CONVERGENCE OF ADAPTIVE FINITE ELEMENT METHODS FOR GENERAL SECOND ORDER LINEAR ELLIPTIC PDE

KHAMRON MEKCHAY * AND RICARDO H. NOCHETTO†

Abstract. We prove convergence of adaptive finite element methods (AFEM) for general (non-symmetric) second order linear elliptic PDE, thereby extending the result of Morin et al [6, 7]. The proof relies on quasi-orthogonality, which accounts for the bilinear form not being a scalar product, together with novel error and oscillation reduction estimates, which now do not decouple. We show that AFEM is a contraction for the sum of energy error plus oscillation. Numerical experiments, including oscillatory coefficients and convection-diffusion PDE, illustrate the theory and yield optimal meshes.

Key words. a posteriori error estimators, quasi-orthogonality, adaptive mesh refinement, error and oscillation reduction estimates, optimal meshes.

AMS subject classifications. 65N12, 65N15, 65N30, 65N50, 65Y20

1. Introduction and Main Result.

Let Ω be a polyhedral bounded domain in $\mathbb{R}^d$, $(d = 2, 3)$. We consider a general second order elliptic Dirichlet boundary value problem

$$\mathcal{L}u = -\nabla\cdot(\mathbf{A}\nabla u) + \mathbf{b}\cdot\nabla u + c\,u = f \qquad \text{in } \Omega, \tag{1.1}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{1.2}$$

with the following assumptions:

- $\mathbf{A} : \Omega \mapsto \mathbb{R}^{d\times d}$ is Lipschitz and symmetric positive definite with smallest eigenvalue a_- and largest eigenvalue a_+, i.e.,

$$a_-(x)\,|\xi|^2 \le \mathbf{A}(x)\xi\cdot\xi \le a_+(x)\,|\xi|^2\,, \quad \forall \xi \in \mathbb{R}^d,\ x \in \Omega; \tag{1.3}$$

- $\mathbf{b} \in [L^\infty(\Omega)]^d$ is divergence free ($\nabla\cdot\mathbf{b} = 0$ in Ω);
- $c \in L^\infty(\Omega)$ is nonnegative ($c \ge 0$ in Ω);
- $f \in L^2(\Omega)$.

The purpose of this paper is to prove the following convergence results for adaptive finite element methods (AFEM) for (1.1-1.2), and document their performance computationally.

THEOREM 1 (Convergence of AFEM). *Let $\{u_k\}_{k\in\mathbb{N}_0}$ be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces $\{\mathbb{V}^k\}_{k\in\mathbb{N}_0}$ produced by the AFEM of §3.5, which involves loops of the form*

$$\mathsf{SOLVE} \to \mathsf{ESTIMATE} \to \mathsf{MARK} \to \mathsf{REFINE}.$$

There exist constants $\sigma, \gamma > 0$, and $0 < \xi < 1$, depending solely on the shape regularity of meshes, the data, the parameters used by AFEM, and a number $0 < s \le 1$ dictated by the angles of $\partial\Omega$, such that if the initial meshsize h_0 satisfies $h_0^s\|\mathbf{b}\|_{L^\infty} < \sigma$, then for any two consecutive iterates k and $k+1$ we have

$$|\!|\!|u - u_{k+1}|\!|\!|^2 + \gamma\ \mathsf{osc}_{k+1}(\Omega)^2 \le \xi^2\left(|\!|\!|u - u_k|\!|\!|^2 + \gamma\ \mathsf{osc}_k(\Omega)^2\right). \tag{1.4}$$

*Department of Mathematics, University of Maryland, College Park, MD 20742, USA (kxm@math.umd.edu). Partially supported by NSF Grant DMS-0204670.

† Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA (rhn@math.umd.edu). Partially supported by NSF Grant DMS-0204670.

Therefore, AFEM converges with a linear rate ξ, namely

$$|||u - u_k|||^2 + \gamma\ \mathsf{osc}_k(\Omega)^2 \leq C_0\ \xi^{2k},$$

where $C_0 := |||u - u_0|||^2 + \gamma\ \mathsf{osc}_0(\Omega)^2$.

