Universität Augsburg

Institut für
Mathematik

Pedro Morin, Kunibert G. Siebert, Andreas Veeser

A Basic Convergence Result for Conforming Adaptive Finite Elements

A BASIC CONVERGENCE RESULT FOR CONFORMING ADAPTIVE FINITE ELEMENTS

PEDRO MORIN, KUNIBERT G. SIEBERT, AND ANDREAS VEESER

ABSTRACT. We consider the approximate solution with adaptive finite elements of a class of linear boundary value problems, which includes problems of 'saddle point' type. For the adaptive algorithm we suppose the following framework: refinement relies on unique quasi-regular element subdivisions and generates locally quasi-uniform grids, the finite element spaces are conforming, nested, and satisfy the inf-sup conditions, the error estimator is reliable as well as locally and discretely efficient, and marked elements are subdivided at least once. Under these assumptions, we give a sufficient and essentially necessary condition on marking for the convergence of the finite element solutions to the exact one. This condition is not only satisfied by Dörfler's strategy, but also by the maximum strategy and the equidistribution strategy.

1. INTRODUCTION

Finite elements are a well-established tool for the numerical solution of boundary value problems. Adaptivity often allows for a much more efficient use of the computational resources; for certain problems, it is even indispensable to their numerical solvability. A quite popular, natural adaptive version of classical finite element methods consists of the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}, \tag{1.1}$$

that is: solve for the finite element solution on the current grid, compute the a posteriori error estimator, mark with its help elements to be subdivided, and refine the current grid into a new, finer one. While there is a rather elaborated theory for the derivation of a posteriori error estimators, see e.g. [1, 5, 33], the theoretical understanding of the convergence and complexity of such methods is in the early stages.

Babuška/Vogelius [7] give an analysis of (1.1) for linear, elliptic, and symmetric problems in 1d. The first multidimensional result is Dörfler [17]. It ensures with pathbreaking ideas that, after a pre-adaptation to data, (1.1) reduces the error below any prescribed tolerance. Proper convergence without conditions on the initial grid was proved by Morin/Nochetto/Siebert [21]. The latter work was generalized in various directions by, e.g., [13, 15, 16, 20, 22, 27, 28, 31]. Relying on techniques from Binev/Dahmen/DeVore [10] and new ideas, Stevenson [30] gives an important complexity result for an algorithm that is quite close to (1.1).

Without any doubts, these results are an important progress in the theoretical understanding of (1.1). Nevertheless, a critical review reveals the following limitations:

- *Symmetric elliptic problems.* Except for [15, 20], all aforementioned results are within a coercive elliptic variational framework. Carstensen/Hoppe [15] treat a mixed formulation for the Poisson problem by exploiting a close relationship to the direct discretization with Crouzeix-Raviart elements. Mekchay/Nochetto

[20] allow for a non-symmetric operator, at the price that the initial grid has to be sufficiently fine – a condition of a priori spirit, the necessity of which is not supported by numerical experiments even for convection-dominated diffusion problems.

- *Conforming discretizations and Galerkin solutions.* In all results the discrete space(s) are included in the continuous one(s) and the discrete solution is the Galerkin approximation.

- *Dörfler's strategy and special marking for oscillation.* Except for [27], all convergence and complexity results assume Dörfler's strategy. However, numerical results strongly suggest that, for instance, the maximum strategy is competitive with Dörfler's one. Results with the standard residual error estimator require a special treatment like marking or pre-adaptation for (data) oscillation, although the standard residual estimator dominates oscillation and algorithms without special treatment of oscillation also seem to work.

- *Artificial conditions for refinement.* All results require a certain, non-minimal subdivision depth for patches associated with marked elements, e.g., the creation of interior nodes is often required. Computational evidence does not support the necessity of these requirements for convergence or quasi-optimal complexity. In addition, such requirements complicate the implementation and, if the required subdivision depth is quite big, yield (1.1) non-practical.

The purpose of this paper is to present a new, basic convergence result for (1.1) applied to linear boundary value problems. To this end, §2 gives general assumptions on the problem itself, the refinement framework, the finite element spaces, the approximate solution, the a posteriori error estimator, the marking strategy and the step REFINE. They ensure the convergence in the following sense:

Let u be the exact solution and let $\| \cdot \|$ be the norm of the trial space. For each iteration k, denote the approximate solution by u_k and the a posteriori error estimator by $\mathcal{E}_k$. Then both error and estimator converge to 0:

$$\|u_k - u\| \to 0 \quad \text{and} \quad \mathcal{E}_k \to 0 \quad \text{as} \quad k \to \infty. \tag{1.2}$$

In order to illustrate the generality of the assumptions, we verify them for various examples in §3. We obtain (1.2) in particular for:

- Saddle point problems, non-symmetric problems without assuming a sufficiently fine initial grid;

- various types of a posteriori error estimators, including the ZZ one and estimators where the (discrete) lower bound requires a very fine, local extension of the finite element trial space;

- the maximum and equidistribution strategy, without extra marking for oscillation;

- a minimal rule for REFINE, which does not necessarily entail the creation of interior nodes.

In §4 we prove the indicated convergence result. The proof differs from the aforementioned ones in the following aspects:

- *Without strict error reduction and without extra regularity.* Except for [27], all previous convergence proofs for linear problems use a quasi-orthogonality result to prove the following: if oscillation is relatively small, then one iteration of (1.1) provides a uniformly strict error reduction. To compensate for the missing quasi-orthogonality result for nonlinear potential operators, Siebert/Veeser [28] and Veeser [31] control the residual by the energy reduction in one iteration of (1.1) and then exploit the fact that the energy reduction vanishes in the limit.

Here, with the help of Lemma 4.2, we use a similar argument that covers also non-potential operators which neither supposes nor implies regularity of the exact solution beyond the trial space.

- *Estimator convergence first.* With the strict error reduction, convergence of error and estimator are shown simultaneously. Here and in [27, 28, 31], the convergence of the estimator is first derived by exploiting the (discrete) local lower but not the upper bound. The upper bound is then used to obtain the convergence of the error.

- *Without residual control in one iteration.* The strict error reduction and the residual control by the energy reduction share the following features: they involve two successive approximate solutions and they split the estimator into marked and non-marked indicators. These features entail the aforementioned artificial conditions for refinement. Here we avoid this by means of a new splitting of the estimator, see (4.9), and a local lower bound for the difference of not necessarily successive approximate solutions, see Remark 4.6.

Since the strict error reduction is used in the complexity result [30], one may consider its absence as a disadvantage. On the other hand, for solutions of non-coercive problems with, e.g., layers, such a result appears to be unrealistic when starting from a coarse initial grid.

Two corollaries in §5 conclude this article. They are rather immediate consequences of our convergence proof. The first one provides necessary and sufficient conditions on marking for (1.2), while the second gives such a condition for reaching any prescribed tolerance in a finite number of iterations. Another corollary from the convergence proof, namely the convergence of algorithms adapting for a (semi)norm that is weaker than the one of trial space, is given in [23].

2. Abstract Convergence

We start by introducing the problem class to be considered. Then we describe the adaptive algorithm, the assumptions on its building blocks and state the convergence result at the end of this section. For the sake of simplicity, we shall stick to a Hilbert space framework; the following results naturally extend to a reflexive Banach space setting, see also [23].

2.1. **Problem Class.** We consider linear boundary value problems that can be reformulated in the following weak form: given real Hilbert spaces $\mathbb{V}$ and $\mathbb{W}$, a continuous bilinear form $\mathcal{B} : \mathbb{V} \times \mathbb{W} \to \mathbb{R}$, and an element $f \in \mathbb{W}^*$ of the (topological) dual space of $\mathbb{W}$, find

$$u \in \mathbb{V}: \qquad \mathcal{B}(u, w) = \langle f, w \rangle \quad \forall w \in \mathbb{W}. \tag{2.1}$$

We suppose that (2.1) is well-posed, that is there exists a unique solution for any $f \in \mathbb{W}^*$. The latter is equivalent to the so-called *inf-sup (or Babuška-Brezzi) condition:* there exists $\alpha > 0$ such that

$$\inf_{\substack{v \in \mathbb{V} \\ \|v\|_{\mathbb{V}}=1}} \sup_{\substack{w \in \mathbb{W} \\ \|w\|_{\mathbb{W}}=1}} \mathcal{B}(v, w) \geq \alpha, \qquad \inf_{\substack{w \in \mathbb{W} \\ \|w\|_{\mathbb{W}}=1}} \sup_{\substack{v \in \mathbb{V} \\ \|v\|_{\mathbb{V}}=1}} \mathcal{B}(v, w) \geq \alpha; \tag{2.2a}$$

see Babuška [2], Babuška/Aziz [3] or Brezzi [12, Corollary 0.1].

In addition, we make the following assumptions that introduce *'local features'* into (2.1). Let Ω be a domain in $\mathbb{R}^d$, $d \geq 2$, that can be meshed. The Hilbert spaces $\mathbb{V} = \mathbb{V}(\Omega)$, $\mathbb{W} = \mathbb{W}(\Omega)$ are subspaces of $L_2(\Omega; \mathbb{R}^m)$ with some $m \in \mathbb{N}$ and such that the squares of their induced norms $\| \cdot \|_{\mathbb{V}(\Omega)}$, $\| \cdot \|_{\mathbb{W}(\Omega)}$ are subadditive

in the underlying domain: if Ω_1, Ω_2 are subdomains such that $\bar{\Omega}_1 \cup \bar{\Omega}_2 = \bar{\Omega}$ and $\Omega_1 \cap \Omega_2 = \emptyset$, then there holds, for instance,

$$\|v\|^2_{\mathbb{V}(\Omega_1)} + \|v\|^2_{\mathbb{V}(\Omega_2)} \leq \|v\|^2_{\mathbb{V}(\Omega)} \tag{2.2b}$$

for all $v \in \mathbb{V}(\Omega)$. Moreover, we suppose that the norm $\|\cdot\|_{\mathbb{V}}$ is absolutely continuous with respect to the Lebesgue measure $|\cdot|$ in that, for any $v \in \mathbb{V}$, there holds

$$\|v\|_{\mathbb{V}(\omega)} \to 0 \quad \text{as} \quad |\omega| \to 0. \tag{2.2c}$$

The last two assumptions allow also for fractional Sobolev spaces.

Finally, we require the following 'local continuity' of $\mathcal{B}$: there is a constant $C_{\mathcal{B}} \geq 0$ such that, if ω is a subdomain of Ω, then we have

$$w = 0 \text{ in } \Omega \setminus \bar{\omega} \quad \Longrightarrow \quad |\mathcal{B}(v, w)| \leq C_{\mathcal{B}} \|v\|_{\mathbb{V}(\omega)} \|w\|_{\mathbb{W}(\Omega)} \qquad \forall v \in \mathbb{V}. \tag{2.2d}$$

2.2. Adaptive Algorithm.

In order to describe the adaptive algorithm, we first present it in a schematic way and then discuss its modules. Doing so, we suppress the dependence on the data Ω, f, and $\mathcal{B}$ of the variational problem (2.1).

The algorithm for approximating u in (2.1) is an iteration of the following main steps:

(1) $u_k := \mathsf{SOLVE}\big(\mathbb{V}(\mathcal{G}_k), \mathbb{W}(\mathcal{G}_k)\big)$.

(2) $\{\mathcal{E}_k(E)\}_{E \in \mathcal{G}_k} := \mathsf{ESTIMATE}\big(u_k, \mathcal{G}_k\big)$.

(3) $\mathcal{M}_k := \mathsf{MARK}\big(\{\mathcal{E}_k(E)\}_{E \in \mathcal{G}_k}, \mathcal{G}_k\big)$.

(4) $\mathcal{G}_{k+1} := \mathsf{REFINE}\big(\mathcal{G}_k, \mathcal{M}_k\big)$, increment k. $\tag{2.3}$

In practice, a stopping test is used after step (2) for terminating the iteration; here we shall ignore it for notational convenience. The realization of these steps requires the following objects and modules:

- An initial grid $\mathcal{G}_0$ of Ω and a refinement procedure REFINE. The refinement procedure has two input arguments: a grid $\mathcal{G}$ and a subset $\mathcal{M} \subset \mathcal{G}$. All elements $E \in \mathcal{M}$ must be 'refined'. The input grid $\mathcal{G}$ can be the initial grid $\mathcal{G}_0$ or the output of a previous application of REFINE. A grid $\mathcal{G}'$ is called a refinement of $\mathcal{G}$ whenever $\mathcal{G}'$ can be produced from $\mathcal{G}$ by a finite number of applications of REFINE. Note that the set

$$\mathbb{G} := \{\mathcal{G} \mid \mathcal{G} \text{ is a refinement of } \mathcal{G}_0\} \tag{2.4}$$

has the property that there always holds $\mathcal{G}_k \in \mathbb{G}$ for the sequence of grids generated by the above iteration, irrespective of the decisions in MARK.

- For any grid $\mathcal{G} \in \mathbb{G}$, there are finite element spaces $\mathbb{V}(\mathcal{G})$ and $\mathbb{W}(\mathcal{G})$ such that the module SOLVE outputs an approximation $u_{\mathcal{G}} \in \mathbb{V}(\mathcal{G})$ of u.

- A module $\mathsf{ESTIMATE}$ that, given a grid $\mathcal{G} \in \mathbb{G}$ and the corresponding output $u_{\mathcal{G}}$ of SOLVE, computes and outputs the a posteriori error estimator $\{\mathcal{E}_{\mathcal{G}}(E)\}_{E \in \mathcal{G}}$, where the so-called indicator $\mathcal{E}_{\mathcal{G}}(E) \geq 0$ is associated with the element $E \in \mathcal{G}$.

- A strategy in the module MARK that, based upon the a posteriori error indicators $\{\mathcal{E}_{\mathcal{G}}(E)\}_{E \in \mathcal{G}}$, collects elements of $\mathcal{G}$ in $\mathcal{M}$, which serve as input for REFINE.

We now state the assumptions on these objects and modules that will be used for the abstract convergence result Theorem 2.1 below. Doing so, we shall write '$\preccurlyeq$' for '$\leq C$' where C may depend on data of (2.1), the class $\mathbb{G}$, and the module $\mathsf{ESTIMATE}$ but not on a particular grid or the iteration number. Similarly, we say that some object is 'fixed' if it only depends on the modules $\mathsf{ESTIMATE}$, MARK and on $\mathbb{G}$. A further discussion of the assumptions will be provided within the applications of Theorem 2.1 in §3.

