Convergence of Adaptive Finite Element Methods with Error-Dominated Oscillation

Christian Kreuzer and Andreas Veeser

Abstract Recently, we devised an approach to a posteriori error analysis, which clarifies the role of oscillation and where oscillation is bounded in terms of the current approximation error. Basing upon this approach, we derive plain convergence of adaptive linear finite elements approximating the Poisson problem. The result covers arbitray H^{-1}-data and characterizes convergent marking strategies.

1 Introduction

Adaptive finite element methods (AFEMs) are well-established and efficient tools for solving boundary value problems. A common form is the iteration of

$$\text{SOLVE} \quad \to \quad \text{ESTIMATE} \quad \to \quad \text{MARK} \quad \to \quad \text{REFINE}, \qquad (1)$$

i.e. solve for the current finite element solution, estimate its error by means of so-called a posteriori indicators, and use this information in a marking strategy to make refinement decisions. The resultant problem-dependent adaptation of the mesh often significantly improves efficiency, compared to classical uniform refinement. There are theoretical results corroborating this practical observation; see the surveys [3, 9] and [5], which even obtains a nonlinear quasi-optimality result with respect to an augmented energy norm error.

Almost all results about the convergence behavior of AFEMs invoke extra regularity for the data which does not fit to the proven convergence rate in the light

C. Kreuzer (✉)
Technische Universität Dortmund, Fakultät für Mathematik, Dortmund, Germany
e-mail: christian.kreuzer@tu-dortmund.de

A. Veeser
Università degli Studi di Milano, Dipartimento di Matematica F. Enriques, Milano, Italy
e-mail: andreas.veeser@unimi.it

of [2, 6]. To exemplify and explain this, let us consider the approximation of the weak solution of the Poisson problem

$$-\Delta u = f \text{ in } \Omega \subset \mathbb{R}^d, \qquad u = 0 \text{ on } \partial\Omega \tag{2}$$

by a realization of (1) with linear finite elements. Here one usually requires the regularity $f \in L^2(\Omega)$ beyond $H^{-1}(\Omega)$, irrespective of the actual regularity of the target function u. One reason for requiring $f \in L^2(\Omega)$ lies in a posteriori error analysis. Aiming at computability, local H^{-1}-norms are replaced with L^2-norms scaled by the local meshsize h_T, leading to terms like the classical oscillation indicator $h_T \min_{f_T \in \mathbb{R}} \|f - f_T\|_{L^2(T)}$, which needs $f \in L^2(\Omega)$ to be defined. Note however that such a term is not yet really computable and may cause serious overestimation; cf. [7]. Another reason lies in many convergence analyses themselves. They often rely on the fact that the scaling factors of the aforementioned L^2-norms strictly reduce under refinement and, e.g., in the case of oscillation terms, this strict reduction hinges on extra regularity of f. It is worth mentioning that the second reason does not apply to the derivations of plain convergence in [8, 10].

In [7] we present a new approach to a posteriori analysis covering arbitrary $f \in H^{-1}(\Omega)$ and prove in particular

$$\|\nabla(u - U_{\mathcal{T}})\|_\Omega \lesssim \mathcal{E}_{\mathcal{T}} + \delta_{\mathcal{T}} \lesssim \|\nabla(u - U_{\mathcal{T}})\|_\Omega,$$

where $U_{\mathcal{T}}$ is the linear finite element solution over the mesh $\mathcal{T}$, the estimator $\mathcal{E}_{\mathcal{T}}$ is computable in a similar sense as $U_{\mathcal{T}}$, while the computation of the oscillation $\delta_{\mathcal{T}}$ requires additional information on f.

Here we prove plain convergence of adaptive linear finite elements for (2) by combining the a posteriori analysis [7] and the convergence analyses [8, 10]. The result forgoes extra regularity and so convergence rates are precluded. Section 2 presents the adaptive finite method in detail, along with an account of [7], and Sect. 3 is concerned with the convergence proof, thereby characterizing convergent marking strategies. Section 4 then concludes with a brief discussion of a few convergent marking strategies.

