1. Introduction

The following is a basic introduction to modern optimality proofs of adaptive mesh refinement algorithms. The material is a simplified version of [CFPP14]. For recent developments in the field concerning non-symmetric, indefinite, and time-dependent problems, see [Fei19, Fei22].

2. Abstract Analysis

2.1. **Mesh refinement.** Throughout this section, we assume that we have a fixed mesh-refinement strategy, e.g., newest-vertex bisection. Then, given an initial mesh $\mathcal{T}_0$, this allows to consider the set of possible meshes

$$\mathbb{T} := \{\mathcal{T} \text{ is a refinement of } \mathcal{T}_0\}.$$

With $|\mathcal{T}|$, we denote the number of elements of $\mathcal{T}$. Note that $\mathbb{T}$ is an infinite but still countable set, since

$$\mathbb{T} = \bigcup_{n\in\mathbb{N}} \Big\{\mathcal{T} \in \mathbb{T} : |\mathcal{T}| \le n\Big\}$$

is the countable union of finite sets.

Example 1. *In one dimension, we consider $\mathcal{T}_0 = \{[0,1]\}$. If we choose bisection as a refinement strategy ($[0,1] \mapsto \{[0,1/2],[1/2,1]\}$), the set of all possible meshes $\mathbb{T}$ consists of all meshes $\mathcal{T}$ with elements $T = [a_T, b_T] \in \mathcal{T}$ which have endpoints of the form $a_T, b_T \in \Big\{j2^{-k} : k \in \mathbb{N},\, j \in \{0,\dots,2^k\}\Big\}$,.*

Assumption 1. *We make the following assumptions on our refinement rule:*

- *For all refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ holds*

$$|\widehat{\mathcal{T}} \setminus \mathcal{T}| \le |\widehat{\mathcal{T}}| - |\mathcal{T}|, \tag{1a}$$

  i.e. each refined element is refined into at least two sons. Moreover $\widehat{\mathcal{T}} = \texttt{refine}(\mathcal{T}, \mathcal{M})$ for some $\mathcal{M} \subseteq \mathcal{T}$, there holds

$$|\widehat{\mathcal{T}}| \le C_{\text{sons}}|\mathcal{T}|, \tag{1b}$$

  i.e. each refined element is split into at most $C_{\text{sons}} > 0$ elements.
- *For given meshes $\mathcal{T}, \mathcal{T}' \in \mathbb{T}$ exists a common refinement $\mathcal{T} \oplus \mathcal{T}' \in \mathbb{T}$ such that the so-called overlay estimate holds*

$$|\mathcal{T} \oplus \mathcal{T}'| \le |\mathcal{T}| + |\mathcal{T}'| - |\mathcal{T}_0|. \tag{2}$$

- *Each sequence $\mathcal{T}_\ell \in \mathbb{T}$ of meshes generated by successive mesh-refinement, i.e. $T_j = \text{refine}(\mathcal{T}_{j-1}, \mathcal{M}_{j-1})$ for all $j = 1,\dots,\ell$ and arbitrary $\mathcal{M}_j \subseteq \mathcal{T}_j$, $j = 0,\dots,\ell-1$, satisfies*

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \le C_{\text{mesh}} \sum_{k=0}^{\ell-1} |\mathcal{M}_k| \quad \text{for all } \ell \in \mathbb{N}. \tag{3}$$

  This means that, up to a multiplicative constant, only the marked elements are refined and the "mesh closure" is negligible. (Note that e.g. newest-vertex bisection avoids hanging nodes by additional bisections, i.e. one refines more elements than only the marked elements)

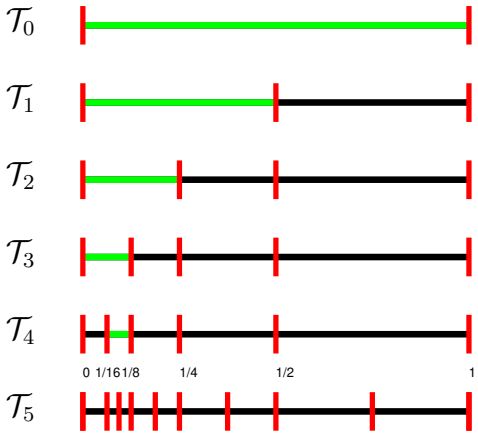
Remark. *The overlay estimate* (2) *has first been observed for newest-vertex bisection by* [Ste07] *in 2D and was generalized to* $\mathbb{R}^d$ *by* [CKNS08]. *The mesh-closure estimate* (3) *has first been observed by* [BDD04] *in 2D and was generalized by* [Ste08] *to* $\mathbb{R}^d$. *Both works require an assumption on* $\mathcal{T}_0$ *which is removed for 2D in* [KPP12]. □

Example 2. *An estimate of the type* $|\mathcal{T}_\ell| - |\mathcal{T}_{\ell-1}| \leq C|\mathcal{M}_{\ell-1}|$ *with* ℓ*-independent constant* $C > 0$ *cannot be expected.*

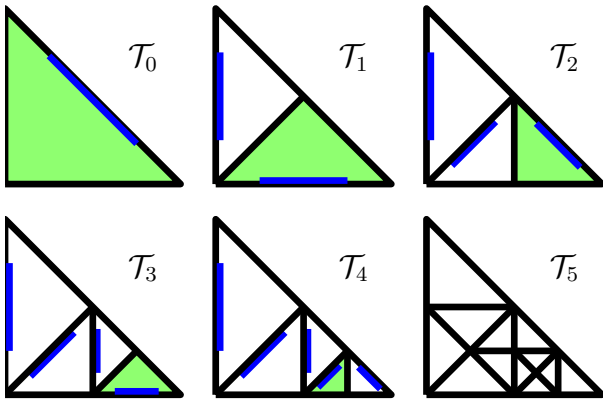
Example in 1D: *Here, the analogue to avoiding hanging nodes is that the quotient of the diameter of neighbouring elements stays bounded, e.g. by factor two*

$$\max\Big\{\operatorname{diam}(T)/\operatorname{diam}(T') \,:\, T, T' \in \mathcal{T},\, T \cap T' \neq \emptyset\Big\} \leq 2. \tag{4}$$

Now consider the situation in the figure below. We start with $\mathcal{T}_0 = \{[0,1]\}$ *and iteratively mark the leftmost elements (green). After 4 steps, this generates the mesh* $\mathcal{T}_4$. *Note that the boundedness* (4) *holds, since each element is half the size of its right neighbour. Next, we mark the second element from the left in* $\mathcal{T}_4$. *To ensure the boundedness* (4), *each element which is located right of the marked element has to be refined. In this case, the marking of one element forces 4 elements to be refined. Obviously, this example can be extended to any* $\ell \in \mathbb{N}$ *and thus shows that there are configurations for which holds* $|\mathcal{T}_\ell| - |\mathcal{T}_{\ell-1}| \geq (\ell-1)|\mathcal{M}_{\ell-1}|$.



Example in 2D: *Consider the situation below, where we iteratively mark the rightmost elements (green). The current labelling of the edges is indicated in blue. After 4 steps of refinement, we end up with* $\mathcal{T}_4$. *Now, we mark the second element from the right. Since the labelled edges are halved, each element which is located left of the marked element has to be refined to avoid hanging nodes. The refinement of one marked element generates 12 new elements.*



Remark. *The proof of the closure estimate is non-trivial, although it looks quite simple and natural. As a byproduct, the proof states that our "counterexamples" are artificial and can only occur finitely often. For newest-vertex bisection, a proof of the three properties* (1)–(3) *will be given later.* □

Remark. *Essentially, newest-vertex bisection is the only refinement strategy which satisfies the assumptions* (1)–(3). *Red-Green-Blue refinement fails to satisfy the overlay estimate* (2) [Pav10], *but satisfies the closure estimate* (3). □

2.2. **Functional setting.** Let $\mathcal{X}$ be a normed space and $u \in \mathcal{X}$ the unknown solution which we aim to approximate. For $\mathcal{T} \in \mathbb{T}$, let $\mathcal{X}_\mathcal{T}$ denote a discrete (finite dimensional) subspace of $\mathcal{X}$ with computable discrete solution $U_\mathcal{T} \in \mathcal{X}_\mathcal{T}$.

Remark. *The goal of this section is to analyze the convergence of the adaptive algorithm in a completely abstract way. To that end, we do not want to consider a particular model problem or to make assumptions on how the approximations are computed. All we need to know is that there is an exact solution $u \in \mathcal{X}$ and there exist some computable (no matter where they come from) approximations $U_\mathcal{T}$.* □

Assumption 2. *For all $\mathcal{T} \in \mathbb{T}$ and for all $\varepsilon > 0$ exists a refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}$ such that*

$$\|u - U_{\widehat{\mathcal{T}}}\|_\mathcal{X} \leq \varepsilon,$$

i.e. uniform mesh-refinement will always lead to convergence.

We start with some facts:

(1) The precise problem formulation is not needed throughout the abstract analysis.
(2) Assumptions on $\mathcal{X}$ and $\mathcal{X}_\mathcal{T}$ can be much weakened to less than quasi-metric spaces and non-conformity $\mathcal{X}_\mathcal{T} \not\subseteq \mathcal{X}$ and non-nestedness $\mathcal{X}_\mathcal{T} \not\subseteq \mathcal{X}_{\widehat{\mathcal{T}}}$ for refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$.
(3) Assumption 2 is necessary, since otherwise one cannot expect that any adaptive algorithm (or mesh-refinement) will prove successful resp. convergent at all.
(4) In practice, Assumption 2 is proved by use of the Céa lemma and the approximation properties of the discrete spaces $\mathcal{X}_\mathcal{T}$, e.g. suppose that $\mathcal{X}_\mathcal{T} \subset H^1(\Omega)$ consists of $\mathcal{T}$-piecewise polynomials. Then, there holds for all $v \in H^2(\Omega)$

$$\inf_{V \in \mathcal{X}_\mathcal{T}} \|v - V\|_{H^1(\Omega)} \lesssim h \|D^2 v\|_{L^2(\Omega)},$$

where $h > 0$ is the maximal mesh-size of $\mathcal{T}$. With this, given $\varepsilon > 0$ and the exact solution $u \in H^1(\Omega)$, we choose $v \in H^2(\Omega)$ with $\|u - v\|_{H^1(\Omega)} < \varepsilon$ and prove by use of the Céa lemma

$$\begin{aligned} \|u - U_\mathcal{T}\|_{H^1(\Omega)} &\lesssim \inf_{V \in \mathcal{X}_\mathcal{T}} \|u - V\|_{H^1(\Omega)} \\ &\leq \inf_{V \in \mathcal{X}_\mathcal{T}} \|v - V\|_{H^1(\Omega)} + \varepsilon \\ &\lesssim h \|D^2 v\|_{L^2(\Omega)} + \varepsilon. \end{aligned}$$

Choosing $\varepsilon, h > 0$ sufficiently small, we prove convergence of $U_\mathcal{T}$ to u as $h \to 0$. Note that this technique does not provide any convergence rates.

Before we proceed, we have to agree on some notation for the error estimator. Note that the structure or type of the estimator is not important. All we need to know is that the estimator $\eta(\cdot)$ consists of elementwise contributions

$$\eta_\mathcal{T}(T, V) \geq 0 \quad \text{for all } T \in \mathcal{T} \text{ and all } V \in \mathcal{X}_\mathcal{T},$$

for which we write

$$\eta_{\mathcal{T}}(T) := \eta_{\mathcal{T}}(T, U_{\mathcal{T}}) \quad \text{for all } \mathcal{T} \in \mathbb{T}$$

if the discrete solution is used as input. Moreover, the global estimator reads

$$\eta_{\mathcal{T}}(V) := \Big(\sum_{T \in \mathcal{T}} \eta_{\mathcal{T}}(T, V)^2 \Big)^{1/2} \quad \text{for all } T \in \mathcal{T} \text{ and all } V \in \mathcal{X}_{\mathcal{T}},$$

and again we write

$$\eta_{\mathcal{T}} := \eta_{\mathcal{T}}(U_{\mathcal{T}}) \quad \text{for all } \mathcal{T} \in \mathbb{T}.$$

For meshes $\mathcal{T}_\ell$ generated by the adaptive algorithm, we use the abbreviate notation $\eta_\ell(\cdot) = \eta_{\mathcal{T}_\ell}(\cdot)$ and $U_\ell = U_{\mathcal{T}_\ell}$ for the associated quantities.

2.3. **First assumptions on estimator.** The following assumptions are used throughout the lecture:

(A1) **Stability on non-refined elements:** There exists $C_{\mathrm{stab}} > 0$ such that for refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ and all subsets of non-refined elements $\mathcal{S} \subseteq \mathcal{T} \cap \widehat{\mathcal{T}}$ there holds

$$\Big| \Big(\sum_{T \in \mathcal{S}} \eta_{\widehat{\mathcal{T}}}(T, \widehat{V})^2 \Big)^{1/2} - \Big(\sum_{T \in \mathcal{S}} \eta_{\mathcal{T}}(T, V)^2 \Big)^{1/2} \Big| \le C_{\mathrm{stab}} \|\widehat{V} - V\|_{\mathcal{X}}$$

for all $V \in \mathcal{X}_{\mathcal{T}}$ and all $\widehat{V} \in \mathcal{X}_{\widehat{\mathcal{T}}}$.

(A2) **Reduction on refined elements:** There exists $C_{\mathrm{red}} > 0$ and $0 < q_{\mathrm{red}} < 1$ such that for all refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ there holds

$$\sum_{T \in \widehat{\mathcal{T}} \setminus \mathcal{T}} \eta_{\widehat{\mathcal{T}}}(T)^2 \le q_{\mathrm{red}} \sum_{T \in \mathcal{T} \setminus \widehat{\mathcal{T}}} \eta_{\mathcal{T}}(T)^2 + C_{\mathrm{red}} \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.$$

(A3) **General quasi-orthogonality**: For all $\varepsilon > 0$ exists $C_{\mathrm{orth}}(\varepsilon) > 0$ such that the adaptive algorithm guarantees

$$\sum_{k=\ell}^{N} \Big(\|U_{k+1} - U_k\|_{\mathcal{X}}^2 - \varepsilon \|u - U_k\|_{\mathcal{X}}^2 \Big) \le C_{\mathrm{orth}}(\varepsilon) \eta_\ell^2$$

for all $\ell, N \in \mathbb{N}$ with $N \ge \ell$. (Recall the abbreviations $U_k = U_{\mathcal{T}_k}$ and $\eta_\ell = \eta_{\mathcal{T}_\ell}$)

(A4) **Reliability:** There exists $C_{\mathrm{rel}} > 0$ such that for all $\mathcal{T} \in \mathbb{T}$ there holds

$$\|u - U_{\mathcal{T}}\|_{\mathcal{X}} \le C_{\mathrm{rel}} \eta_{\mathcal{T}}.$$

Remark. *Assume nestedness $\mathcal{X}_k \subseteq \mathcal{X}_{k+1}$ for all $k \in \mathbb{N}_0$. If $a(\cdot,\cdot)$ is the scalar product on $\mathcal{X}$ with induced norm $\|v\|_{\mathcal{X}}^2 = a(v,v)$, the Galerkin-orthogonality $a(u - U_{k+1}, V) = 0$ for all $V \in \mathcal{X}_{k+1}$, implies the Pythagoras theorem*

$$\|u - U_{k+1}\|_{\mathcal{X}}^2 + \|U_{k+1} - U_k\|_{\mathcal{X}}^2 = \|u - U_k\|_{\mathcal{X}}^2.$$

This gives

$$\sum_{k=\ell}^{N} \|U_{k+1} - U_k\|_{\mathcal{X}}^2 = \sum_{k=\ell}^{N} \Big(\|u - U_k\|_{\mathcal{X}}^2 - \|u - U_{k+1}\|_{\mathcal{X}}^2 \Big) \le \|u - U_\ell\|_{\mathcal{X}}^2 \le C_{\mathrm{rel}}^2 \eta_\ell^2$$

i.e. general quasi-orthogonality (A3) *with $C_{\mathrm{orth}} = C_{\mathrm{rel}}^2$ and $\varepsilon = 0$.* □

2.4. **Linear convergence of adaptive Algorithm.** The following theorem is the main result of this subsection.