Framework for Refinement. We suppose that refinement relies on *unique quasi-regular element subdivisions*. More precisely, there exist constants $q_1, q_2 \in (0, 1)$ such that, irrespective of the grid $\mathcal{G}$, any element $E \in \mathcal{G}$ can be subdivided into $n(E) \geq 2$ subelements $E_1', \ldots, E_{n(E)}'$ such that

$$E = E_1' \cup \cdots \cup E_{n(E)}', \qquad |E| = |E_1'| + \cdots + |E_{n(E)}'|, \tag{2.5a}$$

and

$$q_1 |E| \leq |E_i'| \leq q_2 |E|, \qquad i = 1, \ldots, n(E). \tag{2.5b}$$

These unique element subdivisions generate a 'master forest' $\mathcal{F}$ of infinite trees, where each node corresponds to an element, its direct successors to its subelements, and the roots to the elements of the initial grid $\mathcal{G}_0$. A subforest $\hat{\mathcal{F}} \subset \mathcal{F}$ is called finite if it has a finite number of nodes. Any finite tree may have interior nodes, i.e. nodes with successors, and does have leaf nodes, i.e. nodes without any successor.

Any subdivision $\mathcal{S}$ of the domain Ω that is subordinated to $\mathcal{G}_0$ is uniquely associated with a finite subforest $\mathcal{F}(\mathcal{S})$ of $\mathcal{F}$, where the leaf nodes are the elements of the subdivision. A grid $\mathcal{G}' \in \mathbb{G}$ is a refinement of the grid $\mathcal{G} \in \mathbb{G}$ whenever $\mathcal{F}(\mathcal{G}) \subset \mathcal{F}(\mathcal{G}')$.

For two elements E, E' within a tree $\mathcal{T} \subset \mathcal{F}$, we denote by $\mathrm{dist}(E, E') \in \mathbb{N}_0$ the distance of the nodes E, E' in $\mathcal{T}$. Given $n \in \mathbb{N}$ and a subset $\hat{\mathcal{S}}$ of subdivision $\mathcal{S}$ subordinated to $\mathcal{G}_0$, we denote by $\mathcal{F}^n(\mathcal{S}, \hat{\mathcal{S}})$ the subforest of $\mathcal{F}$ that consists of $\mathcal{F}(\mathcal{S})$ and all successors of elements in $\hat{\mathcal{S}}$ with distance $\leq n$.

We suppose that the class $\mathbb{G}$ is a subclass of the subdivisions of Ω subordinated to $\mathcal{G}_0$ and is *locally quasi-uniform* in that

$$\sup_{\mathcal{G} \in \mathbb{G}} \max_{E \in \mathcal{G}} \# N_{\mathcal{G}}(E) \preccurlyeq 1, \qquad \sup_{\mathcal{G} \in \mathbb{G}} \max_{E' \in N_{\mathcal{G}}(E)} \frac{|E|}{|E'|} \preccurlyeq 1, \tag{2.5c}$$

where $N_{\mathcal{G}}(E) := \{E' \in \mathcal{G} \mid E' \cap E \neq \emptyset\}$ denotes the set of neighbors of E in $\mathcal{G}$. The grids in $\mathbb{G}$ may have additional properties like conformity.

Thus, the master forest $\mathcal{F}$ contains all possible subdivisions, while $\cup_k \mathcal{F}(\mathcal{G}_k)$ consists of all subdivisions associated with a realization of (2.3) and may be a proper subforest of $\mathcal{F}$.

Finite Element Spaces and SOLVE. We suppose that the finite element spaces $\mathbb{V}(\mathcal{G})$ and $\mathbb{W}(\mathcal{G})$ are *conforming*, *nested*, and satisfy a *discrete inf-sup condition*: for any $\mathcal{G}, \mathcal{G}' \in \mathbb{G}$, there hold

$$\mathbb{V}(\mathcal{G}) \subset \mathbb{V} \quad \text{and} \quad \mathbb{W}(\mathcal{G}) \subset \mathbb{W} \tag{2.6a}$$

$$\mathcal{G}' \text{ is a refinement of } \mathcal{G} \implies \mathbb{V}(\mathcal{G}) \subset \mathbb{V}(\mathcal{G}') \tag{2.6b}$$

$$\inf_{\substack{v \in \mathbb{V}(\mathcal{G}) \\ \|v\|_{\mathbb{V}} = 1}} \sup_{\substack{w \in \mathbb{W}(\mathcal{G}) \\ \|w\|_{\mathbb{W}} = 1}} \mathcal{B}(v, w) \geq \beta \quad \text{and} \quad \dim \mathbb{V}(\mathcal{G}) = \dim \mathbb{W}(\mathcal{G}) \tag{2.6c}$$

with some fixed $\beta > 0$. Moreover, we suppose that the output

$$u_{\mathcal{G}} := \mathsf{SOLVE}\big(\mathbb{V}(\mathcal{G}), \mathbb{W}(\mathcal{G})\big)$$

is the *Petrov-Galerkin approximation* of u with respect to $(\mathbb{V}(\mathcal{G}), \mathbb{W}(\mathcal{G}))$:

$$u_{\mathcal{G}} \in \mathbb{V}(\mathcal{G}): \qquad \mathcal{B}(u_{\mathcal{G}}, w) = \langle f, w \rangle, \quad \forall w \in \mathbb{W}(\mathcal{G}). \tag{2.7}$$

Thanks to (2.6c) and (2.6a), the Petrov-Galerkin approximation $u_{\mathcal{G}}$ exists, is unique, and is a $\| \cdot \|_{\mathbb{V}}$-quasi-optimal choice from $\mathbb{V}(\mathcal{G})$.

Notice that condition (2.6c) is equivalent to the discrete counterpart of (2.2a):

$$\inf_{\substack{v \in \mathbb{V}(\mathcal{G}) \\ \|v\|_{\mathbb{V}} = 1}} \sup_{\substack{w \in \mathbb{W}(\mathcal{G}) \\ \|w\|_{\mathbb{W}} = 1}} \mathcal{B}(v, w) = \inf_{\substack{w \in \mathbb{W}(\mathcal{G}) \\ \|w\|_{\mathbb{W}} = 1}} \sup_{\substack{v \in \mathbb{V}(\mathcal{G}) \\ \|v\|_{\mathbb{V}} = 1}} \mathcal{B}(v, w) \geq \beta. \tag{2.8}$$

Moreover, for non-adaptive realizations of (2.3), condition (2.6c) is necessary for the well-posedness of (2.7) and convergence irrespective of $f \in \mathbb{W}^*$; see for instance [11, III.3.9].

ESTIMATE. We suppose that the output

$$\{\mathcal{E}_{\mathcal{G}}(E)\}_{E \in \mathcal{G}} := \mathsf{ESTIMATE}(u_{\mathcal{G}}, \mathcal{G})$$

has the following two properties for any grid $\mathcal{G} \in \mathbb{G}$: First, there holds the following *global upper bound* for the error of the Petrov-Galerkin approximation $u_{\mathcal{G}}$:

$$\|u_{\mathcal{G}} - u\|_{\mathbb{V}(\Omega)} \preccurlyeq \mathcal{E}_{\mathcal{G}}, \tag{2.9a}$$

where, given a subset $\hat{\mathcal{G}} \subset \mathcal{G}$, we define

$$\mathcal{E}_{\mathcal{G}}(\hat{\mathcal{G}}) := \left(\sum_{E \in \hat{\mathcal{G}}} \mathcal{E}_{\mathcal{G}}^2(E) \right)^{1/2}$$

and set $\mathcal{E}_{\mathcal{G}} := \mathcal{E}_{\mathcal{G}}(\mathcal{G})$ as well as $\mathcal{E}_{\mathcal{G}}(\emptyset) := 0$.

Secondly, a *fixed finite subdivision depth implies a local lower bound.* More precisely, there is a fixed $n \in \mathbb{N}$ such that, for any element $E \in \mathcal{G}$ and any finer grid $\mathcal{G}' \in \mathbb{G}$ with $\mathcal{F}(\mathcal{G}') \supset \mathcal{F}^n\big(\mathcal{G}, N_{\mathcal{G}}(E)\big)$, there holds

$$\mathcal{E}_{\mathcal{G}}(E) \preccurlyeq \sup \left\{ \langle \mathcal{R}_{\mathcal{G}}, w \rangle \mid w \in \mathbb{W}\big(\mathcal{G}'; \omega_{\mathcal{G}}(E)\big), \ \|w\|_{\mathbb{W}(\Omega)} \le 1 \right\} + \mathrm{osc}_{\mathcal{G}}(E), \tag{2.9b}$$

where the oscillation indicator satisfies

$$\mathrm{osc}_{\mathcal{G}}(E) \preccurlyeq m(|E|)\Big(\|u_{\mathcal{G}}\|_{\mathbb{V}(\omega_{\mathcal{G}}(E))} + \|D\|_{\mathbb{D}(\omega_{\mathcal{G}}(E))} \Big). \tag{2.9c}$$

Hereafter

- $\mathcal{R}_{\mathcal{G}} \in \mathbb{W}^*$ is the residual defined by

$$\langle \mathcal{R}_{\mathcal{G}}, w \rangle := \mathcal{B}(u_{\mathcal{G}}, w) - \langle f, w \rangle, \qquad \forall w \in \mathbb{W}; \tag{2.10}$$

- $\omega_{\mathcal{G}}(E) \subset \Omega$ is the patch (union) of elements in $N_{\mathcal{G}}(E)$;
- $\mathbb{W}\big(\mathcal{G}'; \omega_{\mathcal{G}}(E)\big)$ is the space of 'local test functions' given by

$$\mathbb{W}\big(\mathcal{G}'; \omega_{\mathcal{G}}(E)\big) := \big\{ w \in \mathbb{W}(\mathcal{G}') \mid \mathrm{supp}\, w \subset \omega_{\mathcal{G}}(E) \big\};$$

- $m : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is a fixed, continuous, and nondecreasing function with $m(0) = 0$;
- $\mathbb{D}$ is another Hilbert space with a norm satisfying (2.2b), (2.2c) and $D \in \mathbb{D}$ is given by the data of (2.1).

The global upper bound (2.9a), which is equivalent to $\|\mathcal{R}_{\mathcal{G}}\|_{\mathbb{W}^*} \preccurlyeq \mathcal{E}_{\mathcal{G}}$ in view of (2.2a) and (2.2d), ensures that the error indicators do not overview any error source. Inequality (2.9b) is a variant of the main step in proving a local lower error bound by Verfürth's constructive argument [33]: indeed, if one inserts (2.1) into (2.10) and recalls (2.2d), then (2.9b) readily yields the local lower error bound

$$\mathcal{E}_{\mathcal{G}}(E) \preccurlyeq \|u_{\mathcal{G}} - u\|_{\mathbb{V}(\omega_{\mathcal{G}}(E))} + \mathrm{osc}_{\mathcal{G}}(E). \tag{2.11}$$

Thus, (2.9b) ensures, up to (2.9c), the sharpness of the upper bound (2.9a) in a local sense. The presence of the oscillation indicator (2.9c) is discussed in Remark 4.7.

MARK. We suppose that the output

$$\mathcal{M} := \mathsf{MARK}\big(\{\mathcal{E}_{\mathcal{G}}(E)\}_{E \in \mathcal{G}}, \mathcal{G}\big) \tag{2.12}$$

of marked elements has the property

$$\forall\, E \in \mathcal{G} \setminus \mathcal{M} \qquad \mathcal{E}_{\mathcal{G}}(E) \le g\big(\mathcal{E}_{\mathcal{G}}(\mathcal{M})\big), \tag{2.13}$$

where $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is a fixed function that is continuous in 0 with $g(0) = 0$.

Most marking strategies used in practice are of the form (2.12) and satisfy (2.13); see §3.2 for examples. Nevertheless in §5 below, we consider more general marking strategies and give a weaker condition that is essentially necessary and sufficient for convergence.

REFINE. We suppose that the output grid

$$\mathcal{G}' := \mathsf{REFINE}(\mathcal{G}, \mathcal{M})$$

satisfies the *minimal requirement*

$$\mathcal{F}(\mathcal{G}') \supset \mathcal{F}^1(\mathcal{G}, \mathcal{M}), \tag{2.14}$$

that is, each marked element of the input grid is subdivided at least once in the output grid. Additional elements in $\mathcal{G} \setminus \mathcal{M}$ may be refined in order to fulfill (2.5c) or to meet some given condition for grids of the class $\mathbb{G}$ like conformity.

2.3. **Convergence.** We now state the main result of this paper, which essentially says that any adaptive algorithm satisfying the assumptions of §2.2 converges for each problem of the class given in §2.1.

Theorem 2.1 (Abstract Convergence). *Let u be the exact solution of (2.1), suppose that there holds (2.2) and that $\{u_k\}_k$ is the sequence of approximate solutions generated by iteration (2.3).*

If the refinement framework, the finite element spaces, the modules SOLVE, ESTIMATE, MARK, and REFINE satisfy, respectively, (2.5), (2.6), (2.7), (2.9), (2.13), and (2.14), then both error and estimator vanish in the limit, that is

$$\|u_k - u\|_{\mathbb{V}(\Omega)} \to 0 \quad and \quad \mathcal{E}_k \to 0 \quad as \quad k \to \infty.$$

Notice that convergence of both error and estimator in Theorem 2.1 are important. The first convergence, $\lim_{k\to\infty} \|u - u_k\|_{\mathbb{V}} = 0$, means that the approximate solutions get arbitrarily close to the exact one. However, this would be of little practical use without the second one, $\lim_{k\to\infty} \mathcal{E}_k = 0$, which is the computable counterpart of the first one and thus allows to recognize the improvement of the approximate solutions. In particular, $\lim_{k\to\infty} \mathcal{E}_k = 0$ ensures that, if one includes a stopping test with a given positive tolerance, then the algorithm stops after a finite number of iterations.

Several applications of Theorem 2.1 are now given in §3, while its proof and supplements are postponed to §4 and §5.

3. Applications

The purpose of this section is to illustrate the general nature of the assumptions in Theorem 2.1 by several examples.

In all examples we shall use the following function $h_{\mathcal{G}}$ associated with any grid $\mathcal{G} \in \mathbb{G}$. Given $x \in \bar{\Omega}$, let $\mathcal{G}(x) := \bigcup \{E \in \mathcal{G} \colon E \ni x\}$ and set

$$h_{\mathcal{G}}(x) := |\mathcal{G}(x)|^{1/d}. \tag{3.1}$$

This meshsize function, or more precisely, its equivalence class in $L_\infty(\Omega)$ will also be used in the proof of Theorem 2.1 in §4 below. Furthermore, we shall write $\|\cdot\|_\omega$ as a shortcut for $\|\cdot\|_{L_2(\omega;\mathbb{R}^m)}$ whenever ω denotes the underlying domain of integration and $m \in \mathbb{N}$.

