Plain convergence of adaptive algorithms without exploiting reliability and efficiency

Gregor Gantner*  
Korteweg–de Vries (KdV) Institute for Mathematics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, The Netherlands  
*Corresponding author: g.gantner@uva.nl  

and  
Dirk Praetorius  
Institute for Analysis and Scientific Computing, TU Wien, Wiedner Hauptstraße 8-10/E101/4, 1040 Vienna, Austria  

[Received on 02 September 2020; revised on 13 January 2021]

We consider $h$-adaptive algorithms in the context of the finite element method and the boundary element method. Under quite general assumptions on the building blocks SOLVE, ESTIMATE, MARK and REFINE of such algorithms we prove plain convergence in the sense that the adaptive algorithm drives the underlying a posteriori error estimator to zero. Unlike available results in the literature, our analysis avoids the use of any reliability and efficiency estimate but relies only on structural properties of the estimator, namely stability on nonrefined elements and reduction on refined elements. In particular, the new framework thus also covers problems involving nonlocal operators like the fractional Laplacian or boundary integral equations, where (discrete) efficiency is (currently) not available.

Keywords: adaptivity; convergence; finite element method; boundary element method.

1. Introduction

A posteriori error estimation and related adaptive mesh refinement via the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE} \quad (1.1)$$

are standard tools in modern scientific computing. Over the last decade, mathematical understanding has matured. Convergence with optimal algebraic rates is mathematically guaranteed for a reasonable class of elliptic model problems and standard discretizations; we refer to the works by Dörfler (1996), Morin et al. (2000), Binev et al. (2004), Stevenson (2007), Manuel Cascon et al. (2008), Cascón & Nochetto (2012) and Feischl et al. (2014) for some important steps, as well as to the state-of-the-art review by Carstensen et al. (2014). However, all these works employ the so-called Dörfler marking strategy proposed in Dörfler (1996) to single out elements for refinement. Moreover, for the two-dimensional Poisson problem, it has recently been shown that a modified maximum criterion does not only lead to optimal convergence rates but also even leads to instance optimal meshes (Diener et al., 2016; Kreuzer & Schedensack, 2016; Innerberger & Praetorius, 2021). As the focus comes to other marking strategies, only plain convergence results are known and the essential works are Morin et al. (2008) and Siebert (2011).
To outline the results of Morin et al. (2008) and Siebert (2011) and the contributions of the present work let us fix some notation. Let $\mathcal{X}$ be a normed space that is linked to some domain (or manifold) $\Omega \subset \mathbb{R}^d$, $d \geq 1$. Let $u \in \mathcal{X}$ be the sought (unknown) solution. Suppose that the discrete subspaces $\mathcal{X}_\ell \subset \mathcal{X}$ are linked to some mesh $\mathcal{T}_\ell$ of $\Omega$ consisting of compact subdomains of $\Omega$. Let $u_\ell \in \mathcal{X}_\ell$ be a computable discrete approximation of $u$. Finally, let $\eta_\ell^2 = \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2$ be a computable error estimator such that $\eta_\ell(T)$ measures, at least heuristically, the error $u - u_\ell$ on $T \in \mathcal{T}_\ell$. We suppose that the sequence of meshes $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ is generated by the adaptive loop (1.1). In such a setting it has already been observed in the seminal work by Babuška & Vogelius (1984) that nestedness $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ of the discrete spaces together with a Céa-type quasi-optimality proves the so-called a priori convergence of adaptive schemes, i.e., there always exists a limit $u_\infty \in \mathcal{X}$ such that

$$\|u_\infty - u_\ell\|_{\mathcal{X}} \to 0 \quad \text{as } \ell \to \infty. \tag{1.2}$$

However, it remains to prove that also $u = u_\infty$.

To explain the abstract notation let us consider the two-dimensional Poisson model problem: in this case $\Omega \subset \mathbb{R}^2$ is a polygonal Lipschitz domain, $f \in L^2(\Omega)$ is some given load, $u \in \mathcal{X} = H^1_0(\Omega)$ solves the two-dimensional Poisson model problem $-\Delta u = f$ in $\Omega$ subject to homogeneous boundary conditions $u = 0$ on $\partial \Omega$, the meshes $\mathcal{T}_\ell$ are conforming triangulations of $\Omega$ into compact triangles $T \in \mathcal{T}_\ell$ and $u_\ell \in \mathcal{X}_\ell = \{v_\ell \in H^1_0(\Omega) : v_\ell|_T \text{ is affine for all } T \in \mathcal{T}_\ell\}$ is the conforming first-order finite element approximation of $u$, which solves

$$\int_{\Omega} \nabla u_\ell \cdot \nabla v_\ell \, dx = \int_{\Omega} f v_\ell \, dx \quad \text{for all } v_\ell \in \mathcal{X}_\ell. \tag{1.3}$$

For this problem the classical residual error estimator reads

$$\eta_\ell^2 = \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \quad \text{with } \eta_\ell(T)^2 = h_T^2 \|f\|_{L^2(T)}^2 + h_T \|\partial_n u_\ell\|_{L^2(\partial T \cap \Omega)}, \tag{1.4}$$

where $[\cdot]$ denotes the jump across interior edges and $h_T = |T|^{1/2}$ denotes the local mesh size; see, e.g., the monographs by Ainsworth & Oden (2000) and Verfürth (2013).

While Morin et al. (2008) formally focuses on conforming Petrov–Galerkin discretizations in the setting of Ladyshenskaja–Babuska–Brezzi (LBB) the actual analysis is more general: besides some assumptions on the locality of the norms of the involved function spaces there are no assumptions on how $u$ or $u_\ell$ is computed. The crucial assumptions in Morin et al. (2008) are local efficiency

$$C_{\text{eff}}^{-1} \eta_\ell(T) \leq \|u - u_\ell\|_{\mathcal{X}(\Omega_{\ell}(T))} + \text{osc}_\ell(\Omega_{\ell}(T)) \quad \text{for all } T \in \mathcal{T}_\ell, \tag{1.5}$$

as well as discrete local efficiency on marked elements,

$$C_{\text{eff}}^{-1} \eta_\ell(T) \leq \|u_{\ell+1} - u_\ell\|_{\mathcal{X}(\Omega_{\ell}(T))} + \text{osc}_\ell(\Omega_{\ell}(T)) \quad \text{for all } T \in \mathcal{M}_\ell, \tag{1.6}$$

where $\Omega_{\ell}(T) = \bigcup \{T' \in \mathcal{T}_\ell : T' \cap T \neq \emptyset\}$ is the patch of $T$ and $\text{osc}_\ell$ are some data oscillation terms. It is known that the latter assumption requires (at least) stronger local refinement, e.g., the local bisections refinement of marked elements in two dimensions to ensure the interior node property; see, e.g.,
The main result of Morin et al. (2008) proves that under these assumptions and for quite general marking strategies (see (2.3) below), the adaptive algorithm ensures that \textit{a priori} convergence \( (1.2) \) already implies estimator convergence

\[ \eta_\ell \to 0 \quad \text{as } \ell \to \infty. \]  

(1.7)

Provided that the error estimator \( \eta_\ell \) additionally satisfies reliability, i.e.,

\[ \|u - u_\ell\|_{\mathcal{X}} \leq C_{\text{rel}} \eta_\ell, \]  

(1.8)

this proves that \( \|u - u_\ell\|_{\mathcal{X}} \to 0 \) as \( \ell \to \infty \). In explicit terms the main result of Morin et al. (2008) reads as follows: if the discrete solutions \( u_\ell \in \mathcal{X}_\ell \) converge \( (1.2) \) then they converge indeed to the correct limit \( u = u_\infty \) — provided that the error estimator satisfies \( (1.5), (1.6) \) and \( (1.8) \).

Conceptually, it is remarkable that the convergence proof of Morin et al. (2008) exploits lower error bounds, although the mesh refinement is driven by the error estimator only. The work of Siebert (2011) thus aimed to prove convergence without using (discrete) lower bounds. This, however, comes at the cost that, first, the analysis exploits the problem setting (and is restricted to Petrov–Galerkin discretizations of operator equations \( Bu = F \)) and, second, the analysis relies on some strengthened reliability estimate (formulated in terms of the residual), which implies \( (1.8) \). The main result of Siebert (2011) then states that under these assumptions and for quite general marking strategies (see (2.4) below), the adaptive algorithm ensures that \textit{a priori} convergence \( (1.2) \) already implies error convergence

\[ \|u - u_\ell\|_{\mathcal{X}} \to 0 \quad \text{as } \ell \to \infty. \]  

(1.9)

In particular, the new proof of Siebert (2011) avoids the discrete local efficiency \( (1.6) \). Surprisingly, however, estimator convergence \( (1.7) \) cannot be proved under the assumptions of Siebert (2011) but requires that the error estimator \( \eta_\ell \) is also locally efficient \( (1.5) \).