Theorem 3 (R-linear convergence)**.** *Under the assumptions* (A1)–(A4)*, there exist constants* $0 < q_{\mathrm{conv}} < 1$ *and* $C_{\mathrm{conv}} > 0$ *such that*

$$\eta_{\ell+k}^2 \le C_{\mathrm{conv}} q_{\mathrm{conv}}^k \eta_\ell^2 \quad \text{for all } \ell, k \in \mathbb{N}, \tag{5}$$

i.e. R-linear convergence of the estimator to zero.

Before we come to the proof, we need some preparation.

Lemma 4 (Estimator reduction)**.** *The axioms stability* (A1) *and reduction* (A2) *imply the existence of* $C_{\mathrm{est}} > 0$ *and* $0 < q_{\mathrm{est}} < 1$ *such that*

$$\eta_{\ell+1}^2 \le q_{\mathrm{est}} \eta_\ell^2 + C_{\mathrm{est}} \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2 \quad \text{for all } \ell \in \mathbb{N}. \tag{6}$$

Proof. First, we recall the Young inequality: We start with $2ab \le a^2 + b^2$ for all $a, b \in \mathbb{R}$. For all $\delta > 0$, we obtain immediately

$$2ab = 2\sqrt{\delta} a \frac{b}{\sqrt{\delta}} \le \delta a^2 + \delta^{-1} b^2,$$

and hence

$$(a+b)^2 = a^2 + 2ab + b^2 \le (1+\delta) a^2 + (1+\delta^{-1}) b^2. \tag{7}$$

Second, the estimator is split into two parts

$$\eta_{\ell+1}^2 = \sum_{\mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell} \eta_{\ell+1}(T)^2 + \sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_{\ell+1}(T)^2. \tag{8}$$

For the first sum, we use reduction (A2) and obtain

$$\sum_{\mathcal{T}_{\ell+1} \setminus \mathcal{T}_\ell} \eta_{\ell+1}(T)^2 \le q_{\mathrm{red}} \sum_{\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} \eta_\ell(T)^2 + C_{\mathrm{red}} \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2.$$

For the second sum in (8), we employ stability (A1) and the Young inequality with $\delta > 0$ to obtain

$$\begin{aligned} \sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_{\ell+1}(T)^2 &\le \Big(\Big(\sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_\ell(T)^2 \Big)^{1/2} + C_{\mathrm{stab}} \|U_{\ell+1} - U_\ell\|_{\mathcal{X}} \Big)^2 \\ &\le (1+\delta) \sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_\ell(T)^2 + (1+\delta^{-1}) C_{\mathrm{stab}}^2 \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2. \end{aligned}$$

Plugging the last two estimates into (8), we end up with

$$\eta_{\ell+1}^2 \le q_{\mathrm{red}} \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} \eta_\ell(T)^2 + (1+\delta) \sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_\ell(T)^2 + ((1+\delta^{-1}) C_{\mathrm{stab}}^2 + C_{\mathrm{red}}) \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2.$$

With $\sum_{T \in \mathcal{T}_{\ell+1} \cap \mathcal{T}_\ell} \eta_\ell(T)^2 = \eta_\ell^2 - \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} \eta_\ell(T)^2$ and $C_{\mathrm{est}} := (1+\delta^{-1}) C_{\mathrm{stab}}^2 + C_{\mathrm{red}}$, we proceed as

$$\begin{aligned} \eta_{\ell+1} &\le (1+\delta) \eta_\ell^2 + (q_{\mathrm{red}} - (1+\delta)) \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} \eta_\ell(T)^2 + C_{\mathrm{est}} \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2 \\ &\le (1+\delta) \eta_\ell^2 + (q_{\mathrm{red}} - (1+\delta)) \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2 + C_{\mathrm{est}} \|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2, \end{aligned}$$

where we used $q_{\mathrm{red}} - (1+\delta) < 0$ and $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$. Finally, we use Dörfler marking $-\sum_{T\in\mathcal{M}_\ell} \eta_\ell(T)^2 \le -\theta\eta_\ell^2$ to conclude

$$\begin{aligned}\eta_{\ell+1} &\le (1+\delta)\eta_\ell^2 + (q_{\mathrm{red}} - (1+\delta))\theta\eta_\ell^2 + C_{\mathrm{est}}\|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2 \\ &\le \Big((1+\delta) + \theta(q_{\mathrm{red}} - (1+\delta))\Big)\eta_\ell^2 + C_{\mathrm{est}}\|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2.\end{aligned}$$

We choose $\delta > 0$ sufficiently small, such that $0 < q_{\mathrm{est}} := (1+\delta) + \theta(q_{\mathrm{red}} - (1+\delta)) = 1 - \theta(1 - q_{\mathrm{red}}) + (1+\theta)\delta < 1$ holds. □

We shall see in the next corollary that in certain situations estimator reduction already proves convergence $U_\ell \to u$.

Corollary 5. *Let $a(\cdot,\cdot)$ be an elliptic bilinear form on $\mathcal{X}$ and assume $f \in \mathcal{X}^*$. Let $u \in \mathcal{X}$ and $U_{\mathcal{T}} \in \mathcal{X}_{\mathcal{T}}$ for all $\mathcal{T} \in \mathbb{T}$ be the solutions of*

$$\begin{aligned}a(u, v) &= f(v) \quad \text{for all } v \in \mathcal{X}, \\ a(U_{\mathcal{T}}, V) &= f(V) \quad \text{for all } V \in \mathcal{X}_{\mathcal{T}}.\end{aligned}$$

Moreover, suppose nestedness $\mathcal{X}_\ell \subseteq \mathcal{X}_{\ell+1}$ for all $\ell \in \mathbb{N}$. Then, the estimator reduction from Lemma 4 implies estimator convergence $\lim_\ell \eta_\ell = 0$. Reliability (A4) *even yields convergence $\lim_\ell \|u - U_\ell\|_{\mathcal{X}} = 0$.*

Proof. We are in the frame of the Lax-Milgram lemma. Hence, each closed subspace $\mathcal{X}_\infty \subseteq \mathcal{X}$ admits a unique Galerkin solution $U_\infty \in \mathcal{X}_\infty$ of

$$a(U_\infty, V) = f(V) \quad \text{for all } V \in \mathcal{X}_\infty.$$

Moreover, the Céa lemma holds

$$\|u - U_\infty\|_{\mathcal{X}} \le C_{\mathrm{Cea}} \inf_{V\in\mathcal{X}_\infty} \|u - V\|.$$

We define $\mathcal{X}_\infty := \overline{\bigcup_{\ell\in\mathbb{N}} \mathcal{X}_\ell} \subseteq \mathcal{X}$. Since $\mathcal{X}_\ell \subseteq \mathcal{X}_\infty$, there holds

$$a(U_\infty, V) = a(u, V) = a(U_\ell, V) \quad \text{for all } V \in \mathcal{X}_\ell.$$

This shows that U_ℓ is also the unique Galerkin approximation to U_∞ in $\mathcal{X}_\ell$. Therefore, the Céa lemma holds also for U_∞, i.e.

$$\|U_\infty - U_\ell\|_{\mathcal{X}} \le C_{\mathrm{Cea}} \inf_{V\in\mathcal{X}_\ell} \|U_\infty - V\|_{\mathcal{X}}^2. \tag{9}$$

For given $\varepsilon > 0$, there exists $\ell_\varepsilon \in \mathbb{N}$ and $V_\varepsilon \in \mathcal{X}_{\ell_\varepsilon}$ such that $\|U_\infty - V_\varepsilon\|_{\mathcal{X}} \le \varepsilon$ by definition of $\mathcal{X}_\infty$. Since $V_\varepsilon \in \mathcal{X}_\ell$ for all $\ell \ge \ell_\varepsilon$, this together with (9) implies

$$\|U_\infty - U_\ell\|_{\mathcal{X}} \le C_{\mathrm{Cea}}\varepsilon.$$

Overall, we thus get $\lim_\ell \|U_\infty - U_\ell\|_{\mathcal{X}} = 0$.

With the notation $\alpha_\ell := C_{\mathrm{est}}\|U_{\ell+1} - U_\ell\|_{\mathcal{X}}^2$, the estimator reduction from Lemma 4 reads $\eta_{\ell+1}^2 \le q_{\mathrm{est}}\eta_\ell^2 + \alpha_\ell$ for all $\ell \in \mathbb{N}$. We just proved that U_ℓ converges and thus is in particular a Cauchy sequence. This implies $\alpha_\ell \to 0$ as $\ell \to \infty$. If we can show $\sup_{\ell\in\mathbb{N}} \eta_\ell^2 < \infty$ which is proved below, we get

$$\limsup_{\ell\in\mathbb{N}} \eta_{\ell+1}^2 \le q_{\mathrm{est}} \limsup_{\ell\in\mathbb{N}} \eta_\ell^2 + \limsup_{\ell\in\mathbb{N}} \alpha_\ell = \limsup_{\ell\in\mathbb{N}} \eta_{\ell+1}^2.$$

Since $0 < q_{\mathrm{est}} < 1$, this shows

$$0 \le \liminf_{\ell\in\mathbb{N}} \eta_\ell^2 \le \limsup_{\ell\in\mathbb{N}} \eta_\ell^2 = 0,$$

and hence $\lim_{\ell\to\infty} \eta_\ell^2 = 0$. With reliability (A4), this also yields convergence

$$\lim_{\ell\to\infty} \|u - U_\ell\|_{\mathcal{X}}^2 \le C_{\mathrm{rel}} \lim_{\ell\to\infty} \eta_\ell^2 = 0.$$

Altogether, it only remains to show $\sup_{\ell\in\mathbb{N}} \eta_\ell^2 < \infty$. By induction on ℓ, we see

$$\eta_\ell^2 \le q_{\mathrm{est}}^\ell \eta_0^2 + \sum_{k=1}^{\ell} q_{\mathrm{est}}^{\ell-k} \alpha_{k-1} \quad \text{for all } \ell \in \mathbb{N}_0. \tag{10}$$

For $\ell = 0$, the statement reads $\eta_0^2 = \eta_0^2$ and is obviously true. For $\ell > 0$, we have by use of the estimator reduction (6) and the induction hypothesis

$$\eta_\ell^2 \le q_{\mathrm{est}} \eta_{\ell-1}^2 + \alpha_{\ell-1} \le q_{\mathrm{est}}^\ell \eta_0^2 + \sum_{k=1}^{\ell-1} q_{\mathrm{est}}^{\ell-k} \alpha_{k-1} + \alpha_{\ell-1}.$$

This proves (10), and we conclude

$$\eta_\ell^2 \le q_{\mathrm{est}}^\ell \eta_0^2 + (1 - q_{\mathrm{est}})^{-1} \sup_{k\in\mathbb{N}} \alpha_k < \infty$$

and hence $\sup_{\ell\in\mathbb{N}} \eta_\ell^2 < \infty$. □

Lemma 6. *The following statements are pairwise equivalent:*

(i) *summability:*

$$\sum_{k=\ell+1}^{\infty} \eta_k^2 \le C_1 \eta_\ell^2 \quad \textit{for all } \ell \in \mathbb{N}_0, \tag{11}$$

(ii) *inverse summability: for all $s > 0$ there exists $C_2 = C_2(s) > 0$ such that*

$$\sum_{k=0}^{\ell-1} \eta_k^{-1/s} \le C_2 \eta_\ell^{-1/s} \quad \textit{for all } \ell \in \mathbb{N}, \tag{12}$$

(iii) *R-linear convergence: there exists $0 < q_3 < 1$ such that*

$$\eta_{\ell+k}^2 \le C_3 q_3^k \eta_\ell^2 \quad \textit{for all } \ell, k \in \mathbb{N}_0. \tag{13}$$

Proof. **(iii) $\Longrightarrow$ (i)**: We start with R-linear convergence (13). By use of the convergence of the geometric series, we see

$$\sum_{k=\ell+1}^{\infty} \eta_k^2 = \sum_{k=1}^{\infty} \eta_{\ell+k}^2 \le C_3 \Big(\sum_{k=1}^{\infty} q_3^k \Big) \eta_\ell^2 \le C_1 \eta_\ell^2,$$

where $C_1 = C_3 \sum_{k=1}^{\infty} q_3^k < \infty$. This shows (i).

(iii) $\Longrightarrow$ (ii): From (13), we get

$$\eta_\ell^{-1/s} \le C_3^{-1/(2s)} q_3^{k/(2s)} \eta_{\ell+k}^{-1/s} \quad \text{for all } k, \ell \in \mathbb{N}_0 \text{ and all } s > 0,$$

and hence

$$\eta_k^{-1/s} \le C_3^{-1/(2s)} q_3^{(\ell-k)/(2s)} \eta_\ell^{-1/s} \quad \text{for all } k, \ell \in \mathbb{N} \text{ with } \ell \ge k \text{ and all } s > 0.$$

This proves

$$\sum_{k=0}^{\ell-1} \eta_k^{-1/s} \le C_3^{-1/(2s)} \Big(\sum_{k=0}^{\ell-1} q_3^{(\ell-k)/(2s)} \Big) \eta_\ell^{-1/s} \le C_2 \eta_\ell^{-1/s},$$

where $C_3^{-1/(2s)} \sum_{k=0}^{\ell-1} q_3^{(\ell-k)/(2s)} \le C_2 := C_3^{-1/(2s)} \sum_{j=1}^{\infty} q_3^{j/(2s)} < \infty$. This shows (ii).

(i) $\Longrightarrow$ (iii): We assume summability (11). There holds

$$(1 + C_1^{-1}) \sum_{k=\ell+1}^{\infty} \eta_k^2 \le \sum_{k=\ell+1}^{\infty} \eta_k^2 + \eta_\ell^2 = \sum_{k=\ell}^{\infty} \eta_k^2.$$

With $q_3 := (1 + C_1^{-1})^{-1}$, we see

$$\sum_{k=\ell+1}^{\infty} \eta_k^2 \le q_3 \sum_{k=\ell}^{\infty} \eta_k^2.$$

By induction on j, this yields

$$\eta_{\ell+j}^2 \le \sum_{k=\ell+j}^{\infty} \eta_k^2 \le q_3^j \sum_{k=\ell}^{\infty} \eta_k^2 = q_3^j \Big(\sum_{k=\ell+1} \eta_k^2 + \eta_\ell^2 \Big) \le q_3^j (C_1 + 1) \eta_\ell^2.$$

This proves (iii) with $C_3 = C_1 + 1$.

