

[MS09] Mario S. Mommer and Rob Stevenson. A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.*, 47(2):861–886, 2009.

[MSV08] Pedro Morin, Kunibert G. Siebert, and Andreas Veeser. A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, 18(5):707–737, 2008.

[Néd01] Jean-Claude Nédélec. *Acoustic and electromagnetic equations*, volume 144 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2001.

[OPD05] J. Tinsley Oden, Serge Prudhomme, and Leszek Demkowicz. A posteriori error estimation for acoustic wave propagation problems. *Arch. Comput. Methods Engrg.*, 12(4):343–389, 2005.

[Pav10] David Pavlicek. Optimalität adaptiver FEM, Bachelor thesis (in German). *TU Wien, Institute for Analysis and Scientific Computing, Wien*, 2010.

[Rud91] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[Saa03] Yousef Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.

[SBA+15] Wojciech Śmigaj, Timo Betcke, Simon Arridge, Joel Phillips, and Martin Schweiger. Solving boundary integral problems with BEM++. *ACM Trans. Math. Software*, 41(2):Art. 6, 40, 2015.

[SH96] James R. Stewart and Thomas J.R. Hughes. A posteriori error estimation and adaptive finite element computation of the Helmholtz equation in exterior domains. *Finite Elem. Anal. Des.*, 22(1):15–24, 1996.

[Sie11] Kunibert G. Siebert. A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.*, 31(3):947–970, 2011.

[SS11] Stefan A. Sauter and Christoph Schwab. *Boundary element methods*, volume 39 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2011.

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

[Ste08a] Olaf Steinbach. *Numerical approximation methods for elliptic boundary value problems*. Springer, New York, 2008.

[Ste08b] Rob Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.

[Ste13] Olaf Steinbach. Boundary integral equations for Helmholtz boundary value and transmission problems. In *Direct and inverse problems in wave propagation and applications*, volume 14 of *Radon Ser. Comput. Appl. Math.*, pages 253–292. De Gruyter, Berlin, 2013.

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

[Tar07] Luc Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin; UMI, Bologna, 2007.

[Tri83] Hans Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.

[Tri92] Hans Triebel. *Theory of function spaces. II*, volume 84 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992.

[Tri95] Hans Triebel. *Interpolation theory, function spaces, differential operators*. Johann Ambrosius Barth, Heidelberg, second edition, 1995.

[Vee02] Andreas Veeser. Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.*, 92(4):743–770, 2002.

[Ver13] Rüdiger Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[vtWGBA15] Elwin van 't Wout, Pierre Gélat, Timo Betcke, and Simon Arridge. A fast boundary element method for the scattering analysis of high-intensity focused ultrasound. *The Journal of the Acoustical Society of America*, 138(5):2726–2737, 2015.

[WC06] Haijun Wu and Zhiming Chen. Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems. *Sci. China Ser. A*, 49(10):1405–1429, 2006.

[XCH10] Xuejun Xu, Huangxin Chen, and Ronald H. W. Hoppe. Optimality of local multilevel methods on adaptively refined meshes for elliptic boundary value problems. *J. Numer. Math.*, 18(1):59–90, 2010.

[Yos80] Kosaku Yosida. *Functional analysis*. Springer-Verlag Berlin Heidelberg, 1980.

# Curriculum Vitae

## Personal Data

| | |
|---|---|
| Name | **Alexander Haberl** |
| Date of birth | 31.12.1987 |
| Place of birth | Waidhofen an der Ybbs |
| Citizenship | Austria |
| Email | `alexander.haberl@asc.tuwien.ac.at` |
| Homepage | `http://www.asc.tuwien.ac.at/~ahaberl/` |

---

## Education

| | |
|---|---|
| since 02/2015 | PhD student, supervised by Dirk Praetorius, Member of FWF research project *Optimal Adaptivity for BEM and FEM-BEM coupling* funded under grant FWF P27005, Institute for Analysis and Scientific Computing, TU Wien. |
| 02/2013–01/2015 | Master studies in Technical Mathematics, TU Wien. |
| 10/2009–02/2013 | Bachelor studies in Technical Mathematics, TU Wien. |
| 06/2008 | Matura (A-level) at HTL1 for engineering, Linz. |

---

## Scientific publications

Gregor Gantner, Alexander Haberl, Dirk Praetorius, Bernhard Stiftner: *Rate optimal adaptive FEM with inexact solver for nonlinear operators*, IMA Journal of Numerical Analysis, online first (2017), DOI: 10.1093/imanum/drx050.

Alex Bespalov, Alexander Haberl, Dirk Praetorius: *Adaptive FEM with coarse initial mesh guarantees optimal convergence rates for compactly perturbed elliptic problems*, Computer Methods in Applied Mechanics and Engineering, 317 (2017), 318–340.

Michael Feischl, Gregor Gantner, Alexander Haberl, Dirk Praetorius: *Optimal convergence for adaptive IGA boundary element methods for weakly-singular integral equations*, Numerische Mathematik, 136 (2017), 147–182.

Gregor Gantner, Alexander Haberl, Dirk Praetorius, Bernhard Stiftner: *Rate optimal adaptive FEM with inexact solver for strongly monotone operators*, Oberwolfach Report, European Mathematical Society, 51 (2016).

Michael Feischl, Thomas Führer, Gregor Gantner, Alexander Haberl, Dirk Praetorius: *Adaptive boundary element methods for optimal convergence of point errors*, Numerische Mathematik, 132 (2016), 541–567.

Michael Feischl, Gregor Gantner, Alexander Haberl, Dirk Praetorius: *Adaptive 2D IGA boundary element methods*, Engineering Analysis with Boundary Elements, 62 (2016), 141–153.

---

## Scientific talks

*Adaptive FEM and adaptive BEM for the Helmholtz equation*, 28.09.2017, ENUMATH, Voss, Norway, 2017.

*Optimal convergence rates for goal-oriented adaptivity*, 14.08.2016, Workshop on Boundary Elements and Adaptivity 2016, Basel, Switzerland, 2016.

*Optimal convergence rates for goal-oriented adaptivity*, 13.01.2013, WONAPDE, Concepción, Chile, 2016.

*Adaptive FEM for elliptic problems with Gårding inequality*, 15.06.2016, MAFELAP, London, UK, 2016.

*Adaptive BEM for optimal convergence of point errors*, 19.09.2015, ENUMATH, Ankara, Turkey, 2015.

*Adaptive BEM for optimal convergence of point errors*, 7.05.2015, Austrian Numerical Analysis Day, Linz, Austria, 2015.

*Adaptive boundary element methods for optimal convergence of point errors*, 17.03.2015, SIAM Conference on Computational Science and Engineering, Salt Lake City, USA, 2015.

Vienna, April 16, 2018

Alexander Haberl