One advantage of the results of Morin et al. (2008) and Siebert (2011) is that they apply to many different \textit{a posteriori} error estimators. In particular, it has recently been shown in Führer & Praetorius (2020) and Gantner & Stevenson (2021) that the assumptions of Siebert (2011) are, in particular, satisfied for a wide range of model problems discretized by least squares finite element methods (FEMs), where adaptivity is driven by the built-in least-squares functional, including even a least-squares space-time discretization of the heat equation. On the other hand, Siebert (2011) excludes adaptive schemes for variational inequalities, and both Morin et al. (2008) and Siebert (2011) need local efficiency of the error estimator, which does not appear to be available for nonlocal operators, e.g., FEMs for the fractional Laplacian (see, e.g., Faustmann et al., 2020) or boundary element methods (BEMs) for elliptic integral equations (see, e.g., Feischl et al., 2013; Gantumur, 2013).

With the latter observations the current paper comes into play. We provide a new proof for plain convergence of adaptive algorithms, which involves \textit{neither} reliability \( (1.8) \) \textit{nor} any kind of (global or local) efficiency \( (1.5) \) and \( (1.6) \). Instead, we exploit that the local contributions of many residual error estimators are weighted by the local mesh size \( (\text{cf. } (1.4) \text{ for the Poisson model problem}) \). With scaling arguments one usually obtains reduction on refined elements

\[ \eta_{\ell+n}(T_{\ell+n}\setminus T_\ell)^2 \leq q \eta_\ell(T_\ell\setminus T_{\ell+n})^2 + C \|u_{\ell+n} - u_\ell\|^2_{\mathcal{X}} \quad \text{for all } \ell, n \in \mathbb{N}_0. \]  

(1.10)
with generic constants $0 < q < 1$ and $C > 0$. Note that $\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+n}$ corresponds to the elements that are going to be refined, while $\mathcal{T}_{\ell+n} \setminus \mathcal{T}_\ell$ corresponds to the generated children. Moreover, stability on nonrefined elements,

$$|\eta_{\ell+n}(\mathcal{T}_{\ell+n} \cap \mathcal{T}_\ell) - \eta_\ell(\mathcal{T}_{\ell+n} \cap \mathcal{T}_\ell)| \leq C \|u_{\ell+n} - u_\ell\|_{X'} \quad \text{for all } \ell, n \in \mathbb{N}_0, \quad (1.11)$$

usually holds. We stress that (1.10) and (1.11) also play a fundamental role in the contemporary proofs of optimal convergence rates for adaptive algorithms; see Carstensen et al. (2014). The main result of the present work (Theorem 3.1) shows that, together with the same marking criterion as in Morin et al. (2008), the structural properties (1.10) and (1.11) suffice to show that a priori convergence (1.2) yields estimator convergence (1.7). Clearly, reliability (1.8) is then finally required to conclude error convergence (1.9).

Outline. The remainder of this work is organized as follows: in Section 2 we provide a formal statement of the adaptive algorithm (Algorithm 2.1) as well as precise assumptions on its four modules from (1.1). The new plain convergence result (Theorem 3.1, Theorem 3.3) is stated and proved in Section 3, before we give some examples, which do not fit the framework of Morin et al. (2008) and Siebert (2011) but are covered by the current analysis.

2. Abstract adaptive algorithm

2.1 Mesh refinement

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain (or a manifold in $\mathbb{R}^d$) with positive measure $|\Omega| > 0$. We say that $\mathcal{T}_H$ is a mesh (of $\Omega$) if

- $\mathcal{T}_H$ is a finite set of compact sets $T \in \mathcal{T}_H$ with positive measure $|T| > 0$;
- for all $T, T' \in \mathcal{T}_H$ with $T \neq T'$, it holds that $|T \cap T'| = 0$;
- $\mathcal{T}_H$ is a covering of $\overline{\Omega}$, i.e., $\overline{\Omega} = \bigcup_{T \in \mathcal{T}_H} T$.

Let refine($\cdot$) be a fixed refinement strategy, i.e., for each mesh $\mathcal{T}_H$ and a set of marked elements $\mathcal{M}_H \subseteq \mathcal{T}_H$, the refinement strategy returns a refined mesh $\mathcal{T}_h := \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$ such that, first, at least the marked elements are refined (i.e., $\mathcal{M}_H \subseteq \mathcal{T}_H \setminus \mathcal{T}_h$) and, second, parents $T \in \mathcal{T}_H$ are the union of their children, i.e.,

$$T = \bigcup \{T' \in \mathcal{T}_h : T' \subseteq T\} \quad \text{for all } T \in \mathcal{T}_H. \quad (2.1)$$

For a mesh $\mathcal{T}_H$ let $\mathcal{T}(\mathcal{T}_H)$ denote the set of all possible refinements of $\mathcal{T}_H$ (as determined by the refinement strategy refine($\cdot$)), i.e., for any $\mathcal{T}_h \in \mathcal{T}(\mathcal{T}_H)$, there exists $n \in \mathbb{N}_0$ and $\mathcal{T}_0', \ldots, \mathcal{T}_n'$ such that $\mathcal{T}_0 = \mathcal{T}_H$, $\mathcal{T}_{j+1}' = \text{refine}(\mathcal{T}_j', \mathcal{M}_j')$ for all $j = 0, \ldots, n - 1$ and appropriate $\mathcal{M}_j' \subseteq \mathcal{T}_j'$, and $\mathcal{T}_n' = \mathcal{T}_h$. Finally, we suppose that we are given a fixed initial mesh $\mathcal{T}_0$ so that it makes sense to call $\mathcal{T} := \mathcal{T}(\mathcal{T}_0)$ the set of all admissible meshes.

2.2 Continuous and discrete setting

Let $X'$ be a normed space (related to $\Omega$) and $u \in X'$ be the (unknown) exact solution. For each mesh $\mathcal{T}_H$ let $X'_H \subseteq X'$ be an associated discrete subspace and $u_H \in X'_H$ be the corresponding (computable) discrete solution.
2.3 Error estimator

For each mesh $\mathcal{T}_H$ and all $T \in \mathcal{T}_H$ let $\eta_H(T) \geq 0$ be a computable quantity that is usually called the refinement indicator. At least heuristically, $\eta_H(T)$ measures the error $u - u_h$ on the element $T$. We abbreviate

$$
\eta_H := \eta_H(\mathcal{T}_H), \quad \text{where} \quad \eta_H(\mathcal{U}_H) := \left( \sum_{T \in \mathcal{U}_H} \eta_H(T)^2 \right)^{1/2} \quad \text{for all } \mathcal{U}_H \subseteq \mathcal{T}_H. \quad (2.2)
$$

We note that $\eta_H$ is usually referred to as the error estimator.

2.4 Adaptive algorithm

Starting from the given initial mesh $\mathcal{T}_0$ we consider the standard adaptive loop (1.1) in the following algorithmic form:

ALGORITHM 2.1 For each $\ell = 0, 1, 2, \ldots$, iterate the following steps (i)–(iv):

(i) **SOLVE**: compute the discrete solution $u_\ell \in \mathcal{X}_\ell$.

(ii) **ESTIMATE**: compute refinement indicators $\eta_\ell(T)$ for all elements $T \in \mathcal{T}_\ell$.

(iii) **MARK**: determine a set of marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$.

(iv) **REFINE**: generate the refined mesh $\mathcal{T}_{\ell+1} := \text{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$.

**Output**: refined meshes $\mathcal{T}_\ell$, corresponding exact discrete solutions $u_\ell$, and error estimators $\eta_\ell$ for all $\ell \in \mathbb{N}_0$.

To analyse Algorithm 2.1 it remains to specify further assumptions on its four modules: as far as **SOLVE** is concerned we shall only assume *a priori* convergence (1.2). While this assumption is guaranteed for many problems (see, e.g., Morin et al., 2008; Siebert, 2011 for problems in the framework of LBB theory, as well as the seminal work by Babuška & Vogelius, 1984 for problems in the Lax–Milgram setting) we stress that, at this point, it is still mathematically unclear whether $u_\infty = u$ holds or not.

As far as **MARK** is concerned let $M : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be continuous at 0 with $M(0) = 0$ and suppose that the sets $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ satisfy the following property from Morin et al. (2008):

$$
\max_{T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(T) \leq M(\eta_\ell(\mathcal{M}_\ell)). \quad (2.3)
$$

We note that the latter assumption is weaker than the following assumption from Siebert (2011):

$$
\max_{T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(T) \leq M\left( \max_{T \in \mathcal{M}_\ell} \eta_\ell(T) \right). \quad (2.4)
$$

Clearly, the marking criteria (2.3) and (2.4) are satisfied with $M(t) = t$ as soon as $\mathcal{M}_\ell$ contains one element with maximal indicator, i.e., there exists $T \in \mathcal{M}_\ell$ such that $\eta_\ell(T) = \max_{T' \in \mathcal{T}_\ell} \eta_\ell(T')$. For instance this is the case for
• the maximum criterion for some fixed $0 \leq \theta \leq 1$, where

$$\mathcal{M}_\ell := \{ T \in \mathcal{T}_\ell : \eta_\ell(T) \geq (1 - \theta) \max_{T' \in \mathcal{T}_\ell} \eta_\ell(T') \};$$ (2.5)

• the equidistribution criterion for fixed $0 \leq \theta \leq 1$, where

$$\mathcal{M}_\ell := \{ T \in \mathcal{T}_\ell : \eta_\ell(T) \geq (1 - \theta) \eta_\ell/\#T\ell \}; \quad \text{(2.6)}$$

Finally, let us consider the Dörfler criterion for some fixed $0 < \theta \leq 1$, i.e.,

$$\theta \eta_\ell^2 \leq \eta_\ell(\mathcal{M}_\ell)^2.$$ (2.7)

While (2.4) cannot be satisfied in general, (2.3) holds with $M(t) := \sqrt{1 - \theta} \theta^{-1} t$. To see this let $T \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell$ and note that

$$\eta_\ell(T)^2 \leq \eta_\ell(\mathcal{T}_\ell \setminus \mathcal{M}_\ell)^2 = \eta_\ell^2 - \eta_\ell(\mathcal{M}_\ell)^2 \leq (1 - \theta) \eta_\ell^2 \leq (1 - \theta) \theta^{-1} \eta_\ell(\mathcal{M}_\ell)^2.$$

However, if the set $\mathcal{M}_\ell$ is constructed via sorting of the indicators, then

$$\max_{T' \in \mathcal{T}_\ell \setminus \mathcal{M}_\ell} \eta_\ell(T') \leq \min_{T' \in \mathcal{M}_\ell} \eta_\ell(T');$$ (2.8)

see Pfeiler & Praetorius (2020) for different algorithms that generate $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ satisfying the Dörfler criterion (2.7) together with (2.8). In the latter case (2.3) and (2.4) hold again with $M(t) := t$.