3.1. **Various Types of Error Estimators.** In this first example we apply iteration (2.3) to the Poisson problem using bisection, Courant elements, Dörfler's marking strategy, and various types of a posteriori error estimators.

Problem. Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded, polyhedral Lipschitz domain, set

$$\mathbb{V} = \mathbb{W} = H_0^1(\Omega), \quad \|\cdot\|_{\mathbb{V}(\Omega)} = \|\cdot\|_{\mathbb{W}(\Omega)} = \|\nabla \cdot\|_\Omega,$$

$$\mathcal{B}(v, w) = \int_\Omega \nabla v \cdot \nabla w, \quad v, w \in \mathbb{V},$$

and suppose $f \in L_2(\Omega)$. Thanks to the Poincaré inequality $\|\cdot\|_\Omega \preccurlyeq \|\nabla \cdot\|_\Omega$ and the coercivity of the bilinear form, $f \in \mathbb{W}^*$ holds and (2.2) is valid with $\alpha = C_{\mathcal{B}} = 1$.

Refinement framework. Let $\mathcal{G}_0$ be a conforming triangulation of Ω into d-simplices. We suppose that $\mathcal{G}_0$ allows to apply iterative (see [8]) or recursive (see [19]) bisection and let $\mathbb{G}$ be the resulting class of all refinements of $\mathcal{G}_0$. Then (2.5) is fulfilled with $n(E) = 2$, and $q_1 = q_2 = \frac{1}{2}$; the hidden constants in (2.5c) depend on $\mathcal{G}_0$. Moreover, $\mathbb{G}$ is a shape-regular family of triangulations.

Finite element spaces and approximate solution. We choose linear finite elements,

$$\mathbb{V}(\mathcal{G}) := \mathbb{W}(\mathcal{G}) := \{v \in C^0(\bar{\Omega}) \mid \forall E \in \mathcal{G} \; v_{|E} \text{ is affine}\} \cap H_0^1(\Omega), \qquad \mathcal{G} \in \mathbb{G},$$

which are included in $\mathbb{V}$. Since coercivity and continuity of $\mathcal{B}$ are inherited by a restriction on a subspace and spaces of piecewise polynomials are nested on nested grids, (2.6) is valid with $\beta = 1$. Moreover, we suppose that SOLVE outputs the Galerkin approximation given by (2.7).

Estimators. Given any Galerkin solution $u_\mathcal{G}$ on a grid $\mathcal{G}$, we consider the following a posteriori error estimators, which will be admissible choices for $\{\mathcal{E}_\mathcal{G}(E)\}_{E \in \mathcal{G}}$:

- the *standard residual error estimator* $\{\mathcal{E}_{\mathcal{G},1}(E)\}_{E \in \mathcal{G}}$ given by

$$\mathcal{E}_{\mathcal{G},1}(E) := \|h_\mathcal{G}^{1/2} j_\mathcal{G}\|_{\partial E \setminus \partial \Omega} + \|h_\mathcal{G} f\|_E,$$

 where $j_\mathcal{G}$ denotes the jump (or interelement) residual;
- a *hierarchical error estimator* $\{\mathcal{E}_{\mathcal{G},2}(E)\}_{E \in \mathcal{G}}$ given by

$$\mathcal{E}_{\mathcal{G},2}(E) := \sum_{i=0}^{d+1} |\langle \mathcal{R}, \phi_{E,i} \rangle| + \|h_\mathcal{G}(f - \bar{f}_E)\|_E,$$

 where $\phi_{E,i}$, $i = 0, \ldots, d+1$ are hat functions over a grid that is finer than $\mathcal{G}$ (for the precise definition, see [31, §2.3]) and $\bar{f}_E$ equals the mean value of f on E.
- in the case $d = 2$ only, an error estimator $\{\mathcal{E}_{\mathcal{G},3}\}_{E \in \mathcal{G}}$ that is computed *with the help of local problems*:

$$\mathcal{E}_{\mathcal{G},3}(E) := \sum_{z \in \mathcal{N} \cap E} \|\xi_z\|_{\mathbb{V}(\omega_z)} + \|h_\mathcal{G}(f - \bar{f}_z)\|_{\omega_z},$$

 where $\mathcal{N}$ denotes the set of all vertices of $\mathcal{G}$, ω_z is the union of all elements containing z, the real number $\bar{f}_z$ stands for the mean value of f on ω_z, and ξ_z solves

$$\xi_z \in \mathbb{V}_z : \qquad \forall \varphi \in \mathbb{V}_z \quad \int_{\omega_z} \nabla \xi_z \cdot \nabla \varphi \phi_z = \langle \mathcal{R}, \varphi \phi_z \rangle$$

 where ϕ_z is the hat function associated with node z and $\mathbb{V}_z$ is a certain finite element space over ω_z; see [22, Definition 2.5] for the precise definition;
- the ZZ estimator $\{\mathcal{E}_{\mathcal{G},4}(E)\}_{E \in \mathcal{G}}$, which is *based upon gradient recovery*:

$$\mathcal{E}_{\mathcal{G},4}(E) := \|\nabla u_\mathcal{G} - G u_\mathcal{G}\|_E + \sum_{z \in \mathcal{N} \cap E} \|h_\mathcal{G}(f - f_z)\|_{\omega_z},$$

where Gu_k is the orthogonal projection of ∇u_k onto the vectorial linear finite element space with respect to the $\mathcal{G}$-lumped $L_2(\Omega)$-scalar product; see [33, §1.5]. The estimators $\mathcal{E}_{\mathcal{G},j}$, $j = 1, \ldots, 4$, verify (2.9) with

$$n = \begin{cases} 3 & \text{if } d = 2, \\ 6 & \text{if } d = 3 \text{ and } j \neq 3, \end{cases} \tag{3.2}$$

$$\mathrm{osc}_{\mathcal{G}}(E) = \begin{cases} \|h_{\mathcal{G}}(f - \bar{f}_E)\|_{\omega_{\mathcal{G}}(E)} & \text{if } j \in \{1,2\}, \\ \displaystyle\sum_{z \in \mathcal{N} \cap E} \|h_{\mathcal{G}}(f - \bar{f}_z)\|_{\omega_z} & \text{if } j \in \{3,4\}, \end{cases}$$

$$m(s) = s^{1/d}, \ s \in [0,\infty), \quad \mathbb{D} = L_2(\Omega), \quad D = f.$$

Indeed, for $\mathcal{E}_{\mathcal{G},1}$ see [21, Lemma 3.1 and Lemma 4.2], for the upper bound of $\mathcal{E}_{\mathcal{G},2}$ see [31, Lemma 3.1] while its lower bound readily follows from the definition, for $\mathcal{E}_{\mathcal{G},3}$ see [22, Theorem 3.6 and Lemma 5.2], for the upper bound of $\mathcal{E}_{\mathcal{G},4}$ see [14, Theorem 4.1], and for its lower bound use first a local version of the last inequality in [33, Proposition 1.21] and then [21, Lemma 4.2].

Marking strategy. We suppose that the output $\mathcal{M}$ of MARK satisfies

$$\mathcal{E}_{\mathcal{G}}(\mathcal{M}) \geq \theta \mathcal{E}_{\mathcal{G}}$$

where $\theta \in (0,1]$ is a fixed parameter. This is Dörfler's original marking (without the feature of collecting the biggest indicators), see [17], and (2.13) is valid with $g(x) := \theta^{-1}\sqrt{1 - \theta^2}\, x$, $x \in [0,\infty)$.

Refinement rule. We only require that each marked triangle ($d = 2$) or tetrahedron ($d = 3$) is bisected at least once and so (2.14) is valid.

Then Theorem 2.1 ensures that both error and estimator converge to 0. In the cases $\mathcal{E}_{\mathcal{G}} = \mathcal{E}_{\mathcal{G},1}, \mathcal{E}_{\mathcal{G},2}, \mathcal{E}_{\mathcal{G},3}$, this result corresponds, respectively, to the convergence results of [17] and [21], of the linear case in [31], and of [22]. Here however, interior nodes for marked elements are not necessarily created and there is neither a special marking nor a pre-adaptation for (data) oscillation if $\mathcal{E}_{\mathcal{G}} = \mathcal{E}_{\mathcal{G},1}$. The case $\mathcal{E}_{\mathcal{G}} = \mathcal{E}_{\mathcal{G},4}$ seems to be the first convergence result for the ZZ estimator.

3.2. An H(div)-elliptic problem and several marking strategies.
In this example we apply (2.3) with Raviart-Thomas and Brezzi-Douglas-Marini elements to an H(div)-elliptic problem, employing the most commonly used marking strategies.

Problem. Given a domain with outer normal $\boldsymbol{n}$ and a right hand side $\boldsymbol{F}$, we consider the following problem: find a vector field $\boldsymbol{u}$ such that

$$\boldsymbol{u} - \nabla \operatorname{div} \boldsymbol{u} = \boldsymbol{F} \quad \text{in } \Omega, \qquad \boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega. \tag{3.3}$$

Its weak form is (2.1) with

$$\mathbb{V} = \mathbb{W} = \{\boldsymbol{v} \in L_2(\Omega; \mathbb{R}^d) \mid \operatorname{div} \boldsymbol{v} \in L_2(\Omega), \ \boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega\}, \tag{3.4}$$

$$\|\boldsymbol{v}\|_{\mathbb{V}} = \mathcal{B}(\boldsymbol{v}, \boldsymbol{v})^{1/2} = \|\boldsymbol{v}\|_{H(\operatorname{div};\Omega)}, \tag{3.5}$$

and

$$\mathcal{B}(\boldsymbol{v}, \boldsymbol{w}) = \int_{\Omega} \boldsymbol{v} \cdot \boldsymbol{w} + \operatorname{div} \boldsymbol{v} \operatorname{div} \boldsymbol{w} \qquad \boldsymbol{v}, \boldsymbol{w} \in \mathbb{V}.$$

In what follows, we suppose that Ω is a bounded, polyhedral Lipschitz domain in $\mathbb{R}^d$, $d \in \{2,3\}$, which has to be topologically equivalent to a ball in the case $d = 3$. Moreover, we suppose that the right hand side $\boldsymbol{F}$ is given by

$$\boldsymbol{F}(\boldsymbol{v}) := \int_{\Omega} (-f_1) \operatorname{div} \boldsymbol{v} + \boldsymbol{f}_2 \cdot \boldsymbol{v}\, dx,$$

where $f_1 \in H^1(\Omega)$ and $\boldsymbol{f}_2 \in H(\mathrm{rot};\Omega)$ for $d = 2$ and $\boldsymbol{f}_2 \in H(\mathbf{curl};\Omega)$ for $d = 3$. In particular, we have $\boldsymbol{F} \in \mathbb{W}^*$ and (2.2) holds with $\alpha = C_{\mathcal{B}} = 1$.

Refinement framework, finite element spaces, and approximate solution. Let $\mathbb{G}$ be the same class as in §3.1 and denote by $\mathbb{P}_\ell$ the set of all polynomials of degree $\leq \ell$. For any $\mathcal{G} \in \mathbb{G}$, let $\mathbb{V}(\mathcal{G})$ be either the *Raviart-Thomas elements* of order $\ell \geq 0$, i.e.

$$\mathbb{V}(\mathcal{G}) = \mathbb{RT}_\ell(\mathcal{G}) := \left\{ \boldsymbol{v} \in \mathbb{V} \mid \forall E \in \mathcal{G} \quad \boldsymbol{v}_{|E} \in \left(\mathbb{P}_\ell(E;\mathbb{R}^d) + \boldsymbol{x}\,\mathbb{P}_\ell(E) \right) \right\},$$

or the *Brezzi-Douglas-Marini elements* of order $\ell \geq 1$, i.e.

$$\mathbb{V}(\mathcal{G}) = \mathbb{BDM}_\ell(\mathcal{G}) := \left\{ \boldsymbol{v} \in \mathbb{V} \mid \forall E \in \mathcal{G} \quad \boldsymbol{v}_{|E} \in \mathbb{P}_\ell(E;\mathbb{R}^d) \right\}.$$

In both cases (2.6) is valid with $\beta = 1$. Again, we suppose that SOLVE outputs the Galerkin approximation $\boldsymbol{u}_\mathcal{G}$ given by (2.7).

Estimator. To define the estimator $\{\mathcal{E}_\mathcal{G}(E)\}_{E \in \mathcal{G}}$ on a given grid $\mathcal{G} \in \mathbb{G}$, we denote by $\boldsymbol{u}_\mathcal{G}$ the Galerkin solution, by $\mathcal{S}$ the interior (or interelement) sides of $\mathcal{G}$ and define element- or side-wise the residuals $\boldsymbol{r}_\mathcal{G}$, $R_\mathcal{G}$, $J_\mathcal{G}$ by

$$\forall E \in \mathcal{G} \qquad \boldsymbol{r}_{\mathcal{G}|E} := \nabla f_1 + \boldsymbol{f}_2 - \boldsymbol{u}_\mathcal{G} + \nabla \operatorname{div} \boldsymbol{u}_\mathcal{G},$$

$$\forall E \in \mathcal{G} \qquad R_{\mathcal{G}|E} := \begin{cases} \operatorname{rot}(\boldsymbol{f}_2 - \boldsymbol{u}_\mathcal{G}), & \text{if } d = 2, \\ \mathbf{curl}(\boldsymbol{f}_2 - \boldsymbol{u}_\mathcal{G}), & \text{if } d = 3, \end{cases}$$

$$\forall S \in \mathcal{S} \qquad J_{\mathcal{G}|S} := \begin{cases} [\![\boldsymbol{u}_\mathcal{G} \cdot \boldsymbol{t}]\!]_S, & \text{if } d = 2, \\ [\![\boldsymbol{u}_\mathcal{G} \times \boldsymbol{n}]\!]_S, & \text{if } d = 3, \end{cases}$$

where $\boldsymbol{t}$ denotes a tangent vector to S in 2d and $\boldsymbol{n}$ a normal vector to S in 3d. For both choices of the finite element spaces, the estimator $\{\mathcal{E}_\mathcal{G}(E)\}_{E \in \mathcal{G}}$ is then given by

$$\mathcal{E}_\mathcal{G}^2(E) := \|h_\mathcal{G}\boldsymbol{r}_\mathcal{G}\|_E^2 + \|h_\mathcal{G}R_\mathcal{G}\|_E^2 + \|h_\mathcal{G}^{1/2}J_\mathcal{G}\|_{\partial E \cap \Omega}^2.$$