2 Adaptive Algorithm with Abstract Marking Strategy

In this section we introduce adaptive finite element methods for approximating the weak solution $u \in H_0^1(\Omega)$ of (2) for some fixed but arbitrary load

$$f \in H^{-1}(\Omega). \tag{3}$$

Here $H_0^1(\Omega)$ is the subspace of Sobolev functions in $H^1(\Omega)$ with vanishing trace and $H^{-1}(\Omega)$ denotes its topological dual space.

Bisection and Conforming Linear Finite Elements Let $\mathcal{T}_0$ be a suitable simplicial, face-to-face (conforming) mesh of Ω and denote by $\mathbb{T}$ the family of its refinements using iterative or recursive bisection; cf. the discussion in [9] and the references therein. In what follows, '$\lesssim$' stands for '$\leq C$', where the generic constant C may depend on the *shape coefficient of* $\mathcal{T}_0$. For example, for any simplex arising in $\mathbb{T}$, we have that its shape coefficient $\lesssim 1$.

Let $\mathcal{T} \in \mathbb{T}$ be a mesh. We denote the set of its vertices by $\mathcal{V}_\mathcal{T}$ and the set of its faces by $\mathcal{F}_\mathcal{T}$. The associated space of continuous piecewise affine functions is

$$\mathbb{V}(\mathcal{T}) := \left\{ V \in C^0(\bar{\Omega}) \mid V_{|T} \in \mathbb{P}_1 \text{ for all } T \in \mathcal{T} \right\} \subset H^1(\Omega).$$

Its *nodal basis* $\{\Phi_{\mathcal{T},z}\}_{z \in \mathcal{V}_\mathcal{T}}$ is given by

$$\Phi_{\mathcal{T},z} \in \mathbb{V}(\mathcal{T}) \quad \text{such that} \quad \Phi_{\mathcal{T},z}(y) = \delta_{zy} \text{ for all } z, y \in \mathcal{V}_\mathcal{T}.$$

Note $\operatorname{supp} \Phi_{\mathcal{T},z} = \bigcup\{T \in \mathcal{T} : T \ni z\} =: \omega_{\mathcal{T},z}$. The finite element functions $\mathring{\mathbb{V}}(\mathcal{T}) := \{V \in \mathbb{V}(\mathcal{T}) \mid \forall z \in \mathcal{V}_\mathcal{T} \cap \partial\Omega \;\; V(z) = 0\} \subset H_0^1(\Omega)$ are *conforming* and the associated *Galerkin approximation* $U_\mathcal{T}$ of (2) is characterized by

$$U_\mathcal{T} \in \mathring{\mathbb{V}}(\mathcal{T}) \quad \text{such that} \quad \forall V \in \mathring{\mathbb{V}}(\mathcal{T}) \quad \int_\Omega \nabla U_\mathcal{T} \cdot \nabla V \, dx = \langle f, V \rangle, \qquad (4)$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. In order to be able to compute $U_\mathcal{T}$, we suppose that

$$\forall \mathcal{T} \in \mathbb{T}, z \in \mathcal{V}_\mathcal{T} \cap \Omega \quad \langle f, \Phi_{\mathcal{T},z} \rangle \text{ is available.} \qquad (5)$$

A Posteriori Error Analysis with Error-Dominated Oscillation We outline the approach to a posteriori error analysis in [7], which prepares the ground for formulating our convergence theorem. To this end, we fix the mesh $\mathcal{T}$ and drop it as subscript in the following discussion. Further, given a subdomain $\omega \subset \Omega$, we choose $\|\nabla \cdot\|_\omega := \|\nabla \cdot\|_{L^2(\omega)}$ as norm on $H_0^1(\omega)$ and denote by $\|\cdot\|_{-1;\omega}$ its dual norm on $H^{-1}(\omega)$.