3. A new plain convergence result

Unlike Morin et al. (2008) and Siebert (2011) we only require the following two structural properties of the error estimator for all $\mathcal{T}_H \in \mathcal{T}$ and all refinements $\mathcal{T}_h \in \mathcal{T}(\mathcal{T}_H)$, where $S, R: \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ are functions that are continuous at 0 with $R(0) = 0 = S(0) \neq 0$ and $0 < q_{red} < 1$:

• **stability on nonrefined elements**, i.e.,

$$\eta_h(\mathcal{T}_h \cap \mathcal{T}_H) \leq \eta_H(\mathcal{T}_H \cap \mathcal{T}_h) + S(\|u_h - u_H\|_X);$$ (3.1)

• **reduction on refined elements**, i.e.,

$$\eta_h(\mathcal{T}_h \setminus \mathcal{T}_H)^2 \leq q_{red} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h)^2 + R(\|u_h - u_H\|_X).$$ (3.2)

We note that (3.1) and (3.2) are implicitly first found in the proof of Manuel Cascon et al. (2008, Corollary 3.4) but in fact seem to go back to Diening & Kreuzer (2008) (used there for the oscillations). In practice, the reduction (3.2) can only be proved if the local contributions $\eta_H(T)$ of the error estimator are weighted by (some positive power of) the local mesh size $h_T$. In the later examples in Section 4 it holds that $S(t) \sim t$ and $R(t) \sim t^2$. 
Under the structural assumptions (3.1) and (3.2) on the estimator the following theorem proves that Algorithm 2.1 leads to estimator convergence. We stress that neither the reliability estimate (1.8) nor any (global or even local) efficiency estimate (e.g., (1.5)–(1.6)) is required.

**Theorem 3.1** Suppose properties (3.1) and (3.2) of the estimator hold and that refinement ensures that each parent is the union of its children (2.1). Consider the output of Algorithm 2.1 with the marking strategy (2.3). Then a priori convergence (1.2) implies estimator convergence

\[ \eta_\ell \to 0 \quad \text{as} \quad \ell \to \infty. \]  

(3.3)

The proof of Theorem 3.1 employs the following elementary result, whose simple proof is included for the convenience of the reader.

**Lemma 3.2** Let \((a_\ell)\ell \in \mathbb{N}_0\) be a sequence with \(a_\ell \geq 0\) for all \(\ell \in \mathbb{N}_0\). Suppose that there exists \(0 < \rho < 1\) and a sequence \((b_\ell)\ell \in \mathbb{N}_0\) with \(b_\ell \to 0\) as \(\ell \to \infty\) such that

\[ a_{\ell+1} \leq \rho a_\ell + b_\ell \quad \text{for all} \quad \ell \in \mathbb{N}_0. \]  

(3.4)

Then it follows that \(a_\ell \to 0\) as \(\ell \to \infty\).

**Proof.** With the convergence of \((b_\ell)\ell \in \mathbb{N}_0\) we note that

\[ 0 \leq \limsup_{\ell \to \infty} a_\ell = \limsup_{\ell \to \infty} a_{\ell+1} \leq \limsup_{\ell \to \infty} (\rho a_\ell + b_\ell) = \rho \limsup_{\ell \to \infty} a_\ell. \]

Thus, it only remains to show that \(\limsup_{\ell \to \infty} a_\ell < \infty\) to conclude that \(0 = \liminf_{\ell \to \infty} a_\ell = \limsup_{\ell \to \infty} a_\ell\) and hence \(\lim_{\ell \to \infty} a_\ell = 0\). Indeed, induction on \(\ell\) proves that

\[ 0 \leq a_\ell \leq \rho^\ell a_0 + \sum_{j=0}^{\ell-1} \rho^{\ell-1-j} b_j \quad \text{for all} \quad \ell \in \mathbb{N}_0. \]

Since \((b_\ell)\ell \in \mathbb{N}_0\) is uniformly bounded the geometric series yields that \(\sup_{\ell \in \mathbb{N}} a_\ell < \infty\). In particular, we thus see that \(\limsup_{\ell \to \infty} a_\ell < \infty\) and conclude the proof. \(\square\)

**Proof of Theorem 3.1.** The proof is split into five steps.

**Step 1:** we prove that \((T_\ell)\ell \in \mathbb{N}_0\) admits a subsequence \((T_{\ell_k})k \in \mathbb{N}_0\) such that \(\eta_{\ell_k} (T_{\ell_k} \setminus T_{\ell_{k+1}}) \to 0\) as \(k \to \infty\). Let \(T_\infty = \bigcup_{\ell \in \mathbb{N}_0} \bigcap_{k \geq \ell} T_\ell\) be the set of all elements that remain unrefined after some (arbitrary) step \(\ell\). We exploit (2.1) and choose a subsequence \((T_{\ell_k})k \in \mathbb{N}_0\) of \((T_\ell)\ell \in \mathbb{N}_0\) such that

\[ T_{\ell_k+1} \cap T_{\ell_k} = T_{\ell_k} \cap T_\infty \quad \text{for all} \quad k \in \mathbb{N}_0, \]  

(3.5)

i.e., only elements \(T \in T_{\ell_k} \cap T_\infty\) remain unrefined if we pass from \(T_{\ell_k}\) to \(T_{\ell_{k+1}}\). Note that the choice of \((T_{\ell_k})k \in \mathbb{N}_0\) guarantees the inclusion

\[ T_{\ell_{k+1}} \cap T_{\ell_k} \supseteq (3.5) T_{\ell_k} \cap T_\infty \subseteq T_{\ell_{k+1}} \cap T_\infty \supseteq (3.5) T_{\ell_{k+2}} \cap T_{\ell_{k+1}} \]
and hence
\[ T_{\ell_{k+1}} \setminus T_{\ell_{k+2}} = T_{\ell_{k+1}} \setminus [T_{\ell_{k+2}} \cap T_{\ell_{k+1}}] \subseteq T_{\ell_{k+1}} \setminus [T_{\ell_{k+1}} \cap T_{\ell_k}] = T_{\ell_{k+1}} \setminus T_{\ell_k}. \]

With this and reduction (3.2) we infer that
\[ \eta_{\ell_{k+1}} (T_{\ell_{k+1}} \setminus T_{\ell_{k+2}})^2 \leq \eta_{\ell_{k+1}} (T_{\ell_{k+1}} \setminus T_{\ell_k})^2 \leq q_{\text{red}} \eta_{\ell_k} (T_{\ell_k} \setminus T_{\ell_{k+1}})^2 + R(\|u_{\ell_{k+1}} - u_{\ell_k}\|_\chi). \]

With \( a_k = \eta_{\ell_k} (T_{\ell_k} \setminus T_{\ell_{k+1}})^2 \), \( \rho := q_{\text{red}} \) and \( b_k := R(\|u_{\ell_{k+1}} - u_{\ell_k}\|_\chi) \), a priori convergence (1.2) proves that
\[ 0 \leq a_{k+1} \leq \rho a_k + b_k \quad \text{for all } k \in \mathbb{N}_0 \quad \text{with } \lim_{k \to \infty} b_k = 0. \]

By use of Lemma 3.2 we conclude that \( a_k = \eta_{\ell_k} (T_{\ell_k} \setminus T_{\ell_{k+1}})^2 \to 0 \) as \( k \to \infty \).

**Step 2:** we prove that the subsequence \( (T_{\ell_k})_{k \in \mathbb{N}_0} \) also guarantees that \( \eta_{\ell_k} (M_{\ell_k}) \to 0 \) as \( k \to \infty \). To this end we first note that \( T \in T_{\ell_k} \setminus T_{\ell_{k+1}} \) implies that \( T \) is refined and hence \( T \not\in T_\infty \). Therefore, we see that
\[ M_{\ell_k} \subseteq T_{\ell_k} \setminus T_{\ell_{k+1}} \subseteq T_{\ell_k} \setminus T_\infty = T_{\ell_k} \setminus [T_{\ell_k} \cap T_\infty] \subseteq T_{\ell_k} \setminus [T_{\ell_{k+1}} \cap T_k] = T_{\ell_k} \setminus T_{\ell_{k+1}}. \]

This implies that \( 0 \leq \eta_{\ell_k} (M_{\ell_k}) \leq \eta_{\ell_k} (T_{\ell_k} \setminus T_{\ell_{k+1}}) \to 0 \) as \( k \to \infty \).