Hereafter, $|||\cdot|||$ denotes the energy norm induced by the operator $\mathcal{L}$ and $\mathsf{osc}(\Omega)$, the oscillation term, stands for information missed by the averaging process associated to FEM. This convergence result extends those of Morin et al. [6, 7] in several ways:
• We deal with a full second order linear elliptic PDE with variable coefficients $\mathbf{A}, \mathbf{b}$ and c, whereas in [6, 7] $\mathbf{A}$ is assumed to be piecewise constant and $\mathbf{b}$ and c to vanish.
• The underlying bilinear form $\mathcal{B}$ is non-symmetric due to the first order term $\mathbf{b} \cdot \nabla u$. Since $\mathcal{B}$ is no longer a scalar product as in [6, 7], the Pythagoras equality relating u, u_k and u_{k+1} fails; we prove a *quasi-orthogonality* property instead.
• The oscillation terms depend on discrete solutions in addition to data. Therefore, oscillation and error cannot be reduced separately as in [6, 7].
• The oscillation terms do not involve the oscillation of the jump residuals. This is achieved by exploiting positivity and continuity of $\mathbf{A}$.
• Since error and oscillation are now coupled, in order to prove convergence we need to handle them together. This leads to a novel argument and result, the contraction property (1.4), according to which both error and oscillation decrease together.

This paper is organized as follows. In section 2 we introduce the bilinear form, the energy norm, recall existence and uniqueness of solutions, and state the quasi-orthogonality property. In section 3 we describe the procedures used in AFEM, namely, SOLVE, ESTIMATE, MARK, and REFINE, state new error and oscillation reduction estimates, present the adaptive algorithm AFEM and prove its convergence. In section 4 we prove the quasi-orthogonality property of section 2 and the error and oscillation reduction estimates of section 3. In section 5 we present three numerical experiments to illustrate properties of AFEM. We conclude in section 6 with extensions to $\mathbf{A}$ piecewise Lipschitz with discontinuities aligned with the initial mesh and non-coercive bilinear form $\mathcal{B}$ due to $\nabla \cdot \mathbf{b} \neq 0$.

2. Discrete Solution and Quasi-Orthogonality. For an open set $G \in \mathbb{R}^d$ we denote by $H^1(G)$ the usual Sobolev space of functions in $L^2(G)$ whose first derivatives are also in $L^2(G)$, endowed with the norm

$$\|u\|_{H^1(G)} := \Big(\|u\|_{L^2(G)} + \|\nabla u\|_{L^2(G)} \Big)^{1/2};$$

we use the symbols $\|\cdot\|_{H^1}$ and $\|\cdot\|_{L^2}$ when $G = \Omega$. Moreover, we denote by $H^1_0(G)$ the space of functions in $H^1(G)$ that vanish on the boundary in the trace sense.

A weak solution of (1.1) and (1.2) is a function u satisfying

$$u \in H^1_0(\Omega): \qquad \mathcal{B}[u, v] = \langle f, v \rangle \qquad \forall v \in H^1_0(\Omega), \tag{2.1}$$

where $\langle u, v \rangle := \int_\Omega uv$ for any $u, v \in L^2(\Omega)$, and the bilinear form is defined on $H^1_0(\Omega) \times H^1_0(\Omega)$ as

$$\mathcal{B}[u, v] := \langle \mathbf{A} \nabla u, \nabla v \rangle + \langle \mathbf{b} \cdot \nabla u + c\, u, v \rangle\,. \tag{2.2}$$

By Cauchy-Schwarz inequality one can easily show the *continuity* of the bilinear form

$$|\mathcal{B}[u, v]| \leq C_B\, \|u\|_{H^1}\, \|v\|_{H^1}\,,$$

where C_B depends only on the data. Combining Poincaré inequality with the divergence free condition $\nabla\cdot\mathbf{b}=0$, one has *coercivity*

$$\mathcal{B}[v,v] \geq \int_\Omega a_- \left|\nabla v\right|^2 + cv^2 \geq c_B \left\|v\right\|_{H^1}^2 ,$$

where c_B depends only on the data. Existence and uniqueness of (2.1) thus follows from Lax-Milgram theorem [5].