(ii) $\Longrightarrow$ (iii): We assume (12). Basically, we repeat the arguments of the previous step. There holds

$$(1 + C_2^{-1}) \sum_{k=0}^{\ell-1} \eta_k^{-1/s} \le \sum_{k=0}^{\ell-1} \eta_k^{-1/s} + \eta_\ell^{-1/s} = \sum_{k=0}^{\ell} \eta_k^{-1/s}.$$

With $\widetilde{q}_3 := (1 + C_2^{-1})^{-1}$, we see

$$\sum_{k=0}^{\ell-1} \eta_k^{-1/s} \le \widetilde{q}_3 \sum_{k=0}^{\ell} \eta_k^{-1/s}.$$

By induction, we obtain

$$\eta_\ell^{-1/s} \le \sum_{k=0}^{\ell} \eta_k^{-1/s} \le \widetilde{q}_3^j \sum_{k=0}^{\ell+j} \eta_k^{-1/s} = \widetilde{q}_3^j \Big(\sum_{k=0}^{\ell+j-1} \eta_k^{-1/s} + \eta_{\ell+j}^{-1/s} \Big) \le \widetilde{q}_3^j (C_2 + 1) \eta_{\ell+j}^{-1/s}.$$

Taking the equation to the power of $-2s$, we end up with

$$\eta_{\ell+j}^2 \le (C_2 + 1)^{2s} \widetilde{q}_3^{2sj} \eta_\ell^2.$$

This shows (iii) with $q_3 = \widetilde{q}_3^{2s}$ and $C_3 = (C_2 + 1)^{2s}$. □

With the collected ingredients, we are able to prove the first convergence result, which actually contains any information about the speed of convergence. The following proposition contains also the result of Theorem 3.

Proposition 7. *There hold the following statements:*

(i) *Result of Theorem 3: The estimator reduction* (6)*, reliability* (A4)*, and general quasi-orthogonality* (A3) *imply R-linear convergence*

$$\eta_{\ell+k}^2 \le C_{\mathrm{conv}} q_{\mathrm{conv}}^k \eta_\ell^2 \quad \text{for all } \ell, k \in \mathbb{N}. \tag{14}$$

(ii) *Conversely, reliability* (A4) *and R-linear convergence* (14) *imply general quasi-orthogonality* (A3) *with* $q_{\mathrm{orth}} = 0$.

Proof of (ii). We use (11) and reliability (A4)

$$\begin{aligned}
\sum_{k=\ell}^{N} \|U_{k+1} - U_k\|_{\mathcal{X}}^2 &\le 2 \sum_{k=\ell}^{N} \Big(\|u - U_{k+1}\|_{\mathcal{X}}^2 + \|u - U_k\|_{\mathcal{X}}^2 \Big) \\
&\le 4 \sum_{k=\ell}^{N+1} \|u - U_k\|_{\mathcal{X}}^2 \\
&\le 4 C_{\mathrm{rel}} \sum_{k=\ell}^{N+1} \eta_k^2 \le 4 C_{\mathrm{rel}} (1 + C_1) \eta_\ell^2.
\end{aligned}$$

This proves general quasi-orthogonality (A3) with $C_{\mathrm{orth}} = 4C_{\mathrm{rel}}(1+C_1)$ and $q_{\mathrm{orth}} = 0$. □

Proof of (i). We prove (11) and use Lemma 6 to obtain (14). With the estimator reduction (6) and reliability (A4), we get

$$\begin{aligned}\sum_{k=\ell+1}^{N} \eta_k^2 &\le \sum_{k=\ell+1}^{N} \Big(q_{\mathrm{est}}\eta_{k-1}^2 + C_{\mathrm{est}}\|U_k - U_{k-1}\|_{\mathcal{X}}^2\Big)\\ &\le \sum_{k=\ell+1}^{N} \Big((q_{\mathrm{est}}+\delta)\eta_{k-1}^2 + C_{\mathrm{est}}\Big(\|U_k - U_{k-1}\|_{\mathcal{X}}^2 - \delta C_{\mathrm{est}}^{-1}\eta_{k-1}^2\Big)\Big)\\ &\le \sum_{k=\ell+1}^{N} \Big((q_{\mathrm{est}}+\delta)\eta_{k-1}^2 + C_{\mathrm{est}}\Big(\|U_k - U_{k-1}\|_{\mathcal{X}}^2 - \delta C_{\mathrm{est}}^{-1}C_{\mathrm{rel}}^{-2}\|u - U_{k-1}\|_{\mathcal{X}}^2\Big)\Big).\end{aligned}$$

Next, we choose sufficiently small $\delta < 1 - q_{\mathrm{est}}$ and sufficiently small $\varepsilon \le \delta C_{\mathrm{est}}^{-1}C_{\mathrm{rel}}^{-2}$. Hence, we may use the general quasi-orthogonality (A3) and obtain

$$\sum_{k=\ell+1}^{N} \eta_k^2 \le (q_{\mathrm{est}}+\delta)\sum_{k=\ell+1}^{N} \eta_{k-1}^2 + C_{\mathrm{est}}C_{\mathrm{orth}}(\varepsilon)\eta_\ell^2.$$

We rearrange the equation as

$$(1 - q_{\mathrm{orth}} - \delta)\sum_{k=\ell+1}^{N} \eta_k^2 \le (q_{\mathrm{est}}+\delta)\eta_\ell^2 + C_{\mathrm{est}}C_{\mathrm{orth}}(\varepsilon)\eta_\ell^2.$$

This proves (11) with $C_1 = (q_{\mathrm{est}} + \delta + C_{\mathrm{est}}C_{\mathrm{orth}}(\varepsilon))/(1 - q_{\mathrm{orth}} - \delta)$. Finally, Lemma 6 implies (14). □

So far, we introduced four axioms (A1)–(A4) and proved

$$\text{(A1)–(A4)} \quad \Longrightarrow \quad R\text{-linear convergence (14)}.$$

Recall that the proof is split into two major substeps:

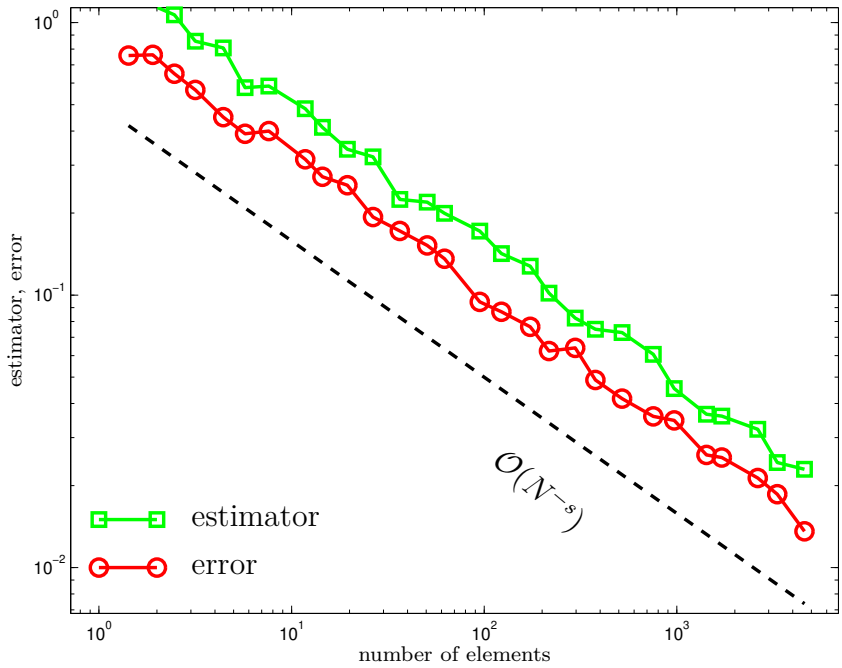
(1) Stability (A1) and reduction (A2) imply the estimator reduction (6).
(2) Estimator reduction (6) combined with the general quasi-orthogonality (A3) and reliability (A4) imply R-linear convergence (14).

Note that the assumptions on the mesh-refinement have not been used, yet. Our next goal will be to improve on the R-linear convergence and show *quasi-optimal* convergence rates.

2.5. **Optimal convergence rates for the estimator.** The fundamental question of the section is if one can prove that the estimator $\eta(\cdot)$ converges to zero with algebraic convergence rates, i.e. if there exist constants $s, C > 0$ such that there holds

$$\eta_\ell \le C(|\mathcal{T}_\ell|)^{-s} \quad \text{for all } \ell \in \mathbb{N}_0. \tag{15}$$

In this case, we say that η_ℓ converges to zero with a rate of N^{-s}. The question is interesting for a reason. Given any numerical example: If we plot the error and the estimator over the number of elements in a graph with logarithmic scaling, we observe something like this:



Here, the error estimator η_ℓ is shown in green, whereas the error $\|u - U_\ell\|_{\mathcal{X}}$ is shown in red. The estimate (15) formulates the fact, that the error estimator converges to zero at least as fast as a certain line with slope $-s$ (dashed black line in the figure above). This has been observed in practical experiments since more than 20 years. However, first proofs became available not before 2004.

In this frame, the natural question is:

What is the largest possible $s \geq 0$ such that (15) still holds true?

To answer that, we first have to study which rate would be achievable if we could choose the best possible meshes in each step. (Note that in (15), the meshes are chosen by the adaptive algorithm and at this point, we do not know if they are optimal in any sense.) To that end, we define the approximation class $\mathbb{B}_s$ for all $s > 0$:

$$u \in \mathbb{B}_s \quad \overset{\text{def.}}{\iff} \quad \|u\|_{\mathbb{B}_s} := \eta_{\mathcal{T}_0} + \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} (N^s \eta_{\mathcal{T}}) < \infty, \tag{16}$$

where $\mathbb{T}(N) := \big\{\mathcal{T} \in \mathbb{T} : |\mathcal{T}| - |\mathcal{T}_0| \leq N\big\}$. If $u \in \mathbb{B}_s$ for some $s > 0$, the definition of $\mathbb{B}_s$ implies that there exists a sequence of meshes $\widetilde{\mathcal{T}}_\ell$, $\ell \in \mathbb{N}$ and a constant $\widetilde{C} > 0$ such that

$$\eta_{\widetilde{\mathcal{T}}_\ell} \leq \widetilde{C}(|\widetilde{\mathcal{T}}_\ell| - |\mathcal{T}_0|)^{-s} \quad \text{for all } \ell \in \mathbb{N}_0. \tag{17}$$

Note that we cannot compute $\widetilde{\mathcal{T}}_\ell$ and that we do not even know if $\widetilde{\mathcal{T}}_{\ell+1}$ is a refinement of $\mathcal{T}_\ell$.

Definition 8. *The adaptive algorithm is optimal if for all $s > 0$, there exists a constant $C_{\text{opt}} > 0$ such that there holds*

$$u \in \mathbb{B}_s \quad \iff \quad \eta_\ell \leq C_{\text{opt}}(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s} \quad \textit{for all } \ell \in \mathbb{N}. \tag{18}$$

Note carefully the difference to the definition of $\mathbb{B}_s$ in (16) *and the interpretation* (17)*. Here, we want the meshes $\mathcal{T}_\ell$ generated by the adaptive algorithm to reveal a certain convergence rate, whereas in* (17)*, theoretically chosen optimal meshes $\widetilde{\mathcal{T}}_\ell$ show the convergence rate.*

Lemma 9. *There holds the $\Longleftarrow$-implication of* (18)*, i.e.*

$$\eta_\ell \le C_{\rm opt}(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s} \quad \text{for all } \ell \in \mathbb{N} \quad \Longrightarrow \quad u \in \mathbb{B}_s \tag{19}$$

for all $s > 0$.

Proof. Let $N \ge N_0 := |\mathcal{T}_1| - |\mathcal{T}_0|$. For each $N \ge N_0$, we choose $\ell = \ell(N) \in \mathbb{N} \setminus \{0\}$ maximal such that $|\mathcal{T}_\ell| - |\mathcal{T}_0| \le N$. Obviously, there holds $|\mathcal{T}_{\ell+1}| - |\mathcal{T}_0| > N$ and since each refined element is split into a bounded number of sons (1b), we see $N < |\mathcal{T}_{\ell+1}| - |\mathcal{T}_0| \le C_{\rm sons}|\mathcal{T}_\ell| - |\mathcal{T}_0|$ for all $\ell \in \mathbb{N}$. With this, we get

$$\eta_\ell N^s \le C_{\rm opt}\Big(\frac{C_{\rm sons}|\mathcal{T}_\ell| - |\mathcal{T}_0|}{|\mathcal{T}_\ell| - |\mathcal{T}_0|}\Big)^{-s}.$$

This shows

$$\begin{aligned}\sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} (N^s\eta_{\mathcal{T}}) &\le \sup_{N\ge N_0} (N^s\eta_{\ell(N)}) + \max_{0<N\le N_0} \min_{\mathcal{T}\in\mathbb{T}(N)} (N^s\eta_{\mathcal{T}}) \\ &\le \sup_{N\ge N_0} \Big(\frac{C_{\rm sons}|\mathcal{T}_{\ell(N)}| - |\mathcal{T}_0|}{|\mathcal{T}_{\ell(N)}| - |\mathcal{T}_0|}\Big)^{-s} + \max_{0<N\le N_0} \min_{\mathcal{T}\in\mathbb{T}(N)} (N^s\eta_{\mathcal{T}}) < \infty.\end{aligned}$$

The boundedness holds since the second term is the maximum of a finite set and

$$\sup_{N\ge N_0} \Big(\frac{C_{\rm sons}|\mathcal{T}_{\ell(N)}| - |\mathcal{T}_0|}{|\mathcal{T}_{\ell(N)}| - |\mathcal{T}_0|}\Big)^{-s} \le C_{\rm sons}^{-s} + \Big(\frac{(C_{\rm sons}-1)|\mathcal{T}_0|}{|\mathcal{T}_1| - |\mathcal{T}_0|}\Big)^{-s}.$$

This shows $u \in \mathbb{B}_s$. □

To prove the $\Longrightarrow$-implication of (18), we have to work harder. The concept of proof goes back to Stevenson (2007) [Ste07] and, it has been simplified by Cascon-Kreuzer-Nochetto-Siebert (2008) [CKNS08]. They considered a so-called total-error quantity for the definition of the approximation class. The drawback of this approach is, that it is well-designed for a particular model problem, but rather hard to generalize. We follow the equivalent approach of Aurada-Feischl-Kemetmüller-Page-Praetorius (2013) [AFK+13] which works with the estimator only.

Either of the approaches needs an additional axiom for the error estimator.