The upper bound (2.9a) is shown in [13, §§5.3 and 7.3], while the proof of the discrete local lower bound in [13, §§5.4 and 7.4] in particular ensures (2.9b) with n as in (3.2) and

$$\operatorname{osc}_\mathcal{G}^2(E) = \|h_\mathcal{G}(\boldsymbol{r}_\mathcal{G} - \bar{\boldsymbol{r}}_\mathcal{G})\|_{\omega_\mathcal{G}(E)}^2 + \|h_\mathcal{G}(R_\mathcal{G} - \bar{R}_\mathcal{G})\|_{\omega_\mathcal{G}(E)}^2 + \|h_\mathcal{G}^{1/2}(J_\mathcal{G} - \bar{J}_\mathcal{G})\|_{\partial E \cap \Omega}^2,$$

where $\bar{\boldsymbol{r}}_\mathcal{G}$, $\bar{R}_\mathcal{G}$, and $\bar{J}_\mathcal{G}$ are suitable local L_2-projections. The oscillation bound (2.9c) holds with

$$m(s) = s^{1/d}, \quad \mathbb{D} = L_2(\Omega), \quad \text{and} \quad D^2 = \begin{cases} |\nabla f_1 + \boldsymbol{f}_2|^2 + |\operatorname{rot}\boldsymbol{f}_2|^2, & \text{if } d = 2, \\ |\nabla f_1 + \boldsymbol{f}_2|^2 + |\mathbf{curl}\,\boldsymbol{f}_2|^2, & \text{if } d = 3. \end{cases}$$

Marking strategies. We consider the most commonly used marking strategies, apart from Dörfler's one which has been already used in §3.1:

- *Maximum strategy:* Given a parameter $0 \leq \theta \leq 1$, we let

$$\mathcal{E}_{\mathcal{G},\max} = \max\{\mathcal{E}_\mathcal{G}(E) \mid E \in \mathcal{G}\} \quad \text{and} \quad \mathcal{M} = \{E \in \mathcal{G} \mid \mathcal{E}_\mathcal{G}(E) \geq \theta\mathcal{E}_{\mathcal{G},\max}\}$$

This strategy was first proposed in Babuška/Rheinboldt [6]. It selects element whose indicators are bigger than a threshold given by the maximal indicator and the parameter θ. The use of the maximal indicator may be motivated by a greedy argument. For larger values of θ, the strategy becomes more selective: the value $\theta = 1$ corresponds to selecting only elements holding the maximal indicator and $\theta = 0$ to selecting all elements for refinement.

- *Modified equidistribution strategy:* Given a parameter $0 \leq \theta \leq 1$, we let

$$\mathcal{M} = \left\{ E \in \mathcal{G} \mid \mathcal{E}_\mathcal{G}(E) \geq \theta \, \mathcal{E}_\mathcal{G}/\sqrt{\#\mathcal{G}} \right\}.$$

Here the threshold for selecting is determined by the mean value $\mathcal{E}_\mathcal{G}/\sqrt{\#\mathcal{G}}$ of the indicators and the parameter θ. The parameter θ again controls the selectivity, while the use of the mean value of the indicators is motivated by the following heuristic characterization in [6]: given a number of elements, a grid has minimal error iff the local error is equidistributed among all elements. The original equidistribution strategy, which only aims at reaching a prescribed tolerance and does not have the form (2.12), is discussed in §5.

It is obvious that both strategies select at least one element holding the maximal indicator. Such strategies fulfill (2.13) with $g(t) := t$, $t \in [0, \infty)$.

Refinement rule. As before, we only require that each marked triangle $(d = 2)$ or tetrahedron $(d = 3)$ is bisected at least once and so (2.14) is valid.

Here Theorem 2.1 yields, apart from non-essential simplifications in the assumptions on the data, a generalization of the convergence results in [13]: the minimal refinement rule and other marking strategies are allowed.

3.3. **A saddle point problem for a non-symmetric operator.** In this section we apply (2.3) to a saddle point problem for a non-symmetric elliptic operator, without an assumption on the fineness of the initial grid.

Problem. We consider the stationary Oseen equations: given a bounded, polyhedral Lipschitz domain Ω in $\mathbb{R}^d$, $d \in \{2, 3\}$, a divergence free vector field $\boldsymbol{b} \in L_\infty(\Omega, \mathbb{R}^d)$, and an external forcing $\boldsymbol{f} \in L_2(\Omega; \mathbb{R}^d)$, find a velocity field $\boldsymbol{u}$ and a pressure p such that

$$-\Delta \boldsymbol{u} + [\nabla \boldsymbol{u}]\boldsymbol{b} + \nabla p = \boldsymbol{f} \quad \text{in } \Omega, \qquad \operatorname{div} \boldsymbol{u} = 0 \quad \text{in } \Omega, \qquad \boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega.$$

Its weak form is (2.1) with

$$\mathbb{V} = \mathbb{W} = H_0^1(\Omega; \mathbb{R}^d) \times \{q \in L_2(\Omega) \mid \textstyle\int_\Omega q \, dx = 0\},$$

$$\|v\|_{\mathbb{V}(\Omega)}^2 = \|\nabla \boldsymbol{v}\|_{H^1(\Omega;\mathbb{R}^d)}^2 + \|q\|_\Omega^2, \quad v = [\boldsymbol{v}, q] \in \mathbb{V},$$

and the bilinear form

$$\mathcal{B}(v, w) = \int_\Omega \nabla \boldsymbol{v} : \nabla \boldsymbol{w} + [\nabla \boldsymbol{v}]\boldsymbol{b} \cdot \boldsymbol{w} - q \operatorname{div} \boldsymbol{w} - \operatorname{div} \boldsymbol{v} \, r, \quad v = [\boldsymbol{v}, q], \, w = [\boldsymbol{w}, r] \in \mathbb{V}.$$

One readily verifies (2.2b), (2.2c), $(\boldsymbol{f}, 0) \in \mathbb{V}^*$, and that the bilinear form $\mathcal{B}$ is continuous with $C_\mathcal{B} \leq 3 + \|\boldsymbol{b}\|_{L_\infty(\Omega;\mathbb{R}^d)}$ in (2.2d). Moreover, $\mathcal{B}$ satisfies the inf-sup condition (2.2a) thanks to $\operatorname{div} \boldsymbol{b} = 0$, the Poincaré inequality on $H_0^1(\Omega; \mathbb{R}^d)$, and the solvability of the divergence equation [18, Theorem III.3.1].

Refinement framework, finite element spaces, and approximate solution. Again, we use the class $\mathbb{G}$ of the first example in §3.1. For any grid $\mathcal{G} \in \mathbb{G}$, let $\mathbb{V}(\mathcal{G})$ be the space of *Taylor-Hood elements* of order $\ell \geq 2$ over $\mathcal{G}$, i.e. $[\boldsymbol{v}, q] \in \mathbb{V}(\mathcal{G})$ means that $\boldsymbol{v}$ is a piecewise polynomial of degree $\leq \ell$ in $C(\bar{\Omega}; \mathbb{R}^d)$ and q a piecewise polynomial of degree $\leq \ell - 1$ in $C(\bar{\Omega})$. Then there holds (2.6); for the inf-sup condition (2.6c) see [24]. As before, we suppose that the output of SOLVE is the Galerkin solution given by (2.7).

Estimator. Let us write $u_\mathcal{G} = [\boldsymbol{u}_\mathcal{G}, p_\mathcal{G}] \in \mathbb{V}(\mathcal{G})$ for the Galerkin solution on any grid $\mathcal{G}$ and define the residuals $\boldsymbol{r}_\mathcal{G}$, $\boldsymbol{j}_\mathcal{G}$ by

$$\forall E \in \mathcal{G} \qquad \boldsymbol{r}_{\mathcal{G}|E} := -\Delta \boldsymbol{u}_\mathcal{G} + [\nabla \boldsymbol{u}_\mathcal{G}]\boldsymbol{b} + \nabla p_\mathcal{G} - \boldsymbol{f},$$

$$\forall S \in \mathcal{S} \qquad \boldsymbol{j}_{\mathcal{G}|S} := [\![\nabla \boldsymbol{u}_\mathcal{G}\boldsymbol{n}]\!]_S,$$

where $\mathcal{S}$ is the set of interior (or interelement) sides of $\mathcal{G}$. The estimator $\{\mathcal{E}_{\mathcal{G}}(E)\}_{E\in\mathcal{G}}$ is then given by

$$\mathcal{E}_{\mathcal{G}}^2(E) := \|h_{\mathcal{G}}\boldsymbol{r}_{\mathcal{G}}\|_E^2 + \|h_{\mathcal{G}}^{1/2}\boldsymbol{j}_{\mathcal{G}}\|_{\partial E\cap\Omega}^2.$$

Using the techniques for the Stokes problem in [32], the upper bound

$$\|u_{\mathcal{G}} - u\|_{\mathbb{V}(\Omega)} \preccurlyeq \mathcal{E}_{\mathcal{G}}(\mathcal{G}) + \|\operatorname{div}\boldsymbol{u}_{\mathcal{G}}\|_\Omega$$

follows also for the Oseen equations. The computed velocity $\boldsymbol{u}_{\mathcal{G}}$ is discretely divergence free, hence [9, Proposition 5.4] implies

$$\|\operatorname{div}\boldsymbol{u}_{\mathcal{G}}\|_\Omega^2 \preccurlyeq \sum_{E\in\mathcal{G}} \|h_{\mathcal{G}}^{1/2}\,[\![\operatorname{div}\boldsymbol{u}_{\mathcal{G}}]\!]\,\|_{\partial E\cap\Omega}^2 \leq \sum_{E\in\mathcal{G}} \|h_{\mathcal{G}}^{1/2}\boldsymbol{j}_{\mathcal{G}}\|_{\partial E\cap\Omega}^2$$

and the upper bound (2.9a) emerges. Similar to [9, Lemma 5.2], one can prove the discrete lower bound (2.9b) with n as in (3.2) and

$$\operatorname{osc}_{\mathcal{G}}^2(E) = \|h_{\mathcal{G}}(\boldsymbol{r} - \bar{\boldsymbol{r}}_{\mathcal{G}})\|_{\omega_{\mathcal{G}}(E)}^2,$$

where $\bar{\boldsymbol{r}}_{\mathcal{G}}$ is the orthogonal L_2-projection of $\boldsymbol{r}$ onto the piecewise polynomials of degree $\ell - 1$ over $\mathcal{G}$. Hence, (2.9c) holds with $m(s) = s^{1/d}$, $\mathbb{D} = L_2(\Omega; \mathbb{R}^d)$, $D = \boldsymbol{f}$, and the hidden constant depends also on $\|\boldsymbol{b}\|_{L_\infty(\Omega;\mathbb{R}^d)}$.

Marking strategy and refinement rule. Let us take Dörfler's strategy, the maximum strategy or the modified equidistribution strategy (see §3.2) and only require that each marked simplex is bisected at least once; consequently, (2.13) and (2.14) are valid.

To our best knowledge, Theorem 2.1 applied to this example is the first convergence result of (1.1) for a saddle point problem. The alternative approach in [9] tackles the saddle point structure by an approximate infinite-dimensional Uzawa algorithm, which uses (1.1) for the approximate solution of the arising linear elliptic problems.

3.4. General elliptic equations using rectangular grids.

In this example we apply (2.3) to the Dirichlet problem for a general elliptic equation with variable coefficients, using rectangular/cuboidal grids with hanging nodes.

Problem. Given a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, meshed by $\mathcal{G}_0$ (see below), we consider the following problem: find $u : \Omega \to \mathbb{R}$ such that

$$-\operatorname{div}(\boldsymbol{A}\nabla u) + \boldsymbol{b}\cdot\nabla u + cu = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega,$$

where $\boldsymbol{A} \colon \Omega \to \mathbb{R}^{d\times d}$ is piecewise Lipschitz over $\mathcal{G}_0$, for $x \in \Omega$ the matrix $A(x)$ is symmetric and positive definite with smallest eigenvalue uniformly bounded away from 0, the vector field $\boldsymbol{b} \in L_\infty(\Omega; \mathbb{R}^d)$ is divergence free, $c \in L_\infty(\Omega)$ is nonnegative, and $f \in L_2(\Omega)$. Its weak form is (2.1) with

$$\mathbb{V} = \mathbb{W} = H_0^1(\Omega) \quad \|\cdot\|_{\mathbb{V}(\Omega)} = \|\cdot\|_{\mathbb{W}(\Omega)} = \|\cdot\|_{H^1(\Omega)},$$

$$\mathcal{B}(v,w) = \int_\Omega \boldsymbol{A}\nabla v\cdot\nabla w + \boldsymbol{b}\cdot\nabla v\,w + c\,v\,w, \quad v,w \in \mathbb{V}.$$

In view of the assumptions on $\boldsymbol{A}$, $\boldsymbol{b}$, c and the Poincaré inequality $\|\cdot\|_\Omega \preccurlyeq \|\nabla\cdot\|_\Omega$, the bilinear form is coercive on $H_0^1(\Omega)$ and (2.2) is valid with $\alpha, C_{\mathcal{B}}$ depending on $\boldsymbol{A}, \boldsymbol{b}, c$.

Refinement framework. Suppose that the initial grid $\mathcal{G}_0$ is a finite subdivision of the domain Ω into rectangular ($d = 2$) or cuboidal ($d = 3$) elements such that each element side is a union of sides from other elements. Moreover, suppose that the master forest $\mathcal{F}$ is generated by the following subdivision rule: any element is split into 2^d subelements of the same size which are equivalent to the parent element. Then (2.5a) and (2.5b) hold with $n(E) = 2^d$ and $q_1 = q_2 = 2^{-d}$.

