4. Lecture 4. Complexity of AFEM

Despite the overwelming computational evidence that AFEM lead to optimal meshes, the mathematical theory started very recently with Binev et al [4] and Stevenson [44]. Paper [4] discusses geometric properties of bisection, reported in section 4.1, which turn out to be crucial for optimal complexity; the AFEM in [4] includes a *coarsening* step which is not necessary in practice for linear elliptic PDE. Paper [44], instead, analyzes the AFEM of Lecture 2 and shows optimal complexity without coarsening. We describe the main ideas of [44] section 4.2 and 4.3. We conclude in sections 4.4 and 4.5 with optimality bounds taken from [10] for general elliptic PDE such as (1.3).

4.1. **Procedure REFINE: Properties of Bisection.** We now study geometric properties of bisection [2, 4, 31, 33]. We already described the algorithm in section 2.3 as is the one used by ALBERT for mesh refinement [43].

Let $\mathcal{T}_k$ be the conforming triangulation generated AFEM in the k-th step. We denote $\mathcal{T}_{k,*}$ a refinement of $\mathcal{T}_k$ (in general non-conforming) satisfying the *Interior Node Property*, and $\mathcal{M}(\mathcal{T}_{k,*})$ the set of elements of $\mathcal{T}_k$ that were refined in $\mathcal{T}_{k,*}$. Finally, $\overline{\mathcal{M}}(\mathcal{T}_{k,*})$ is the set of elements of $\mathcal{T}_k$ with nonempty intersection with $\mathcal{M}(\mathcal{T}_{k,*})$; in particular, note that $\mathcal{M}(\mathcal{T}_{k,*}) \subset \overline{\mathcal{M}}(\mathcal{T}_{k,*})$ and $\#\mathcal{M}(\mathcal{T}_{k,*}) \le C^* \#\overline{\mathcal{M}}(\mathcal{T}_{k,*})$ with C^* a universal constant solely depending on the shape regularity constant γ^* and the dimension d. This is a consequence of $\#N(T) \le C^*$ for all $T \in \mathcal{T}_k$.

If an adjacent element to one marked for refinement happens to have a refinement edge other than the common edge, then the adjacent element must be refined first recursively. This generates a compatible refinement edge of the original element and its (new) neighbour. This recursion is guaranteed to stop for every element of a refined triangulation if the recursive refinement does not create cycles on the (usually very coarse) initial triangulation $\mathcal{T}_0$ [33]; this can be easily checked for a triangulation with a given choice of refinement edges. For three dimensions, this is proved only under some additional assumptions on the initial triangulation [31]. In Figure 4.1 we show a situation where recursion is needed. For all triangles, the longest edge is the refinement edge. Let us assume that triangles A and B are marked for refinement. Triangle A can be refined at once, as its refinement edge is a boundary edge. For refinement of triangle B, we have to recursively refine triangles C and D. Again, triangle D can be directly refined, so recursion stops there. This is shown in the second part of the figure. Back in triangle C, this can now be refined together with its neighbour. After this, also triangle B can be refined together with its neighbour.

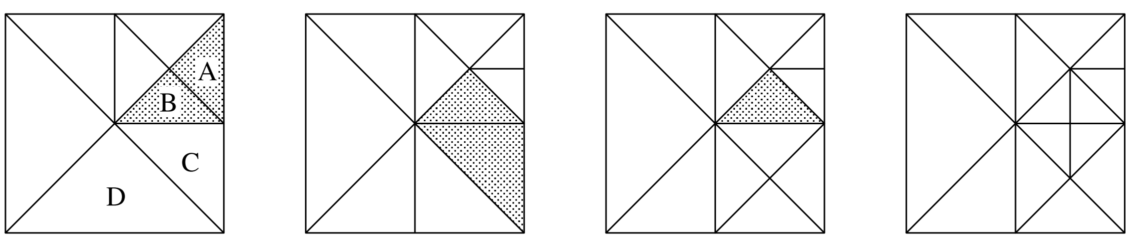

Figure 4.1. Recursive bisection refinement: Triangles A, B are initially marked for refinement. Triangle B yields a concatenation of refinements that stops in triangle D.

As shown in Figure 4.1, the concatenation of refinements of $\mathcal{T}_k$ to maintain conformity may involve several refinement levels, perhaps as many as k. This hints at the difficulties in bounding the number of degree of freedom $\#\mathcal{T}_{k+1}$ in terms of those marked $\#\mathcal{M}(\mathcal{T}_{k,*})$ with a constant independent of k. The following geometric estimate of Binev et al [4] and Stevenson [45] is crucial in this respect.