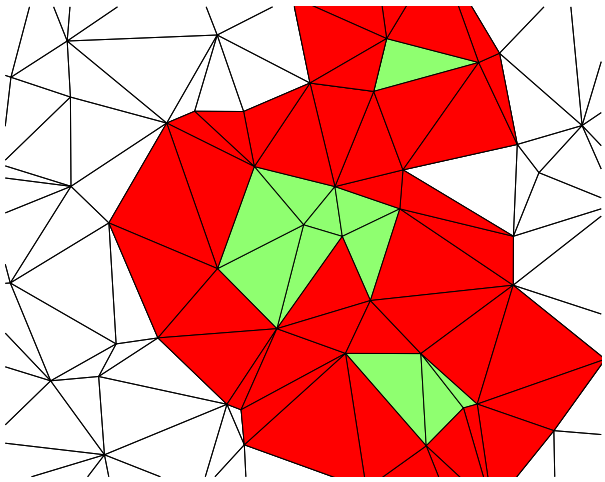
(A5) **Discrete reliability:** For all $\varepsilon > 0$, there exist constants $C_{\rm drel}(\varepsilon), C_{\rm ref}(\varepsilon) > 0$ such that for all refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ there exists a subset $\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}}) \subseteq \mathcal{T}$ with

$$\|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2 \le \varepsilon\eta_{\mathcal{T}} + C_{\rm drel}(\varepsilon)^2 \sum_{T\in\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2.$$

as well as

$$\mathcal{T}\setminus\widehat{\mathcal{T}} \subseteq \mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}}) \quad \text{and} \quad |\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})| \le C_{\rm ref}(\varepsilon)|\mathcal{T}\setminus\widehat{\mathcal{T}}|. \tag{20}$$

Remark. *Basically,* (20) *states that $\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})$ is* close *to the set of refined elements $\mathcal{T}\setminus\widehat{\mathcal{T}}$. In practical examples $\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})$ is $\mathcal{T}\setminus\widehat{\mathcal{T}}$ plus possibly a certain fixed number of element layers. In the figure below, we show a typical example. The set of refined elements is marked in green, whereas the set $\mathcal{R}(\varepsilon,\mathcal{T},\widehat{\mathcal{T}})$ is the union of the elements in red and green. The number of extra layers may depend on ε.*



In this example, $|\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})|$ *seems to be much larger than* $|\mathcal{T}\setminus\widehat{\mathcal{T}}|$*. However,* γ*-shape regularity guarantees that the element layers have essentially the same number of elements as* $|\mathcal{T}\setminus\widehat{\mathcal{T}}|$*.*

The formulation with $\varepsilon = 0$ *goes back to* [Ste07] *whereas the generalization* (A5) *is motivated from* [BDK12] *and firstly introduced in* [CFPP14]. □

We start with a simple observation

Lemma 10. *Discrete reliability* (A5) *implies reliability* (A4) *with* $C_{\mathrm{rel}} = \inf_{\varepsilon>0}(\varepsilon + C_{\mathrm{drel}}(\varepsilon)^2)^{1/2}$.

Proof. Assumption 2 states that for all $\mathcal{T}\in\mathbb{T}$ and for all $\delta>0$ exists a refinement $\widehat{\mathcal{T}}\in\mathbb{T}$ of $\mathcal{T}$ such that

$$\|u-U_{\widehat{\mathcal{T}}}\|_{\mathcal{X}} \le \delta.$$

For given $\mathcal{T}\in\mathbb{T}$ and $\delta>0$ choose $\widehat{\mathcal{T}}\in\mathbb{T}$ as above. Then, by use of discrete reliability (A5)

$$\begin{aligned}\|u-U_{\mathcal{T}}\|_{\mathcal{X}} &\le \|u-U_{\widehat{\mathcal{T}}}\|_{\mathcal{X}} + \|U_{\widehat{\mathcal{T}}}-U_{\mathcal{T}}\|_{\mathcal{X}}\\ &\le \delta + \Big(\varepsilon\eta_{\mathcal{T}}^2 + C_{\mathrm{drel}}(\varepsilon)^2 \sum_{T\in\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2\Big)^{1/2}\\ &\le \delta + \Big(\varepsilon\eta_{\mathcal{T}}^2 + C_{\mathrm{drel}}(\varepsilon)^2 \sum_{T\in\mathcal{T}} \eta_{\mathcal{T}}(T)^2\Big)^{1/2}.\end{aligned}$$

Since $\delta,\varepsilon>0$ are arbitrary, we get

$$\|u-U_{\mathcal{T}}\|_{\mathcal{X}} \le \inf_{\varepsilon>0}(\varepsilon + C_{\mathrm{drel}}(\varepsilon)^2)^{1/2}\eta_{\mathcal{T}}.$$

This concludes the proof. □

The next proposition can be interpreted as follows: So far, we have seen that Dörfler marking in the adaptive algorithm implies R-linear convergence (14) of $\eta(\cdot)$. The next proposition shows that if you already know that there holds R-linear convergence, Dörfler marking holds after finitely many steps, independently of how the meshes were refined. In other words, Dörfler marking is not only sufficient for R-linear convergence, it is in some sense even necessary.

Proposition 11 (Optimality of Dörfler marking)**.** *Let* $\eta(\cdot)$ *satisfy stability* (A1) *and discrete reliability* (A5)*. Define*

$$\theta_\star := \sup_{\varepsilon>0} \frac{1 - C_{\rm stab}^2\varepsilon}{1 + C_{\rm stab}^2 C_{\rm drel}(\varepsilon)^2}. \tag{21}$$

Then, there holds $0 < \theta_\star \le 1$*. Moreover, for all* $0 < \theta_0 < \theta_\star$ *there exists* $0 < q_0 < 1$ *and* $\varepsilon_0 > 0$ *such that for all refinements* $\widehat{\mathcal{T}} \in \mathbb{T}$ *of* $\mathcal{T} \in \mathbb{T}$ *the following implication is true*

$$\eta_{\widehat{\mathcal{T}}}^2 \le q_0 \eta_{\mathcal{T}}^2 \quad\Longrightarrow\quad \theta\eta_{\mathcal{T}}^2 \le \sum_{\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 \quad \text{for all } 0 < \theta \le \theta_0. \tag{22}$$

Proof. By definition, there holds $\theta_\star < 1$ and since $1 - C_{\rm stab}^2\varepsilon > 0$ for sufficiently small $\varepsilon > 0$, there also holds $\theta_\star > 0$.

To prove (22), we work with the free parameters q_0, ε_0 which are fixed later. With $\delta > 0$ and stability (A1), the Young inequality (7) gives

$$\begin{aligned}
\eta_{\mathcal{T}}^2 &= \sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \eta_{\mathcal{T}}(T)^2 + \sum_{T\in\mathcal{T}\cap\widehat{\mathcal{T}}} \eta_{\mathcal{T}}(T)^2 \\
&\le \sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}} \eta_{\mathcal{T}}(T)^2 + (1+\delta) \sum_{T\in\mathcal{T}\cap\widehat{\mathcal{T}}} \eta_{\widehat{\mathcal{T}}}(T)^2 + (1+\delta^{-1}) C_{\rm stab}^2 \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.
\end{aligned}$$

We use $\mathcal{T} \setminus \widehat{\mathcal{T}} \subseteq \mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})$ to get

$$\eta_{\mathcal{T}}^2 \le \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 + (1+\delta) \sum_{T\in\mathcal{T}\cap\widehat{\mathcal{T}}} \eta_{\widehat{\mathcal{T}}}(T)^2 + (1+\delta^{-1}) C_{\rm stab}^2 \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.$$

With $\eta_{\widehat{\mathcal{T}}}^2 \le q_0 \eta_{\mathcal{T}}^2$, there holds

$$\begin{aligned}
\eta_{\mathcal{T}}^2 &\le \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 + (1+\delta)\eta_{\widehat{\mathcal{T}}}^2 + (1+\delta^{-1}) C_{\rm stab}^2 \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2 \\
&\le \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 + (1+\delta) q_0 \eta_{\mathcal{T}}^2 + (1+\delta^{-1}) C_{\rm stab}^2 \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.
\end{aligned}$$

Finally, by use of discrete reliability (A5), we end up with

$$\eta_{\mathcal{T}}^2 \le \big(1 + (1+\delta^{-1}) C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2\big) \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 + (1+\delta) q_0 \eta_{\mathcal{T}}^2 + (1+\delta^{-1}) C_{\rm stab}^2 \varepsilon_0 \eta_{\mathcal{T}}^2,$$

which can be rearranged to

$$\frac{1 - (1+\delta^{-1}) C_{\rm stab}^2 \varepsilon_0 - (1+\delta) q_0}{1 + (1+\delta^{-1}) C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2} \eta_{\mathcal{T}}^2 \le \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2. \tag{23}$$

For arbitrary $0 < \theta_0 < \theta_\star$, choose $\varepsilon_0 > 0$ such that

$$\theta_0 < \frac{1 - C_{\rm stab}^2\varepsilon_0}{1 + C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2} \le \sup_{\varepsilon>0} \frac{1 - C_{\rm stab}^2\varepsilon}{1 + C_{\rm stab}^2 C_{\rm drel}(\varepsilon)^2} = \theta_\star.$$

Next, choose $\delta > 0$ sufficiently large such that

$$\theta_0 < \frac{1 - (1+\delta^{-1}) C_{\rm stab}^2 \varepsilon_0}{1 + (1+\delta^{-1}) C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2} < \frac{1 - C_{\rm stab}^2\varepsilon_0}{1 + C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2}.$$

Finally, choose $q_0 > 0$ such that

$$\theta_0 = \frac{1 - (1+\delta) q_0 - (1+\delta^{-1}) C_{\rm stab}^2 \varepsilon_0}{1 + (1+\delta^{-1}) C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2} < \frac{1 - (1+\delta^{-1}) C_{\rm stab}^2 \varepsilon_0}{1 + (1+\delta^{-1}) C_{\rm stab}^2 C_{\rm drel}(\varepsilon_0)^2}.$$

With this, (23) becomes

$$\theta_0 \eta_{\mathcal{T}}^2 \leq \sum_{T \in \mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2,$$

and the claim follows. □

The next theorem is the main result of this section.

Theorem 12 (Optimal convergence of adaptive algorithm)**.** *Let $\eta(\cdot)$ satisfy stability* (A1)*, reduction* (A2)*, general quasi-orthogonality* (A3)*, reliability* (A4)*, and discrete reliability* (A5)*. Let $0 < \theta < \theta_\star$, where $\theta_\star$ is defined in* (21)*. Then, it holds for all $s > 0$:*

$$u \in \mathbb{B}_s \quad \Longleftrightarrow \quad \eta_\ell \leq C_{\mathrm{opt}}(|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s} \quad \text{for all } \ell \in \mathbb{N}, \tag{24}$$

i.e. the adaptive algorithm is optimal in the sense of Definition 8.

Recall the previous steps of proof:

- The axioms (A1)–(A4) prove R-linear convergence (14).
- The axioms (A1) and (A5) prove the optimality of the Dörfler marking in Proposition 11.

In the following, we will also employ the assumptions on the mesh-refinement to establish a connection between convergence speed and number of refined elements.

First, we prove a quasi-monotonicity property of the error estimator. Recall that the Céa lemma and the nestedness $\mathcal{X}_{\mathcal{T}} \subseteq \mathcal{X}_{\widehat{\mathcal{T}}}$ for $\widehat{\mathcal{T}} \in \mathbb{T}$ a refinement of $\mathcal{T} \in \mathbb{T}$ imply $\|u - U_{\widehat{\mathcal{T}}}\|_{\mathcal{X}} \leq C_{\mathrm{cea}} \|u - U_{\mathcal{T}}\|_{\mathcal{X}}$, which is the analogue quasi-monotonicity of the error.

Lemma 13. *Let $\eta(\cdot)$ satisfy stability* (A1)*, reduction* (A2)*, and discrete reliability* (A5)*. Then, $\eta(\cdot)$ is quasi-monotone, i.e. it exists $C_{\mathrm{mon}} > 0$ such that for all refinements $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T} \in \mathbb{T}$ there holds*

$$\eta_{\widehat{\mathcal{T}}}^2 \leq C_{\mathrm{mon}} \eta_{\mathcal{T}}^2. \tag{25}$$

Proof. Similarly to the proof of Lemma 4, we split the estimator into two parts.

$$\eta_{\widehat{\mathcal{T}}}^2 = \sum_{T \in \widehat{\mathcal{T}} \setminus \mathcal{T}} \eta_{\widehat{\mathcal{T}}}(T)^2 + \sum_{T \in \widehat{\mathcal{T}} \cap \mathcal{T}} \eta_{\widehat{\mathcal{T}}}(T)^2. \tag{26}$$

For the first term on the right-hand side, we employ reduction (A2) to see

$$\sum_{T \in \widehat{\mathcal{T}} \setminus \mathcal{T}} \eta_{\widehat{\mathcal{T}}}(T)^2 \leq q_{\mathrm{red}} \sum_{T \in \mathcal{T} \setminus \widehat{\mathcal{T}}} \eta_{\mathcal{T}}(T)^2 + C_{\mathrm{red}} \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.$$

For the second term on the right-hand side of (26), we use stability (A1) and Young's inequality (7) for $\delta > 0$

$$\sum_{T \in \widehat{\mathcal{T}} \cap \mathcal{T}} \eta_{\widehat{\mathcal{T}}}(T)^2 \leq (1+\delta) \sum_{T \in \widehat{\mathcal{T}} \cap \mathcal{T}} \eta_{\mathcal{T}}(T)^2 + (1+\delta^{-1}) C_{\mathrm{stab}}^2 \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.$$

Plugging everything together, we see

$$\eta_{\widehat{\mathcal{T}}}^2 \leq (q_{\mathrm{red}} + (1+\delta)) \eta_{\mathcal{T}}^2 + (C_{\mathrm{red}} + (1+\delta^{-1}) C_{\mathrm{stab}}^2) \|U_{\widehat{\mathcal{T}}} - U_{\mathcal{T}}\|_{\mathcal{X}}^2.$$

Finally, we choose some $\varepsilon > 0$ and apply discrete reliability (A5) to estimate the last term

$$\begin{aligned}\eta_{\widehat{\mathcal{T}}}^2 &\le (q_{\mathrm{red}} + (1+\delta) + (C_{\mathrm{red}} + (1+\delta^{-1})C_{\mathrm{stab}}^2)\varepsilon)\eta_{\mathcal{T}}^2 \\ &\qquad + (C_{\mathrm{red}} + (1+\delta^{-1})C_{\mathrm{stab}}^2)C_{\mathrm{drel}}(\varepsilon)^2 \sum_{T\in\mathcal{R}(\varepsilon;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2 \\ &\le \Big(q_{\mathrm{red}} + (1+\delta) + (C_{\mathrm{red}} + (1+\delta^{-1})C_{\mathrm{stab}}^2)(C_{\mathrm{drel}}(\varepsilon)^2 + \varepsilon)\Big)\eta_{\mathcal{T}}^2.\end{aligned}$$