**Step 3:** for all fixed \( \ell' \in \mathbb{N}_0 \) we prove that
\[ \eta_{\ell_k} (T_{\ell'} \cap T_\infty) \to 0 \quad \text{as } \ell' \leq \ell_k \to \infty \text{ together with } k \to \infty. \quad (3.6) \]

To see this we exploit the marking strategy (2.3) and note with Step 2 that
\[ \max_{T \in T_{\ell_k} \setminus M_{\ell_k}} \eta_{\ell_k} (T) \leq M(\eta_{\ell_k} (M_{\ell_k})) \to 0 \quad \text{as } k \to \infty. \]

For \( \ell' \leq \ell_k \), it holds that \( T_{\ell'} \cap T_\infty \subseteq T_{\ell_k} \setminus M_{\ell_k} \) and hence
\[ \eta_{\ell_k} (T) \to 0 \quad \text{as } k \to \infty \quad \text{for all } T \in T_{\ell'} \cap T_\infty. \]

Since \( T_{\ell'} \cap T_\infty \subseteq T_{\ell_k} \) is a fixed finite set we conclude the proof of (3.6).

**Step 4:** we prove that \( (T_{\ell_k})_{k \in \mathbb{N}_0} \) admits a subsequence \( (T_{\ell_j})_{j \in \mathbb{N}_0} \) such that \( \eta_{\ell_j} \to 0 \) as \( j \to \infty \). To this end first let \( T_H \in T \) and \( T_h \in T(T_H) \). Reduction (3.2) proves that
\[
\eta_h^2 = \eta_h (T_h \setminus T_H)^2 + \eta_h (T_h \cap T_H)^2 \\
\leq q_{\text{red}} \eta_h (T_H \setminus T_h)^2 + \eta_h (T_h \cap T_H)^2 + R(\|u_h - u_H\|_\chi) \quad (3.7a) \\
\leq q_{\text{red}} \eta_h^2 + \eta_h (T_h \cap T_H)^2 + R(\|u_h - u_H\|_\chi).}
\]
If $\eta_H = 0$, stability (3.1) and reduction (3.2) prove that
\[
\eta^2_h = \eta_h(T_h \setminus T_H)^2 + \eta_h(T_h \cap T_H)^2
\leq \eta^2_h(R(\eta_H \setminus T_H)^2 + 2 \eta_h(T_h \cap T_H)^2 + [R + 2S^2]\|u_h - u_H\|_{X}^2) \tag{3.7b}
\]
where $[R + 2S^2](t) = R(t) + 2S(t)^2$. For the subsequence $(T_k)_{k \in \mathbb{N}_0}$ we recall from (3.5) that $T_{k+n} \cap T_k = T_k \cap T_\infty$. Hence, the estimates (26) read, for all $k \in \mathbb{N}_0$ and $n \in \mathbb{N}$,
\[
\eta^2_{k+n} \leq \eta^2_{k} + \eta_{k+n}(T_k \cap T_\infty)^2 + R(\|u_{k+n} - u_k\|_{X}) \quad \text{if} \quad \eta^2_k \neq 0,
\]
resp.
\[
\eta^2_{k+n} \leq [R + 2S^2](\|u_{k+n} - u_k\|_{X}) \quad \text{if} \quad \eta^2_k = 0.
\]
Let $0 < \eta_{\text{red}} < \eta'_{\text{red}} < 1$. Given $k \in \mathbb{N}_0$ with $\eta^2_k \neq 0$, convergence (3.6) allows us to pick some $n(k) \in \mathbb{N}$ such that
\[
\eta^2_{k+n} \leq \eta^2_{k} + \eta_{k+n}(T_k \cap T_\infty)^2 \leq \eta^2_{k}.
\]
In particular, we can choose a further subsequence $(T_j)_{j \in \mathbb{N}_0}$ of $(T_k)_{k \in \mathbb{N}_0}$ such that
\[
\eta^2_{j+1} \leq \eta^2_{j} + [R + 2S^2](\|u_{j+1} - u_j\|_{X}) \quad \text{for all} \quad j \in \mathbb{N}_0.
\]
With $a_j = \eta^2_{j}$, $\rho = \eta'_{\text{red}}$ and $b_j = [R + 2S^2](\|u_{j+1} - u_j\|_{X})$, a priori convergence (1.2) proves that
\[
0 \leq a_{j+1} \leq \rho a_j + b_j \quad \text{for all} \quad j \in \mathbb{N}_0 \quad \text{with} \quad \lim_{j \to \infty} b_j = 0.
\]
By use of Lemma 3.2 we see that $a_j = \eta^2_{j} \to 0$ as $j \to \infty$.

**Step 5:** we prove that convergence of the subsequence $\eta_{\ell_j} \to 0$ already implies convergence of the full sequence $\eta_{\ell} \to 0$ as $\ell \to \infty$. To this end we argue as in (3.7c) and use $\eta_{\text{red}} \leq 2$ to see that, for all $T_H \in T$ and all $T_h \in T(T_H)$,
\[
\eta^2_h \leq 2 \eta^2_H + [R + 2S^2](\|u_h - u_H\|_{X}). \tag{3.9}
\]
Given $\epsilon > 0$, there exists an index $\ell_j$ such that $\eta_{\ell_j} \leq \epsilon$ and $[R + 2S^2](\|u_{\ell_j} - u_{\ell_j}\|_{X}) \leq \epsilon$ for all $\ell \geq \ell_j$.
For $\ell \geq \ell_j$ estimate (3.9) thus proves that
\[
\eta^2_{\ell} \leq 2\epsilon^2 + \epsilon.
\]
This concludes the proof. \qed
For the Dörfler marking criterion (2.7) the refinement assumption (2.1) exploited in the proof of Theorem 3.1 can even be dropped. The following result (together with its very simple proof) is essentially the key argument in Aurada et al. (2012).

**Theorem 3.3** Suppose properties (3.1) and (3.2) of the estimator. Consider the output of Algorithm 1.2, where the marked elements $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ satisfy the Dörfler criterion (2.7) for some fixed marking parameter $0 < \theta \leq 1$. Then a priori convergence (1.2) yields estimator convergence (3.3).

**Proof.** Let $\mathcal{T}_H \in \mathbb{T}$ and $\mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H)$. Arguing as for (3.7b) and exploiting the Young inequality for some arbitrary $\delta > 0$ (instead of $\delta = 1$ above) we see that

$$
\eta_h^2 \leq q_{\text{red}} \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h)^2 + (1 + \delta) \eta_H(\mathcal{T}_h \cap \mathcal{T}_H)^2 + [R + (1 + \delta^{-1})S^2](\|u_h - u_H\|_{\mathcal{X}})
$$

$$
= (1 + \delta) \eta_h^2 - [(1 + \delta - q_{\text{red}})] \eta_H(\mathcal{T}_H \setminus \mathcal{T}_h)^2 + [R + (1 + \delta^{-1})S^2](\|u_h - u_H\|_{\mathcal{X}}).
$$

For the sequence $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$, the Dörfler criterion (2.7) and $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ yield

$$
\theta \eta_{\ell+1}^2 \leq \eta_{\ell}(\mathcal{M}_\ell)^2 \leq \eta_{\ell}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1})^2.
$$

Combining the latter estimates we obtain

$$
\eta_{\ell+1}^2 \leq (1 + \delta) \eta_{\ell}^2 - [(1 + \delta) - q_{\text{red}}] \eta_{\ell}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1})^2 + [R + (1 + \delta^{-1})S^2](\|u_{\ell+1} - u_\ell\|_{\mathcal{X}})
$$

$$
\leq (1 + \delta) - [(1 + \delta) - q_{\text{red}}] \theta \eta_{\ell}^2 + [R + (1 + \delta^{-1})S^2](\|u_{\ell+1} - u_\ell\|_{\mathcal{X}}).
$$

We define $\rho = (1 + \delta) - [(1 + \delta) - q_{\text{red}}] \theta > 0$ as well as $a_\ell = \eta_{\ell}^2$ and $b_\ell = [R + (1 + \delta^{-1})S^2](\|u_{\ell+1} - u_\ell\|_{\mathcal{X}})$. Choosing $\delta > 0$ sufficiently small we observe that $0 < \rho < 1$ and that a priori convergence (1.2) proves that

$$
0 \leq a_{\ell+1} \leq \rho a_\ell + b_\ell \quad \text{for all } \ell \in \mathbb{N}_0 \quad \text{with} \quad \lim_{\ell \to \infty} b_\ell = 0.
$$

By use of Lemma 3.2 we conclude that $a_\ell = \eta_\ell^2 \to 0$ as $\ell \to \infty$. $\square$