We define the energy norm on $H_0^1(\Omega)$ by $\left|\!\left|\!\left|v\right|\!\right|\!\right|^2 := \mathcal{B}[v,v]$, which is equivalent to $H_0^1(\Omega)$-norm $\left\|\cdot\right\|_{H^1}$. In fact we have

$$c_B \left\|v\right\|_{H^1}^2 \leq \left|\!\left|\!\left|v\right|\!\right|\!\right|^2 \leq C_B \left\|v\right\|_{H^1}^2 \qquad \forall\, v \in H_0^1(\Omega). \tag{2.3}$$

2.1. Discrete Solutions on Nested Meshes. Let $\{\mathcal{T}_H\}$ be a shape regular family of nested conforming meshes over Ω: that is there exists a constant γ^* such that

$$\frac{H_T}{\rho_T} \leq \gamma^* \qquad \forall\, T \in \bigcup_H \mathcal{T}_H, \tag{2.4}$$

where, for each $T \in \mathcal{T}_H$, H_T is the diameter of T, and ρ_T is the diameter of the biggest ball contained in T; the global meshsize is $h_H := \max_{T\in\mathcal{T}_H} H_T$.

Let $\{\mathbb{V}_H\}$ be a corresponding family of nested finite element spaces consisting of continuous piecewise polynomials over $\mathcal{T}_H$ of fixed degree $n \geq 1$, that vanish on the boundary. Let u_H be a discrete solution of (2.1) satisfying

$$u_H \in \mathbb{V}_H : \qquad \mathcal{B}[u_H, v_H] = \langle f, v_H\rangle \qquad \forall\, v_H \in \mathbb{V}_H. \tag{2.5}$$

Existence and uniqueness of this problem follows from Lax-Milgram theorem, since $\mathbb{V}_H \subset H_0^1(\Omega)$.

2.2. Quasi-Orthogonality. Consider two consecutive nested meshes $\mathcal{T}_H \subset \mathcal{T}_h$, i.e. $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. For the corresponding spaces $\mathbb{V}_H \subset \mathbb{V}_h \subset H_0^1(\Omega)$, let $u_h \in \mathbb{V}_h$ and $u_H \in \mathbb{V}_H$ be the discrete solutions. Since the bilinear form is non-symmetric, it is not a scalar product and the orthogonality relation between $u - u_H$ and $u_h - u_H$, the so-called Pythagoras equality, fails to hold. We have instead a perturbation result referred to as quasi-orthogonality provided that the initial mesh is fine enough. This result is stated below and the proof is given in section 4.

LEMMA 2.1 (Quasi-orthogonality). *Let $f \in L^2(\Omega)$. There exist a constant $C^* > 0$, solely depending on the shape regularity constant γ^* and coercivity constant c_B, and a number $0 < s \leq 1$ dictated only by the angles of $\partial\Omega$, such that if the meshsize h_0 of the initial mesh satisfies $C^* h_0^s \left\|\mathbf{b}\right\|_{L^\infty} < 1$, then*

$$\left|\!\left|\!\left|u - u_h\right|\!\right|\!\right|^2 \leq \Lambda_0 \left|\!\left|\!\left|u - u_H\right|\!\right|\!\right|^2 - \left|\!\left|\!\left|u_h - u_H\right|\!\right|\!\right|^2 , \tag{2.6}$$

where $\Lambda_0 := (1 - C^* h_0^s \left\|\mathbf{b}\right\|_{L^\infty})^{-1}$. The equality holds provided $\mathbf{b} = 0$ in Ω.

3. Adaptive Algorithm. The Adaptive procedure consists of loops of the form

$$\boxed{\mathsf{SOLVE} \rightarrow \mathsf{ESTIMATE} \rightarrow \mathsf{MARK} \rightarrow \mathsf{REFINE}.}$$

The procedure SOLVE solves (2.5) for the discrete solution u_H. The procedure ESTIMATE determines the element indicators $\eta_H(T)$ and oscillation $\mathsf{osc}_H(T)$ for all

elements $T \in \mathcal{T}_H$. Depending on their relative size, these quantities are later used by the procedure MARK to mark elements T, and thereby create a subset $\widehat{\mathcal{T}}_H$ of $\mathcal{T}_H$ of elements to be refined. Finally, procedure REFINE partitions those elements in $\widehat{\mathcal{T}}_H$ and a few more to maintain mesh conformity. These procedures are discussed more in detail below.