Proposition 4.1 (Complexity of REFINE). *Let $\{\mathcal{T}_k\}_{k\ge0}$ be a sequence of conforming nested partitions generated by REFINE starting from $\mathcal{T}_0$. Let $\mathcal{M}(\mathcal{T}_{k,*})$ be the set of elements of $\mathcal{T}_k$ marked for refinement and $\mathcal{T}_{k,*}$ be the (nonconforming) partition created by refinement of elements just in $\mathcal{M}(\mathcal{T}_{k,*})$. Then* *there is a constant C_0 solely dependent on $\mathcal{T}_0$ such that*

$$\#\mathcal{T}_{k+1} - \#\mathcal{T}_0 \le C_0 \sum_{i=1}^{k} \#\mathcal{T}_{i,*} - \#\mathcal{T}_i. \tag{4.1}$$

In what follows we will repeatedly used the notation $a \preccurlyeq b$ to indicate $a \le Cb$ with a constant independent of the main parameters involved.

4.2. **Localized Upper Bound and Optimal Marking.** The next key estimate is due to Stevenson [44] and gives an upper bound between two consecutive discrete solutions in terms of the error indicators of the coarser solution, but restricted to the elements in which the partition differ. We refer to Lemma 2.1 for a global upper bound.

Proposition 4.2 (Localized Upper Estimate)**.** *Let $u_k \in \mathbb{V}_k$ and $u_{k,*} \in \mathbb{V}_{k,*}$ be a discrete solutions over a conforming mesh $\mathcal{T}_k$ and its nonconforming refinement $\mathcal{T}_{k,*}$ with marked elements $\mathcal{M}(\mathcal{T}_{k,*})$. Then*

$$\|u_k - u_{k,*}\|_{\mathbb{V}} \preccurlyeq \sum_{T \in \overline{\mathcal{M}}(\mathcal{T}_{k,*})} \eta_k(T)^2. \tag{4.2}$$

Proof. Since $\mathbb{V}_{k,*} \subset \mathbb{V}_k$ are nested subspaces of $\mathbb{V}$, we can proceed as in Lemma 2.1 with test function $v = u_k - u_{k,*} \in \mathbb{V}_{k,*}$. Since the definition of Clement interpolation operator I_k of Proposition 1.6 on an element T involves its neighborhood $N(T)$, we realize that $I_k v = v$ for all $T \notin \overline{\mathcal{M}}(\mathcal{T}_{k,*})$. Therefore, the same argument as in the proof of Lemma 2.1 applies with $\phi = v - I_k v$ and leads to (4.2). □

The following result of Stevenson is a consequence of Proposition 4.2. It allows us to relate an optimal partition in terms of degrees of freedom with the Marking Strategy E of Dörfler [21], which has optimal properties in terms of energy. We assume

- The PDE is (2.19) with $\mathbf{A} = \mathbf{I}$: $-\Delta u = f$
- The forcing function f is piecewise constant over $\mathcal{T}_0$, whence $\mathrm{osc}_k = 0$ for all $k \ge 1$
- The polynomial degree is 1.

A by-product of these assumptions is the equivalence of energy error and estimator (see Lemmas 2.1 and 2.2); note that constant C_2 is proportional to that in (2.7):

$$C_2 \sum_{T \in \mathcal{T}_k} \eta_k(T)^2 \le \|u - u_k\|^2 \le C_1 \sum_{T \in \mathcal{T}_k} \eta_k(T)^2. \tag{4.3}$$

The ratio $\frac{C_1}{C_2} \ge 1$ is a measure of the precision of the indicators $\{\eta_k(T)\}_T$: the closer to 1 the better!

Corollary 4.3 (Optimal Marking)**.** *Let the previous assumptions be valid. Let $u_k \in \mathbb{V}_k$ and $u_{k,*} \in \mathbb{V}_{k,*}$ be as in Proposition 4.2. Suppose that they satisfy the energy decrease property*

$$\|u - u_{k,*}\|_{\mathbb{V}} \le \lambda \|u - u_k\|_{\mathbb{V}}, \tag{4.4}$$

with $0 < \lambda < 1$. Then the set $\mathcal{M}(\mathcal{T}_{k,*})$ of marked elements satisfies the Dörfler property

$$\sum_{T \in \overline{\mathcal{M}}(\mathcal{T}_{k,*})} \eta_k(T)^2 \ge \theta^2 \sum_{T \in \mathcal{T}_k} \eta_k(T)^2, \tag{4.5}$$

with $\theta^2 = \frac{C_2}{C_1}(1 - \lambda^2)$.