We then have the global relationship $\|\nabla(u - U)\|_\Omega = \|f + \Delta U\|_{-1;\Omega}$ between error and residual as well as, locally,

$$\forall z \in \mathcal{V} \quad \|f + \Delta U\|_{-1;\omega_z} \leq \|\nabla(u - U)\|_{\omega_z}. \qquad (6)$$

Exploiting Galerkin orthogonality and the fact that $\{\Phi_z\}_{z \in \mathcal{V}}$ provides a partition of unity, one derives that, for any $v \in H_0^1(\Omega)$ and $V \in \mathring{\mathbb{V}}(\mathcal{T})$,

$$\langle f + \Delta U, v \rangle = \langle f + \Delta U, v - V \rangle \lesssim \sum_{z \in \mathcal{V}} \|f + \Delta U\|_{-1,\omega_z} \|\nabla v\|_{\omega_z}; \qquad (7)$$

see, e.g., [4, 7]. Applying the Cauchy-Schwarz inequality on the sum and recalling (6), one obtains the error localization

$$\|\nabla(u - U)\|_{\Omega}^2 \approx \sum_{z \in \mathcal{V}} \|f + \Delta U\|_{-1;\omega_z}^2.$$

In general, upper bounds for the local residual norms $\|f + \Delta U\|_{-1;\omega_z}$, as for the global one $\|f + \Delta u\|_{-1;\Omega}$, cannot be computed by a finite number of operations from the information in (5). The reason lies in the possibly infinite-dimensional nature of f; see [7]. One therefore may isolate this difficulty by inserting a finite-dimensional approximation of f, e.g., some piecewise constant function $\bar{f}$. Unfortunately, the arising terms $\|f - \bar{f}\|_{-1;\omega_z}$, or the larger classical *oscillation* terms $\|h(f - \bar{f})\|_{\omega_z}$, may overestimate the error by far on a given mesh, see [7], and even lead to worse convergence rates, see [4].

To overcome these drawbacks, we have developed in [7] a projection operator $\mathcal{P}$: $H^{-1}(\Omega) \to \mathbb{D}(\mathcal{T})$ with the following properties. The evaluation $\mathcal{P}f$ is computable under (5) and the range

$$\mathbb{D}(\mathcal{T}) := \Big\{ \ell \in H^{-1}(\Omega) \mid \langle \ell, v \rangle = \sum_{T \in \mathcal{T}} \int_T c_T v \, \mathrm{d}x + \sum_{F \in \mathcal{F}} \int_F c_F v \, \mathrm{d}s$$

$$\text{for all } v \in H_0^1(\Omega) \text{ with } c_T, c_F \in \mathbb{R} \text{ for } T \in \mathcal{T}, F \in \mathcal{F} \Big\}$$

contains $\Delta U \in H^{-1}(\Omega)$. Moreover, $\mathcal{P}$ is locally stable and invariant in that, for any $\ell \in H^{-1}(\Omega)$, $z \in \mathcal{V}$, and $D \in \mathbb{D}(\mathcal{T})$, we have

$$\|\mathcal{P}\ell\|_{-1;\omega_z} \lesssim \|\ell\|_{-1;\omega_z}, \tag{8a}$$

as well as

$$\ell = \mathcal{P}\ell \text{ in } H^{-1}(\omega_z) \quad \text{whenever} \quad \ell = D \text{ in } H^{-1}(\omega_z). \tag{8b}$$

Consequently, we have $\mathcal{P}(\Delta U) = \Delta U$ and the local splittings

$$\|f + \Delta U\|_{-1;\omega_z} \approx \|\mathcal{P}f + \Delta U\|_{-1;\omega_z} + \|f - \mathcal{P}f\|_{-1;\omega_z}, \tag{9}$$

where $\|\mathcal{P}f + \Delta U\|_{-1;\omega_z}$ is quantifiable under (5), but the computation or bounding of $\|f - \mathcal{P}f\|_{-1;\omega_z}$ requires information beyond. To sum up, we formulate the following theorem.

Theorem 1 (A Posteriori Bounds) *For any mesh $\mathcal{T} \in \mathbb{T}$, we have*

$$\|\nabla(u - U_{\mathcal{T}})\|_{\Omega}^2 \lesssim \sum_{z \in \mathcal{V}_{\mathcal{T}}} \|\mathcal{P}_{\mathcal{T}}f + \Delta U_{\mathcal{T}}\|_{-1;\omega_{\mathcal{T},z}}^2 + \|f - \mathcal{P}_{\mathcal{T}}f\|_{-1;\omega_{\mathcal{T},z}}^2$$

as well as, for any $z \in \mathcal{V}_{\mathcal{T}}$,

$$\|\mathcal{P}_{\mathcal{T}}f + \Delta U_{\mathcal{T}}\|^2_{-1;\omega_{\mathcal{T},z}} + \|f - \mathcal{P}_{\mathcal{T}}f\|^2_{-1;\omega_{\mathcal{T},z}} \lesssim \|\nabla(u - U_{\mathcal{T}})\|^2_{\omega_{\mathcal{T},z}}.$$