3.1. Procedure SOLVE : Linear Solver. We employ linear solvers, either direct or iterative methods, such as preconditioned GMRES, CG, and BICG, to solve linear system (2.5). In other words, given a mesh $\mathcal{T}_k$, an initial guest u_{k-1} for the solution, and the data $\mathbf{A}, \mathbf{b}, c, f$, SOLVE computes the discrete solution

$$\boxed{u_k := \mathsf{SOLVE}(\mathcal{T}_k, u_{k-1}, \mathbf{A}, \mathbf{b}, c, f)}$$

3.2. Procedure ESTIMATE : A Posteriori Error Estimate. Subtracting (2.5) from (2.1), we have the Galerkin orthogonality

$$\mathcal{B}[u - u_H, v_H] = 0 \qquad \forall\, v_H \in \mathbb{V}_H. \tag{3.1}$$

In addition to $\mathcal{T}_H$, let $\mathcal{S}_H$ denote the set of interior faces (edges or sides) of the mesh (triangulation) $\mathcal{T}_H$. We consider the *residual* $\mathcal{R}(u_H) \in H^{-1}(\Omega)$ defined by

$$\mathcal{R}(u_H) := f + \nabla\cdot(\mathbf{A}\nabla u_H) - \mathbf{b}\cdot\nabla u_H - c\, u_H,$$

and its relation to the error $\mathcal{L}(u - u_H) = \mathcal{R}(u_H)$. It is then clear that to estimate $|\!|\!| u - u_H |\!|\!|$ we can equivalently deal with $\|\mathcal{R}(u_H)\|_{H^{-1}(\Omega)}$. To this end, we integrate by parts elementwise the bilinear form $\mathcal{B}[u - u_H, v]$ to obtain the *error representation formula*

$$\mathcal{B}[u - u_H, v] = \sum_{T\in\mathcal{T}_H} \int_T R_T(u_H)v + \sum_{S\in\mathcal{S}_H} \int_S J_S(u_H)v \qquad \forall\, v \in H^1_0(\Omega), \tag{3.2}$$

where the *element residual* $R_T(u_H)$ and the *jump residual* $J_S(u_H)$ are defined as

$$R_T(u_H) := f + \nabla\cdot(\mathbf{A}\nabla u_H) - \mathbf{b}\cdot\nabla u_H - c\, u_H \qquad \text{in } T \in \mathcal{T}_H, \tag{3.3}$$

$$J_S(u_H) := -\mathbf{A}\nabla u_H^+ \cdot \nu^+ - \mathbf{A}\nabla u_H^- \cdot \nu^- := [\![\mathbf{A}\nabla u_H]\!]_S \cdot \nu_S \qquad \text{on } S \in \mathcal{S}_H, \tag{3.4}$$

where S is the common side of elements T^+ and T^- with unit outward normals ν^+ and ν^-, respectively, and $\nu_S = \nu^-$. Whenever convenient, we will use the abbreviations $R_T = R_T(u_H)$ and $J_S = J_S(u_H)$.

3.2.1. Upper Bound. For $T \in \mathcal{T}_H$, we define the *local error indicator* $\eta_H(T)$ by

$$\eta_H(T)^2 := H_T^2 \, \|R_T(u_H)\|^2_{L^2(T)} + \sum_{S\subset\partial T} H_S \, \|J_S(u_H)\|^2_{L^2(S)}\,. \tag{3.5}$$

Given a subset $\omega \subset \Omega$, we define the *error estimator* $\eta_H(\omega)$ by

$$\eta_H(\omega)^2 := \sum_{T\in\mathcal{T}_H,\, T\subset\omega} \eta_H(T)^2.$$

Hence, $\eta_H(\Omega)$ is the error estimator of Ω with respect to the mesh $\mathcal{T}_H$. Using (3.1),(3.2) and properties of Clément interpolation, as shown in [1, 3, 12], we obtain the upper bound of the error in terms of the estimator,

$$|\!|\!| u - u_H |\!|\!|^2 \le C_1 \eta_H(\Omega)^2, \tag{3.6}$$

where the constant $C_1 > 0$ depends only on the shape regularity γ^*, coercivity constant c_B and continuity constant C_B of the bilinear form.