With

$$C_{\mathrm{mon}} := \inf_{\varepsilon>0,\,\delta>0} \Big(q_{\mathrm{red}} + (1+\delta) + (C_{\mathrm{red}} + (1+\delta^{-1})C_{\mathrm{stab}}^2)(C_{\mathrm{drel}}(\varepsilon)^2 + \varepsilon)\Big),$$

we prove (25). □

Lemma 14. *Let $\eta(\cdot)$ satisfy stability* (A1)*, reduction* (A2)*, and discrete reliability* (A5)*. Then, for $s > 0$ and $u \in \mathbb{B}_s$, there holds the following: For all $0 < \theta < \theta_\star$, there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all meshes $\mathcal{T} \in \mathbb{T}$ there exists a refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}$ with*

$$|\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})| \le C\|u\|_{\mathbb{B}_s}^{1/s}\eta_{\mathcal{T}}^{-1/s} \tag{27a}$$

as well as

$$\theta\eta_{\mathcal{T}}^2 \le \sum_{T\in\mathcal{R}(\varepsilon_0;\mathcal{T},\widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2. \tag{27b}$$

Proof. Without loss of generality, we may assume $\|u\|_{\mathbb{B}_s} > 0$. Otherwise, $\eta_{\mathcal{T}_0} \le \|u\|_{\mathbb{B}_s} = 0$ and Lemma 13 implies $\eta_{\mathcal{T}} = 0$ for all $\mathcal{T} \in \mathbb{T}$ and (27a)–(27b) hold with $\widehat{\mathcal{T}} = \mathcal{T}$. Let $\lambda := C_{\mathrm{mon}}^{-1}q_0$ with $0 < q_0 < 1$ defined in Proposition 11 and set $\varepsilon^2 := \lambda\eta_{\mathcal{T}}^2$. The quasi-monotonicity from Lemma 13 shows

$$\varepsilon^2 \le q_0\eta_{\mathcal{T}_0}^2 \le \|u\|_{\mathbb{B}_s}^2.$$

Step 1: The first statement we want to prove, is the following: For $\varepsilon^2 := \lambda\eta_{\mathcal{T}}^2$, there exists $\mathcal{T}_\varepsilon \in \mathbb{T}$ with

$$\eta_{\mathcal{T}_\varepsilon} \le \varepsilon \tag{28a}$$

$$\mathcal{T}_\varepsilon| - |\mathcal{T}_0| \le 2\|u\|_{\mathbb{B}_s}^{-1/s}\varepsilon^{-1/s}. \tag{28b}$$

Basically, this says that if we want the error estimator to be smaller than ε, we have to pay with $\varepsilon^{-1/s}$ elements if we choose the optimal mesh $\mathcal{T}_\varepsilon$. To see this statement, let $N \in \mathbb{N}$ be minimal with $N^{-s}\|u\|_{\mathbb{B}_s} \le \varepsilon$. For $N = 1$, we have $\|u\|_{\mathbb{B}_s} = \varepsilon$ and hence

$$N = 1 = \|u\|_{\mathbb{B}_s}^{1/s}\varepsilon^{-1/s}.$$

For $N > 1$, minimality of N yields $(N-1)^{-s}\|u\|_{\mathbb{B}_s} > \varepsilon$ and hence

$$N \le 2(N-1) < 2\|u\|_{\mathbb{B}_s}^{1/s}\varepsilon^{-1/s}.$$

Here, we used $N \le 2(N-1)$ which holds for all $N > 1$. Next, we choose $\mathcal{T}_\varepsilon \in \mathbb{T}(N)$ such that

$$\eta_{\mathcal{T}_\varepsilon} = \min_{\mathcal{T}\in\mathbb{T}(N)} \eta_{\mathcal{T}}.$$

By definition of the approximation class (16) and the choice of $N \in \mathbb{N}$

$$\eta_{\mathcal{T}_\varepsilon} \le N^{-s}\|u\|_{\mathbb{B}_s} \le \varepsilon.$$

Step 2: We consider the overlay $\widehat{\mathcal{T}} := \mathcal{T} \oplus \mathcal{T}_\varepsilon \in \mathbb{T}$ (cf. (2)). Since $\widehat{\mathcal{T}}$ is a refinement of $\mathcal{T}_\varepsilon$, we may employ quasi-monotonicity (25) of $\eta(\cdot)$ to see

$$\eta_{\widehat{\mathcal{T}}}^2 \le C_{\mathrm{mon}} \eta_{\mathcal{T}_\varepsilon}^2 \le C_{\mathrm{mon}} \varepsilon^2 = q_0 \eta_{\mathcal{T}}^2, \tag{29}$$

by definition of $\varepsilon > 0$.

Step 3: Finally, we employ the assumptions on the refinement strategy. First, we use the overlay estimate (2) as well as the result (28) from Step 1 and obtain

$$|\widehat{\mathcal{T}}| - |\mathcal{T}| \le (|\mathcal{T}_\varepsilon| + |\mathcal{T}| - |\mathcal{T}_0|) - |\mathcal{T}| = |\mathcal{T}_\varepsilon| - |\mathcal{T}_0| \le 2\|u\|_{\mathbb{B}_s}^{1/s} \varepsilon^{-1/s}.$$

Second, the discrete reliability (A5) and (1a) show for all $\varepsilon_0 > 0$

$$|\mathcal{R}(\varepsilon_0; \mathcal{T}, \widehat{\mathcal{T}})| \le C_{\mathrm{ref}} |\mathcal{T} \setminus \widehat{\mathcal{T}}| \le C_{\mathrm{ref}} (|\widehat{\mathcal{T}}| - |\mathcal{T}|).$$

Altogether, we have

$$|\mathcal{R}(\varepsilon_0; \mathcal{T}, \widehat{\mathcal{T}})| \le 2C_{\mathrm{ref}} \|u\|_{\mathbb{B}_s}^{1/s} \varepsilon^{-1/s} \le 2C_{\mathrm{ref}} \|u\|_{\mathbb{B}_s}^{1/s} \lambda^{-1/(2s)} \eta_{\mathcal{T}}^{-1/s},$$

which proves (27a) with $C = 2C_{\mathrm{ref}} \|u\|_{\mathbb{B}_s}^{1/s} \lambda^{-1/(2s)}$. The estimate (29) allows to apply (22). This yields

$$\theta \eta_{\mathcal{T}}^2 \le \sum_{T \in \mathcal{R}(\varepsilon_0; \mathcal{T}, \widehat{\mathcal{T}})} \eta_{\mathcal{T}}(T)^2$$

and concludes the proof. □

With all preparations made, the proof of our main theorem is peanuts.

Proof of Theorem 12. Let $u \in \mathbb{B}_s$ for some $s > 0$ and let $\ell \in \mathbb{N}$. We apply Lemma 14 to choose a refinement $\widehat{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}_\ell$ which satisfies (27) for $\mathcal{T} = \mathcal{T}_\ell$. According to (27b), the set $\mathcal{R}(\varepsilon_0; \mathcal{T}_\ell, \widehat{\mathcal{T}})$ satisfies the Dörfler marking. Since we chose $\mathcal{M}_\ell$ in the adaptive algorithm to be a set of minimal cardinality to satisfy the Dörfler marking, this together with (27a) yields

$$|\mathcal{M}_\ell| \le |\mathcal{R}(\varepsilon_0; \mathcal{T}_\ell, \widehat{\mathcal{T}})| \le C \|u\|_{\mathbb{B}_s}^{1/s} \eta_{\mathcal{T}_\ell}^{-1/s}.$$

With the closure estimate (3), we conclude

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \le C_{\mathrm{mesh}} \sum_{j=0}^{\ell-1} |\mathcal{M}_j| \le C \|u\|_{\mathbb{B}_s}^{1/s} \sum_{j=0}^{\ell-1} \eta_{\mathcal{T}_j}^{-1/s}.$$

R-linear convergence (14) and its equivalent formulation (12) thus show

$$|\mathcal{T}_\ell| - |\mathcal{T}_0| \le C_2 C \|u\|_{\mathbb{B}_s}^{1/s} \eta_\ell^{-1/s}.$$

Put differently, we have

$$\eta_\ell \le C_{\mathrm{opt}} (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s},$$

where $C_{\mathrm{opt}} := C_2^s C^s \|u\|_{\mathbb{B}_s}$. This proves the $\Longrightarrow$-implication of (24). The $\Longleftarrow$-implication has already been proved in Lemma 9. □

2.6. **Characterization of the approximation class.** So far, we have characterized optimal convergence of the estimator with the approximation class $\mathbb{B}_s$ from (16). On the one hand, this is very natural, since the adaptive algorithm has no other information than the error estimator to steer the mesh-refinement. On the other hand, however, we are interested in optimal convergence rates of the error $\|u - U_\ell\|_{\mathcal{X}}$ instead of the error estimator. To that end, define the approximation class $\mathbb{A}_s$ for all $s > 0$

$$u \in \mathbb{A}_s \quad \Longleftrightarrow \quad \|u\|_{\mathbb{A}_s} := \|u - U_{\mathcal{T}_0}\|_{\mathcal{X}} + \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \Big(N^s \|u - U_{\mathcal{T}}\|_{\mathcal{X}} \Big) < \infty. \tag{30}$$

Moreover, one could even consider the approximability of the unknown solution and define the approximation class $\widetilde{\mathbb{A}}_s$ for all $s > 0$

$$u \in \widetilde{\mathbb{A}}_s \quad \Longleftrightarrow \quad \|u\|_{\widetilde{\mathbb{A}}_s} := \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \min_{V \in \mathcal{X}_{\mathcal{T}}} \Big(N^s \|u - V\|_{\mathcal{X}} \Big) < \infty. \tag{31}$$

We start with some observations which are formulated in the next lemma.

Lemma 15. *There hold (i)–(iii):*

(i) *For all $s > 0$, we have $\mathbb{A}_s \subseteq \widetilde{\mathbb{A}}_s$.*

(ii) *Under reliability* (A4)*, it holds $\mathbb{B}_s \subseteq \mathbb{A}_s$ for all $s > 0$.*

(iii) *Under the Céa lemma*

$$\|u - U_{\mathcal{T}}\|_{\mathcal{X}} \le C_{\mathrm{cea}} \min_{V \in \mathcal{X}_{\mathcal{T}}} \|u - V\|_{\mathcal{X}} \tag{32}$$

with $C_{\mathrm{cea}} > 0$ independent of $\mathcal{T} \in \mathbb{T}$, it holds $\widetilde{\mathbb{A}}_s = \mathbb{A}_s$ for all $s > 0$.

Proof. To prove (i), we suppose $u \in \mathbb{A}_s$. With

$$\min_{V \in \mathcal{X}_{\mathcal{T}}} \|u - V\|_{\mathcal{X}} \le \|u - U_{\mathcal{T}}\|_{\mathcal{X}} \quad \text{for all } \mathcal{T} \in \mathbb{T}$$

we get $\|u\|_{\widetilde{\mathbb{A}}_s} \le \|u\|_{\mathbb{A}_s} < \infty$ and hence $u \in \widetilde{\mathbb{A}}_s$. To prove (iii), we use the Céa lemma (32) to show $\|u\|_{\widetilde{\mathbb{A}}_s} \ge C_{\mathrm{cea}}^{-1} \|u\|_{\mathbb{A}_s}$ and hence $\widetilde{\mathbb{A}}_s \subseteq \mathbb{A}_s$. Finally, to see (ii), we use reliability (A4)

$$\|u\|_{\mathbb{A}_s} = \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \Big(N^s \|u - U_{\mathcal{T}}\|_{\mathcal{X}} \Big) \le C_{\mathrm{rel}} \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \Big(N^s \eta_{\mathcal{T}} \Big) = C_{\mathrm{rel}} \|u\|_{\mathbb{B}_s}.$$

This shows $\mathbb{B}_s \subseteq \mathbb{A}_s$ and concludes the proof. □

The interesting question now is: What do we need to show $\mathbb{A}_s = \mathbb{B}_s$ or even $\widetilde{\mathbb{A}}_s = \mathbb{B}_s$ for all $s > 0$. In other words, under which assumptions will the adaptive algorithm lead to optimal convergence rates for the error. The answer is the final axiom:

(A6) **Efficiency:** There exists a constant $C_{\mathrm{eff}} > 0$ and for all $\mathcal{T} \in \mathbb{T}$ exists a mapping $\mathrm{osc}_{\mathcal{T}}(\cdot) : \mathcal{X}(\mathcal{T}) \to \mathbb{R}_{\ge 0}$ such that

$$C_{\mathrm{eff}}^{-2} \eta_{\mathcal{T}}^2 \le \|u - U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \mathrm{osc}_{\mathcal{T}}(U_{\mathcal{T}})^2.$$

To abbreviate the notation, we write $\mathrm{osc}_{\mathcal{T}} = \mathrm{osc}_{\mathcal{T}}(U_{\mathcal{T}})$.

Remark. *For the Poisson model problem*

$$-\Delta u = f \quad \text{in } \Omega,$$

we shall see that efficiency (A6) *follows from inverse estimates with*

$$\mathrm{osc}_{\mathcal{T}} = \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)},$$

where $f_{\mathcal{T}} \in \mathcal{P}^q(\mathcal{T})$ is a $\mathcal{T}$-elementwise polynomial best-approximation of f with arbitrary but fixed polynomial degree $q \ge 0$. In this case, $\mathrm{osc}_{\mathcal{T}}$ depends only on the smoothness of the data and measures how good the data f is resolved on a given mesh $\mathcal{T} \in \mathbb{T}$. A special case is when f itself is $\mathcal{T}$-elementwise polynomial. Then, we have $\mathrm{osc}_{\mathcal{T}} = 0$. □

Lemma 16. *Suppose that $\eta(\cdot)$ satisfies efficiency* (A6). *Then, there holds (i)–(ii):*

(i) *Under reliability* (A4) *and* $\operatorname{osc}_{\mathcal{T}} \le C_{\rm osc}\eta_{\mathcal{T}}$ *for all* $\mathcal{T} \in \mathbb{T}$, *it holds*

$$u \in \mathbb{B}_s \quad \Longleftrightarrow \quad \|u\|_{\overline{\mathbb{B}}_s} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} N^s \Big(\|u-U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}^2\Big)^{1/2} < \infty.$$

(ii) *Suppose that additionally the Céa lemma* (32) *and the stability*

$$C_{\rm osc}^{-2}\operatorname{osc}_{\mathcal{T}}(V)^2 \le \operatorname{osc}_{\mathcal{T}}(W)^2 + \|V-W\|_{\mathcal{X}}^2 \quad \text{for all } V,W \in \mathcal{X}_{\mathcal{T}} \text{ and all } \mathcal{T}\in\mathbb{T} \tag{33}$$

hold. Then, it follows

$$u \in \mathbb{B}_s \quad \Longleftrightarrow \quad \|u\|_{\overline{\overline{\mathbb{B}}}_s} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} \min_{V\in\mathcal{X}_{\mathcal{T}}} N^s \Big(\|u-V\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}(V)^2\Big)^{1/2} < \infty.$$

Proof. We start with (i): All we have to show is that $\|u\|_{\mathbb{B}_s} < \infty \Longleftrightarrow \|u\|_{\overline{B}_s} < \infty$ which is particularly true, if $C^{-1}\|u\|_{\mathbb{B}_s} \le \|u\|_{\overline{\mathbb{B}}_s} \le C\|u\|_{\mathbb{B}_s}$ for some constant $C>0$. By assumption we have $\operatorname{osc}_{\mathcal{T}} \le C_{\rm osc}\eta_{\mathcal{T}}$ and together with reliability (A4), this yields

$$\|u\|_{\overline{\mathbb{B}}_s} \le (C_{\rm osc}^2 + C_{\rm rel}^2)^{1/2}\|u\|_{\mathbb{B}_s}.$$