**Remark 3.4** The proof of Theorem 3.3 shows that in the case of the Dörfler marking (2.7), it is sufficient to have stability (3.1) and reduction (3.2) for one-level refinements, i.e., (3.1) and (3.2) are only required for all $\mathcal{T}_H \in \mathbb{T}$, all $\mathcal{M}_H \subseteq \mathcal{T}_H$ and $\mathcal{T}_h = \text{refine}(\mathcal{T}_H, \mathcal{M}_H)$. Assuming additionally the refinement assumption (2.1) and abbreviating $\mathcal{T}_h|_{\mathcal{U}_H} := \{T' \in \mathcal{T}_H : T' \subseteq \bigcup \mathcal{U}_H\}$ for all $\mathcal{U}_H \subseteq \mathcal{T}_H$, (3.1) and (3.2) can further be replaced by stability on nonmarked elements,

$$
\eta_h(\mathcal{T}_h|_{\mathcal{U}_H \setminus \mathcal{M}_H}) \leq \eta_H(\mathcal{T}_H \setminus \mathcal{M}_H) + S(\|u_h - u_H\|_{\mathcal{X}}), \quad (3.10)
$$

and reduction on marked elements,

$$
\eta_h(\mathcal{T}_h|_{\mathcal{M}_H})^2 \leq q_{\text{red}} \eta_H(\mathcal{M}_H)^2 + R(\|u_h - u_H\|_{\mathcal{X}}). \quad (3.11)
$$
4. **Examples**

4.1 **Laplace obstacle problem**

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain. Let $\chi$ be an affine function with $\chi \leq 0$ on $\partial \Omega$. Denote the set of admissible functions by

$$A := \{ v \in \mathcal{X} : v \geq \chi \text{ a.e. on } \Omega \} \quad \text{with } \mathcal{X} := H^1_0(\Omega)$$

and note that $A \neq \emptyset$. The minimization problem reads as follows: given a continuous linear functional $f \in H^{-1}(\Omega)$ find $u \in A$ such that

$$E(u) = \min_{v \in A} E(v), \quad \text{where } E(v) := \frac{1}{2} \| \nabla v \|_{L^2(\Omega)}^2 - \int_{\Omega} f v \, dx.$$  (4.2)

Since $A \neq \emptyset$ is convex and closed it is well known (Kinderlehrer & Stampacchia, 2000, Theorem II.2.1) that the minimization problem (4.2) admits a unique solution $u \in A$.

We consider regular triangulations $T_H$ of $\Omega$ into nondegenerate compact simplices and the corresponding first-order Courant finite element space

$$X_H := \{ v_H \in H^1_0(\Omega) : v_H|_T \text{ is affine for all } T \in T_H \}. \quad (4.3)$$

Then $A_H := A \cap X_H \neq \emptyset$ is closed and convex. As in the continuous case there hence exists a unique discrete minimizer $u_H \in A_H$ such that

$$E(u_H) = \min_{v_H \in A_H} E(v_H).$$ \quad (4.4)

Under additional regularity $f \in L^2(\Omega)$ one can argue as in Braess et al. (2007) to show reliability

$$\| \nabla (u - u_H) \|_{L^2(\Omega)}^2 \leq 2[E(u_H) - E(u)] \leq C_{rel}^2 \eta_H^2$$ \quad (4.5)

for the residual error estimator with local contributions,

$$\eta_H(T)^2 = h_T \| [n u_H] \|_{L^2(\partial T \cap \Omega)}^2 + h_T^2 \sum_{E \in \mathcal{E}_H(T) \cap \partial \Omega} \| f \|_{L^2(T_e)}^2 + h_T^2 \sum_{E \notin \mathcal{E}_H(T) \cap \partial \Omega} \| f - f_E \|_{L^2(T_e)}^2,$$ \quad (4.6)

where $h_T := |T|^{1/d}$ and $\mathcal{E}_H(T)$ is the set of all $(d - 1)$-dimensional facets (i.e., edges for $d = 2$) and $f_E \in \mathbb{R}$ is the integral mean of $f$ over the corresponding patch $\omega_E = T \cup T'$ with $E = T \cap T'$. In addition to (2.1) we suppose uniform contraction of the mesh size on refined elements, i.e., there exists $0 < q_{ctr} < 1$,

$$|T'| \leq q_{ctr}|T| \quad \text{for all } T \in T_H \subset \mathbb{T} \text{ and all } T' \subset T \in T_H \cap \mathbb{T}(T_H) \text{ with } T' \subsetneq T.$$ \quad (4.7)

Then stability (3.1) and reduction (3.2) with $S(t) = C_{\text{stab}} t$ and $R(t) = C_{\text{red}} t^2$ follow as for the linear case; see Manuel Cascon et al. (2008) (or Page & Praetorius, 2013 for the obstacle problem). The
constants $C_{\text{rel}}, C_{\text{stab}} > 0$ depend only on $\Omega, d$ and uniform $\gamma$-shape regularity of the admissible meshes $T_H \in \mathbb{T}$ in the sense of

$$\gamma := \sup_{T_H \in \mathbb{T}} \max_{T \in T_H} \frac{\text{diam}(T)}{|T|^{1/d}} < \infty,$$

(4.8)

while $C_{\text{red}} > 0$ and $0 < q_{\text{red}} < 1$ depend additionally on $q_{\text{ctr}}$. The a priori convergence (1.2) follows essentially as in the seminal work by Babuška & Vogelius (1984); assumption (2.1) on the mesh refinement implies that refinement leads to nested spaces, i.e., Algorithm 2.1 leads to

$$X_\ell \subseteq X_{\ell+1}$$

and hence

$$A_\ell \subseteq A_{\ell+1}$$

for all $\ell \in \mathbb{N}_0$. Therefore,

$$A_\infty := \text{closure}(\bigcup_{\ell \in \mathbb{N}_0} A_\ell) \neq \emptyset$$

is a closed and convex subset of $X$ and thus gives rise to a unique minimizer $u_\infty \in A_\infty$ such that

$$E(u_\infty) = \min_{v_\infty \in A_\infty} E(v_\infty).$$

(4.9)

Based on estimates for the equivalent variational inequalities (see Kinderlehrer & Stampacchia, 2000, Theorem II.2.1) it follows that

$$\|\nabla (u_\infty - u_\ell)\|_{L^2(\Omega)}^2 \lesssim \inf_{v_\ell \in A_\ell} \|\nabla (u_\infty - v_\ell)\|_{L^2(\Omega)} \to 0 \text{ as } \ell \to \infty,$$

(4.10)

where we stress the different powers of the norms, which are due to the lack of Galerkin orthogonality; see, e.g., Falk (1974). Overall, we thus get the following plain convergence result, where we note that for Dörfler marking (2.7) with sufficiently small marking parameter $0 < \theta \ll 1$, Carstensen & Hu (2015) even prove rate-optimal convergence for $d = 2$.

**Proposition 4.1** As long as the mesh-refinement strategy guarantees regular simplicial triangulations satisfying (2.1), (4.7) and (4.8) and as long as the marking strategy satisfies (2.3), Algorithm 2.1 for the Laplace obstacle problem (4.2) driven by the indicators (4.6) yields convergence

$$C_{\text{rel}}^{-1} \|\nabla (u - u_\ell)\|_{L^2(\Omega)} \leq \eta_\ell \to 0 \text{ as } \ell \to \infty.$$ 

(4.11)

### 4.2 Fractional Laplacian

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $d \geq 2$, and $0 < s < 1$. Given $f \in H^{-s}(\Omega)$ we consider the Dirichlet problem of the fractional Laplacian

$$(-\Delta)^s u = f \text{ in } \Omega \text{ subject to } u = 0 \text{ in } \mathbb{R}^d \setminus \overline{\Omega}.$$ 

(4.12)

There are several different ways to define $(-\Delta)^s$, e.g., in terms of the Fourier transformation (Bonito et al., 2018), via semigroup theory (Kwaśnicki, 2017) or as a Dirichlet-to-Neumann map of a half-space extension problem (Caffarelli & Silvestre, 2007). For the latter a convenient representation of the fractional Laplacian is given in terms of a principal value integral,

$$((-\Delta)^s u)(x) := C(d, s) \text{ p.v.} \int_{\mathbb{R}^d} \frac{u(x) - u(y)}{|x - y|^{d+2s}} \text{ dy} \quad \text{with} \quad C(d, s) := -2^{2s} \frac{\Gamma(s + d/2)}{\pi^{d/2} \Gamma(-s)},$$

(4.13)
where \( \Gamma(\cdot) \) denotes the Gamma function. According to Kwaśnicki (2017, Theorem 1.1) the weak formulation of (4.12) reads as follows: find \( u \in \mathcal{X} := \tilde{H}^s(\Omega) \) (being the closure of the space of distributions \( D(\Omega) \) in \( H^s(\mathbb{R}^d) \) equipped with the same norm) such that

\[
a(u, v) := \frac{C(d, s)}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{[u(x) - u(y)][v(x) - v(y)]}{|x - y|^{d + 2s}} \, dx \, dy = \int_{\Omega} fv \, dx \quad \text{for all} \ v \in \mathcal{X}. \tag{4.14}
\]

The Lax–Milgram lemma proves existence and uniqueness of \( u \in \mathcal{X} \).