3.2.2. Lower Bound. Using the explicit construction of Verfürth [1, 12] via bubble functions and positivity and continuity of A, we can get a local lower bound of the error in terms of local indicators and oscillation. That is, there exist constants $C_2, C_3 > 0$, depending only on the shape regularity γ^*, C_B , and c_B, such that

$$C_2\ \eta_H(T)^2 - C_3 \sum_{T\subset\omega_T} H_T^2 \left\|R_T - \overline{R_T}\right\|^2_{L^2(T)} \leq \|u-u_H\|^2_{H^1(\omega_T)}\,, \tag{3.7}$$

where the domain ω_T consists of all elements sharing at least a side with T, and $\overline{R_T}$ is a polynomial approximation of R_T on T. We define the *oscillation* on the elements $T \in \mathcal{T}_H$ by

$$\mathsf{osc}_H(T)^2 \coloneqq H_T^2 \left\|R_T - \overline{R_T}\right\|^2_{L^2(T)}\,, \tag{3.8}$$

and for a subset $\omega \subset \Omega$, we define

$$\mathsf{osc}_H(\omega)^2 \coloneqq \sum_{T\in\mathcal{T}_H,\, T\subset\omega} \mathsf{osc}_H(T)^2\,.$$

Remark 3.2.1. We see from (3.7) that if the oscillation $\mathsf{osc}_H(\omega_T)$ is small compared to the indicator $\eta_H(T)$, then the size of the indicator $\eta_H(T)$ will give reliable information about the size of the error $\|u-u_H\|_{H^1(\omega_T)}$. This explains why refining elements with large indicators usually tends to equi-distribute the errors, which is an ultimate goal of adaptivity. This idea is employed by the procedure MARK of §3.3.

Remark 3.2.2. The oscillation $\mathsf{osc}_H(T)$ does not involve oscillation of the jump residual $J_S(u_H)$ as is customary [1, 12]. This result follows from the positivity and continuity of $\mathbf{A}$, and is explained in §4.2.

Remark 3.2.3. The oscillation $\mathsf{osc}_H(T)$ depends on $R_T = R_T(u_H)$ which in turn depends on the discrete solution u_H. This is a fundamental difference with Morin et al. [6, 7], where the oscillation is purely data oscillation. It is not clear now that the oscillation will decrease when the mesh $\mathcal{T}_H$ will be refined because u_H will also change. Controlling the decay of $\mathsf{osc}_H(T)$ is thus a major challenge addressed in this work; see §3.3 and §3.4. It is not possible to show that oscillation will always decrease as the mesh gets refined as in [6, 7].

For a given mesh $\mathcal{T}_H$ and discrete solution u_H, along with input data $\mathbf{A}, \mathbf{b}, c$ and f, the procedure ESTIMATE computes indicators $\eta_H(T)$ and oscillations $\mathsf{osc}_H(T)$ for all elements $T \in \mathcal{T}_H$ according to (3.5) and (3.8):

$$\boxed{\{\eta_H(T), \mathsf{osc}_H(T)\}_{T\in\mathcal{T}_H} = \mathsf{ESTIMATE}(\mathcal{T}_H, u_H, \mathbf{A}, \mathbf{b}, c, f)}$$

3.3. Procedure MARK . Our goal is to devise a marking procedure, namely to identify a subset $\widehat{\mathcal{T}}_H$ of the mesh $\mathcal{T}_H$ such that after refining, both error and oscillation will be reduced. We use two strategies for this: Marking Strategy E deals with the error estimator, and Marking Strategy O does so with the oscillation.

3.3.1. Marking Strategy E : Error Reduction. This strategy was introduced by Dörfler [4] to enforce error reduction:

`Marking Strategy E` : Given a parameter $0 < \theta < 1$, construct a subset $\widehat{\mathcal{T}}_H$ of $\mathcal{T}_H$ such that

$$\sum_{T\in\widehat{\mathcal{T}}_H} \eta_H(T)^2 \geq \theta^2 \eta_H(\Omega)^2, \tag{3.9}$$

and mark all elements in $\widehat{\mathcal{T}}_H$ for refinement.

We will see later that Marking Strategy E guarantees error reduction in the absence of oscillation terms. Since the latter account for information missed by the averaging process associated with the finite element method, we need a separate procedure to guarantee oscillation reduction.

3.3.2. Marking Strategy O : Oscillation Reduction. This procedure was introduced by Morin et al. [6, 7] as a separate means for reducing oscillation:

`Marking Strategy O` : Given a parameter $0 < \hat{\theta} < 1$ and the subset $\widehat{\mathcal{T}}_H \subset \mathcal{T}_H$ produced by Marking Strategy E, enlarge $\widehat{\mathcal{T}}_H$ such that

$$\sum_{T \in \widehat{\mathcal{T}}_H} \mathsf{osc}_H(T)^2 \geq \hat{\theta}^2 \mathsf{osc}_H(\Omega)^2 \,, \tag{3.10}$$

and marked all elements in $\widehat{\mathcal{T}}_H$ for refinement.