Adaptive Algorithm The bounds in Theorem 1 are vertex-wise, while bisection is applied element-wise. Therefore, we assume that there are element indicators $\mathcal{E}_{\mathcal{T}}(T), T \in \mathcal{T}$, such that

$$\|\mathcal{P}_{\mathcal{T}}f + \Delta U_{\mathcal{T}}\|^2_{-1;\omega_{\mathcal{T},z}} \lesssim \sum_{T \subset \omega_{\mathcal{T},z}} \mathcal{E}_{\mathcal{T}}(T)^2 \tag{10a}$$

and

$$\mathcal{E}_{\mathcal{T}}(T)^2 \lesssim \sum_{z \in T \cap \mathcal{V}_{\mathcal{T}}} \|\mathcal{P}_{\mathcal{T}}f + \Delta U_{\mathcal{T}}\|^2_{-1;\omega_{\mathcal{T},z}}. \tag{10b}$$

Such local equivalences hold for various indicators resembling the ones from the standard residual estimator, the hierarchical estimator, or estimators based upon discrete local problems; cf. [7].

In order to achieve a similar grouping for the oscillation, we introduce, for any $T \in \mathcal{T}$, its minimal ring $\omega(T) := \bigcap\{\omega_{\mathcal{T}'}(T) : \mathcal{T}' \in \mathbb{T}, \mathcal{T}' \ni T\}$, where $\omega_{\mathcal{T}}(T) := \cup\{T' \in \mathcal{T} : T' \cap T \neq \emptyset\}$. Then

$$\delta(T) := \|f - \mathcal{P}_{\mathcal{T}}f\|_{-1;\omega(T)} \tag{11}$$

does not depend on the surrounding mesh, and we can derive the inequalities

$$\|f - \mathcal{P}_{\mathcal{T}}f\|^2_{-1;\omega_{\mathcal{T},z}} \lesssim \sum_{T \subset \omega_{\mathcal{T},z}} \delta(T)^2, \quad \delta(T)^2 \lesssim \sum_{z \in T \cap \mathcal{V}_{\mathcal{T}}} \|f - \mathcal{P}_{\mathcal{T}}f\|^2_{-1;\omega_{\mathcal{T},z}}. \tag{12}$$

Algorithm (AFEM) Starting from $\mathcal{T}_0$ and $k = 0$, compute $(\mathbb{V}_k)_k$, $(U_k)_k$, $(\mathcal{T}_k)_k$ iteratively as follows, where step 3 will be specified further below.

1. Set $\mathbb{V}_k := \mathring{\mathbb{V}}(\mathcal{T}_k)$ and let $U_k \in \mathbb{V}_k$ be the solution of (4) with $\mathcal{T} = \mathcal{T}_k$.
2. Compute the indicators of estimator and oscillation in (10) and (11), respectively, and write $\mathcal{E}_k$ as an abbreviation for $\mathcal{E}_{\mathcal{T}_k}$.
3. Choose a subset $\mathcal{M}_k \subset \mathcal{T}_k$ with the help of the values of the indicators $\{\mathcal{E}_k(T)\}_{T \in \mathcal{T}_k}$ and $\{\delta(T)\}_{T \in \mathcal{T}_k}$.
4. Let $\mathcal{T}_{k+1}$ be the smallest refinement of $\mathcal{T}_k$ in $\mathbb{T}$ such that $\mathcal{M}_k \cap \mathcal{T}_{k+1} = \emptyset$.
5. Increment k and go to step 1.

3   Convergence and Marking Strategy

The above adaptive method generalizes global uniform refinement, which is the special case corresponding to $\mathcal{M}_k = \mathcal{T}_k$ for all k. The new key feature is that certain elements may not be refined anymore. In other words: it may happen that

$$\mathcal{T}^* := \{T \mid \exists m \in \mathbb{N}_0 \, \forall k \geq m \; T \in \mathcal{T}_k\} \neq \emptyset \quad \text{and so} \quad \mathbb{V}_\infty := \overline{\bigcup_k \mathbb{V}_k} \neq H_0^1(\Omega).$$

It is then the task of step 3, the marking strategy, to preclude $u \in H_0^1(\Omega) \setminus \mathbb{V}_\infty$, which is equivalent to non-convergence.