Proof. We invoke the orthogonality relation of Lemma 2.11, which is valid for (2.19), to write

$$\begin{aligned} C_1 \sum_{T \in \overline{\mathcal{M}}(\mathcal{T}_{k,*})} \eta_k(T)^2 &\ge \|u_{k,*} - u_k\|_{\mathbb{V}}^2 = \|u - u_{k,*}\|_{\mathbb{V}}^2 - \|u - u_k\|_{V}^2 \\ &\ge (1 - \lambda^2)\|u - u_k\|_{\mathbb{V}}^2 \ge C_2(1 - \lambda^2) \sum_{T \in \mathcal{T}_k} \eta_k(T)^2. \end{aligned}$$

This implies (4.5) with the asserted parameter θ. □

We immediately see from Corollary 4.3 that the range of θ is limited by the precision of the indicators $\{\eta_k(T)\}_T$. Since $\lambda > 0$, we need a more conservative marking with θ satisfying

$$0 < \theta < \frac{C_2}{C_1}. \tag{4.6}$$

4.3. **Optimal Complexity I: The Simplest Case.** We now show that (4.6) is the bridge between the minimality property of Dörfler marking and optimal partitioning for a given accuracy. To this end, we need to introduce an approximation class

$$\mathcal{A}_s := \Big\{ v \in \mathbb{V} : |v|_s = \sup_{\varepsilon > 0} \ \varepsilon \inf_{\mathcal{T} \subset \mathcal{T}_0 :\ \inf \|u - u_{\mathcal{T}}\|_{\mathbb{V}} \le \varepsilon} \Big(\#\mathcal{T} - \#\mathcal{T}_0 \Big)^s < \infty \Big\}. \tag{4.7}$$

and equip it with the norm $\|v\|_s = \|v\|_{\mathbb{V}} + |v|_s$. So $\mathcal{A}_s$ is the class of functions that can be approximated within a given tolerance ε by continuous piecewise linear functions on a partition $\mathcal{T}$ with degrees of freedom $\#\mathcal{T}_k - \#\mathcal{T}_0 \preccurlyeq \varepsilon^{-1/s} |v|_s^{1/s}$. Classical estimates show that for $s \le \frac{1}{2}$ we have $H^{1+2s}(\Omega) \cap \mathbb{V} \subset \mathcal{A}_s$, in which case uniform refinement is optimal. On the other hand, the class is much larger and related to Besov regularity [5].

Lemma 4.4 (Upper Bound of DOF). *In addition to the previous assumptions, let $u \in \mathcal{A}_s$. Let $\mathcal{T}_k$ be a conforming partition obtained from $\mathcal{T}_0$. Let $\mathcal{T}_{k,*}$ be a nonconforming mesh created from $\mathcal{T}_k$ upon marking the set $\mathcal{M}(\mathcal{T}_{k,*})$ according with the Dörfler marking with θ_E satisfying* (4.6). *Then*

$$\#\mathcal{T}_{k,*} - \#\mathcal{T}_k \preccurlyeq \|u - u_k\|_{\mathbb{V}}^{-1/s} |u|_s^{1/s}, \tag{4.8}$$

where the hidden constant depends on the discrepancy between θ_E and $\frac{C_2}{C_1}$.

Proof. Let λ and ε satisfy $0 < \lambda^2 = 1 - \frac{C_1}{C_2}\theta_E < 1$ and $\varepsilon = \lambda \|u - u_k\|_V$. Let $\mathcal{T}_\varepsilon$ be a refinement of $\mathcal{T}_0$ with minimal degrees of freedom satisfying $\|u - u_{\mathcal{T}_\varepsilon}\|_{\mathbb{V}} \le \varepsilon$ and

$$\#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \preccurlyeq \varepsilon^{-1/s} |u|_s^{1/s} \le \lambda^{-1/s} \|u - u_k\|_{\mathbb{V}}^{-1/s} |u|_s^{1/s}.$$

Let now $\mathcal{T}_k^+$ be the smallest (nonconforming) common refinement of $\mathcal{T}_k$ and $\mathcal{T}_\varepsilon$. Since both $\mathcal{T}_k$ and $\mathcal{T}_\varepsilon$ are refinements of $\mathcal{T}_0$, the number of triangles in $\mathcal{T}_k^+$ that are not in $\mathcal{T}_k$ is less than the number of triangles that must be added to go from $\mathcal{T}_0$ to $\mathcal{T}_\varepsilon$, namely,

$$\#\mathcal{T}_k^+ - \#\mathcal{T}_k \le \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0.$$

Since the energy error is monotone and $\mathcal{T}_k^+$ is a refinement of $\mathcal{T}_\varepsilon$, we see that

$$\|u - u_{\mathcal{T}_k^+}\|_{\mathbb{V}} \le \|u - u_{\mathcal{T}_\varepsilon}\|_{\mathbb{V}} \le \varepsilon = \lambda \|u - u_k\|_V,$$

whence $\mathcal{T}_k^+$ satisfies a Dörfler property according to Corollary 4.3. On the other hand, Marking Strategy E selects a minimal set $\mathcal{M}(\mathcal{T}_{k,*}) = \hat{\mathcal{T}}_k$ with the Dörfler property, which implies that the ensuing nonconforming partition $\mathcal{T}_{k,*}$ satisfies

$$\#\mathcal{T}_{k,*} - \#\mathcal{T}_k \preccurlyeq \#\mathcal{T}_k^+ - \#\mathcal{T}_k \le \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \preccurlyeq \lambda^{-1/s} \|u - u_k\|_{\mathbb{V}}^{-1/s} |u|_s^{1/s}.$$

This is the desired (4.8) with an explicit dependence on the discrepance between θ_E and $\frac{C_2}{C_1}$ via λ. $\square$

We are now in the position to show that AFEM possesses optimal complexity. To this end we have to accumulate via Proposition 4.1 the effect of each adaptive loop quantified in Lemma 4.4.