Given a mesh $\mathcal{T}_H$ and all information about the local error indicators $\eta_H(T)$, and oscillation $\mathsf{osc}_H(T)$, together with user parameters θ and $\hat{\theta}$, MARK generates a subset $\widehat{\mathcal{T}}_H$ of $\mathcal{T}_H$

$$\widehat{\mathcal{T}}_H = \mathsf{MARK}(\theta, \hat{\theta}\,;\, \mathcal{T}_H, \{\eta_H(T), \mathsf{osc}_H(T)\}_{T \in \mathcal{T}_H})$$

3.4. Procedure REFINE. The following Interior Node Property, due to Morin et al [6, 7], is known to be necessary for error and oscillation reduction:

`Interior Node Property` : Refine each marked element $T \in \widehat{\mathcal{T}}_H$ to obtain a new mesh $\mathcal{T}_h$ compatible with $\mathcal{T}_H$ such that
T and the $d+1$ adjacent elements $T' \in \mathcal{T}_H$ of T, as well as their common sides, contain a node of the finer mesh $\mathcal{T}_h$ in their interior.

In addition to the Interior Node Property, we assume that the refinement is done in such a way that the new mesh $\mathcal{T}_h$ is conforming, which guarantees that both $\mathcal{T}_H$ and $\mathcal{T}_h$ are nested. With this property, we have a reduction factor $\gamma_0 < 1$ of element size, i.e. if $T \in \mathcal{T}_h$ is obtained by refining $T' \in \widehat{\mathcal{T}}_H$, then $h_T \leq \gamma_0 H_{T'}$. For example, when $d = 2$ with triangular elements, to have Interior Node Property we can use 3 newest bisections for each single refinement step, whence $\gamma_0 \leq 1/2$.

Given a mesh $\mathcal{T}_H$ and a marked set $\widehat{\mathcal{T}}_H$, REFINE constructs the refinement $\mathcal{T}_h$ satisfying the Interior Node Property:

$$\mathcal{T}_h = \mathsf{REFINE}(\mathcal{T}_H, \widehat{\mathcal{T}}_H)$$

Combining the marking strategies of §3.3 with the Interior Node Property, we obtain the following two crucial results whose proofs are given in §4.

LEMMA 3.1 (Error Reduction). *There exist constants C_4 and C_5, only depending on the shape regularity constant γ^* and θ, such that*

$$\eta_H(T)^2 \leq C_4 \left\| u_h - u_H \right\|^2_{H^1(\omega_T)} + C_5 \mathsf{osc}_H(\omega_T)^2 \qquad \forall\, T \in \widehat{\mathcal{T}}_H. \tag{3.11}$$

We realize that the local energy error between consecutive discrete solutions is bounded below by the local indicators for elements in the marked set $\widehat{\mathcal{T}}_H$, provided the oscillation term is relatively small.

LEMMA 3.2 (Oscillation Reduction). *There exist constants* $0 < \rho_1 < 1$ *and* $0 < \rho_2$, *only depending on* γ^* *and* $\hat{\theta}$, *such that*

$$\mathsf{osc}_h(\Omega)^2 \leq \rho_1 \mathsf{osc}_H(\Omega)^2 + \rho_2 \left|\!\left|\!\left| u_h - u_H \right|\!\right|\!\right|^2 . \tag{3.12}$$

We have that oscillation reduces with a factor $\rho_1 < 1$ provided the energy error between consecutive discrete solutions is relatively small.

REMARK 3.4.1 (Coupling of error and oscillation). Lemmas 3.1 and 3.2 seem to lead to conflicting demands on the relative sizes of error and oscillation. These two concepts are indeed coupled, which contrasts with [6, 7] where oscillation just depends on data and reduces separately from the error. This suggests that we must handle them together, this being the main contribution of this paper. We make this assertion explicit in Theorem 1 below.

3.5. Adaptive Algorithm AFEM. The adaptive algorithm consists of the loops of procedures SOLVE, ESTIMATE, MARK, and REFINE, consecutively, given that the parameters θ and $\hat{\theta}$ are chosen according to Marking Strategies E and O.