Efficiency shows

$$\|u\|_{\mathbb{B}_s} \le C_{\rm eff}\|u\|_{\overline{\mathbb{B}}_s}$$

and concludes the proof of (i). To prove (ii), we first observe

$$\|u\|_{\overline{\overline{\mathbb{B}}}_s} \le \|u\|_{\overline{\mathbb{B}}_s} \le (C_{\rm osc}^2 + C_{\rm rel}^2)^{1/2}\|u\|_{\mathbb{B}_s}.$$

To see the converse estimate, let $V \in \mathcal{X}_{\mathcal{T}}$. With (33)

$$\operatorname{osc}_{\mathcal{T}}^2 \le C_{\rm osc}^2\operatorname{osc}_{\mathcal{T}}(V)^2 + C_{\rm osc}^2\|U_{\mathcal{T}}-V\|_{\mathcal{X}}^2.$$

For the last term, we apply the Young inequality (7) with $\delta = 1$ to see

$$\operatorname{osc}_{\mathcal{T}}^2 \le C_{\rm osc}^2\operatorname{osc}_{\mathcal{T}}(V)^2 + 2C_{\rm osc}^2(\|u-U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \|u-V\|_{\mathcal{X}}^2).$$

This together with the Céa lemma (32) shows

$$\begin{aligned}\|u-U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}^2 &\le C_{\rm osc}^2\operatorname{osc}_{\mathcal{T}}(V)^2 + 3C_{\rm osc}^2\|u-U_{\mathcal{T}}\|_{\mathcal{X}}^2 + 2C_{\rm osc}^2\|u-V\|_{\mathcal{X}}^2\\ &\le C_{\rm osc}^2\operatorname{osc}_{\mathcal{T}}(V)^2 + (3C_{\rm osc}^2C_{\rm cea}^2 + 2C_{\rm osc}^2)\|u-V\|_{\mathcal{X}}^2.\end{aligned}$$

Since $V \in \mathcal{X}_{\mathcal{T}}$ is arbitrary, we get with $C = 3C_{\rm osc}^2C_{\rm cea}^2 + 2C_{\rm osc}^2$

$$\eta_{\mathcal{T}}^2 \le C_{\rm eff}^2(\|u-U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}^2) \le C_{\rm eff}C \inf_{V\in\mathcal{X}_{\mathcal{T}}} \Big(\|u-V\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}(V)^2\Big)$$

and hence $\|u\|_{\mathbb{B}_s} \le C_{\rm eff}C\|u\|_{\overline{\overline{\mathbb{B}}}_s}$. □

Remark. *The approximation class $\overline{\overline{\mathbb{B}}}_s$ is usually found in the literature, e.g. Stevenson (2007), Cascon-Kreuzer-Nochetto-Siebert, and additionally $\theta_\star$ usually hinges also on $C_{\rm eff}$. Our approach considers the more general approximation class $\mathbb{B}_s$, and $\theta_\star$ is independent of $C_{\rm eff}$.* □

Next, we want to characterize the approximation class $\mathbb{B}_s$ in terms of the Galerkin error only. To that end, we need to quantify the quality of the oscillation term $\operatorname{osc}(\cdot)$. We define

$$u \in \mathbb{O}_s \quad \overset{\rm def.}{\Longleftrightarrow} \quad \|u\|_{\mathbb{O}_s} := \operatorname{osc}_{\mathcal{T}_0} + \sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} (N^s\operatorname{osc}_{\mathcal{T}}) < \infty. \tag{34}$$

Proposition 17. *Assume reliability* (A4)*, efficiency* (A6) *and quasi-monotonicity of oscillations and error in the sense that there exists a constant* $C_{\mathrm{mon}} > 0$ *such that for all refinements* $\widehat{\mathcal{T}} \in \mathbb{T}$ *of all* $\mathcal{T} \in \mathbb{T}$ *holds*

$$\operatorname{osc}_{\widehat{\mathcal{T}}} \le C_{\mathrm{mon}} \operatorname{osc}_{\mathcal{T}}, \tag{35a}$$

$$\|u - U_{\widehat{\mathcal{T}}}\|_{\mathcal{X}} \le C_{\mathrm{mon}} \|u - U_{\mathcal{T}}\|_{\mathcal{X}}. \tag{35b}$$

Then, there holds (i)–(ii)*.*

(i) $u \in \mathbb{A}_s$ *and* $u \in \mathbb{O}_s \implies u \in \mathbb{B}_s$

(ii) $u \in \mathbb{B}_s$ *and* $\operatorname{osc}_{\mathcal{T}} \le C_{\mathrm{osc}} \eta_{\mathcal{T}}$ *for all* $\mathcal{T} \in \mathbb{T} \implies u \in \mathbb{O}_s$ *and* $u \in \mathbb{A}_s$

Proof. (ii): The statement follows from the definition of the approximation classes $\mathbb{A}_s, \mathbb{B}_s, \mathbb{O}_s$ and efficiency (A6).

(i): Let $N \in \mathbb{N}$ be even. The definition of $\mathbb{A}_s$ and $\mathbb{O}_s$ provides meshes $\mathcal{T}_{\mathrm{err}} \in \mathbb{T}(N/2)$ and $\mathcal{T}_{\mathrm{osc}} \in \mathbb{T}(N/2)$ with

$$(N/2)^s \|u - U_{\mathcal{T}_{\mathrm{err}}}\|_{\mathcal{X}} \le \|u\|_{\mathbb{A}_s},$$
$$(N/2)^s \operatorname{osc}_{\mathcal{T}_{\mathrm{osc}}} \le \|u\|_{\mathbb{O}_s}.$$

Now, we consider the overlay $\mathcal{T} := \mathcal{T}_{\mathrm{err}} \oplus \mathcal{T}_{\mathrm{osc}}$ and employ the overlay estimate (2) to see $|\mathcal{T}| \le |\mathcal{T}_{\mathrm{err}}| + |\mathcal{T}_{\mathrm{osc}}| - |\mathcal{T}_0|$ and hence $\mathcal{T} \in \mathbb{T}(N)$. With efficiency and monotonicity (35), this implies

$$N^{2s} \eta_{\mathcal{T}}^2 \le C_{\mathrm{eff}}^2 N^{2s} \big(\|u - U_{\mathcal{T}}\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}}^2 \big) \tag{36}$$

$$\le C_{\mathrm{eff}}^2 C_{\mathrm{mon}}^2 N^{2s} \big(\|u - U_{\mathcal{T}_{\mathrm{err}}}\|_{\mathcal{X}}^2 + \operatorname{osc}_{\mathcal{T}_{\mathrm{osc}}}^2 \big) \tag{37}$$

$$\le C_{\mathrm{eff}}^2 C_{\mathrm{mon}}^2 4^s (\|u\|_{\mathbb{A}_s}^2 + \|u\|_{\mathbb{O}_s}^2). \tag{38}$$

Altogether, we show

$$\begin{aligned}
\|u\|_{\mathbb{B}_s} &\le \eta_{\mathcal{T}_0} + \sup_{N \text{ even}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \eta_{\mathcal{T}} + \sup_{N \text{ uneven}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \eta_{\mathcal{T}} \\
&\le \eta_{\mathcal{T}_0} + \sup_{N \text{ even}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \eta_{\mathcal{T}} + \sup_{N > 1 \text{ uneven}} \left(\frac{N}{N-1} \right)^s \min_{\mathcal{T} \in \mathbb{T}(N)} (N-1)^s \eta_{\mathcal{T}} + \min_{\mathcal{T} \in \mathbb{T}(1)} \eta_{\mathcal{T}} \\
&\le \eta_{\mathcal{T}_0} + (1 + 2^s) \sup_{N \text{ even}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \eta_{\mathcal{T}} + \min_{\mathcal{T} \in \mathbb{T}(1)} \eta_{\mathcal{T}} < \infty
\end{aligned}$$

since the second term is bounded in (36) and the third term is a minimum over a finite set. □

To conclude the section, we provide an overview of the optimality proof in form of a roadmap in Figure 1.

3. Verification of the Axioms

The goal of this section is to verify the introduced axioms (A1)–(A6) for a certain model problem.

3.1. Model problem.

We consider a general second-order linear elliptic PDE in divergence form

$$\mathcal{L}u := -\operatorname{div}\big(\boldsymbol{A} \nabla u \big) + \boldsymbol{b} \cdot \nabla u + cu = f \quad \text{in } \Omega, \tag{39a}$$

$$u = 0 \quad \text{on } \Gamma := \partial\Omega. \tag{39b}$$

We pose the following regularity assumptions on the coefficients. $\boldsymbol{A}(x) \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$ is a symmetric matrix with $\boldsymbol{A} \in W^{1,\infty}(\Omega)$. The vector $\boldsymbol{b} \in \mathbb{R}^d$ satisfies $\boldsymbol{b} \in L^\infty(\Omega)$ and the

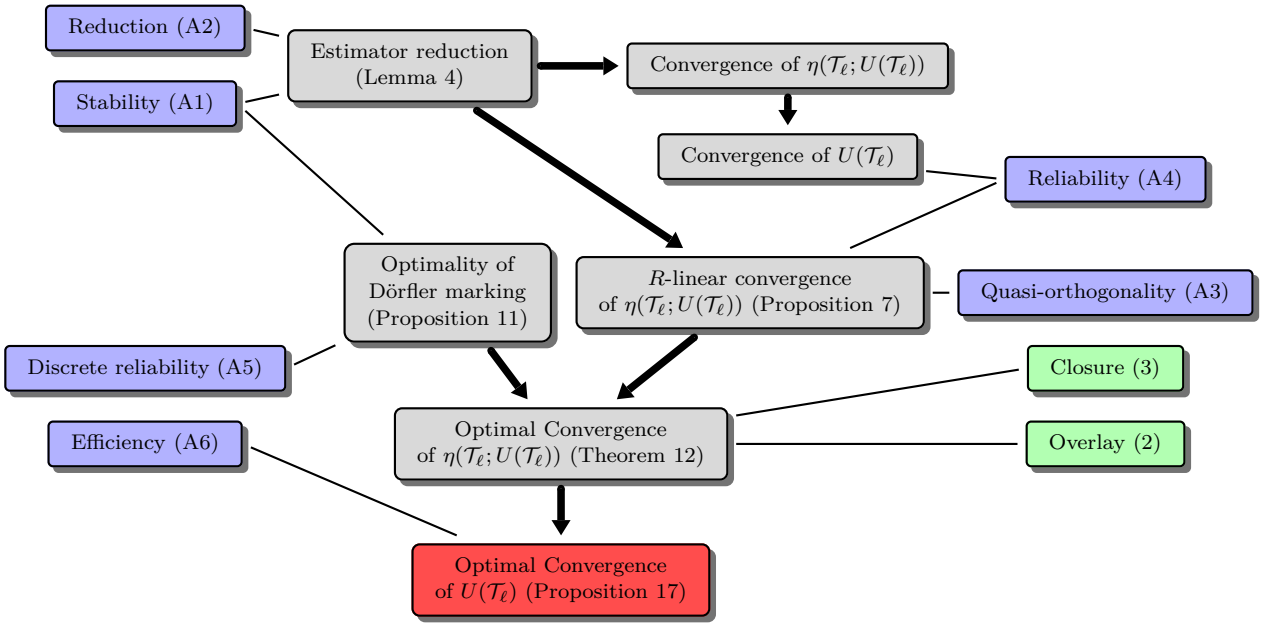


FIGURE 1. Map of the quasi-optimality proof.

scalar $c \in \mathbb{R}$ satisfies also $c \in L^\infty(\Omega)$. The operator $\mathcal{L}$ is interpreted in its weak form and defines a bilinear form

$$b(u,v) := \langle \mathcal{L}u\,,\, v\rangle = \int_\Omega \boldsymbol{A}(x)\nabla u(x)\cdot\nabla v(x) + \boldsymbol{b}(x)\cdot\nabla u(x)\,v(x) + c(x)u(x)\,v(x)\,dx.$$

The Cauchy-Schwartz inequality shows immediately the continuity of $b(\cdot,\cdot)$. There holds for all $v, w \in H^1(\Omega)$

$$b(v,w) \le \Big(\|A\|_{L^\infty(\Omega} + \|b\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\Big)\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}.$$

Additionally, we assume ellipticity of $b(\cdot,\cdot)$ on $H^1_0(\Omega)$ in the sense

$$b(v,v) \ge C_{\mathrm{ell}}\|v\|^2_{H^1(\Omega)} \quad \text{for all } v \in H^1_0(\Omega). \tag{40}$$

Remark. *The assumption* (40) *can be guaranteed e.g. if* $c \ge 0$ *and the smallest eigenvalue* $\lambda(x) \in \mathbb{R}$ *of* $\boldsymbol{A}(x)$ *is bounded from zero since then there holds with the Young inequality and the Friedrich's inequality*

$$\begin{aligned} b(v,v) &\ge \int_\Omega \lambda(x)|\nabla v(x)|^2 - |\boldsymbol{b}(x)||\nabla v(x)|\,v(x)\,dx \\ &\ge \int_\Omega (\lambda(x) - |\boldsymbol{b}(x)|^2/2)|\nabla v(x)|^2 - v(x)^2/2\,dx \\ &\ge \Big(\inf_{x\in\Omega}(\lambda(x) - |\boldsymbol{b}(x)|^2/2) - C_{\mathrm{F}}(\Omega)/2\Big)\|\nabla v\|^2_{L^2(\Omega)}. \end{aligned}$$

□

The problem fits into the framework of the Lax-Milgram lemma. We choose piecewise polynomial ansatz spaces $\mathcal{X}_\mathcal{T} := \mathcal{S}^p_0(\mathcal{T})$ for all $\mathcal{T} \in \mathbb{T}$ and define the weighted residual

error estimator for all $T \in \mathcal{T}$, all $\mathcal{T} \in \mathbb{T}$, and all $V \in \mathcal{S}^p_0(\mathcal{T})$

$$\eta_{\mathcal{T}}(T,V)^2 := |T|^{2/d}\|f + \operatorname{div}\boldsymbol{A}\nabla V - \boldsymbol{b}\cdot\nabla V - cV\|^2_{L^2(\Omega)} + |T|^{1/d}\|[\boldsymbol{A}\nabla V\cdot n]\|^2_{L^2(\partial T\cap\Omega)}. \tag{41}$$

The first term measures the so-called *volume residual*, whereas the second term measures the so-called *normal jumps*.

3.2. **Shape regularity.**

Definition 18. *A triangulation $\mathcal{T}$ is regular (or: conforming) if*

- *$\mathcal{T}$ is a finite set of compact simplices $T \subseteq \mathbb{R}^d$*
- *$\bigcup_{T\in\mathcal{T}} T = \overline{\Omega}$*
- *For all $T, T' \in \mathcal{T}$ with $T \neq T'$ holds*
  - *$T \cap T' = \emptyset$*
  - *$T \cap T'$ is some $(d-k)$ dimensional hyperface for $k = 1, \dots, d$*
    - *node for $d-k=0$*
    - *edge for $d-k=1$*
    - *face for $d-k=2$*
    - *etc.*

Definition 19. *A triangulation $\mathcal{T}$ is γ-shape regular (or: locally γ-quasi uniform) if $\mathcal{T}$ is regular in the sense of Definition 18 and*

$$\max_{T\in\mathcal{T}} \frac{h_T^d}{|T|} \leq \gamma \tag{42}$$

with $h_T := \operatorname{diam}(T) := \max_{x,y\in T}|x-y|$, and $|T| = \int_T 1\,dx$, where dx denotes the surface measure.