To define the class $\mathbb{G}$, let us introduce some notation associated with a subdivision $\mathcal{S}$ of Ω given by $\mathcal{F}$. We denote by $\mathcal{V}(\mathcal{S})$ the set of vertices of $\mathcal{S}$. A vertex $z \in \mathcal{V}(\mathcal{S})$ may or may not have the following property: if z is contained in an element $E \in \mathcal{S}$, then z is a vertex of E; we collect these nodes in $\mathcal{V}^*(\mathcal{S})$. The so-called hanging nodes in $\mathcal{V}(\mathcal{S}) \setminus \mathcal{V}^*(\mathcal{S})$ do not correspond to degrees of freedom of $\mathbb{RE}_1(\mathcal{S})$, the space of the continuous functions that are piecewise in $\otimes_{i=1}^d \mathbb{P}_1$. Thus, the functions $(\varphi_{\mathcal{S},z})_{z \in \mathcal{V}^*(\mathcal{S})}$ satisfying $\varphi_{\mathcal{S},z} \in \mathbb{RE}_1(\mathcal{S})$ and $\varphi_{\mathcal{S},z}(y) = \delta_{yz}$ for all $y \in \mathcal{V}^*(\mathcal{S})$ form a Lagrange basis of $\mathbb{RE}_1(\mathcal{S})$. Following Babuška/Miller [4], we call $\mathcal{S}$ a K-subdivision iff the 'spread' of the basis functions is limited in the following way:

$$\forall z \in \mathcal{V}^*(\mathcal{S}) \quad |\operatorname{supp} \varphi_{\mathcal{S},z}| \leq K \min \left\{ |E| \mid E \in \mathcal{S} \text{ and } E \subset \operatorname{supp} \varphi_{\mathcal{S},z} \right\}. \qquad (3.6)$$

Now let $K > 0$ be such that the initial grid $\mathcal{G}_0$ is a K-subdivision and define $\mathbb{G}$ as the set of those K-subdivisions that are given by $\mathcal{F}$. Then (2.5c) follows from the d-dimensional versions of [4, Corollary 1.3.4 and Lemma 1.4.1]. Moreover, there exists $k \in \mathbb{N}_0$ such that any grid $\mathcal{G} \in \mathbb{G}$ is k-irregular in that any side of an element is the union of at most $k + 1$ other element sides. Finally, $\mathbb{G}$ is a shape-regular family of grids.

One practical way to ensure (3.6) with K depending on $\mathcal{G}_0$ is to restrict to 1-irregular grids; 2-irregular grids do not imply (3.6) in general. The 1-irregular grids correspond to so-called balanced quadtrees ($d = 2$) and octrees ($d = 3$), which can be obtained from non-balanced ones by an algorithm proposed in [29].

Finite element spaces and approximate solution. For a fixed degree $\ell \in \mathbb{N}$, we denote by $\mathbb{Q}_\ell = \otimes_{i=1}^d \mathbb{P}_\ell$ the space of tensor products of 1-dimensional polynomials of degree $\leq \ell$ and define the finite element spaces by

$$\mathbb{V}(\mathcal{G}) = \mathbb{W}(\mathcal{G}) = \left\{ v \in C^0(\bar{\Omega}) \mid \forall E \in \mathcal{G} \; v_{|E} \in \mathbb{Q}_\ell(E) \right\} \cap H_0^1(\Omega)$$

for any $\mathcal{G} \in \mathbb{G}$. This definition and the coercivity of $\mathcal{B}$ readily imply (2.6). As before, we suppose that SOLVE outputs the Galerkin approximation given by (2.7).

Estimator. Let $u_\mathcal{G}$ denote the Galerkin solution on the grid $\mathcal{G} \in \mathbb{G}$ and define the interior residual $r_\mathcal{G}$ and the jump residual $j_\mathcal{G}$ by

$$\forall E \in \mathcal{G} \qquad r_{\mathcal{G}|E} := -\operatorname{div}(A \nabla u_\mathcal{G}) + \boldsymbol{b} \cdot \nabla u_\mathcal{G} + c u_\mathcal{G} - f,$$

$$\forall S \in \mathcal{S} \qquad j_{\mathcal{G}|S} := [\![A \nabla u_\mathcal{G} \cdot \boldsymbol{n}]\!]_S \,,$$

where $\mathcal{S}$ denotes the set of interior sides of $\mathcal{G}$. The standard residual estimator $\{\mathcal{E}_\mathcal{G}(E)\}_{E \in \mathcal{G}}$ is then given by

$$\mathcal{E}_\mathcal{G}^2(E) := \|h_\mathcal{G} r_\mathcal{G}\|_E^2 + \|h_\mathcal{G}^{1/2} j_\mathcal{G}\|_{\partial E \cap \Omega}^2.$$

In view of (3.6), [4, Corollary 1.3.4 and Lemma 1.4.1], the upper upper bound (2.9a) can be established with the help of a Clément-type interpolation operator into $\mathbb{RE}_1(\mathcal{S}) \cap H_0^1(\Omega)$ that averages on the supports of the basis functions; see also [4, Theorem 2.4.2] for an alternative approach.

Before embarking on the sketch of the proof of the lower bound, let us define $\operatorname{osc}_\mathcal{G}(E)$ and verify (2.9c). Note that taking a partial derivative is not an operator of the form $\mathbb{Q}_\ell \to \mathbb{Q}_{\ell-1}$ (while it is of the form $\mathbb{P}_\ell \to \mathbb{P}_{\ell-1}$). We therefore project the residuals in the 'common' space $\mathbb{Q}_\ell$ and define the oscillation by

$$\operatorname{osc}_\mathcal{G}(E) = \|h_\mathcal{G}(r_\mathcal{G} - \bar{r}_\mathcal{G})\|_{\omega_\mathcal{G}(E)}^2 + \|h_\mathcal{G}^{1/2}(j_\mathcal{G} - \bar{\jmath}_\mathcal{G})\|_{\partial E \cap \Omega}^2,$$

where $\bar{r}_{\mathcal{G}|E}$ is the $L_2(E)$-projection of $r_\mathcal{G}$ and $\bar{\jmath}_{\mathcal{G}|S}$ is the $L_2(S)$-projection of $j_\mathcal{G}$ onto $\mathbb{Q}_\ell$. Denoting by $\boldsymbol{A}_E$ the mean value of $\boldsymbol{A}$ in E, the identity $\boldsymbol{A}_E D^2 u_\mathcal{G} = \overline{\boldsymbol{A}_E D^2 u_\mathcal{G}}$

and an inverse inequality imply that

$$\begin{aligned}
\|r_{\mathcal{G}} - \bar{r}_{\mathcal{G}}\|_E &\le \|\operatorname{div} \boldsymbol{A} \cdot \nabla u_{\mathcal{G}}\|_E + \|\boldsymbol{A} - \boldsymbol{A}_E\|_{L_\infty(E)}\|D^2 u_{\mathcal{G}}\|_E \\
&\quad + \|\boldsymbol{b} \cdot \nabla u_{\mathcal{G}}\|_E + \|c u_{\mathcal{G}}\|_E + \|f\|_E \\
&\preccurlyeq \|\operatorname{div} \boldsymbol{A} \cdot \nabla u_{\mathcal{G}}\|_E + h_{\mathcal{G}|E}\|\boldsymbol{A}\|_{W^1_\infty(E)}\frac{1}{h_{\mathcal{G}|E}}\|\nabla u_{\mathcal{G}}\|_E \\
&\quad + \|\boldsymbol{b} \cdot \nabla u_{\mathcal{G}}\|_E + \|c u_{\mathcal{G}}\|_E + \|f\|_E \\
&\preccurlyeq \left(\|\boldsymbol{A}\|_{W^1_\infty(E)} + \|\boldsymbol{b}\|_{L_\infty} + \|c\|_{L_\infty}\right)\|u_{\mathcal{G}}\|_{H^1(E)} + \|f\|_E.
\end{aligned}$$

After an application of a scaled trace theorem, an analogous argument gives

$$\|j_{\mathcal{G}} - \bar{\jmath}_{\mathcal{G}}\|_S \preccurlyeq h_{\mathcal{G}|S}^{1/2}\max\{\|\boldsymbol{A}\|_{W^1_\infty(E_1)}, \|\boldsymbol{A}\|_{W^1_\infty(E_2)}\}\|u_{\mathcal{G}}\|_{H^1(E_1 \cup E_2)},$$

where S is a side that is contained in the two elements E_1, $E_2 \in \mathcal{G}$. Since the $\partial E \cap \Omega$ can be covered by such sides, the last two estimates imply (2.9c) with

$$m(s) = s^{1/d}, \quad \mathbb{D} = L_2(\Omega), \quad \text{and} \quad D = f.$$

We finally sketch the proof of the discrete local lower bound (2.9b) with $n = k+2$. The requirement of k subdivision levels ensures that, for two elements sharing a side, the coarse element is subdivided at least to the level of the fine element. The additional two subdivision levels allow the application of the following lemma, to which we resort instead of multiplying $\bar{r}_{\mathcal{G}}$ and $\bar{\jmath}_{\mathcal{G}}$ with a fine base function, which provides a test function with local support but is not necessarily piecewise in $\mathbb{Q}_\ell$.

Lemma 3.1. *For $m \in \mathbb{N}$ let $Q \subset \mathbb{R}^m$ be a cube and consider it decomposed into 2^{2m} cubes $Q_1, Q_2, \ldots, Q_{2^{2m}}$, obtained by two regular refinements. Let Π be the $L_2(Q)$-orthogonal projection onto the space*

$$\left\{v \in C(\bar{Q}) : v_{|Q_i} \in \mathbb{Q}_\ell(Q_i), i = 1, \ldots, 2^{2m}\right\} \cap H_0^1(Q).$$

Then $\|R\|_Q^2 \preccurlyeq \int_Q R\,\Pi R$, for all $R \in \mathbb{Q}_\ell(Q)$.

Proof. The case $m = 1$ can be established in the same manner as [23, Lemma 3.1]. The general case follows by induction exploiting the tensor product structure of the spaces $\mathbb{Q}_\ell$. $\square$

Using Lemma 3.1 to replace $\bar{r}_{\mathcal{G}}$ and $\bar{\jmath}_{\mathcal{G}}$ by discrete test functions with local support and deriving properties of the replacements in the spirit of [33, (3.3)-(3.5)], one establishes (2.9b) with $n = k + 2$ along the lines of the argument by Verfürth; see, e.g. [33, §1.2].

Marking strategy and refinement rule. We choose any marking strategy from §3.2 and require that any marked element is subdivided at least once. Thus, (2.13) and (2.14) are satisfied.

To our best knowledge, Theorem 2.1 applied to this example is the first convergence proof of (1.1) for general elliptic equations with rectangular elements. Convergence holds without marking due to oscillation and starting from any coarse initial mesh, even in the convection-dominated case.

In [23] we present two further examples within a more general framework for the error notion. For both examples, (2.3) approximates the solution of the Poisson problem and uses bisection. However, discretization and error notion differ: the fist example employs Lagrange elements and an a posteriori error estimator for the mean square error, while the second one employs Raviart-Thomas or Brezzi-Douglas-Marini elements in a mixed formulation and an a posteriori error estimator for the mean square error of the flux.

4. Proof of Abstract Convergence

In this section we prove the abstract convergence result. The proof of Theorem 2.1 is divided into two main steps. First, we prepare the proper convergence proof by defining and studying certain limits of sequences that are generated by iteration (2.3). Secondly, we establish convergence of the approximate solutions to the exact one by showing that the error estimator vanishes in the limit.

For the simplicity of notation, we replace the subscript (or also the argument) $\mathcal{G}_k$ by the iteration counter k in what follows.

4.1. **Three limits.** Iteration (2.3) generates a sequence of grids $\{\mathcal{G}_k\}_k$. This sequence is accompanied in particular by the following ones:

$$\{\mathcal{F}_k\}_k, \qquad \{u_k\}_k, \qquad \{h_k\}_k, \tag{4.1}$$

where $\mathcal{F}_k$ indicates the finite forest associated with $\mathcal{G}_k$, u_k is the approximate solution on $\mathcal{G}_k$ given by (2.7), and h_k denotes the meshsize function of $\mathcal{G}_k$ as defined in (3.1). We now study limits of these sequences.

We may consider the forest

$$\mathcal{F}_\infty := \bigcup_{k \in \mathbb{N}_0} \mathcal{F}_k \tag{4.2}$$

as a limit of the first sequence in (4.1). This subforest of the master forest $\mathcal{F}$ contains all element subdivisions occurring in the execution of iteration (2.3). In view of the adaptive marking, $\mathcal{F}_\infty = \mathcal{F}$ does not hold in general and $\mathcal{F}_\infty$ may have leaf nodes. The set $\mathcal{G}^+$ of the leaf nodes can be obtained also directly from $\{\mathcal{G}_k\}_{k \in \mathbb{N}}$:

$$\mathcal{G}^+ = \bigcup_{k \geq 0} \bigcap_{\ell \geq k} \mathcal{G}_\ell. \tag{4.3}$$

Moreover, the assumptions (2.5) on the refinement framework entail that neighbors of leaf nodes are eventually also leaf nodes.

Lemma 4.1 (Leaf nodes of $\mathcal{F}_\infty$). *Let $E \in \mathcal{F}_\infty$ be a leaf node, that is, $E \in \mathcal{G}^+$. Then there exists an iteration number $K = K(E) \in \mathbb{N}_0$ such that*

$$\forall k \geq K \qquad N_k(E) = N_K(E) \subset \mathcal{G}^+.$$

Proof. Let $E \in \mathcal{F}_\infty$ be a leaf node and let k be the smallest index with $E \in \mathcal{G}_k$. For $\ell \geq k$ let $E' \in N_\ell(E)$ be an arbitrary neighbor of E. Then, there is an element $R \in N_k(E)$ being the root of a subtree $\mathcal{T}_R \subset \mathcal{F}_\infty$ starting at R with $E' \in \mathcal{T}_R$. By local quasi-uniformity (2.5c) and quasi-regular subdivision (2.5b), we have

$$|E| \preccurlyeq |E'| \leq q_2^{\mathrm{dist}(E',R)} |R| \preccurlyeq q_2^{\mathrm{dist}(E',R)} |E|,$$

which implies that there exists a fixed constant M such that $\mathrm{dist}(E', R) \leq M$. The set $\cup_{\ell \geq k} N_\ell(E)$ of all possible neighbors of E is thus contained in $\mathcal{F}^M\big(E, N_k(E)\big)$. Therefore, if we choose K such that the finite forest $\mathcal{F}^M\big(E, N_k(E)\big) \cap \mathcal{F}_\infty \subset \mathcal{F}_K$, then all neighbors of E are contained in $\mathcal{F}_K$. Moreover, in view of (2.5a), subdividing a neighbor of E creates at least one new neighbor. Hence, all elements in $N_K(E)$ are leaf nodes of $\mathcal{F}_\infty$ and $N_\ell(E) = N_K(E) \subset \mathcal{G}^+$ is valid for all $\ell \geq K$. $\square$

Next, we turn to the sequence $\{u_k\}_k$ of approximate solutions. In [7] Babuška and Vogelius observed that this sequence converges for symmetric elliptic problems. Exploiting the assumptions (2.6) on the finite element spaces, we generalize that

result to non-coercive forms and additionally show that the limit is a solution with
respect to $(\mathbb{V}_\infty, \mathbb{W}_\infty)$ defined by

$$\mathbb{V}_\infty := \overline{\bigcup_{k \geq 0} \mathbb{V}_k} \quad \text{and} \quad \mathbb{W}_\infty := \overline{\bigcup_{k \geq 0} \mathbb{W}_k}.$$

Lemma 4.2 (Convergence of Petrov-Galerkin Approximations)**.** *The sequence* $\{u_k\}_k$
of approximate solutions converges in $\mathbb{V}$ *to the solution* u_∞ *with respect to the pair*
$(\mathbb{V}_\infty, \mathbb{W}_\infty)$*, which is characterized by*

$$u_\infty \in \mathbb{V}_\infty : \qquad \mathcal{B}(u_\infty, w) = f(w) \quad \forall w \in \mathbb{W}_\infty. \tag{4.4}$$

Proof. ⌐1⌐ Let us first prove that the pair $(\mathbb{V}_\infty, \mathbb{W}_\infty)$ satisfies the inf-sup condition

$$\inf_{\substack{v \in \mathbb{V}_\infty \\ \|v\|_\mathbb{V}=1}} \sup_{\substack{w \in \mathbb{W}_\infty \\ \|w\|_\mathbb{W}=1}} \mathcal{B}(v,w) \geq \beta, \quad \inf_{\substack{w \in \mathbb{W}_\infty \\ \|w\|_\mathbb{W}=1}} \sup_{\substack{v \in \mathbb{V}_\infty \\ \|v\|_\mathbb{V}=1}} \mathcal{B}(v,w) \geq \beta \tag{4.5}$$

with $\beta > 0$ from (2.6c).