AFEM

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Pick an initial mesh $\mathcal{T}_0$, initial guest $u_{-1} = 0$, and set $k = 0$.
2. $u_k = \mathsf{SOLVE}(\mathcal{T}_k, u_{k-1}, \mathbf{A}, \mathbf{b}, c, f)$.
3. $\{\eta_k(T), \mathsf{osc}_k(T)\}_{T\in\mathcal{T}_k} = \mathsf{ESTIMATE}(\mathcal{T}_k, u_k, \mathbf{A}, \mathbf{b}, c, f))$.
4. $\widehat{\mathcal{T}_k} = \mathsf{MARK}(\theta, \hat{\theta}\ ;\ \mathcal{T}_k, \{\eta_k(T), \mathsf{osc}_k(T)\}_{T\in\mathcal{T}_k})$.
5. $\mathcal{T}_{k+1} = \mathsf{REFINE}(\mathcal{T}_k, \widehat{\mathcal{T}_k})$.
6. Set $k = k + 1$ and go to Step 2.

THEOREM 1 (Convergence of AFEM). *Let* $\{u_k\}_{k\in\mathbb{N}_0}$ *be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces* $\{\mathbb{V}^k\}_{k\in\mathbb{N}_0}$ *produced by AFEM. There exist constants* $\sigma, \gamma > 0$, *and* $0 < \xi < 1$, *depending solely on the mesh regularity constant* γ^*, *data, parameters* θ *and* $\hat{\theta}$, *and a number* $0 < s \leq 1$ *dictated by angles of* $\partial\Omega$, *such that if the initial meshsize* h_0 *satisfies* $h_0^s \|\mathbf{b}\|_{L^\infty} < \sigma$, *then for any two consecutive iterates* k *and* $k+1$, *we have*

$$\left|\!\left|\!\left| u - u_{k+1} \right|\!\right|\!\right|^2 + \gamma\ \mathsf{osc}_{k+1}(\Omega)^2 \leq \xi^2 \left(\left|\!\left|\!\left| u - u_k \right|\!\right|\!\right|^2 + \gamma\ \mathsf{osc}_k(\Omega)^2 \right). \tag{3.13}$$

Therefore AFEM converges with a linear rate ξ, *namely,*

$$\left|\!\left|\!\left| u - u_k \right|\!\right|\!\right|^2 + \gamma\ \mathsf{osc}_k(\Omega)^2 \leq C_0\ \xi^{2k},$$

where $C_0 := \left|\!\left|\!\left| u - u_0 \right|\!\right|\!\right|^2 + \gamma\ \mathsf{osc}_0(\Omega)^2$.

Proof. We just prove the contraction property (3.13), which obviously implies the decay estimate. For convenience, we introduce the notation

$$e_k := \left|\!\left|\!\left| u - u_k \right|\!\right|\!\right| , \qquad \varepsilon_k := \left|\!\left|\!\left| u_{k+1} - u_k \right|\!\right|\!\right| , \qquad \mathsf{osc}_k := \mathsf{osc}_k(\Omega) .$$

The idea is to use the quasi-orthogonality (2.6) and replace the term $\left|\!\left|\!\left| u_{k+1} - u_k \right|\!\right|\!\right|^2$ using new results of error and oscillation reduction estimates (3.11) and (3.12). We proceed in three steps as follows.

1. We first get a lower bound for ε_k in terms of e_k. To this end, we use Marking Strategy E and the upper bound (3.6) to write

$$\theta^2 e_k^2 \le C_1 \theta^2 \eta_k(\Omega)^2 \le C_1 \sum_{T\in\widehat{\mathcal{T}_k}} \eta_k(T)^2.$$

Adding (3.11) of Lemma 3.1 over all marked elements $T \in \widehat{\mathcal{T}_k}$, and observing that each element can be counted at most $D := d+2$ times due to overlap of the sets ω_T, together with $\|v\|_{H^1}^2 \le c_B^{-1} \|\!|v|\!\|^2$ for all $v \in H_0^1(\Omega)$, we arrive at

$$\theta^2 e_k^2 \le \frac{DC_1C_4}{c_B}\, \varepsilon_k^2 + DC_1C_5\, \mathsf{osc}_k^2.$$