Theorem 4.5 (Optimal Complexity 1). *Let f be piecewise constant in $\mathcal{T}_0$ and let $u \in \mathcal{A}_s$ for $0 < s \le 1/2$ be the solution of $-\Delta u = f$ in a polyhedral domain Ω of $\mathbb{R}^d$ with vanishing Dirichlet boundary condition. Then, the k-th iterate $u_k \in \mathbb{V}_k$ of AFEM satisfies the optimal bound*

$$\|u - u_k\|_V \preccurlyeq (\#\mathcal{T}_k - \#\mathcal{T}_0)^{-1/s} |u|_s^{1/s}, \tag{4.9}$$

where the hidden constant depends on the discrepancy between θ_E and $\frac{C_2}{C_1}$.

Proof. We first note that since $\mathrm{osc}_k = 0$ Theorem 2.12 implies the error reduction

$$\|u - u_{k+1}\|_{\mathbb{V}} \le \alpha \|u - u_k\|_{\mathbb{V}},$$

whence for $0 < k \le n$

$$\|u - u_k\|_{\mathbb{V}}^{-1/s} \le \alpha^{(n-k)/s} \|u - u_n\|_{\mathbb{V}}^{-1/s}.$$

We next employ Proposition 4.1 in conjunction with Lemma 4.4 to deduce

$$\begin{aligned}
\#\mathcal{T}_n - \#\mathcal{T}_0 &\preccurlyeq \sum_{k=0}^{n-1} \#\mathcal{T}_{k,*} - \#\mathcal{T}_k \\
&\preccurlyeq |u|_s^{1/s} \sum_{k=0}^{n-1} \|u - u_k\|_{\mathbb{V}}^{-1/s} \\
&\preccurlyeq \|u - u_n\|_{\mathbb{V}}^{-1/s} |u|_s^{1/s} \sum_{k=1}^{n} \alpha^{\frac{k}{s}} \\
&\preccurlyeq \|u - u_n\|_{\mathbb{V}}^{-1/s} |u|_s^{1/s},
\end{aligned}$$

because $\alpha < 1$ makes the sum bounded independently of n. This is the desired estimate in disguise. □

We note that Theorem 4.5 asserts that the error decay is the optimal one expected for a membership in the approximation class $\mathcal{A}_s$. Theorem 4.5 does not account for possible degeneracies of u and corresponding higher rates.

4.4. **Marking Revisited: Towards Optimal Complexity.** If $\mathrm{osc}_k \ne 0$, then the above proof have to be modified considerably. Instead of the approach proposed by Stevenson in [44], we follow here the more recent approach of Cascón and Nochetto [10] that simplifies that in [44] and applies to general operators $\mathcal{L}$ as in (1.3).

We start by showing that Marking Strategy O, even though effective in reducing oscillation, may lead to suboptimal meshes. The following counterexample is a simple modification of Example 1.3 studied computationally in section 2.6.1

$$u = u_K + u_S.$$

Here u_K is the function of Kellogg [29] discussed in Example 1.3 with $\gamma = 0.1$, and u_S is the oscillatory function

$$u_s = 10^{-6} a_i^{-1} \sin^2(10\pi x) \sin^2(10\pi y)$$

where a_i, being constant in each quadrant, is the diffusion coefficient given in Example 1.3. We observe that u_S is much smaller in magnitude than u_K and is also smooth, but it leads to a small amount of data oscillation. Since Marking Strategy O reduces data oscillation in step $k+1$ relative to the preceeding value in step k, the absolute magnitude of data oscillation relative to the error is inmaterial. This fact is an early indication that Marking Strategy O may yield suboptimal meshes and this is confirmed computationally and depicted in Figures 4.2 and 4.3.

4.5. **Optimal Complexity II: The general Case.** In order to decrease the oscillation relative to the error size we need to make a few important modifications in AFEM. We start by defining a new notion of oscillation which majorizes the previous one osc_k and incorporates the variation of all data within an element $T \in \mathcal{T}_k$:

$$\begin{aligned}
\mathrm{osc}(f,T)^2 &:= h_T^2 \left\| f - \bar{f}_T \right\|_{L^2(T)}^2 \\
\mathrm{osc}(\mathbf{A},\mathbf{b},c;T)^2 &:= h_T^2 \,\mathrm{osc}(D\mathbf{A},T)^2 + h_T^2 \,\mathrm{osc}(\mathbf{b},T)^2 + h_T^2 \,\mathrm{osc}(c,T)^2 \\
\mathrm{osc}(\mathbf{A},\mathbf{b},c,u_k;T)^2 &:= \mathrm{osc}(\mathbf{A},\mathbf{b},c;T)^2 \,|\!|\!| u_k |\!|\!|_T^2 \,,
\end{aligned}$$