To this end, fix first any $v \in \mathbb{V}_\infty$ and choose a sequence $\{v_k\}_k$ of functions
$v_k \in \mathbb{V}_k$ such that $v_k \to v$ in $\mathbb{V}$ as $k \to \infty$. Moreover, with the help of (2.6c) choose
a sequence $\{w_k\}_k$ of functions $w_k \in \mathbb{W}_k$ such that

$$\|w_k\|_\mathbb{W} = 1 \qquad \text{and} \qquad \mathcal{B}(v_k, w_k) \geq \left(\beta - \frac{1}{k}\right) \|v_k\|_\mathbb{V}. \tag{4.6}$$

Thanks to (2.6a), the sequence $\{w_k\}_k$ is in $\mathbb{W}$. Since the latter is reflexive, there
exists a subsequence $\{w_{k_j}\}_j$ and a function $w \in \mathbb{W}$ such that $w_{k_j} \rightharpoonup w$ weakly
in $\mathbb{W}$ as $j \to \infty$. Since $\mathbb{W}_\infty$ is closed and convex as well as $\|\cdot\|_\mathbb{W}$ weakly lower
semicontinuous, we have $w \in \mathbb{W}_\infty$ and $\|w\|_\mathbb{W} \leq \lim_{j\to\infty} \|w_{k_j}\|_\mathbb{W} = 1$. Combing this
with (4.6) yields

$$\mathcal{B}(v, w) \geq \beta\|v\|_\mathbb{V} \geq \beta\|v\|_\mathbb{V}\|w\|_\mathbb{W}.$$

In view of the first inequality, $w \neq 0$ and the first part of (4.5) is proved.

In the same way, but using (2.8) instead of (2.6c), we show that for any $w \in \mathbb{W}_\infty$
there exists $v \in \mathbb{V}_\infty$ such that $\mathcal{B}(v, w) \geq \beta\|v\|_\mathbb{V}\|w\|_\mathbb{W}$.

⌐2⌐ By (4.5), (2.6c) and (2.7), the Babuška-Brezzi theory implies that there is a
unique $u_\infty \in \mathbb{V}_\infty$ with $\mathcal{B}(u_\infty, w) = f(w)$ for all $w \in \mathbb{W}_\infty$ and that each u_k is a
$\|\cdot\|_\mathbb{V}$-quasi-optimal choice in $\mathbb{V}_k$ with respect to u_∞, i.e. there holds

$$\|u_k - u_\infty\|_\mathbb{V} \leq \left(1 + \frac{C_\mathcal{B}}{\beta}\right) \inf_{v \in \mathbb{V}_k} \|v - u_\infty\|_\mathbb{V}.$$

By construction $\bigcup_{k \geq 0} \mathbb{V}_k$ is dense in $\mathbb{V}_\infty$ and, thus, $\lim_{k\to\infty} \|u_\infty - u_k\|_\mathbb{V} = 0$. □

As a consequence of Lemma 4.2, the proof of $\lim_{k\to\infty} \|u - u_k\|_\mathbb{V} = 0$ reduces
to verifying that $u_\infty = u$. In view of (2.6a) and (4.4), the latter is equivalent to
$u \in \mathbb{V}_\infty$. It is worth noticing that $\mathbb{V}_\infty$ depends on $\{\mathcal{G}_k\}_k$, which in turn depends on
the given problem and on the modules SOLVE, ESTIMATE, MARK, and REFINE. In
particular, if $\mathcal{F}_\infty$ has leaf nodes, that is $\mathcal{G}^+ \neq \emptyset$, then $\mathbb{V}_\infty$ is a proper subset of $\mathbb{V}$.
The validity of $u \in \mathbb{V}_\infty$ thus hinges on properties of the aforementioned modules.

We turn to the last sequence in (4.1), the one of the meshsize functions $\{h_k\}_k$.
Recall that in the classical, non-adaptive setting convergence is ensured with the
help of

$$\lim_{k\to\infty} \|h_k\|_{L_\infty(\Omega)} = 0. \tag{4.7}$$

However, (4.7) may not hold in the adaptive setting and we therefore have to
generalize it appropriately. For a first generalization of (4.7), we observe that the
skeleton $\Sigma_k := \bigcup\{\partial E \cap \Omega : E \in \mathcal{G}_k\}$ of $\mathcal{G}_k$ has d-dimensional Lebesgue measure zero
and that $\Omega \setminus \Sigma_k = \{x \in \Omega \mid \#\mathcal{G}_k(x) = 1\}$. We may thus interpret h_k as a piecewise

constant function in $L_\infty(\Omega)$. Moreover, the limiting skeleton $\Sigma_\infty := \cup_{k\geq0}\Sigma_k$ has d-dimensional Lebesgue measure zero, too. Since, for every $x \in \Omega \setminus \Sigma_\infty$, the sequence $h_k(x)$ is monotonically decreasing and bounded from below by 0,

$$h_\infty(x) := \lim_{k\to\infty} h_k(x) \tag{4.8}$$

is well-defined for $x \in \Omega \setminus \Sigma_\infty$ and defines a function in $L_\infty(\Omega)$. As the next lemma shows, the pointwise convergence in (4.8) holds actually in $L_\infty(\Omega)$. Another generalization of (4.7), where the limit function is 0, will be provided in Corollary 4.5 below.

Lemma 4.3 (Uniform Convergence of Meshsize Functions). *The sequence $\{h_k\}_k$ of meshsize functions converges uniformly in $\Omega \setminus \Sigma_\infty$ to h_∞. In other words:*

$$\lim_{k\to\infty} \|h_k - h_\infty\|_{L_\infty(\Omega)} = 0.$$

Proof. For arbitrary $\varepsilon > 0$, let $m = m(\varepsilon) \in \mathbb{N}$ be the smallest number such that

$$m \geq \log(\varepsilon^d/M)/\log(q_2)$$

with $M = \max\{|E| \mid E \in \mathcal{G}_0\}$ and q_2 from (2.5a). Obviously, $\hat{\mathcal{F}} := \mathcal{F}_\infty \cap \mathcal{F}^m(\mathcal{G}_0, \mathcal{G}_0)$ is a subdivision of Ω and has finite depth, i.e. any node of $\hat{\mathcal{F}}$ is created by at most m subdivisions of an element in $\mathcal{G}_0$. This implies $\#\hat{\mathcal{F}} < \infty$ and since $\hat{\mathcal{F}} \subset \mathcal{F}_\infty$ there exists $k = k(\varepsilon) \geq 0$ with $\hat{\mathcal{F}} \subset \mathcal{F}_k$.

Let $E \in \mathcal{G}_k$ be any leaf node of $\mathcal{F}_k$ and let E be generated by subdividing $E_0 \in \mathcal{G}_0$. To estimate $h_k - h_\infty$ on E, we distinguish the following two cases:

Case 1: $\mathrm{dist}(E, E_0) < m$. This implies that E is a leaf node of $\mathcal{F}_\infty$ and thus $h_{k|E} = h_{\infty|E}$ or, equivalently, $(h_k - h_\infty)_{|E} = 0$.

Case 2: $\mathrm{dist}(E, E_0) \geq m$. Hence, E is generated by at least m subdivisions of E_0. By (2.5a), the monotonicity of the meshsize function, and the choice of m, we obtain

$$0 \leq (h_k - h_\infty)_{|E} \leq h_{k|E} = |E|^{1/d} \leq q_2^{m/d}|E_0|^{1/d} \leq q_2^{m/d}M^{1/d} \leq \varepsilon.$$

Combining the two cases we end up with $0 \leq (h_k - h_\infty)_{|E} \leq \varepsilon$ for all $E \in \mathcal{G}_k$. Since ε was arbitrary and we have $0 \leq h_\ell - h_\infty \leq h_k - h_\infty$ in Ω for all $\ell \geq k$, this finishes the proof. $\square$

4.2. **Convergence of Estimator.** In order to show that the error estimator vanishes in the limit, we write

$$\mathcal{E}_k = \left[\mathcal{E}_k^2(\mathcal{G}_k^0) + \mathcal{E}_k^2(\mathcal{G}_k^*) + \mathcal{E}_k^2(\mathcal{G}_k^+)\right]^{1/2}, \tag{4.9}$$

where

$$\mathcal{G}_k^0 := \left\{E \in \mathcal{G}_k \mid \mathcal{F}^n\big(\mathcal{G}_k, N_k(E)\big) \subset \mathcal{F}_\infty\right\}, \tag{4.10a}$$

$$\mathcal{G}_k^+ := \left\{E \in \mathcal{G}_k \mid N_k(E) \subset \mathcal{G}^+\right\}, \tag{4.10b}$$

$$\mathcal{G}_k^* := \mathcal{G}_k \setminus (\mathcal{G}_k^0 \cup \mathcal{G}_k^+) \tag{4.10c}$$

is a decomposition of $\mathcal{G}_k$ with the following properties:

- for indicators associated with $\mathcal{G}_k^0$, we can apply the local lower bound (2.9b) with a successive finite element space;
- for indicators associated with $\mathcal{G}_k^+$, we will have to exploit properties of the marking strategy since they belong to elements that are not marked anymore;
- the remaining indicators belong to elements that are between the preceding two extreme cases.

We now estimate the three corresponding estimator contributions separately.

First, we investigate the limit behavior of the estimator contribution $\{\mathcal{E}_k(\mathcal{G}_k^0)\}_k$ associated with (4.10a). For this purpose, we exploit the problem properties (2.2b), (2.2d), the local quasi-uniformity (2.5c) of the grids, the estimator properties (2.9b), (2.9c), and the auxiliary results Lemma 4.2 and Lemma 4.3; the marking strategy does not play any role.

Proposition 4.4 (Convergence of Estimator – I)**.** *The estimator contribution associated with* $\mathcal{G}_k^0$ *vanishes in the limit:*

$$\lim_{k\to\infty} \mathcal{E}_k(\mathcal{G}_k^0) = 0.$$

Proof. $\boxed{1}$ Let $E \in \mathcal{G}_k^0$ be arbitrary and choose $\ell > k$ with $\mathcal{F}_\ell \supset \mathcal{F}^n\big(\mathcal{G}_k, N_k(E)\big)$. Then there holds (2.9b) with $\mathcal{G}' = \mathcal{G}_\ell$. Moreover, for $w \in \mathbb{W}(\mathcal{G}_\ell) \subset \mathbb{W}_\infty$, inserting (4.4) into (2.10) and using (2.2d) yields

$$\langle \mathcal{R}_k, w \rangle = \mathcal{B}(u_k - u_\infty, w) \leq C_\mathcal{B} \|u_k - u_\infty\|_{\mathbb{V}(\mathrm{supp}(w))} \|w\|_\mathbb{W}.$$

Hence, exploiting also (2.9c), we obtain the following bound for all indicators associated with $\mathcal{G}_k^0$:

$$\mathcal{E}_k(E) \precsim \|u_k - u_\infty\|_{\mathbb{V}(\omega_k(E))} + m(|E|)\Big[\|u_k\|_{\mathbb{V}(\omega_k(E))} + \|D\|_{\mathbb{D}(\omega_k(E))}\Big]. \qquad (4.11)$$

$\boxed{2}$ To prepare the summation of (4.11), denote by M the maximal number of neighbors; cf. assumption (2.5c). Then we can split $\mathcal{G}_k^0$ into M^2+1 subsets $\mathcal{G}_{k,0}^0, \ldots, \mathcal{G}_{k,M^2}^0$ such that each $\mathcal{G}_{k,i}^0$, $i = 0, \ldots, M^2$, consists of elements for which the patches $\omega_\mathcal{G}(E)$ are pairwise disjoint. The subadditivity (2.2b) of $\|\cdot\|_{\mathbb{V}(\Omega)}$ then implies, for $v \in \mathbb{V}$,

$$\sum_{E \in \mathcal{G}_k^0} \|v\|_{\mathbb{V}(\omega_k(E))}^2 \leq \sum_{i=0}^{M^2} \sum_{E \in \mathcal{G}_{k,i}^0} \|v\|_{\mathbb{V}(\omega_k(E))}^2 \leq (M^2 + 1)\|v\|_{\mathbb{V}(\Omega_k^0)}^2 \qquad (4.12)$$

with

$$\Omega_k^0 := \bigcup\{\omega_k(E) : E \in \mathcal{G}_k^0\}. \qquad (4.13)$$

A similar statement holds for $\|\cdot\|_{\mathbb{D}(\Omega)}$ instead of $\|\cdot\|_{\mathbb{V}(\Omega)}$.

$\boxed{3}$ Summing (4.11) with the help of Step 2, we arrive at

$$\sum_{E \in \mathcal{G}_k^0} \mathcal{E}_k^2(E) \precsim \|u_k - u_\infty\|_{\mathbb{V}(\Omega)}^2 + \|m(h_k^d)\|_{L_\infty(\Omega_k^0)}^2 \Big(\|u_k\|_{\mathbb{V}(\Omega)}^2 + \|D\|_{\mathbb{D}(\Omega)}^2\Big). \qquad (4.14)$$

In Corollary 4.5 below we prove $\|h_k\|_{L_\infty(\Omega_k^0)} \to 0$ as $k \to \infty$, which by monotonicity and continuity of m implies

$$\lim_{k\to\infty} \|m(h_k^d)\|_{L_\infty(\Omega_k^0)} \leq \lim_{k\to\infty} m\Big(\|h_k\|_{L_\infty(\Omega_k^0)}^d\Big) = m(0) = 0.$$

Combining this with Lemma 4.2 and (4.14) yields $\mathcal{E}_k^2(\mathcal{G}_k^0) = \sum_{E \in \mathcal{G}_k^0} \mathcal{E}_k^2(E) \to 0$ as $k \to \infty$. $\qquad \square$

In order to complete the proof of Proposition 4.4, we need the following generalization of (4.7), which quite directly follows from Lemma 4.3.