Example 20. *Consider the mesh $\mathcal{T} := \{T_\varepsilon\}$ with $T_\varepsilon \subseteq \mathbb{R}^2$ the triangle which is defined by the nodes $(0,0),(1,0),(0,\varepsilon) \in \mathbb{R}^2$. There holds*

$$\frac{h^2_{T_\varepsilon}}{|T_\varepsilon|} = \frac{1}{\varepsilon/2} \to \infty \text{ as } \varepsilon \to 0.$$

Remark. *Note that each triangulation $\mathcal{T}$ which consists of non-degenerate simplices is γ-shape regular, since the maximum in (42) is taken over a finite set. However, the notion of γ-shape regularity is of importance if one considers a sequence of meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$. Then, γ-shape regularity of $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ ensures that the simplices to not degenerate as $\ell \to \infty$.* □

Lemma 21. *Let $\mathcal{T}$ denote a γ-shape regular triangulation. For $T \in \mathcal{T}$, define the patch $\omega_T := \bigcup\{T' \in \mathcal{T} : T \cap T' \neq \emptyset\} \subset \mathbb{R}^d$. Then,*

$$h_T = \operatorname{diam}(T) \leq \operatorname{diam}(\omega_T) \leq C(\gamma) h_T$$

and the number of elements $T' \in \mathcal{T}$ with $T' \subseteq \omega_T$ is bounded by $C(\gamma) > 0$.

Proof. T.B.D. □

The result of the following lemma is that each simplex can be transformed to a reference simplex by an affine function.

Lemma 22. *Let $T := conv\{z_0, \dots, z_d\}$*
in$\mathcal{T}$ denote a d-dimensional simplex and let $\mathcal{T}$ denote a γ-shape regular triangulation.

Define the reference element $\widehat{T} := textconv\{0, e_1, \dots, e_d\}$*, where* $e_i \in \mathbb{R}^d$ *denotes the* i*-th unit vector. Define the function*

$$\Phi_T : \widehat{T} \to T, \quad \Phi_T(x) := z_0 + Bx \quad \text{with } B := (z_1 - z_0, \dots, z_d - z_0) \in \mathbb{R}^{d\times d}.$$

Then, there holds (i)–(iii)

(i) $|\mathrm{det}B| = \frac{|T|}{|\widehat{T}|} \simeq |T| \simeq h_T^d$,
(ii) $\|B\|_F \simeq h$,
(iii) $\|B^{-1}\|_F \simeq h^{-1}$,

where the hidden constants depend on γ *and the dimension* d *only.*

Proof. There holds $\partial_i \Phi_T(x) = z_i - z_0$. Hence, the Jacobian of Φ_T reads $D\Phi = B$. With this and integration by substitution, we obtain

$$|T| = \int_T 1\,dx = \int_{\widehat{T}} 1|\mathrm{det}D\Phi_T|\,dx = |\mathrm{det}B| \int_{\widehat{T}} 1\,dx = |\mathrm{det}B||\widehat{T}|,$$

where the volume $|\widehat{T}|$ depends only on the dimension d. Moreover, we know that $|z_i - z_0| \leq h_T$ for all $i = 1, \dots, d$. Therefore,

$$\|B\|_F^2 = \sum_{i=1}^{d} |z_i - z_0|^2 \leq dh_T^2.$$

The proof of (iii) is rather technical for $d > 2$ and therefore only presented for $d = 2$. In this case, we have

$$B^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\mathrm{det}B} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

and hence

$$\|B^{-1}\|_F = \frac{1}{\mathrm{det}B}\|B\|_F \simeq \frac{h_T}{|T|} \simeq h_T^{-1}.$$

□

3.3. **Scaling arguments.** This section consists of several technical, but very useful results. The following lemma is an implication of integration by substitution and Lemma 22.

Lemma 23 (Transformation formula)**.** *Let* $\widehat{T}, T \subseteq \mathbb{R}^d$ *denote Lipschitz domains and let* $\Phi(x) := Bx + y$ *define a mapping with regular* $B \in \mathbb{R}^{d\times d}$*,* $y \in \mathbb{R}^d$*, and* $\Phi(\widehat{T}) = T$*. Then, for* $u \in H^m(T)$*, it holds*

$$u \circ \Phi \in H^m(\widehat{T}) \quad \text{with} \quad \|D^m(u \circ \Phi)\|_{L^2(\widehat{T})} \leq |\mathrm{det}B|^{-1/2}\|B\|_F^m \|D^m u\|_{L^2(T)}.$$

Proposition 24 (Poincaré estimate)**.** *Let* $\mathcal{T}$ *denote a* γ*-shape regular triangulation. Given* $T \in \mathcal{T}$*, let* $u \in H^1(T)$ *with integral mean* $\overline{u} := |T|^{-1} \int_T u\,dx$*. Then, there exists a constant* $C_{pc} > 0$ *which does not depend on* T *and* u *such that*

$$\|u - \overline{u}\|_{L^2(T)} \leq C_{pc} h_T \|\nabla u\|_{L^2(T)}. \tag{43}$$

Proof. On the reference element, there holds according to the PDE lecture (Poincaré inequality) for all $\widehat{u} \in H^1(\widehat{T})$

$$\|\widehat{u} - \overline{\widehat{u}}\|_{L^2(\widehat{T})} \leq C\|\nabla \widehat{u}\|_{L^2(\widehat{T})}.$$

Now, we choose $\widehat{u} := u \circ \Phi_T$, where Φ_T is defined in Lemma 22. There holds

$$\overline{u} = |T|^{-1} \int_T u\,dx |T|^{-1} \int_{\widehat{T}} \widehat{u} |\mathrm{det}D\Phi_T|\,dx = |T|^{-1} \frac{|T|}{|\widehat{T}|} \int_{\widehat{T}} \widehat{u}\,dx = \overline{\widehat{u}}.$$

Let $w = u - \overline{u}$ and $\widehat{w} = w \circ \Phi_T$ as well as $w = \widehat{w} \circ \Phi_T^{-1}$. Note that $\Phi_T(x) = Bx + y$ and Φ_T^{-1} is also an affine function with linear part B^{-1}. Therefore, we may apply Lemma 23 with w and $\widehat{w}$, to obtain

$$\begin{aligned}\|u - \overline{u}\|_{L^2(T)} &= |\mathrm{det} B^{-1}|^{-1/2} \|\widehat{u} - \overline{\widehat{u}}\|_{L^2(\widehat{T})} \\ &\leq C |\mathrm{det} B^{-1}|^{-1/2} \|\nabla \widehat{u}\|_{L^2(\widehat{T})} \\ &\leq C |\mathrm{det} B^{-1}|^{-1/2} |\mathrm{det} B|^{-1/2} \|B\|_F \|\nabla u\|_{L^2(T)} \\ &= C \|B\|_F \|\nabla u\|_{L^2(T)}.\end{aligned}$$

With Lemma 22, we conclude the proof. □

Lemma 25. *Let $\mathcal{X}$ denote a finite dimensional vector space with seminorms $|\cdot|_1, |\cdot|_2 : \mathcal{X} \to \mathbb{R}$, i.e. for $i = 1, 2$ there holds*

$$\begin{aligned}|\lambda| |v|_i = |\lambda v|_i \quad &\text{for all } \lambda \in \mathbb{R} \text{ and all } v \in \mathcal{X}, \\ |v + w|_i \leq |v|_i + |w|_i \quad &\text{for all } v, w \in \mathcal{X}.\end{aligned}$$

Then, (i)–(ii) *are equivalent*

(i) *There exists $C > 0$ such for all $v \in \mathcal{X}$ holds $|v|_1 \leq C |v|_2$*
(ii) $\Big\{ v \in \mathcal{X} \,:\, |v|_1 = 0 \Big\} \supseteq \Big\{ v \in \mathcal{X} \,:\, |v|_2 = 0 \Big\}$

Proof. The implication $(i) \implies (ii)$ is trivial. It remains to prove $(ii) \implies (i)$. Let $|\cdot| : \mathcal{X} \to \mathbb{R}$ denote a seminorm and let $\mathcal{Y} \subseteq \mathcal{X}$ be a subspace with $\mathcal{Y} \supseteq \Big\{ v \in \mathcal{X} \,:\, |v| = 0 \Big\}$. Then, on the factor space $\mathcal{X}/\mathcal{Y}$, define

$$|v + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}} := \inf_{y \in \mathcal{Y}} |v + y| \quad \text{for all } v + \mathcal{Y} \in \mathcal{X}/\mathcal{Y}.$$

We will check that $|\cdot|_{\mathcal{X}/\mathcal{Y}}$ is a norm. First, homogeneity follows since $\lambda \mathcal{Y} = \mathcal{Y}$ and

$$|\lambda(v + \mathcal{Y})|_{\mathcal{X}/\mathcal{Y}} = |\lambda v + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}} = \inf_{y \in \mathcal{Y}} |\lambda v + y| = \inf_{y \in \mathcal{Y}} |\lambda (v + y)| = |\lambda| |v + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}}.$$

Second, we prove the triangle inequality

$$\begin{aligned}|v + \mathcal{Y} + w + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}} &= \inf_{y \in \mathcal{Y}} |v + w + y| = \inf_{y_1 \in \mathcal{Y}} \inf_{y_2 \in \mathcal{Y}} |v + w + y_1 + y_2| \\ &\leq \inf_{y_1 \in \mathcal{Y}} |v + w + y_1| + \inf_{y_2 \in \mathcal{Y}} |v + w + y_2| \\ &= |v + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}} + |w + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}}.\end{aligned}$$

Third, definiteness follows since $\mathcal{Y}$ contains the kernel of the seminorm, i.e. $|v + \mathcal{Y}|_{\mathcal{X}/\mathcal{Y}} = 0$ implies the existence of $y \in \mathcal{Y}$ with $|v + y| = 0$. This shows $v + y \in \mathcal{Y}$ and particularly implies $v \in \mathcal{Y}$. Therefore $v + \mathcal{Y} = \mathcal{Y} \equiv 0$ in $\mathcal{X}/\mathcal{Y}$.

We define $\mathcal{Y} := \Big\{ v \in \mathcal{X} \,:\, |v|_1 = 0 \Big\}$ and obtain that $|\cdot|_{i, \mathcal{X}/\mathcal{Y}}$ defines a norm on $\mathcal{X}/\mathcal{Y}$ for $i = 1, 2$. Since $\mathcal{X}/\mathcal{Y}$ is finite dimensional, all norms are equivalent, and we obtain that there exists a constant $C > 0$ such that for all $v \in \mathcal{X}$

$$|v|_1 = |v|_{1, \mathcal{X}/\mathcal{Y}} \leq C |v|_{2, \mathcal{X}/\mathcal{Y}} \leq C |v|_2.$$

This concludes the proof. □

Remark. *Suppose a triangulation $\mathcal{T}$. With the previous result, we are able to prove*

$$\|\nabla V\|_{L^2(\Omega)} \leq C \|V\|_{L^2(\Omega)} \quad \text{for all } V \in \mathcal{S}^p(\mathcal{T}),$$

since the vector space $\mathcal{X} = \mathcal{S}^p(\mathcal{T})$ is finite dimensional and both sides of the inequality define seminorms. However, the constant C may depend on the dimension of $\mathcal{S}^p(\mathcal{T})$ and

therefore depends on $|\mathcal{T}|$. As a result, we cannot guarantee that C remains bounded if $|\mathcal{T}| \to \infty$. To fix this issue, we will employ scaling arguments *in the following.* □

Proposition 26 (Inverse estimate)**.** *Suppose a γ-shape regular triangulation $\mathcal{T}$, a polynomial degree $p \geq 1$, as well as integers $m, n \in \mathbb{N}$ with $n \leq m$. Then, there exists a constant $C_{\mathrm{ie}} > 0$ such that for all $T \in \mathcal{T}$ all $\alpha \in \mathbb{R}$ and all $V \in \mathcal{S}^p(\mathcal{T})$ holds*

$$\|h_T^{m-n+\alpha} D^m V\|_{L^2(T)} \leq C_{\mathrm{ie}} \|h_T^\alpha D^n V\|_{L^2(T)}. \tag{44}$$

Particularly, this implies

$$\|h_{\mathcal{T}}^{m-n+\alpha} D^m V\|_{L^2(\Omega)} \leq C_{\mathrm{ie}} \|h_{\mathcal{T}}^\alpha D^n V\|_{L^2(\Omega)}. \tag{45}$$

The constant C_{ie} depends only on p, m, n and γ.