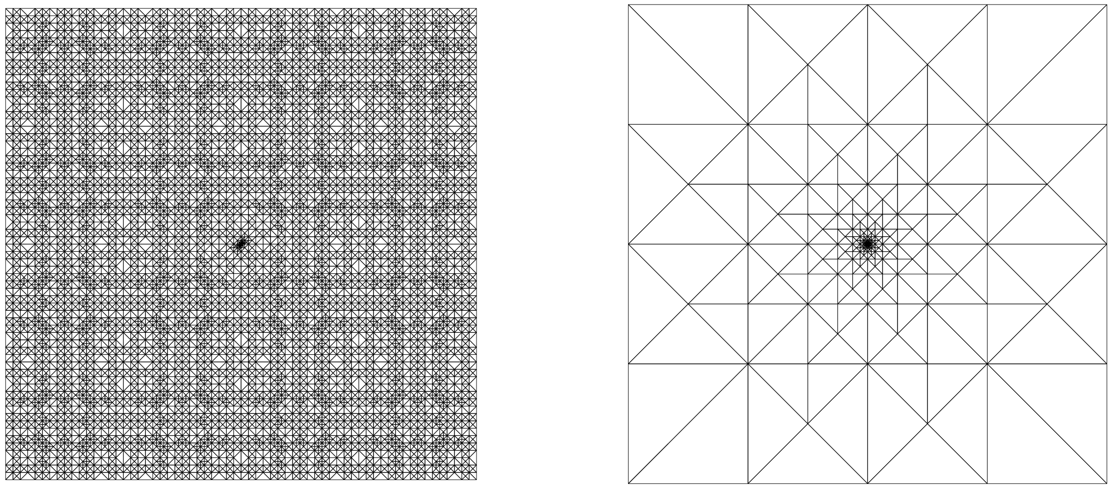

FIGURE 4.2. Refined mesh after $k = 15$ iterations for the AFEM MNS of Morin-Nochetto-Siebert of Lecture 2 (left: 8077 degrees of freedom) and by the new AFEM (right: 363 degrees of freedom) with $\theta_E = \theta_0 = 0.4$. It is quite evident that Marking Strategy O yields a suboptimal performance in terms of degrees of freedom.

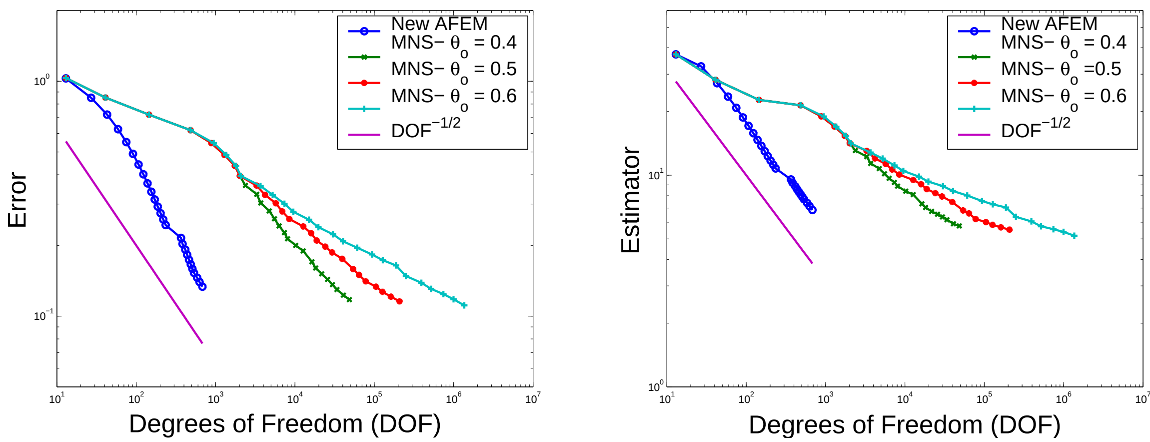

FIGURE 4.3. Decay rates of error (left) and estimator (right) for the new AFEM vs the AFEM MNS of Morin-Nochetto-Siebert of Lecture 2 for several values of parameter $\theta_0 = 0.4, 0.5, 0.6$ and $\theta_E = 0.4$. As the marking parameter θ_0 increases from 0.4 to 0.6 the curve flattens out thereby showing lack of optimality and its sensitivity to θ_0. On the other hand, the new marking exhibits optimal decay.