Corollary 4.5. *The sequence* $\{\Omega_k^0\}_k$ *given by* (4.13) *satisfies*

$$\lim_{k\to\infty} \big\|h_k \chi_{\Omega_k^0}\big\|_{L_\infty(\Omega)} = 0,$$

where $\chi_{\Omega_k^0}$ *stands for the characteristic function of* Ω_k^0.

Proof. For any $E \in \mathcal{G}_k^0$ all elements of the patch $\omega_k(E)$ are subdivided at least n times by definition (4.10a) of $\mathcal{G}_k^0$. Hence (2.5a), definitions (3.1) and (4.8) imply $h_\infty \leq q_2^{n/d} h_k$ in $\omega_k(E)$, whence

$$h_k \leq \alpha(h_k - h_\infty) \quad \text{in} \quad \omega_k(E) \quad \text{with} \quad \alpha = \frac{1}{1 - q_2^{n/d}}.$$

Since $E \in \mathcal{G}_k^0$ is arbitrary, this implies

$$\left\| h_k \, \chi_{\Omega_k^0} \right\|_{L_\infty(\Omega)} \leq \left\| \alpha(h_k - h_\infty) \, \chi_{\Omega_k^0} \right\|_{L_\infty(\Omega)} \leq \alpha \| h_k - h_\infty \|_{L_\infty(\Omega)} \to 0$$

for $k \to \infty$ by Lemma 4.3. $\qquad\square$

Before embarking on the limit behavior of the other estimator contributions, we comment on the relationship of the local lower bound (2.9b) and the proof of Proposition 4.4.

Remark 4.6 (Using Local Lower Bound and Subdivision Depth n). It is worth observing that the lower bound (2.9b) is used to locally bound from below not the correction $\|u_{k+1} - u_k\|_{\mathbb{V}}$ (as in the existing convergence proofs) but rather $\|u_\ell - u_k\|_{\mathbb{V}}$ where ℓ is *some* iteration after k. This allows, without losing the practicability of the algorithm, for big values of the parameter n. Recall that n denotes the subdivision depth that ensures the validity of (2.9b). A concrete example with big n is given in §3.4.

Remark 4.7 (Need of Oscillation). The use of Corollary 4.5, which partially has the taste of a classical, non-adaptive argument, is limited to the terms arising from the oscillation $\mathrm{osc}_k(E)$ in the local lower bound (2.9b). Moreover, $\mathrm{osc}_k(E)$ 'spoils' the local lower bound. For these two reasons, it is worth arguing about the necessity of the oscillation indicator.

One reason for the presence of oscillation is that the error estimator provides a non-sharp upper bound (2.9a). We illustrate this with the help of the standard residual estimator. In this case the local lower bound presents an oscillation term. This cannot be avoided because the use of integral norms in the standard residual indicator causes an overestimation in the upper bound and, consequently, a 'clean', oscillation-free lower bound is not possible; see [25, Remark 3.1].

However, the oscillation will appear even if the error estimator provides a sharp upper bound. The reason for this lies in the finite-dimensional nature of the first term on the right hand side of (2.9b). In fact, this term is a dual norm, where the test functions are taken from a finite-dimensional space. Therefore, a finite number of such terms cannot alone bound the whole estimator, which in turn bounds the error from an infinite-dimensional space.

We next investigate the limiting behavior of $\{\mathcal{E}_k(\mathcal{G}_k^*)\}_k$ associated with (4.10c). Similarly to Proposition 4.4, the following proposition exploits the problem properties (2.2b), (2.2c), (2.2d), the refinement framework (2.5), the estimator properties (2.9b), (2.9c), and the auxiliary result Lemma 4.2; also here, the marking strategy does not play any role.

Proposition 4.8 (Convergence of Estimator – II). *The estimator contribution associated with $\{\mathcal{G}_k^*\}_k$ vanishes in the limit:*

$$\lim_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^*) = 0.$$

Proof. $\boxed{1}$ We first show that

$$\lim_{k \to \infty} |\Omega_k^*| = 0, \tag{4.15}$$

where $\Omega_k^* := \bigcup \{\omega_k(E) : E \in \mathcal{G}_k^*\}$ is the subdomain of Ω associated with $\mathcal{G}_k^*$.

To this end, consider any $E \in \mathcal{G}_k^*$ and distinguish two cases with respect to the location of $N_k(E)$ in $\mathcal{F}_\infty$.

Case 1: $N_k(E) \cap \mathcal{G}^+ = \emptyset$, that is all elements of $N_k(E)$ are interior nodes of $\mathcal{F}_\infty$. Since $E \notin \mathcal{G}_k^0$ and thus $\mathcal{F}^n(\mathcal{G}_k, N_k(E)) \setminus \mathcal{F}_\infty \neq \emptyset$, there is an element $E' \in N_k(E)$ and a leaf node $S_E \subset E'$ of $\mathcal{F}_\infty$ with $0 < \text{dist}(S_E, E') \leq n$. Moreover, there holds

$$|\omega_k(E)| \preccurlyeq |E'| \leq q_1^{-n}|S_E|$$

by (2.5c) and (2.5b). In other words, S_E has the following three properties:

$$S_E \in \mathcal{G}^+ \setminus \mathcal{G}_k, \quad S_E \subset \omega_k(E), \quad \text{and} \quad |\omega_k(E)| \preccurlyeq |S_E|. \qquad (4.16)$$

Case 2: $N_k(E) \cap \mathcal{G}^+ \neq \emptyset$, that is at least one element in $N_k(E)$ is a leaf node of $\mathcal{F}_\infty$. In view of $E \notin \mathcal{G}_k^*$, not all elements of the patch $\omega_k(E)$ are leaf nodes of $\mathcal{F}_\infty$. Thus, we can choose elements E', E'' of the patch $\omega_k(E)$ such that

$$E' \in \mathcal{G}^+, \quad E'' \notin \mathcal{G}^+, \quad \text{and } E' \cap E'' \neq \emptyset.$$

According to Lemma 4.1, there exists a successor S_E of E'' such that $S_E \in \mathcal{G}^+$ and $S_E \cap E' \neq \emptyset$. Furthermore, two applications of (2.5c) yield $|\omega_k(E)| \preccurlyeq |E'| \preccurlyeq |S_E|$. Therefore also in this case, S_E verifies (4.16).

Using (4.16) in both cases and the fact that the first part of (2.5c) entails $\#\{E \in \mathcal{G}_k^* \mid S_E = S\} \preccurlyeq 1$ for any $S \in \mathcal{G}^+ \setminus \mathcal{G}_k$, we obtain

$$|\Omega_k^*| \leq \sum_{E \in \mathcal{G}_k^*} |\omega_k(E)| \preccurlyeq \sum_{E \in \mathcal{G}_k^*} |S_E| \preccurlyeq \sum_{S \in \mathcal{G}^+ \setminus \mathcal{G}_k} |S|.$$

The last term in the previous inequality is a tail of the series $\sum_{S \in \mathcal{G}^+} |S|$. The latter is convergent because it has positive terms and all partial sums are bounded by $|\Omega|$. Consequently, we have proved (4.15).

[2] Recalling the lower bound (2.11) with respect to the true error, (2.9c), and using the bound $m(|E|) \leq m(|\Omega|)$, we derive

$$\mathcal{E}_k(E) \preccurlyeq \|u_k - u\|_{\mathbb{V}(\omega_k(E))} + m(|E|)\Big(\|u_k\|_{\mathbb{V}(\omega_k(E))} + \|D\|_{\mathbb{D}(\omega_k(E))}\Big)$$

$$\preccurlyeq \|u_k - u_\infty\|_{\mathbb{V}(\omega_k(E))} + \|u_\infty\|_{\mathbb{V}(\omega_k(E))} + \|u\|_{\mathbb{V}(\omega_k(E))} + \|D\|_{\mathbb{D}(\omega_k(E))}$$

in particular for all $E \in \mathcal{G}_k^*$. Arguing in a similar manner as in Step 2 of Proposition 4.4, we sum over $\mathcal{G}_k^*$ and obtain

$$\mathcal{E}_k^2(\mathcal{G}_k^*) = \sum_{E \in \mathcal{G}_k^*} \mathcal{E}_k^2(E)$$

$$\preccurlyeq \sum_{E \in \mathcal{G}_k^*} \|u_k - u_\infty\|_{\mathbb{V}(\omega_k(E))}^2 + \|u_\infty\|_{\mathbb{V}(\omega_k(E))}^2 + \|u\|_{\mathbb{V}(\omega_k(E))}^2 + \|D\|_{\mathbb{D}(\omega_k(E))}^2$$

$$\preccurlyeq \|u_k - u_\infty\|_{\mathbb{V}(\Omega)}^2 + \|u_\infty\|_{\mathbb{V}(\Omega_k^*)}^2 + \|u\|_{\mathbb{V}(\Omega_k^*)}^2 + \|D\|_{\mathbb{D}(\Omega_k^*)}^2.$$

As $k \to \infty$ the right hand side goes to 0 by Lemma 4.2 and $\lim_{k \to \infty} |\Omega_k^*| = 0$ of Step 1. $\qquad \square$

It is worth noticing the following difference in the use of (2.9b): in Proposition 4.4, we estimate with the help of the 'auxiliary' solution u_∞, while in Proposition 4.8 with the help of the true solution u and then introduce u_∞ in a second step.

Finally, we turn to $\{\mathcal{E}_k(\mathcal{G}_k^+)\}_k$ associated with (4.10a). Here, in addition to the already exploited properties, we use the property (2.13) of the marking strategy and the requirement (2.14) of the refine step.

Proposition 4.9 (Convergence of Estimator – III). *The estimator contribution associated with $\{\mathcal{G}_k^+\}_k$ vanishes in the limit:*

$$\lim_{k\to\infty} \mathcal{E}_k(\mathcal{G}_k^+) = 0.$$

Proof. $\boxed{1}$ Since elements in $\mathcal{G}_k^+$ are not subdivided and (2.14) requires that marked elements are subdivided, we have $\mathcal{M}_k \subset \mathcal{G}_k \setminus \mathcal{G}_k^+ = \mathcal{G}_k^0 \cup \mathcal{G}_k^*$. This entails

$$0 \le \lim_{k\to\infty} \mathcal{E}_k(\mathcal{M}_k) \le \lim_{k\to\infty} \mathcal{E}_k(\mathcal{G}_k^0 \cup \mathcal{G}_k^*) = 0$$

in view of Propositions 4.4 and 4.8. The assumption (2.13) on the marking strategy then implies the element-wise convergence

$$\forall\, E \in \mathcal{G}^+ \qquad \lim_{k\to\infty} \mathcal{E}_k(E) = 0. \tag{4.17}$$

To conclude that $\lim_{k\to\infty} \mathcal{E}_k(\mathcal{G}_k^+) = 0$, we reformulate this element-wise convergence as a pointwise one in a integral framework and use a generalized version of the dominated convergence theorem.

$\boxed{2}$ In order to prepare the application of dominated convergence, we observe the following facts for any given element $E \in \mathcal{G}_k^+$. First, in view of the definition (4.10b) of $\mathcal{G}_k^+$, there holds $\omega_k(E) = \omega_\ell(E)$ for all $\ell \ge k$ and we may call this set $\omega(E)$. Next, exploiting the lower bound (2.11) with respect to the true error and (2.9c) as in Step 2 of the proof of Proposition 4.8, we obtain

$$\mathcal{E}_k^2(E) \precsim \|u_k - u_\infty\|_{\mathbb{V}(\omega(E))}^2 + \|u_\infty\|_{\mathbb{V}(\omega(E))}^2 + \|u\|_{\mathbb{V}(\omega(E))}^2 + \|D\|_{\mathbb{D}(\omega(E))}^2$$
$$=: \|u_k - u_\infty\|_{\mathbb{V}(\omega(E))}^2 + C_E^2, \tag{4.18}$$

where C_E does not depend on k. Notice that, as a consequence of Lemma 4.2, the right hand side tends to C_E as $k \to \infty$. Moreover, arguing a further time as in Step 2 of the proof of Proposition 4.4, we obtain for the sum of the asymptotic bounds:

$$\sum_{E\in\mathcal{G}_k^+} C_E^2 = \sum_{E\in\mathcal{G}_k^+} \|u_\infty\|_{\mathbb{V}(\omega(E))}^2 + \|u\|_{\mathbb{V}(\omega(E))}^2 + \|D\|_{\mathbb{D}(\omega(E))}^2$$
$$\precsim \|u_\infty\|_{\mathbb{V}(\Omega)}^2 + \|u\|_{\mathbb{V}(\Omega)}^2 + \|D\|_{\mathbb{D}(\Omega)}^2 \precsim 1. \tag{4.19}$$

$\boxed{3}$ We turn to the integral reformulation. Let $\Omega^+ := \bigcup\{E : E \in \mathcal{G}^+\}$ be the underlying domain and notice that, thanks to Lemma 4.1, there holds $\mathcal{G}^+ = \bigcup_{k\in\mathbb{N}_0} \mathcal{G}_k^+$, where the sequence $\{\mathcal{G}_k^+\}_k$ is nested. Given $x \in \Omega^+$, choose the smallest iteration number $\ell = \ell(x)$ such that there is an $E \in \mathcal{G}_\ell^+$ with $x \in E$. Now define

$$\varepsilon_k(x) = M_k(x) = 0 \qquad \text{for } k < \ell$$

and

$$\varepsilon_k(x) := \frac{1}{|E|}\mathcal{E}_k^2(E), \quad M_k(x) := \frac{1}{|E|}\left(\|u_k - u_\infty\|_{\mathbb{V}(\omega(E))}^2 + C_E^2\right) \qquad \text{for } k \ge \ell.$$

Then, for any $k \in \mathbb{N}_0$, there holds

$$\mathcal{E}_k^2(\mathcal{G}_k^+) = \int_{\Omega^+} \varepsilon_k(x)\, dx$$

and (4.17) implies the pointwise convergence in Ω^+ of ε_k to 0:

$$\varepsilon_k(x) = \frac{1}{|E|}\mathcal{E}_k^2(E) \to 0 \qquad \text{for } k \to \infty.$$

Moreover, (4.18), $\sum_{E\in\mathcal{G}_k^+}\|u_k - u_\infty\|_{\mathbb{V}(\omega(E))}^2 \precsim \|u_k - u_\infty\|_{\mathbb{V}}$, and (4.19) ensure that each M_k is a summable majorant for ε_k.