Let us first derive a necessary condition for the marking strategy. If $\lim_{k\to\infty} \|\nabla(u - U_k)\|_\Omega = 0$, then, for any element $T \in \mathcal{T}^*$, the local lower bounds in (10), (11), and Theorem 1 imply $\lim_{k\to\infty} \mathcal{E}_k(T) + \delta(T) = 0$ for the associated indicators. Hence it is necessary for convergence that the marking strategy ensures

$$\forall T \in \mathcal{T}^* \quad \delta(T) = 0 \text{ and } \lim_{k\to\infty} \mathcal{E}_k(T) = 0, \tag{13}$$

which specifies [8, (5.1)]. It turns out that this condition is also sufficient for convergence, as it complements the following property of our refinement procedure. Let h_k denote the meshsize function of $\mathcal{T}_k$ given by $h_{k|T} = |T|^{1/d}$ for all $T \in \mathcal{T}_k$ and let χ_k the characteristic function of $\bigcup\{T \in \mathcal{T}_k : T \notin \mathcal{T}_{k+1}\}$. Then [9, Lemma 9] shows

$$\lim_{k\to\infty} \|h_k \chi_k\|_{L^\infty(\Omega)} = 0, \tag{14}$$

which generalizes $\max_\Omega h_k \to 0$ for non-adaptive, uniform refinement.

Theorem 2 (Adaptive Convergence) *Let $f \in H^{-1}(\Omega)$ be arbitrary and assume that the indicators of the estimator satisfy* (10). *Then the approximate solutions $(U_k)_k$ of the AFEM converge to the exact solution u if and only if the marking strategy ensures* (13).

Proof We adopt the proof of [9, Theorem 8], which essentially follows [10]. Using the Lax-Milgram theorem, let $U_\infty \in \mathbb{V}_\infty$ be such that

$$\int_\Omega \nabla U_\infty \cdot \nabla V \, dx = \langle f, V \rangle \quad \text{for all } V \in \mathbb{V}_\infty.$$

Thanks to [9, Lemma 7], we have

$$\lim_{k\to\infty} \|\nabla(U_\infty - U_k)\|_\Omega = 0 \tag{15}$$

and it remains to show that $U_\infty = u$ or, as an alternative, that its residual vanishes:

$$0 = \langle R_\infty, v \rangle := \langle f, v \rangle - \int_\Omega \nabla U_\infty \cdot \nabla v \, dx \quad \text{for all } v \in H_0^1(\Omega). \tag{16}$$

Here we can take the test functions only from $C_0^\infty(\Omega)$, because $C_0^\infty(\Omega)$ is a dense subset of $H_0^1(\Omega)$. Consequently, the convergence (15) shows that (16) follows from

$$\forall \varphi \in C_0^\infty(\Omega) \qquad \lim_{k \to \infty} \langle R_k, \varphi \rangle = 0, \tag{17}$$

where $R_k := f + \Delta U_k \in H^{-1}(\Omega)$ is the residual of U_k.

In order to verify this, let $\varphi \in C_0^\infty(\Omega)$ and $k, \ell \in \mathbb{N}_0$ with $k \geq \ell$ and introduce the set $\mathcal{T}_\ell^* := \mathcal{T}_\ell \cap \mathcal{T}^*$. The inclusion $\mathbb{V}_\ell \subset \mathbb{V}_k$, the abstract error bound (7), the local equivalences (9), as well as the upper bounds in (10) and (11) imply

$$\langle R_k, \varphi \rangle = \langle R_k, \varphi - I_\ell \varphi \rangle \lesssim S_{\ell,k} + S_{\ell,k}^*, \tag{18}$$

where I_ℓ denotes Lagrange interpolation onto $\mathbb{V}_\ell$ and we expect that