Following Ainsworth & Glusa (2017) and Faustmann et al. (2020) we consider regular triangulations \( \mathcal{T}_H \) of the domain \( \Omega \) into nondegenerate compact simplices and the corresponding first-order Courant finite element space \( \mathcal{X}_H \) from (4.3). Let \( u_H \in \mathcal{X}_H \) be the corresponding Galerkin solution, i.e.,

\[
a(u_H, v_H) = \int_{\Omega} fv_H \, dx \quad \text{for all} \ v_H \in \mathcal{X}_H. \tag{4.15}
\]

Under the additional regularity assumption \( f \in L^2(\Omega) \), the local contributions of the error estimator from Faustmann et al. (2020) read

\[
\eta_H(T) := \| h^s_T[f - (-\Delta)^s u_H] \|_{L^2(T)} \quad \text{for all} \ T \in \mathcal{T}_H, \tag{4.16a}
\]

with the modified local mesh width

\[
h^s_T := \begin{cases} 
|T|^{s/d} & \text{for } 0 < s \leq 1/2, \\
|T|^{1/(2d)} \text{dist}(\cdot, \partial T)^{s-1/2} & \text{for } 1/2 < s < 1. \tag{4.16b}
\end{cases}
\]

According to Faustmann et al. (2020, Theorem 2.3) the error estimator is reliable:

\[
\| u - u_H \|_{\tilde{H}^s(\Omega)} \leq C_{\text{rel}} \eta_H. \tag{4.17}
\]

Provided (2.1) and (4.7), stability (3.1) and reduction (3.2) are proved in Faustmann et al. (2020, Proposition 3.1) with \( S(t) = C_{\text{stab}} t \) and \( R(t) = C_{\text{red}} t^2 \). The constants \( C_{\text{rel}}, C_{\text{stab}} > 0 \) depend only on \( \Omega, d, s \) and uniform \( \gamma \)-shape regularity (4.8), while \( C_{\text{red}} > 0 \) and \( 0 < q_{\text{ctr}} < 1 \) depend additionally on \( q_{\text{ctr}} \). Finally, a priori convergence (1.2) follows as in the seminal work by Babuška & Vogelius (1984) (and essentially with the same arguments as in the previous section): assumption (2.1) on the mesh refinement implies that refinement leads to nested spaces, i.e., Algorithm 2.1 leads to \( \mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1} \) for all \( \ell \in \mathbb{N}_0 \). Therefore, \( \mathcal{X}_\infty := \text{closure} \left( \bigcup_{\ell \in \mathbb{N}_0} \mathcal{X}_\ell \right) \) is a closed subspace of \( \mathcal{X} \) and the Lax–Milgram lemma guarantees existence and uniqueness of \( u_\infty \in \mathcal{X}_\infty \) such that

\[
a(u_\infty, v_\infty) = \int_{\Omega} fv_\infty \, dx \quad \text{for all} \ v_\infty \in \mathcal{X}_\infty. \tag{4.18}
\]

With the Galerkin orthogonality and the resulting Céa lemma it follows that

\[
\| u_\infty - u_\ell \|_{\tilde{H}^s(\Omega)} \lesssim \min_{v_\ell \in \mathcal{X}_\ell} \| u_\infty - v_\ell \|_{\tilde{H}^s(\Omega)} \to 0 \quad \text{as} \ \ell \to \infty. \tag{4.19}
\]
Overall, we thus get the following plain convergence result, where we note that for Dörfler marking (2.7) with sufficiently small marking parameter $0 < \theta \ll 1$, Faustmann et al. (2020, Theorem 2.6) proves even rate-optimal convergence.

**Proposition 4.2** As long as the mesh-refinement strategy guarantees regular simplicial triangulations satisfying (2.1), (4.7) and (4.8) and as long as the marking strategy satisfies (2.3), Algorithm 2.1 for the fractional Laplacian (4.14) driven by the indicators (44) yields convergence

$$C_{rel}^{-1} \| u - u_\ell \|_{H^s(\Omega)} \leq \eta_\ell \rightarrow 0 \text{ as } \ell \rightarrow \infty. \quad (4.20)$$

### 4.3 Weakly singular integral equations

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a Lipschitz domain with compact boundary $\partial \Omega$ such that $\text{diam}(\Omega) < 1$ if $d = 2$. On a (relatively) open subset $\Gamma \subseteq \partial \Omega$ and given $f \in H^{1/2}(\Gamma)$ the weakly singular integral equation

$$(Vu)(x) := \int_{\Gamma} G(x - y) u(y) \, dy = f(x) \quad \text{for all } x \in \Gamma \quad (4.21)$$

seeks the unknown integral density $u \in X := \tilde{H}^{-1/2}(\Omega)$. Here, $G(\cdot)$ is the fundamental solution of the Laplacian, i.e., $G(z) = -\frac{1}{2\pi} \log |z|$ for $d = 2$ resp. $G(z) = \frac{1}{4\pi} |z|^{-1}$ for $d = 3$. We note that for $\Gamma = \partial \Omega$, (4.21) is equivalent to the Dirichlet problem

$$-\Delta U = 0 \text{ in } \Omega \quad \text{subject to } U = f \text{ on } \partial \Omega,$$

supplemented by the appropriate radiation condition if $\Omega$ is unbounded; see McLean (2000). The weak formulation reads

$$a(u, v) := \iint_{\Gamma \times \Gamma} G(x - y) u(y) v(x) \, dy \, dx = \int_{\Gamma} fv \, dx \quad \text{for all } v \in X, \quad (4.22)$$

and the Lax–Milgram lemma yields existence and uniqueness of the solution $u \in X$.

For a fixed polynomial degree $p \geq 0$ and a regular triangulation $T_H$ of $\Gamma$ into nondegenerate compact surface simplices we consider standard boundary element spaces $X_H = \mathbb{P}_p(T_H)$ consisting of $T_H$-piecewise polynomials of degree $\leq p$ (w.r.t. the boundary parametrization). Let $u_H \in X_H$ be the corresponding Galerkin solution, i.e.,

$$a(u_H, v_H) = \int_{\Gamma} fv_H \, dx \quad \text{for all } v_H \in X_H. \quad (4.23)$$

According to the seminal work by Carstensen et al. (2001) and under additional regularity $f \in H^1(\Gamma)$, the local contributions of the residual error estimator read

$$\eta_H(T) := h_T^{-1/2} \| \nabla_T (f - Vu_H) \|_{L^2(T)} \quad \text{with} \quad h_T := |T|^{1/(d-1)}, \quad (4.24)$$

where $\nabla_T(\cdot)$ is the surface gradient and $| \cdot |$ is the surface measure. While Carstensen et al. (2001) prove reliability (4.17) stability (3.1) and reduction (3.2) were first proved in Feischl et al. (2013) and
Gantumur (2013) with $S(t) = C_{\text{stab}} t$ and $R(t) = C_{\text{red}} t^2$ provided that (2.1) and (4.7) are satisfied. The constants $C_{\text{rel}}, C_{\text{stab}} > 0$ depend only on $\Gamma, d, p$ and uniform $\gamma$-shape regularity of the admissible meshes $\mathcal{T}_H \in \mathcal{T}$, i.e.,

$$\gamma := \sup_{T_H \in \mathcal{T}} \max_{T, T' \in T_H} \frac{|T'|}{|T|} < \infty \quad \text{if } d = 2,$$

(4.25a)

resp.

$$\gamma := \sup_{T_H \in \mathcal{T}} \max_{T \in T_H} \frac{\text{diam}(T)}{|T|^{1/(d-1)}} < \infty \quad \text{if } d = 3,$$

(4.25b)

while $C_{\text{red}} > 0$ and $0 < q_{\text{red}} < 1$ depend additionally on $q_{\text{ctr}}$. The a priori convergence follows as in the previous section. Overall, we thus get the subsequent plain convergence result, where we note that for Dörfler marking (2.7) with sufficiently small marking parameter $0 < \theta \ll 1$, Feischl et al. (2013) and Gantumur (2013) even prove rate-optimal convergence. We also refer to our recent work Gantner & Praetorius (2020), which proves well-posedness of the residual estimator (4.24) together with reliability, stability and reduction for a large class of second-order elliptic PDEs with constant coefficients and the related weakly singular integral operator $V$, as well as general mesh-refinement strategies.