where

$$\operatorname{osc}(D\mathbf{A}, T) := \max_{1\le i,j\le d} \max_{x,y\in T} |\partial_i \{a_{ij}(x) - a_{ij}(y)\}|$$

$$\operatorname{osc}(\mathbf{b}, T) := \max_{1\le i\le d} \max_{x,y\in T} |\mathbf{b}_i(x) - \mathbf{b}_i(y)|$$

$$\operatorname{osc}(c, T) := \max_{x\in T} |c(x)| + h^{-\frac{4}{p}} \left\{ \frac{1}{|T|} \int_T \left(\frac{1}{|T|} \int_T [c(x) - c(y)]^2 \, dx \right)^{\frac{q}{2}} dy \right\}^{\frac{2}{q}}.$$

We now define the new element oscillation $\operatorname{osc}(u_k, T)$ and element indicator $\zeta_k(T)$ to be

$$\operatorname{osc}(u_k, T)^2 := \operatorname{osc}(f, T)^2 + \operatorname{osc}(\mathbf{A}, \mathbf{b}, c, u_k; T)^2$$
$$\zeta_k(T)^2 := \eta_k(T)^2 + \operatorname{osc}(u_k, T)^2,$$

where $\eta_k(T)$ is the error indicator already. introduced in (2.4). For their global contributions we employ the following notation:

$$\eta_k^2 = \sum_{T\in\mathcal{T}_k} \eta_k(T)^2,$$
$$\operatorname{osc}(u_k, \mathcal{T}_k)^2 = \sum_{T\in\mathcal{T}_k} \operatorname{osc}(u_k, T)^2,$$
$$\operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_k)^2 = \max_{T\in\mathcal{T}_k} \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; T)^2,$$
$$\zeta_k^2 = \sum_{T\in\mathcal{T}_k} \zeta_k(T)^2.$$

4.5.1. *New Marking Strategy and AFEM.* We still utilize the Dörfler marking as follows:

- MARK_1: This procedure selects a set $\mathcal{M}(\mathcal{T}_{k,1}) \subset \mathcal{T}_k$ satisfing the following property for the element indicator ζ_k^2,
$$\sum_{T\in\mathcal{M}(\mathcal{T}_{k,1})} \zeta(T)^2 \geq \theta_1^2 \sum_{T\in\mathcal{T}_k} \zeta(T)^2.$$
This marking gives rise to a (nonconforming) partition $\mathcal{T}_{k,1}$ of $\mathcal{T}_k$:
$$\mathcal{T}_{k,1} := \mathsf{MARK}_1(\theta_1, \zeta_k).$$
- MARK_2: Given parameters $\theta_2, \varepsilon_k \in (0,1)$, this procedure performs several steps of Dörfler marking with θ_2 on $\{\operatorname{osc}(u_k, T)\}_T$, where u_k is taken as a constant, and generates a (nonconforming) refinement $\mathcal{T}_{k,2}$
$$\mathcal{T}_{k,2} := \mathsf{MARK}_2(\theta_2, \operatorname{osc}(u_k, \mathcal{T}_k), \varepsilon_k)$$
satisfying the condition,
$$\operatorname{osc}(u_k, \mathcal{T}_{k,2})^2 \leq \varepsilon_k.$$

Note that the tolerance for MARK_2 depends on u_k which is kept fixed in this process. So execution of MARK_2 does not require any call to SOLVE to update u_k in the subsequent refinements of $\mathcal{T}_k$ until the final partition $\mathcal{T}_{k,2}$ is reached.

We are now in a position to write the new AFEM. We assume that the initial partition $\mathcal{T}_0$ satisfies the following two restrictions:

$$C^* h_0^s \|\mathbf{b}\|_{L^\infty(\Omega)} < 1 \tag{4.10}$$
$$8\Lambda_0 C_1 \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_0)^2 < \delta \tag{4.11}$$

where $\delta > 0$ is a (rather technical) constant to be determined in Theorem 4.6 and $\Lambda_0 = (1 - C^* h_0^s \|\mathbf{b}\|_{L^\infty(\Omega)})^{-1} > 1$ is the constant in (3.1).

AFEM

Choose parameters $0 < \theta_1, \theta_2, \delta < 1$, and initial mesh $\mathcal{T}_0$ satisfying (4.10) and (4.10). Set $k = 0$.

(1) $u_k := \mathsf{SOLVE}(\mathcal{T}_k, u_{k-1})$.
(2) $\{\eta_k^2, \operatorname{osc}(u_k, \mathcal{T}_k)^2\} := \mathsf{ESTIMATE}(\mathcal{T}_k, u_k)$.
(3) $\mathcal{T}_{k,1} := \mathsf{MARK}_1(\theta_1, \zeta_k^2)$.
(4) If $\left(\operatorname{osc}(u_k, \mathcal{T}_k)^2 > \varepsilon_k := \frac{\delta}{8}\zeta_k^2\right)$ then $\mathcal{T}_{k,2} := \mathsf{MARK}_2(\theta_2, \operatorname{osc}(u_k, \mathcal{T}_k)^2, \varepsilon_k)$.
(5) $\mathcal{T}_{k+1} := \mathsf{REFINE}(\mathcal{T}_{k,1}, \mathcal{T}_{k,2})$.
(6) Update $k \leftarrow k + 1$, and go to step (1).