$$S_{\ell,k} := \sum_{T \in \mathcal{T}_k \backslash \mathcal{T}_\ell^*} \big(\mathcal{E}_k(T) + \delta(T) \big) \|\nabla(\varphi - I_\ell \varphi)\|_{\omega_k(T)}$$

gets small because of decreasing meshsize, whereas

$$S_{\ell,k}^* := \sum_{T \in \mathcal{T}_\ell^*} \big(\mathcal{E}_k(T) + \delta(T) \big) \|\nabla(\varphi - I_\ell \varphi)\|_{\omega_k(T)}$$

gets small thanks to condition (13) on the marking strategy. Here $\omega_k(T) = \omega_{\mathcal{T}_k}(T)$ is the ring around T in $\mathcal{T}_k$.

We first deal with $S_{\ell,k}$. The lower bounds in (10), (11), and Theorem 1 entail that

$$\sum_{T \in \mathcal{T}_k \backslash \mathcal{T}_\ell^*} \big(\mathcal{E}_k(T) + \delta(T) \big)^2 \lesssim \|\nabla(U_k - U_\infty)\|_\Omega^2 + \|\nabla(U_\infty - u)\|_\Omega^2 \lesssim 1 \tag{19}$$

is uniformly bounded thanks to (15). Furthermore, standard error bounds for Lagrange interpolation on $\mathcal{T}_\ell$ yield

$$\sum_{T \in \mathcal{T}_k \backslash \mathcal{T}_\ell^*} \|\nabla(\varphi - I_\ell \varphi)\|_{\omega_k(T)}^2 \lesssim \sum_{T \in \mathcal{T}_\ell \backslash \mathcal{T}_\ell^*} \|\nabla(\varphi - I_\ell \varphi)\|_{\omega_\ell(T)}^2$$

$$\lesssim \|h_\ell \chi_\ell\|_{L^\infty(\Omega)}^2 \|D^2 \varphi\|_{L^\infty(\Omega)}^2 |\Omega|.$$

Hence the Cauchy-Schwarz inequality on the sum $S_{\ell,k}$ and (14) imply

$$S_{\ell,k} \to 0 \text{ as } \ell \to \infty \text{ uniformly in } k. \tag{20}$$

For $S^*_{\ell,k}$, condition (13) leads to

$$\sum_{T \in \mathcal{T}^*_\ell} \left(\mathcal{E}_k(T) + \delta(T)\right)^2 = \sum_{T \in \mathcal{T}^*_\ell} \mathcal{E}_k(T)^2 \to 0 \quad \text{as} \quad k \to \infty$$

and, since $\sum_{T \in \mathcal{T}^*_\ell} \|\nabla(\varphi - I_\ell \varphi)\|^2_{L^2(\omega_k(T))}$ is uniformly bounded, the Cauchy-Schwarz inequality provides

$$S^*_{\ell,k} \to 0 \text{ as } k \to \infty \text{ for any fixed } \ell. \tag{21}$$

To conclude, let $\epsilon > 0$ be arbitrary. We exploit (20) and (21) by first choosing ℓ so that $S_{\ell,k} \leq \epsilon/2$ and next $k \geq \ell$ so that $S^*_{\ell,k} \leq \epsilon/2$. Inserting this into (18) yields the desired convergence (17) and finishes the proof. $\square$

4 Convergent Marking Strategies

In this concluding section, we discuss how to ensure condition (13) on the marking strategy, which is necessary and sufficient for the convergence of the AFEM.

To this end, the following observation about a vanishing limit is helpful. If $(T_k)_{k \in \mathbb{N}_0}$ is a sequence of elements satisfying $T_k \in \mathcal{T}_k \setminus \mathcal{T}_{k+1}, k \in \mathbb{N}_0$, then

$$
\begin{aligned}
\mathcal{E}_k(T_k) + \delta(T_k) &\lesssim \|\nabla(u - U_k)\|_{\omega_k(T_k)} \\
&\leq \|\nabla(U_\infty - U_k)\|_{\omega_k(T_k)} + \|\nabla(u - U_\infty)\|_{\omega_k(T_k)} \to 0
\end{aligned}
\tag{22}
$$

as $k \to \infty$. In fact, the first term on the right-hand side vanishes thanks to (15) and the second term vanishes in view of $\omega_k(T_k) \lesssim |T_k| \lesssim \|h_k \chi_k\|^d_{L^\infty(\Omega)}$ and property (14).