Proof. First, we prove (44) on the reference element. According to Lemma 25, there holds

$$\|D^m \widehat{V}\|_{L^2(\widehat{T})} \leq C \|D^n \widehat{V}\|_{L^2(\widehat{T})} \quad \text{for all } \widehat{V} \in \mathcal{S}^p(\widehat{T}).$$

Second, we transfer the result to $T \in \mathcal{T}$. Let $V \in \mathcal{S}^p(T)$ and define $\widehat{V} := V \circ \Phi_T$ with $\Phi_T(x) = Bx + y$ from Lemma 23. We also have $V = \widehat{V} \circ \Phi_T^{-1}$. Lemma 23 states

$$\|D^m V\|_{L^2(T)} = \|D^m(\widehat{V} \circ \Phi_T^{-1})\|_{L^2(T)} \leq |\mathrm{det} B^{-1}|^{-1/2} \|B^{-1}\|_F^m \|D^m \widehat{V}\|_{L^2(\widehat{T})}$$

as well as

$$\|D^n \widehat{V}\|_{L^2(\widehat{T})} \leq |\mathrm{det} B|^{-1/2} \|B\|_F^n \|D^n V\|_{L^2(T)}.$$

Plugging everything together, we end up with

$$\|D^m V\|_{L^2(T)} \leq C \|B^{-1}\|_F^m \|B\|_F^n \|D^n V\|_{L^2(T)}.$$

Lemma 22 shows that $\|B\|_F \simeq h_T$ and $\|B^{-1}\|_F \simeq h_T^{-1}$. We multiply by h_T^α to obtain (44). To prove (45), we sum over all elements, i.e.

$$\begin{aligned} \|h_{\mathcal{T}}^{m-n+\alpha} D^m V\|_{L^2(\Omega)}^2 &= \sum_{T \in \mathcal{T}} \|h_T^{m-n+\alpha} D^m V\|_{L^2(T)}^2 \\ &\leq C^2 \sum_{T \in \mathcal{T}} \|h_T^\alpha D^n V\|_{L^2(T)}^2 = \|h_{\mathcal{T}}^\alpha D^n V\|_{L^2(\Omega)}^2. \end{aligned}$$

This concludes the proof. □

3.4. **Trace inequality.** We start with an auxiliary result

Lemma 27 (Trace identity)**.** *Suppose a non-degenerate simplex $T = \mathrm{conv}\{z_0, \dots, z_d\} \subseteq \mathbb{R}^d$. Let $E = \mathrm{conv}\{z_1, \dots, z_d\}$ denote a hyperface of T (Note that the permutation of $z_0, \dots, z_d$ does not affect T). Then, there holds*

$$\frac{1}{|E|} \int_E w \, d\Gamma = \frac{1}{|T|} \int_T w \, dx + \frac{1}{d|T|} \int_T (x - z_0) \cdot \nabla w(x) \, dx \quad \textit{for all } w \in W^{1,1}(T). \tag{46}$$

Proof. Define the function $f(x) := w(x)(x - z_0)$. Then, there holds

$$\mathrm{div} f(x) = \sum_{j=1}^d \partial_j f_j(x) = \sum_{j=1}^d \Big(\partial_j w(x)(x - z_0)_j + w(x) \Big) = \nabla w(x) \cdot (x - z_0) + d w(x).$$

With integration by parts, we obtain from this

$$d \int_T w \, dx + \int_T (x - z_0) \cdot \nabla w(x) \, dx = \int_T \mathrm{div} f \, dx = \int_{\partial T} f \cdot n \, d\Gamma. \tag{47}$$

Consider the unique hyperplane $H \subseteq \mathbb{R}^d$ with $E \subseteq H$. Next, we aim to prove

$$(x - z_0) \cdot n = 0 \quad \text{for all } x \in \partial T \setminus E, \tag{48a}$$

$$(x - z_0) \cdot n = \operatorname{diam}(z_0, H) \quad \text{for all } x \in E. \tag{48b}$$

To see (48a), we observe that $x \in E' \neq E$ for some hyperface of E' of T. Since $z_0 \neq E$, it holds $z_0 \in E'$ and therefore $x - z_0 \perp n$, since $n \perp E$. To see (48b), consider the normal projection $x_0 \in H$ of z_0 onto the hyperplane H along n. The vectors n and $x_0 - z_0$ are parallel and therefore it holds $(x_0 - z_0) \cdot n = |n||x_0 - z_0| = \operatorname{dist}(z_0, H)$. Since $x, x_0 \in H$, there holds $x - x_0 \perp n$ and thus

$$(x - z_0) \cdot n = (x_0 - z_0) \cdot n + (x - x_0) \cdot n = \operatorname{dist}(z_0, H).$$

The combination of (47) and (48) shows

$$d \int_T w \, dx + \int_T (x - z_0) \cdot \nabla w(x) \, dx = \int_{\partial T} w(x)(x - z_0) \cdot n \, d\Gamma = \operatorname{dist}(z_0, H) \int_E w \, d\Gamma$$

Rearrangement of the equation shows

$$\frac{\operatorname{dist}(z_0, H)|E|}{d|T|} \frac{1}{|E|} \int_E w \, d\Gamma = \frac{1}{|T|} \int_T w \, dx + \frac{1}{d|T|} \int_T (x - z_0) \cdot \nabla w(x) \, dx.$$

By choosing $w = 1$, we see that $C := \frac{\operatorname{dist}(z_0, H)|E|}{d|T|} = 1$ and prove the statement. □

Proposition 28 (Trace inequality)**.** *Suppose a non-degenerate simplex $T = \operatorname{conv}\{z_0, \dots, z_d\} \subseteq \mathbb{R}^d$. Let $E = \operatorname{conv}\{z_1, \dots, z_d\}$ denote a hyperface of T. Then, there holds for all $v \in H^1(T)$*

$$\|v\|_{L^2(E)}^2 \leq \frac{|E|}{|T|} \Big(\|v\|_{L^2(T)}^2 + \frac{2}{d} h_T \|v\|_{L^2(T)} \|\nabla v\|_{L^2(T)} \Big) \tag{49}$$

as well as

$$\|v - v_E\|_{L^2(E)}^2 \leq \|v - v_T\|_{L^2(E)}^2 \leq \Big(C_{\mathrm{pc}}^2 + \frac{2}{d} C_{\mathrm{pc}} \Big) \frac{|E|}{|T|} h_T^2 \|\nabla v\|_{L^2(T)}^2, \tag{50}$$

where $v_E := |E|^{-1} \int_E v \, d\Gamma$ and $v_T := |T|^{-1} \int_T v \, dx$.

Proof. We set $w = v^2$ and use the trace identity (46) to obtain

$$\begin{aligned} \frac{1}{|E|} \int_E v^2 \, d\Gamma &= \frac{1}{|T|} \int_T v^2 \, dx + \frac{1}{d|T|} \int_T (x - z_0) \cdot 2v \nabla v \, dx \\ &\leq \frac{1}{|T|} \|v\|_{L^2(T)}^2 + \frac{2}{d|T|} h_T \|v\|_{L^2(T)} \|\nabla v\|_{L^2(T)}, \end{aligned}$$

since $|x - z_0| \leq h_T$. Multiplying by $|E|$, we prove (49). To see (50), we use the best approximation property of the integral mean, i.e.

$$\|v - v_E\|_{L^2(E)}^2 = \inf_{c \in \mathbb{R}} \|v - c\|_{L^2(E)}^2 \leq \|v - v_T\|_{L^2(E)}^2.$$

The trace inequality (49) then shows

$$\|v - v_E\|_{L^2(E)}^2 \leq \frac{|E|}{|T|} \Big(\|v - v_T\|_{L^2(T)}^2 + \frac{2}{d} h_T \|v - v_T\|_{L^2(T)} \|\nabla v\|_{L^2(T)} \Big).$$

Finally, we use the Poincaré estimate (43) to bound the L^2-norms with the H^1-seminorm. This results in

$$\|v - v_E\|_{L^2(E)}^2 \leq \frac{|E|}{|T|} \Big(C_{\mathrm{pc}}^2 h_T^2 \|\nabla v\|_{L^2(T)}^2 + \frac{2}{d} h_T^2 C_{\mathrm{pc}} \|\nabla v\|_{L^2(T)}^2 \Big).$$

□

3.5. **Axiom** (A1)**: stability.** The first axiom is verified for the residual error estimator $\eta(\cdot)$

Proposition 29. *The residual error estimator*

$$\eta_{\mathcal{T}}(T,V)^2 := |T|^{2/d}\|f + div\boldsymbol{A}\nabla V - \boldsymbol{b}\cdot\nabla V - cV\|^2_{L^2(\Omega)} + |T|^{1/d}\|[\boldsymbol{A}\nabla V\cdot n]\|^2_{L^2(\partial T\cap\Omega)}$$

satisfies the stability axiom (A1)*: There exists* $C_{\mathrm{stab}} > 0$ *such that for refinements* $\widehat{\mathcal{T}} \in \mathbb{T}$ *of* $\mathcal{T} \in \mathbb{T}$ *and all subsets of non-refined elements* $\mathcal{S} \subseteq \mathcal{T}\cap\widehat{\mathcal{T}}$ *there holds*

$$\Big|\Big(\sum_{T\in\mathcal{S}}\eta_{\widehat{\mathcal{T}}}(T,\widehat{V})^2\Big)^{1/2} - \Big(\sum_{T\in\mathcal{S}}\eta_{\mathcal{T}}(T,V)^2\Big)^{1/2}\Big| \le C_{\mathrm{stab}}\|\widehat{V}-V\|_{H^1(\Omega)} \tag{51}$$

for all $V\in\mathcal{X}_{\mathcal{T}}$ *and all* $\widehat{V}\in\mathcal{X}_{\widehat{\mathcal{T}}}$*. The constant* C_{stab} *depends only on the* γ*-shape regularity of* $\mathcal{T}$ *and* $\widehat{\mathcal{T}}$*.*

Proof. First, note that there holds

$$\eta_{\mathcal{T}}(T,V) = \eta_{\widehat{\mathcal{T}}}(T,V) \quad \text{for all } T\in\mathcal{T}\cap\widehat{\mathcal{T}} \text{ and all } V\in\mathcal{S}^p(\mathcal{S}).$$

Second, to abbreviate notation, define

$$|||V||| := \Big(\sum_{T\in\mathcal{S}}\eta_{\mathcal{T}}(T,V)^2\Big)^{1/2} \quad \text{for all } V\in\mathcal{S}^p(\mathcal{S}).$$

We may rewrite the left-hand side of (51) as $|||V||| - |||\widehat{V}|||$| and thus, by definition of $|||\cdot|||$, we get

$$\big|\,|||V||| - |||\widehat{V}|||\,\big| = \Big(\sum_{T\in\mathcal{S}}\alpha(T)^2\Big)^{1/2},$$

where

$$\begin{aligned}\alpha(T) &= |T|^{1/d}\|\mathrm{div}\boldsymbol{A}\nabla(V-\widehat{V}) - \boldsymbol{b}\cdot\nabla(V-\widehat{V}) - c(V-\widehat{V})\|_{L^2(T)}\\ &\quad + |T|^{1/(2d)}\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|_{L^2(\partial T\cap\Omega)}\\ &\le |T|^{1/d}\Big(\|\mathrm{div}\boldsymbol{A}\nabla(V-\widehat{V})\|_{L^2(\Omega)} + \big(\|\boldsymbol{b}\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\big)\|V-\widehat{V}\|_{H^1(\Omega)}\Big)\\ &\quad + |T|^{1/(2d)}\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|_{L^2(\partial T\cap\Omega)}\end{aligned} \tag{52}$$

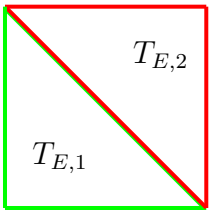
The first term on the right-hand side is further estimated by use of the product rule and the inverse estimate (44) with $m=2, n=1$, and $\alpha=0$

$$\begin{aligned}\|\mathrm{div}\boldsymbol{A}\nabla(V-\widehat{V})\|_{L^2(\Omega)} &\le \|\nabla\boldsymbol{A}\|_{L^\infty(\Omega)}\|\nabla(V-\widehat{V})\|_{L^2(T)} + \|\boldsymbol{A}\|_{L^\infty(\Omega)}\|\Delta(V-\widehat{V})\|_{L^2(\Omega)}\\ &\le \big(\|\nabla\boldsymbol{A}\|_{L^\infty(\Omega)} + C_{\mathrm{ie}}h_T^{-1}\|\boldsymbol{A}\|_{L^\infty(\Omega)}\big)\|\nabla(V-\widehat{V})\|_{L^2(T)}\end{aligned}$$

The last term on the right-hand side of (52) is split further

$$\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|^2_{L^2(\partial T\cap\Omega)} = \sum_{E \text{ hyperface of } \partial T\cap\Omega}\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|^2_{L^2(E)}.$$

Each interior hyperface $E\subseteq\partial T\cap\Omega$ is the intersection of two elements $T_{E,1}, T_{E,2}\in\mathcal{T}$ with $T_{E,1}\cap T_{E,2} = E$ as sketched below



For each hyperface E, this leads to

$$\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|_{L^2(T_{E,1}\cap T_{E,2})} \le \|\boldsymbol{A}\|_{L^\infty(\Omega)}\Big(\|\nabla(V-\widehat{V})\|_{L^2(\partial T_{E,1})} + \|\nabla(V-\widehat{V})\|_{L^2(\partial T_{E,2})}\Big).$$

The trace inequality (46) and the Young inequality show

$$\begin{aligned}
&\|[\boldsymbol{A}\nabla(V-\widehat{V})\cdot n]\|^2_{L^2(T_{E,1}\cap T_{E,2})}\\
&\quad\le \|\boldsymbol{A}\|^2_{L^\infty(\Omega)}\max\left\{\frac{|E|}{|T_{E,1}|},\frac{|E|}{|T_{E,2}|}\right\}\Big(\|\nabla(V-\widehat{V})\|^2_{L^2(T_{E,1}\cup T_{E,2})}\\
&\qquad\qquad + \frac{2}{d}\max\{h_{T_{E,1}},h_{T_{E,2}}\}\|V-\widehat{V}\|_{L^2(T_{E,1}\cup T_{E,2})}\|\nabla(V-\widehat{V})\|_{L^2(T_{E,1}\cup T_{E,2})}\Big)\\
&\quad\le C\|\boldsymbol{A}\|^2_{L^\infty(\Omega)}\max\{h^{-1}_{T_{E,1}},h^{-1}_{T_{E,2}}\}\|V-\widehat{V}\|^2_{H^1(T_{E,1}\cup T_{E,2})},
\end{aligned}$$

where $C>0$ depends only on $\|h_{\mathcal{T}}\|_{L^\infty(\Omega)}\le \operatorname{diam}(\Omega)$ as well as the γ-shape regularity which ensures $|E|/|T|\simeq h_T^{-1}$. Plugging this into (52), we end up with

$$\begin{aligned}
\alpha(T) &\le |T|^{1/d}\Big(\|\nabla\boldsymbol{A}\|_{L^\infty(\Omega)} + C_{\mathrm{ie}}h_T^{-1}\|\boldsymbol{A}\|_{L^\infty(\Omega)}\Big)\|\nabla(V-\widehat{V})\|_{L^2(T)}\\
&\quad + |T|^{1/d}\Big(\|\boldsymbol{b}\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)}\Big)\|V-\widehat{V}\|_{H^1(T)}\\
&\quad + C^{1/2}\|\boldsymbol{A}\|_{L^\infty(\Omega)}|T|^{1/(2d)}\\
&\qquad\Big(\sum_{E\text{ hyperface of }\partial T\cap\Omega}\max\{h^{-1}_{T_{E,1}},h^{-1}_{T_{E,2}}\}\|V-\widehat{V}\|^2_{H^1(T_{E,1}\cup T_{E,2})}\Big)^{1/2}.
\end{aligned}$$

The γ-shape regularity shows $|T|\simeq|T_{E,1}|\simeq|T_{E,2}|$ (see Lemma 21) and therefore, we prove

$$\alpha(T)\le (C'+h_T)\|V-\widehat{V}\|_{H^1(T)},$$

for some constant $C'>0$ which depends only on the γ-shape regularity of $\mathcal{T}$ and $\widehat{\mathcal{T}}$. We sum over all $T\in\mathcal{S}$ to get

$$\Big|\|\!|V|\!\| - \|\!|\widehat{V}|\!\|\Big| \le (C'+\operatorname{diam}(\Omega))\|V-\widehat{V}\|_{H^1(\Omega)}.$$

□

References

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

[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.

[CFPP14] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[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.

[Fei19] Michael Feischl. Optimality of a standard adaptive finite element method for the Stokes problem. *SIAM J. Numer. Anal.*, 57(3):1124–1157, 2019.

[Fei22] Michael Feischl. Inf-sup stability implies quasi-orthogonality. *Math. Comp.*, 91(337):2059–2094, 2022.

[KPP12] Michael Karkulik, David Pavlicek, and Dirk Praetorius. On 2D newest vertex bisection: Optimality of mesh-closure and H^1-stability of L_2-projection. *ASC Report*, 10/2012, *Institute for Analysis and Scientific Computing, Vienna University of Technology*, 2012.

[Pav10] David Pavlicek. Optimalität adaptiver FEM, Bachelor thesis (in german). *Institute for Analysis and Scientific Computing, Vienna University of Technology*, 2010.

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

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