We stress that in step (4) of AFEM we reduce oscillation $\operatorname{osc}(u_k, \mathcal{T}_k)$ only relative to the estimator ζ_k as suggested by the counterexample of section 4.4

4.5.2. *Convergence and Complexity of AFEM.* The next challenge is to prove convergence of AFEM. The difficulty arises from the fact that

Theorem 4.6 (Convergence of AFEM)**.** *Let $\{u_k\}_k$ be the sequence of discrete solutions produced by AFEM There exist constants $0 < \alpha < 1$ and $\beta > 0$, such that two consecutive iterates satisfy*

$$\|u - u_{k+1}\|_\Omega^2 + \beta \operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le \alpha^2 \left(\|u - u_k\|_\Omega^2 + \beta \operatorname{osc}(u_k, \mathcal{T}_k)^2 \right).$$

Proof. As in Theorems 2.13 and 3.4, we use the notation

$$e_k^2 := \|u - u_k\|_\Omega^2, \qquad \epsilon_k^2 := \|u_{k+1} - u_k\|_\Omega^2.$$

In view of the upper bound (2.1), with $C_1 \ge 1$, the definition of ζ_k, and procedure MARK_1, we get the following expresion for the error in terms of two consecutive discrete solutions and oscillation

$$\begin{aligned} e_k^2 + \operatorname{osc}(u_k, \mathcal{T}_k)^2 &\le C_1 \eta_k^2 + \operatorname{osc}(u_k, \mathcal{T}_k)^2 \le C_1 \zeta_k^2 \\ &\le \frac{C_1}{\theta_1^2} \sum_{T \in \mathcal{M}(\mathcal{T}_{k,1})} \zeta_k(T)^2 \\ &\le \frac{C_1}{\theta_1^2 C_2} \epsilon_k^2 + \left(\frac{C_1}{\theta_1^2 C_2} + \frac{C_1}{\theta_1^2} \right) \operatorname{osc}(u_k, \mathcal{T}_k)^2. \end{aligned}$$

We define the constants $\Lambda_1 := \frac{\theta_1^2 C_2}{C_1}$, and $\Lambda_2 := 1 + C_2 \left(1 - \frac{\theta_1^2}{C_1}\right)$, and thereby obtain,

$$\epsilon_k^2 \ge \Lambda_1 e_k^2 - \Lambda_2 \operatorname{osc}(u_k, \mathcal{T}_k)^2.$$

Using Lemma 3.1 about quasi-orthogonality, we easily see that

$$e_{k+1}^2 \le (\Lambda_0 - \Lambda_1) e_k^2 + \Lambda_2 \operatorname{osc}(u_k, \mathcal{T}_k)^2. \tag{4.12}$$

On the other hand, the following expression is valid for all $\mu > 0$ between consecutive oscillations (the proof is similar to that of Lemma 3.3)

$$\operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le (1 + \mu) \operatorname{osc}(u_k, \mathcal{T}_k)^2 + (1 + \mu^{-1}) \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_k)^2 \varepsilon_k^2.$$

We choose $\mu = 1$ and use the upper bound (2.5) to write

$$\operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le 2 \operatorname{osc}(u_k, \mathcal{T}_k)^2 + 2C_1 \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_k)^2 \eta_k^2.$$

We next argue according to whether the conditional of step (4) of AFEM is verified or not. If $\operatorname{osc}(u_k, \mathcal{T}_k)^2 \le \varepsilon_k$, then the mesh is no longer modified and

$$\operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le \frac{\delta}{4} \zeta_k^2 + \frac{\delta}{8} \zeta_k^2 < \delta \zeta_k^2.$$

If instead $\operatorname{osc}(u_k, \mathcal{T}_k)^2 > \varepsilon_k$, then MARK_2 reduces the oscillation so that on exit we have $\operatorname{osc}(u_k, \mathcal{T}_{k,2})^2 \le \varepsilon_k$. Consequently, after repeated use of the relation between consecutive oscillations, we obtain

$$\begin{aligned} \operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le& 2 \operatorname{osc}(u_{k,2}, \mathcal{T}_{k,2})^2 + 2 \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_{k,2})^2 \|u_{k+1} - u_{k,2}\|_\Omega^2 \\ \le& 4 \operatorname{osc}(u_k, \mathcal{T}_{k,2})^2 + 4 \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_{k,2})^2 \left(\|u_{k+1} - u_{k,2}\|_\Omega^2 + \|u_{k,2} - u_k\|_\Omega^2 \right), \end{aligned}$$

whence

$$\operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le 4\varepsilon_k + 4\Lambda_0 C_1 \operatorname{osc}(\mathbf{A}, \mathbf{b}, c; \mathcal{T}_k)^2 \eta_k^2 \le \frac{\delta}{2} \zeta_k^2 + \frac{\delta}{2} \eta_k^2 \le \delta \zeta_k^2.$$