Condition (13) is thus satisfied if we have

$$\max_{T' \in \mathcal{T}_k} \left(\delta(T') + \mathcal{E}_k(T')\right) \leq \max_{T' \in \mathcal{T}_k \setminus \mathcal{T}_{k+1}} \left(\delta(T') + \mathcal{E}_k(T')\right),$$

which, thanks to $\mathcal{M}_k \subset \mathcal{T}_k \setminus \mathcal{T}_{k+1}$, can be achieved by requiring that at least one element with maximal combined indicators is marked:

$$\mathcal{M}_k \cap \left\{T \in \mathcal{T}_k \mid \delta(T) + \mathcal{E}_k(T) \geq \max_{T' \in \mathcal{T}_k} \left(\delta(T') + \mathcal{E}_k(T')\right)\right\} \neq \emptyset. \tag{23}$$

This property is verified by most of the common marking strategies applied to the combined indicators $\delta(T) + \mathcal{E}_k(T)$, $T \in \mathcal{T}_k$. Important examples are maximum-, Dörfler-, and equal distribution strategy; cf. [8].

As an alternative to combined marking, one may mark the two indicators separately. Similarly as before, requiring (23) for the single indicator $\mathcal{E}_k(T)$ instead of the combined indicator $\mathcal{E}_k(T) + \delta(T)$ then implies

$$\forall T \in \mathcal{T}^* \quad \lim_{k \to \infty} \mathcal{E}_k(T) = 0.$$

To ensure also $\delta(T) = 0$ for all $T \in \mathcal{T}^*$, one may employ again the respective counterpart of (23) or a different approach, which, as fast tree approximation [1], capitalizes on the locality and history of the oscillation indicator. The following simple consequence of Theorem 2, which is also of interest by its own, is useful in this context.

Remark 3 (Modified Oscillation) For any simplex T in $\mathbb{T}$, let $\tilde{\delta}(T)$ be a modification or approximation of $\delta(T)$ such that $\tilde{\delta}(T) = 0 \implies \delta(T) = 0$. Then the AFEM, where each $\delta(T)$ is replaced by $\tilde{\delta}(T)$, converges if and only if $\lim_{k \to \infty} \mathcal{E}_k(T) = 0$ and $\tilde{\delta}(T) = 0$ for all $T \in \mathcal{T}^*$.

References

1. P. Binev, R. DeVore, Fast computation in adaptive tree approximation. Numer. Math. **97**(2), 193–217 (2004)
2. P. Binev, W. Dahmen, R. DeVore, P. Petrushev, Approximation classes for adaptive methods. Serdica Math. J. **28**(4), 391–416 (2002). Dedicated to the memory of Vassil Popov on the occasion of his 60th birthday
3. C. Carstensen, M. Feischl, D. Praetorius, Axioms of adaptivity. Comput. Math. Appl. **67**(6), 1195–1253 (2014)
4. A. Cohen, R. DeVore, R.H. Nochetto, Convergence rates of AFEM with H^{-1} data. Found. Comput. Math. **12**(5), 671–718 (2012)
5. L. Diening, C. Kreuzer, R. Stevenson, Instance optimality of the adaptive maximum strategy. Found. Comput. Math. **16**(1), 33–68 (2016)
6. F.D. Gaspoz, P. Morin, Approximation classes for adaptive higher order finite element approximation. Math. Comput. **83**(289), 2127–2160 (2014)
7. C. Kreuzer, A. Veeser, Oscillation in a posteriori error estimation. (in preparation)
8. P. Morin, K.G. Siebert, A. Veeser, A basic convergence result for conforming adaptive finite elements. Math. Models Methods Appl. Sci. **18**(5), 707–737 (2008)
9. R.H. Nochetto, A. Veeser, Primer of adaptive finite element methods, in *Multiscale and Adaptivity: Modeling, Numerics and Applications*. Lecture Notes in Mathematics, vol. 2040 (Springer, Heidelberg, 2012), pp. 125–225
10. K.G. Siebert, A convergence proof for adaptive finite elements without lower bound. IMA J. Numer. Anal. **31**(3), 947–970 (2011)