**Proposition 4.3** As long as the mesh-refinement strategy guarantees regular simplicial triangulations satisfying (2.1), (4.7) and (52), and as long as the marking strategy satisfies (2.3), Algorithm 1 for the weakly singular integral equation (4.22) driven by the indicators (4.24) yields convergence

$$C_{\text{rel}}^{-1} \| u - u_\ell \|_{H^{-\frac{1}{2}}(\Gamma)} \leq \eta_\ell \to 0 \quad \text{as } \ell \to \infty.$$

(4.26)

### 4.4 Hyper-singular integral equations

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a Lipschitz domain with compact boundary $\partial \Omega$. On a (relatively) open and connected subset $\Gamma \subseteq \partial \Omega$ and given $f \in H^{-1/2}(\Gamma)$ with $\int_{\Gamma} f \, dx = 1$, in the case of $\Gamma = \partial \Omega$ the hyper-singular integral equation

$$(Wu)(x) := \text{p.v.} \int_{\Gamma} \frac{\partial_x}{\partial \nu(x)} G(x - y) u(y) \, dy = f(x) \quad \text{for all } x \in \Gamma$$

(4.27)

seeks the unknown integral density $u \in \mathcal{X} := \{ v \in H^{1/2}(\Omega) : \int_{\Gamma} v \, dx = 0 \}$ if $\Gamma = \partial \Omega$ resp. $u \in \mathcal{X} := \tilde{H}^{1/2}(\Gamma)$ if $\Gamma \subsetneq \partial \Omega$. Here $\nu(\cdot)$ denotes the exterior normal vector and $G(\cdot)$ is again the fundamental solution of the Laplacian, i.e., $G(z) = -\frac{1}{2\pi} \log |z|$ for $d = 2$ resp. $G(z) = \frac{1}{4\pi} |z|^{-1}$ for $d = 3$. We note that for $\Gamma = \partial \Omega$, (4.27) is equivalent to the Neumann problem

$$-\Delta U = 0 \text{ in } \Omega \quad \text{subject to } \frac{\partial U}{\partial \nu} = f \text{ on } \partial \Omega,$$

supplemented by the appropriate radiation condition if $\Omega$ is unbounded; see McLean (2000). The weak formulation reads

$$a(u, v) := \int_{\Gamma \times \Gamma} G(x - y) \text{curl}_\Gamma u(y) \cdot \text{curl}_\Gamma v(x) \, dy \, dx = \int_{\Gamma} fv \, dx \quad \text{for all } v \in \mathcal{X},$$

(4.28)
where $\text{curl}_G(\cdot)$ denotes the arclength derivative for $d = 2$ resp. the surface curl for $d = 3$; see McLean (2000). The Lax–Milgram lemma yields existence and uniqueness of the solution $u \in X$.

For a fixed polynomial degree $p \geq 1$ and a regular triangulation $T_H$ of $\Gamma$ into nondegenerate compact surface simplices, we consider standard boundary element spaces $X_H = S^p(T_H)$ consisting of globally continuous $T_H$-piecewise polynomials of degree $\leq p$ (w.r.t. the boundary parametrization). Let $u_H \in X_H$ be the corresponding Galerkin solution, i.e.,

$$a(u_H, v_H) = \int_{\Gamma} f v_H \, dx \quad \text{for all } v_H \in X_H.$$  \hfill (4.29)

According to the seminal work by Carstensen et al. (2004) and under additional regularity $f \in L^2(\Gamma)$ the local contributions of the residual error estimator read

$$\eta_H(T) := h_T^{1/2} \| f - W u_H \|_{L^2(T)} \quad \text{with} \quad h_T := |T|^{1/(d-1)},$$  \hfill (4.30)

where $| \cdot |$ is the surface measure. While Carstensen et al. (2004) prove reliability (4.17), stability (3.1) and reduction (3.2) were first proved in Gantumur (2013) and Feischl et al. (2015) with $S(t) = C_{\text{stab}} t$ and $R(t) = C_{\text{red}} t^2$ provided that (2.1) and (4.7) are satisfied. The constants $C_{\text{rel}}, C_{\text{stab}} > 0$ depend only on $\Gamma, d, p$ and uniform $\gamma$-shape regularity (52) of the admissible meshes $T_H \in \mathcal{T}$, while $C_{\text{red}} > 0$ and $0 < q_{\text{ctr}} < 1$ depend additionally on $q_{\text{ctr}}$. The $a$ priori convergence follows as in Section 4.2. Overall, we thus get the subsequent plain convergence result, where we note that for Dörfler marking (2.7) with sufficiently small marking parameter $0 < \theta \ll 1$, Gantumur (2013) and Feischl et al. (2015) even prove rate-optimal convergence.

**Proposition 4.4** As long as the mesh-refinement strategy guarantees regular simplicial triangulations satisfying (2.1), (4.7) and (52), and as long as the marking strategy satisfies (2.3), Algorithm 2.1 for the hyper-singular integral equation (4.28) driven by the indicators (4.30) yields convergence

$$C_{\text{rel}}^{-1} \| u - u_\ell \|_{H^{1/2}(\Gamma)} \leq \eta_\ell \to 0 \quad \text{as } \ell \to \infty.$$  \hfill (4.31)

4.5 \textit{Nonlinear interface problems}

Let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded Lipschitz domain with compact boundary $\Gamma := \partial \Omega$ and exterior domain $\Omega^{\text{ext}} := \mathbb{R}^d \setminus \overline{\Omega}$ such that $\text{diam}(\Omega) < 1$ if $d = 2$. Further, let $A : \mathbb{R}^d \to \mathbb{R}^d$ be a Lipschitz continuous and strongly monotone coefficient function in the sense that there exist constants $C_{\text{lip}}, C_{\text{mon}} > 0$ such that

$$|Ax - Ay| \leq C_{\text{lip}}|x - y| \quad \text{for all } x, y \in \mathbb{R}^d,$$  \hfill (4.32)

$$C_{\text{mon}} \| \nabla u - \nabla v \|_{L^2(\Omega)}^2 \leq \int_\Omega (A \nabla u - A \nabla v) \cdot ( \nabla u - \nabla v ) \, dx \quad \text{for all } u, v \in H^1(\Omega).$$  \hfill (4.33)

For given data $f \in L^2(\Omega)$, $u_D \in H^{1/2}(\Gamma)$ and $\phi_N \in H^{-1/2}(\Gamma)$ with additional compatibility condition

$$\int_\Omega f \, dx + \int_\Gamma \phi_N \, dx = 0 \quad \text{in the case of } d = 2,$$  \hfill (4.34)
we consider the nonlinear interface problem
\begin{align}
-\text{div}(A\nabla u) &= f \quad \text{in } \Omega, \tag{4.35a} \\
-\Delta u^{\text{ext}} &= 0 \quad \text{in } \Omega^{\text{ext}}, \tag{4.35b} \\
u - u^{\text{ext}} &= u_D \quad \text{on } \Gamma, \tag{4.35c} \\
(A\nabla u - \nabla u^{\text{ext}}) \cdot n &= \phi_N \quad \text{on } \Gamma, \tag{4.35d} \\
u^{\text{ext}} &= \mathcal{O}(|x|^{-1}) \quad \text{as } |x| \to \infty. \tag{4.35e}
\end{align}

We seek a weak solution \((u, u^{\text{ext}}) \in H^1(\Omega) \times H^1_{\text{loc}}(\Omega^{\text{ext}})\), where \(H^1_{\text{loc}}(\Omega^{\text{ext}}) = \{v : v \in H^1(\omega) \text{ for all open and bounded } \omega \subseteq \Omega^{\text{ext}}\}\). There are different ways to equivalently reformulate (62) as an FEM–BEM coupling. To ease presentation we restrict ourselves to the Bielak–MacCamy coupling (Bielak & MacCamy, 1983), but we stress that Proposition 4.5 holds accordingly for the Johnson–Nédélec coupling (Johnson & Claude Nédélec, 1980), as well as Costabel’s symmetric coupling (Costabel, 1988); see Aurada et al. (2013) for details. Recalling the single-layer operator \(V\) from (4.21) and defining the adjoint double layer operator
\begin{equation}
(K' \phi)(x) := \int_{\Gamma} \frac{\partial}{\partial n(x)} G(x, y) \phi(y) \, dy \quad \text{for all } \phi \in H^{-1/2}(\Gamma) \text{ and all } x \in \Gamma,
\end{equation}

the variational formulation resulting from the Bielak–MacCamy coupling seeks some \(u = (u, \phi) \in X := H^1(\Omega) \times H^{-1/2}(\Gamma)\) such that
\begin{align}
a(u, v) := & \int_{\Omega} (A\nabla u) \cdot \nabla v \, dx + \int_{\Gamma} (1/2 - K') \phi v \, dx + \int_{\Gamma} (u - V \phi) \psi \, dx \\
= & \int_{\Gamma} f v \, dx + \int_{\Gamma} \phi_N v \, dx - \int_{\Gamma} u_D \psi \, dx := F(v) \quad \text{for all } v = (v, \psi) \in X. \tag{4.37}
\end{align}

According to Aurada et al. (2013), (4.37) is uniquely solvable provided that \(C_{\text{mon}} > 1/4\).