④ Next, we show that the majorants $\{M_k\}_k$ converge in $L_1(\Omega^+)$ to M defined by

$$M(x) := \frac{1}{|E|} C_E^2 \qquad \text{for } x \in E \text{ and } E \in \mathcal{G}^+.$$

The monotone convergence theorem [34, Appendix, (19b), p. 1016] and the definition of M_k imply

$$\|M_k - M\|_{L_1(\Omega^+)} = \sum_{E \in \mathcal{G}_k^+} \|M_k - M\|_{L_1(E)} + \sum_{E \in \mathcal{G}^+ \setminus \mathcal{G}_k^+} \|M\|_{L_1(E)}.$$

Since the first term satisfies

$$\sum_{E \in \mathcal{G}_k^+} \|M_k - M\|_{L_1(E)} = \sum_{E \in \mathcal{G}_k^+} \|u_k - u_\infty\|_{\mathbb{V}(\omega(E))}^2 \preccurlyeq \|u_k - u_\infty\|_{\mathbb{V}(\Omega)}^2 \to 0$$

for $k \to \infty$ thanks to Lemma 4.2 and the second term is a tail of the series $\sum_{E \in \mathcal{G}^+} \|M\|_{L_1(E)} = \sum_{E \in \mathcal{G}^+} C_E^2$, which is finite thanks to (4.19), we conclude that $M_k \to M$ in $L_1(\Omega^+)$.

⑤ The properties of $\{\varepsilon_k\}_k$, $\{M_k\}_k$, and M derived in Steps 3 and 4 allow to apply the generalized majorized convergence theorem [34, Appendix, (19a), p. 1015] and we conclude that $\lim_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^+)^2 = \lim_{k \to \infty} \int_{\Omega^+} \varepsilon_k = 0$. □

Combining the preceding propositions with the upper bound (2.9a) readily provides the

Proof of Theorem 2.1. Inserting the three Propositions 4.4, 4.8, and 4.9 in the splitting (4.9), we easily arrive at $\lim_{k \to \infty} \mathcal{E}_k = 0$. This, in view of the upper bound (2.9a), entails $\lim_{k \to \infty} \|u - u_k\|_{\mathbb{V}} = 0$. □

5. Supplements

We derive further results that are readily available through the proof of Theorem 2.1 in §4.

We first consider general marking strategies for which the output of marked elements does not depend only on the triangulation and estimator of the current iteration but also on those of previous iterations. Marking strategies of this type have been proposed, e.g., in [6, §6].

In order to derive an appropriate generalization of (2.13), let us first deduce a necessary condition for convergence. Obviously,

$$\forall E \in \mathcal{G}^+ \qquad \lim_{k \to \infty} \mathcal{E}_k(E) = 0 \tag{5.1}$$

is necessary for the 'practical' convergence $\lim_{k \to \infty} \mathcal{E}_k = 0$. In addition, if there holds $\lim_{k \to \infty} \mathrm{osc}_k(E) = 0$ for all $E \in \mathcal{G}^+$, then the lower bound (2.9b) ensures that (5.1) is necessary also for the 'theoretical' convergence $\lim_{k \to \infty} \|u - u_k\|_{\mathbb{V}} = 0$. The assumption on the data oscillation is verified, e.g., in the example of §3.1 whenever f is piecewise constant over the initial triangulation $\mathcal{G}_0$. In this context, recall also the various reasons for the need of oscillation indicators in Remark 4.7.

Condition (5.1) can neither be checked with the marking strategy at hand only nor generalizes (2.13). We therefore consider the implication

$$\lim_{k \to \infty} \mathcal{E}_k(\mathcal{M}_k) = 0 \implies \forall E \in \mathcal{G}^+ \quad \lim_{k \to \infty} \mathcal{E}_k(E) = 0, \tag{5.2}$$

which is necessary for $\lim_{k \to \infty} \mathcal{E}_k = 0$ or $\lim_{k \to \infty} \|u - u_k\|_{\mathbb{V}} = 0$ in the same way like (5.1). Condition (5.2) is weaker than (2.13) but can replace (2.13) in Step 1 of the proof of Proposition 4.9 — the only place where (2.13) is used. We thus have a condition for marking that is essentially necessary and sufficient for convergence.

Theorem 5.1 (Characterization of Convergent Marking)**.** *Suppose that u is the exact solution of (2.1) and that there hold (2.2), (2.5), (2.6), (2.7), (2.9), as well as (2.14).*

The sequence of approximate solutions $\{u_k\}_k$ generated by iteration (2.3) with a general marking strategy converges in sense that

$$\|u_k - u\|_{\mathbb{V}(\Omega)} \to 0 \quad and \quad \mathcal{E}_k \to 0 \quad as \quad k \to \infty$$

if and only if the marking satisfies (5.2).

Next, we investigate marking strategies with the property that, for some given tolerance $\mathsf{TOL} > 0$, there holds

$$\mathcal{E}_k < \mathsf{TOL} \implies \text{the output of } \mathsf{MARK} \text{ is void,} \tag{5.3}$$

that is, they only aim at reaching $\mathcal{E}_k < \mathsf{TOL}$ and, in that case, if we suppose that

$$\mathcal{G} = \mathsf{REFINE}(\mathcal{G}, \emptyset), \tag{5.4}$$

one can exit iteration (2.3). Such strategies are of interest, e.g., for the solution of a time step in evolution problems. Examples can be obtained from the previous strategies by requiring (2.13) only if $\mathcal{E}_k \geq \mathsf{TOL}$ and otherwise marking no element. An example of different nature is the equidistribution strategy for which

$$\mathsf{MARK}\big(\{\mathcal{E}_k(E)\}_{E \in \mathcal{G}_k}, \mathcal{G}_k, \mathsf{TOL}\big) = \left\{ E \in \mathcal{G}_k \mid \mathcal{E}_k(E) \geq \frac{\theta \mathsf{TOL}}{\sqrt{\#\mathcal{G}_k}} \right\}$$

with a fixed $\theta \in (0, 1]$ and, in the case $\theta < 1$, θ may be seen as a stabilization parameter. As we shall see in a moment, the aforementioned strategies satisfy the condition

$$\forall k \in \mathbb{N}_0 \ \ \mathcal{E}_k \geq \mathsf{TOL} \quad \implies \quad \liminf_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^+) < \mathsf{TOL}. \tag{5.5}$$

For modified strategies with (2.13), this directly follows from Proposition 4.9. For the equidistribution strategy, we first observe that $\mathcal{E}_k \geq \mathsf{TOL}$ in all iterations entails $\lim_{k \to \infty} \#\mathcal{G}_k = \infty$ and so $\lim_{k \to \infty} \mathcal{E}_k(E) = 0$ for all $E \in \mathcal{G}^+$ and then the proof of Proposition 4.9 ensures again $\lim_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^+) = 0 < \mathsf{TOL}$.

Theorem 5.2 (Termination of Marking for Tolerance)**.** *Suppose that u is the exact solution of (2.1) and that there hold (2.2), (2.5), (2.6), (2.7), (2.9), (2.14) as well as (5.3) and (5.4). Then the algorithm reaches*

$$\mathcal{E}_k < \mathsf{TOL} \quad (and \ so \ \|u_k - u\|_{\mathbb{V}(\Omega)} \preccurlyeq \mathsf{TOL}) \tag{5.6}$$

after a finite number of iterations if and only if the marking satisfies (5.5).

Proof. For any iteration $k \in \mathbb{N}_0$, we may write

$$\mathcal{E}_k^2 = \mathcal{E}_k(\mathcal{G}_k^0 \cup \mathcal{G}_k^*)^2 + \mathcal{E}_k(\mathcal{G}_k^+)^2. \tag{5.7}$$

If (5.6) holds for some iteration k, then (5.3), (5.4), and (5.7) imply

$$\liminf_{\ell \to \infty} \mathcal{E}_\ell(\mathcal{G}_\ell^+) = \mathcal{E}_k(\mathcal{G}_k^+) = \mathcal{E}_k < \mathsf{TOL},$$

which means the necessity of (5.5).

In order to show its sufficiency, suppose that (5.6) is never satisfied. Then (5.5) implies $\liminf_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^+) < \mathsf{TOL}$. Inserting this, Proposition 4.4, and Proposition 4.8 in the limit of (5.7) gives

$$\liminf_{k \to \infty} \mathcal{E}_k^2 = \lim_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^0 \cup \mathcal{G}_k^*)^2 + \liminf_{k \to \infty} \mathcal{E}_k(\mathcal{G}_k^+)^2 < \mathsf{TOL},$$

which contradicts the assumption that (5.6) is never satisfied. $\qquad \square$

References

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

[2] I. BABUŠKA, *Error-Bounds for Finite Element Method*, Numer. Math. 16 (1971), 322–333.

[3] I. BABUŠKA, A. K. AZIZ, *Lectures on the mathematical foundations of the finite element method*, University of Maryland, College Park, MD, 1972. Technical Note BN-748.

[4] I. BABUŠKA, A. MILLER, *A feedback finite element method with a posteriori error estimation: I. The finite element method and some basic properties of the a posteriori error estimator*, Comput. Methods Appl. Mech. Engrg. 61 (1987), 1–40.

[5] I. BABUŠKA, T. STROUBOULIS, *The finite element method and its reliability*, The Clarendon Press Oxford University Press, New York, 2001.

[6] I. BABUŠKA, W. RHEINBOLDT, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. 15 (1978), 736–754.

[7] I. BABUŠKA, M. VOGELIUS, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numer. Math. 44 (1984), 75-102.

[8] E. BÄNSCH, *Local mesh refinement in 2 and 3 dimensions*, IMPACT Comput. Sci. Engrg. 3 (1991), 181–191.

[9] E. BÄNSCH, P. MORIN, R. H. NOCHETTO, *An adaptive Uzawa FEM for the Stokes Problem. Convergence without the inf-sup condition*, SIAM J. Numer. Anal. 40 (2002), 1207–1229.

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

[11] D. BRAESS, *Finite elements. Theory, fast solvers, and applications in solid mechanics*, 2nd ed., Cambridge University Press, 2001.

[12] F. BREZZI, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers*, R.A.I.R.O. Anal. Numer. **R2**, T 129–151 (1974).

[13] J. M. CASCÓN, R. H. NOCHETTO, K. G. SIEBERT, *Design and convergence of AFEM in H(div)*, preprint Augsburg 2006.

[14] C. CARSTENSEN, S. BARTELS, *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. Part I: Low order conforming, non conforming, and mixed FEM*, Math. Comp. 71 (2002), 945–969.

[15] C. CARSTENSEN, R. H. W. HOPPE, *Error reduction and convergence for an adaptive mixed finite element method*, Math. Comp. 75 (2006), 1033–1042.

[16] Z. CHEN, J. FENG, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*, Math. Comp. 73 (2004), 1167–1193.

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

[18] G. P. GALDI, *An introduction to the mathematical theory of the Navier-Stokes equations. Vol. 1: Linearized steady problems*, Springer Tracts in Natural Philosophy, 38 (1994).

[19] I. KOSSACZKÝ, *A recursive approach to local mesh refinement in two and three dimensions*, J. Comput. Appl. Math. 55 (1994), 275–288.

[20] K. MEKCHAY, R. H. NOCHETTO, *Convergence of adaptive finite element methods for general second order linear elliptic PDE*, SIAM J. Numer. Anal. 43 (2005), 1803–1827.

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

[22] P. MORIN, R. H. NOCHETTO, K. G. SIEBERT, *Local problems on stars: a posteriori error estimators, convergence, and performance*, Math. Comp. 72 (2003), 1067–1097.

[23] P. MORIN, K. G. SIEBERT, A. VEESER, *Convergence of finite elements adapted for weak norms*, preprint no. 2/2007, Dipartimento di Matematica "F. Enriques", Via C. Saldini 50, 20133 Milano, Italy.

[24] F. OTTO, *On the Babuška–Brezzi condition for the Taylor–Hood element*, (in German) Master thesis, University of Bonn (1990).

[25] R. SACCHI, A. VEESER, *Locally efficient and reliable a posteriori error estimators for Dirichlet problems*, Math. Mod. Meth. Appl. Sci. 16 (2006), 319–346.

[26] A. Schmidt, K. G. Siebert, *Design of adaptive finite element software. The finite element toolbox ALBERTA*, Springer, 2005.

[27] K. G. Siebert, A. Veeser, *Convergence of the equidistribution strategy*, Oberwolfach Reports 2 (2005), 2129–2131.

[28] K. G. Siebert, A. Veeser, *A unilaterally constrained quadratic minimization with adaptive finite elements*, to appear in SIAM J. Optim.

[29] R. Sivan, *Surface Modeling using Quadtrees*, PhD Thesis, Center for Automation, University of Maryland, 1996.

[30] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math. Published online: 5 July 2006. DOI 10.1007/s10208-005-0183-0.

[31] A. Veeser, *Convergent adaptive finite elements for the nonlinear Laplacian*, Numer. Math. 92 (2002), 743-770.

[32] R. Verfürth, *A posteriori error estimators for the Stokes equations*, Numer. Math. 55 (1989), 309–325.

[33] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Adv. Numer. Math., John Wiley, Chichester, UK, 1996.

[34] E. Zeidler, *Nonlinear functional analysis and its applications. II/B: Nonlinear monotone operators*, Springer, New York etc., 1990.

Pedro Morin, Departamento de Matemática, Facultad de Ingeniería Química, Instituto de Matemática Aplicada del Litoral Universidad Nacional del Litoral, CONICET, Güemes 3450, S3000GLN Santa Fe, Argentina

URL: `http://math.unl.edu.ar/~pmorin/`

E-mail address: `pmorin@math.unl.edu.ar`

Kunibert G. Siebert, Institut für Mathematik, Universität Augsburg, Universitätsstrasse 14, D-86159 Augsburg, Germany

URL: `http://scicomp.math.uni-augsburg.de/Siebert/`

E-mail address: `siebert@math.uni-augsburg.de`

Andreas Veeser, Dipartimento di Matematica, Università degli Studi di Milano, Via C. Saldini 50, I-20133 Milano, Italy

URL: `http://mat.unimi.it/users/veeser/`

E-mail address: `veeser@mat.unimi.it`