Note that we have resorted here to the quasi-orthogonality relation (3.1) and the upper bound (2.5)

$$\|u_{k+1} - u_{k,2}\|_\Omega^2 + \|u_{k,2} - u_k\|_\Omega^2 \le \Lambda_0 \|u_{k+1} - u_k\|_\Omega^2 \le \Lambda_0 C_1 \eta_k^2.$$

Since $\operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le \delta \zeta_k^2$ in either case, we have for all $\beta > 0$

$$\beta \operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le \beta \delta \zeta_k^2 \le \frac{\beta \delta}{C_2} e_k^2 + \frac{\beta \delta}{C_2} (1 + C_2) \operatorname{osc}(u_k, \mathcal{T}_k)^2.$$

Adding this expression to (4.12), we obtain

$$e_{k+1}^2 + \beta \operatorname{osc}(u_{k+1}, \mathcal{T}_{k+1})^2 \le \left(\Lambda_0 - \Lambda_1 + \frac{\beta\delta}{C_2}\right) e_k^2 + \left(\Lambda_2 + \frac{\beta\delta(1+C_2)}{C_2}\right) \operatorname{osc}(u_k, \mathcal{T}_k)^2.$$

The asserted contraction property follows at once provided we can choose $\alpha \in (0,1)$ such that,

$$\Lambda_0 - \Lambda_1 + \frac{\beta\delta}{C_2} = \alpha^2 = \frac{\Lambda_2 + \frac{\beta\delta(1+C_2)}{C_2}}{\beta}.$$

To see that this is possible, we first h_0 sufficiently small so that

$$h_0^s \|\mathbf{b}\|_{L^\infty(\Omega)} \le \frac{\Lambda_1}{C^*(1+2\Lambda_1)} \quad \Rightarrow \quad \Lambda_0 - \Lambda_1 \le 1 - \frac{\Lambda_1^2}{1+\Lambda_1} < 1,$$

and then $\alpha^2 \in (1 - \frac{\Lambda_1^2}{1+\Lambda_1}, 1)$. Having chosen α, a simple calculation shows that β and δ are given by

$$\beta = \frac{\Lambda_2 + (1+C_2)(\alpha^2 - \Lambda_0 + \Lambda_1)}{\alpha^2}, \qquad \delta = \frac{(\alpha^2 - \Lambda_0 + \Lambda_1) C_2 \alpha^2}{\Lambda_2 + (1+C_2)(\alpha^2 - \Lambda_0 + \Lambda_1)}.$$

These explicit expressions for β and δ conclude the proof. □

This Theorem is the first key ingredient to prove optimal complexity of AFEM. The next step is to identify an approximation class, but turns out to be much more involved than $\mathcal{A}_s$ in section (4.7) because now we have to approximate not only the solution but also the coefficients and they interact in a nonlinear fashion. We consider the following quantity

$$|(u, f, \mathbf{A}, \mathbf{b}, c)|_s := \sup_{\epsilon > 0} \epsilon \Big(\inf_{\|u - u_{\mathcal{T}}\|_\Omega + \operatorname{osc}(u, \mathcal{T}) < \epsilon} (\#\mathcal{T} - \#\mathcal{T}_0)^s \Big). \tag{4.13}$$

This defines the concept of optimality of a partition $\mathcal{T}$ in terms of number of degrees of freedom $\#\mathcal{T}$ to approximate the combined quantity $\|u - u_{\mathcal{T}}\|_\Omega + \operatorname{osc}(u, \mathcal{T})$ within tolerance ϵ. We note that Theorem 4.6 establishes geometric reduction of a related quantity, namely $\|u - u_{\mathcal{T}}\|_\Omega + \operatorname{osc}(u_{\mathcal{T}}, \mathcal{T})$.

We are now in a position to state, but not prove, the complexity result alluded to earlier. The proof along with applications in given in [10].

Theorem 4.7 (Optimal Complexity 2)**.** *Let* $\{u_k, \mathcal{T}_k\}_{k>0}$ *be the sequence of finite element solutions and nested meshes generated by AFEM. Then*

$$\|u - u_k\|_\Omega + \operatorname{osc}(u_k, \mathcal{T}_k) \preccurlyeq (\#\mathcal{T}_k - \#\mathcal{T}_0)^{-s} \, |(u, f, \mathbf{A}, \mathbf{b}, c)|_s.$$

4.6. **Exercises.**

4.6.1. *Exercise: Relation between Oscillations.* Show that the new oscillation $\operatorname{osc}(\mathbf{A}, \mathbf{b}, c, u_k; T)^2$ bounds from above the old oscillation of Lecture 2.