For a regular triangulation \(T_H\) of \(\Omega\) into nondegenerate compact simplices and the induced regular triangulation \(T_{\text{loc}}\) of \(\Gamma\) into nondegenerate compact surface simplices we consider globally continuous \(T_H\)-piecewise affine functions \(S^1(T_H)\) to discretize \(H^1(\Omega)\), and \(T_{\text{loc}}\)-piecewise constant functions \(P^0(T_{\text{loc}})\) to discretize \(H^{-1/2}(\Gamma)\), i.e., \(X_H = S^1(T_H) \times P^0(T_{\text{loc}})\). Let \(u_H \in X_H\) be the corresponding Galerkin solution, i.e.,
\begin{equation}
a(u_H, v_H) = F(v_H) \quad \text{for all } v_H \in X_H. \tag{4.38}
\end{equation}

According to Aurada et al. (2013) the local contributions of the residual error estimator read
\begin{align}
\eta_H(T)^2 &:= h_T^2 \|f\|^2_{L^2(T)} + h_T \left(\|[(A\nabla u_H) \cdot v]\|^2_{L^2(\partial T \cap \Omega)} \\
&\quad + \|\phi_N + (K' - 1/2) \phi_H - (A\nabla u_H) \cdot v\|^2_{L^2(\partial T \cap \Gamma)} + \|\nabla\Gamma(u_H - u_D - V \phi_H)\|^2_{L^2(\partial T \cap \Gamma)}\right) \tag{4.39}
\end{align}
with the surface gradient $\nabla T(\cdot)$ and the mesh size $h_T := |T|^{1/d} \text{ for all } T \in \mathcal{T}_H$. Indeed, Aurada et al. (2013) prove reliability (4.17), stability (3.1) and reduction (3.2) (where the terms $u_H, u_h$ are replaced by $u_H, u_h$) with $S(t) = C_{\text{stab}} t$ and $R(t) = C_{\text{red}} t^2$ provided that (2.1) and (4.7) are satisfied. The constants $C_{\text{rel}}, C_{\text{stab}} > 0$ depend only on $\Gamma, d, C_{\text{lip}}$ and uniform $\gamma$-shape regularity (4.8) of the admissible meshes $\mathcal{T}_H \in \mathcal{T}$, while $C_{\text{red}} > 0$ and $0 < q_{\text{red}} < 1$ depend additionally on $q_{\text{ctr}}$. The $a \text{ priori}$ convergence follows as in Section 4.2, where the required Céa lemma is given in Aurada et al. (2013, Corollary 12). Overall, we thus get the subsequent plain convergence result.

**Proposition 4.5** As long as the mesh-refinement strategy guarantees regular simplicial triangulations satisfying (2.1), (4.7) and (4.8), and as long as the marking strategy satisfies (2.3), Algorithm 2.1 for the Bielak–MacCamy coupling (4.37) driven by the indicators (4.39) yields convergence

$$C_{\text{rel}}^{-1} \| u - u_\ell \|_{H^1(\Omega) \times H^{-1/2}(\Gamma)} \leq \eta_\ell \to 0 \text{ as } \ell \to \infty. \quad (4.40)$$

**Funding**

Austrian Science Fund (FWF) through the SFB Taming Complexity in Partial Differential Systems (SFB F65); Austrian Science Fund (FWF) through the stand-alone project Computational Nonlinear PDEs (P33216); Austrian Science Fund (FWF) through the Erwin Schrödinger Fellowship Optimal Adaptivity for Space-Time Methods (J4379).

**References**

Ainsworth, M. & Glusa, C. (2017) Aspects of an adaptive finite element method for the fractional Laplacian: a priori and a posteriori error estimates, efficient implementation and multigrid solver. Comput. Methods Appl. Mech. Engrg., 327, 4–35.

Ainsworth, M. & Oden, J. T. (2000) A Posteriori Error Estimation in Finite Element Analysis. Pure and Applied Mathematics. New York: Wiley-Interscience.

Aurada, M., Feischl, M., Führer, T., Karkulik, M., Meink, J. M. & Praetorius, D. (2013) Classical FEM-BEM coupling methods: nonlinearities, well-posedness, and adaptivity. Comput. Mech., 51, 399–419.

Aurada, M., Ferraz-Leite, S. & Praetorius, D. (2012) Estimator reduction and convergence of adaptive BEM. Appl. Numer. Math., 62, 787–801.

Babuška, I. & Vogelius, M. (1984) Feedback and adaptive finite element solution of one-dimensional boundary value problems. Numer. Math., 44, 75–102.

Bielak, J. & MacCamy, R. C. (1983) An exterior interface problem in two-dimensional elastodynamics. Quart. Appl. Math., 41, 143–159.

Binev, P., Dahmen, W. & DeVore, R. (2004) Adaptive finite element methods with convergence rates. Numer. Math., 97, 219–268.

Bonito, A., Borthagaray, J. P., Nochetto, R. H., Otárola, E. & Salgado, A. J. (2018) Numerical methods for fractional diffusion. Comput. Vis. Sci., 19, 19–46.

Braess, D., Carstensen, C. & Hoppe, R. H. W. (2007) Convergence analysis of a conforming adaptive finite element method for an obstacle problem. Numer. Math., 107, 455–471.

Carstensen, C., Feischl, M., Page, M. & Praetorius, D. (2014) Axioms of adaptivity. Comput. Math. Appl., 67, 1195–1253.

Caffarelli, L. & Silvestre, L. (2007) An extension problem related to the fractional Laplacian. Comm. Partial Differential Equations, 32, 1245–1260.

Carstensen, C. & Hu, J. (2015) An optimal adaptive finite element method for an obstacle problem. Comput. Methods Appl. Math, 15, 259–277.
CARSTENSEN, C., MAISCHAK, M., PRAETORIUS, D. & STEPHAN, E. P. (2004) Residual-based a posteriori error estimate for hypersingular equation on surfaces. *Numer. Math.*, 97, 397–425.

CARSTENSEN, C., MAISCHAK, M. & STEPHAN, E. P. (2001) A posteriori error estimate and h-adaptive algorithm on surfaces for Symm’s integral equation. *Numer. Math.*, 90, 197–213.

CASCON, J. M., KREUZER, C., NOCHETTO, R. H. & SIEBERT, K. G. (2008) Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46, 2524–2550.

CASCÓN, J. M. & NOCHETTO, R. H. (2012) Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA J. Numer. Anal.*, 32, 1–29.

COSTABEL, M. (1988) A symmetric method for the coupling of finite elements and boundary elements. *The Mathematics of Finite Elements and Applications IV*. London: Academic Press, pp. 281–288.

DIENING, L. & KREUZER, C. (2008) Linear convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM J. Numer. Anal.*, 46, 614–638.

DIENING, L., KREUZER, C. & STEVENSON, R. (2016) Instance optimality of the adaptive maximum strategy. *Found. Comput. Math.*, 16, 33–68.

DÖRFLER, W. (1996) A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33, 1106–1124.

FALK, R. S. (1974) Error estimates for the approximation of a class of variational inequalities. *Math. Comp.*, 28, 963–971.

FAUSTMANN, M., MELENK, J. M. & PRAETORIUS, D. (2020) Quasi-optimal convergence rate for an adaptive method for the integral fractional Laplacian. Accepted for publication in *Math. Comp.*

FEISCHL, M., FÜHRER, T., KARKULIK, M., MELENK, J. M. & PRAETORIUS, D. (2015) Quasi-optimal convergence rates for adaptive boundary element methods with data approximation, part II: hyper-singular integral equation. *Electron. Trans. Numer. Anal.*, 44, 153–176.

FEISCHL, M., FÜHRER, T. & PRAETORIUS, D. (2014) Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems. *SIAM J. Numer. Anal.*, 52, 601–625.

FEISCHL, M., KARKULIK, M., MELENK, J. M. & PRAETORIUS, D. (2013) Quasi-optimal convergence rate for an adaptive boundary element method. *SIAM J. Numer. Anal.*, 51, 1327–1348.

FÜHRER, T. & PRAETORIUS, D. (2020) A short note on plain convergence of adaptive least-squares finite element methods. *Comput. Math. Appl.*, 80, 1619–1632.

GANTNER, G. & PRAETORIUS, D. (2020) Adaptive BEM for elliptic PDE systems. Part I: abstract framework for weakly-singular integral equations. Accepted for publication in *Appl. Anal.*

GANTNER, G. & STEVENSON, R. (2021) Further results on a space-time FOSLS formulation of parabolic PDEs. *ESAIM Math. Model. Numer. Anal.*, 55, 283–299.

GANTUMUR, T. (2013) Adaptive boundary element methods with convergence rates. *Numer. Math.*, 124, 471–516.

INNERBERGER, M. & PRAETORIUS, D. (2021) Instance-optimal goal-oriented adaptivity. *Comp. Methods Appl. Math.*, 4, 109–126.

JOHNSON, C. & NÉDÉLEC, J. C. (1980) On the coupling of boundary integral and finite element methods. *Math. Comp.*, 1063–1079.

KINDERLEHRER, D. & STAMPACCHIA, G. (2000) An *Introduction to Variational Inequalities and Their Applications*. Classics in Applied Mathematics, vol. 31. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM): Reprint of the 1980 Original.

KREUZER, C. & SCHEDENSACK, M. (2016) Instance optimal Crouzeix–Raviart adaptive finite element methods for the Poisson and Stokes problems. *IMA J. Numer. Anal.*, 36, 593–617.

KWAŚNICKI, M. (2017) Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.*, 20, 7–51.

MCLEAN, W. (2000) *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge: Cambridge University Press.

MORIN, P., NOCHETTO, R. H. & SIEBERT, K. G. (2000) Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38, 466–488.

MORIN, P., SIEBERT, K. G. & VEESER, A. (2008) A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, 18, 707–737.
Page, M. & Praetorius, D. (2013) Convergence of adaptive FEM for some elliptic obstacle problem. *Appl. Anal.*, 92, 595–615.

Pfeiler, C.-M. & Praetorius, D. (2020) Dörfler marking with minimal cardinality is a linear complexity problem. *Math. Comp.*, 89, 2735–2752.

Siebert, K. G. (2011) A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.*, 31, 947–970.

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

Verfürth, R. (2013) *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press.