If $\Lambda_1 := \frac{\theta^2 c_B}{DC_1C_4}, \Lambda_2 := \frac{C_5 c_B}{C_4}$, then this implies the lower bound for ε_k^2,

$$\varepsilon_k^2 \ge \Lambda_1 e_k^2 - \Lambda_2 \mathsf{osc}_k^2. \tag{3.14}$$

2. If h_0 is sufficiently small so that the quasi-orthogonality (2.6) of Lemma 2.1 holds, then

$$e_{k+1}^2 \le \Lambda_0 e_k^2 - \varepsilon_k^2,$$

where $\Lambda_0 = (1 - C^* h_0^s \|\mathbf{b}\|_{L^\infty})^{-1}$. Replacing the fraction $\beta\varepsilon_k^2$ of ε_k^2 via (3.14) we obtain

$$e_{k+1}^2 \le (\Lambda_0 - \beta\Lambda_1) e_k^2 + \beta\Lambda_2 \mathsf{osc}_k^2 - (1-\beta)\varepsilon_k^2,$$

where $0 < \beta < 1$ is a constant to be chosen suitably. We now assert that it is possible to chose h_0 compatible with Lemma 2.1 and also that

$$0 < \alpha := \Lambda_0 - \beta\Lambda_1 < 1.$$

A simple calculation shows that this is the case provided

$$h_0^s \|\mathbf{b}\|_{L^\infty} < \frac{\beta\Lambda_1}{C^*(1+\beta\Lambda_1)} < \frac{1}{C^*},$$

i.e., $h_0^s \|\mathbf{b}\|_{L^\infty} < \sigma$ with $\sigma := \frac{\beta\Lambda_1}{C^*(1+\beta\Lambda_1)}$. Consequently

$$e_{k+1}^2 \le \alpha e_k^2 + \beta\Lambda_2 \mathsf{osc}_k^2 - (1-\beta)\varepsilon_k^2. \tag{3.15}$$

3. To remove the last term of (3.15) we resort to the oscillation reduction estimate of Lemma 3.2

$$\mathsf{osc}_{k+1}^2 \le \rho_1 \mathsf{osc}_k^2 + \rho_2 \varepsilon_k^2.$$

We multiply it by $(1-\beta)/\rho_2$ and add it to (3.15) to deduce

$$e_{k+1}^2 + \frac{1-\beta}{\rho_2}\mathsf{osc}_{k+1}^2 \le \alpha\, e_k^2 + \left(\beta\Lambda_2 + \frac{\rho_1}{\rho_2}(1-\beta)\right) \mathsf{osc}_k^2.$$

If $\gamma := \frac{1-\beta}{\rho_2}$, then we would like to choose $\beta < 1$ in such a way that

$$\beta\Lambda_2 + \rho_1\gamma = \mu\gamma$$

for some $\mu < 1$. A simple calculation yields

$$\beta = \frac{\frac{\mu-\rho_1}{\rho_2}}{\Lambda_2 + \frac{\mu-\rho_1}{\rho_2}},$$

and shows that $\rho_1 < \mu < 1$ guarantees that $0 < \beta < 1$. Therefore,

$$e_{k+1}^2 + \gamma\, \mathsf{osc}_{k+1}^2 \leq \alpha\, e_k^2 + \mu\gamma\, \mathsf{osc}_k^2,$$

and the asserted estimate (3.13) follows upon taking $\xi = \max(\alpha, \mu) < 1$. □

REMARK 3.5.1 (Comparison with [6, 7]). In [6, 7] the oscillation is independent of discrete solutions, i.e. $\rho_2 = 0$, and is reduced by the factor $\rho_1 < 1$ in (3.12). Consequently, Step 3 above is avoided by setting $\beta = 1$ and the decay of e_k and osc_k is monitored separately. Since this is no longer possible, e_k and osc_k are now combined and decreased together.

REMARK 3.5.2 (Splitting of ε_k). The idea of splitting ε_k is already used by Chen and Jia [2] in examining one time step for the heat equation. This is because a mass (zero order) term naturally occurs, which did not take place in [6, 7]. The elliptic operator is just the Laplacian in [2].

REMARK 3.5.3 (Effect of Convection). Assuming that $h_0^s \|\mathbf{b}\|_{L^\infty} < \sigma$ implies that the local Peclet number is sufficiently small for the Galerkin method not to exhibit oscillations. This appears to be essential for u_0 to contain relevant information and guide correctly the adaptive process. This restriction is difficult to verify in practice because it involves unknown constants. However, starting from a coarse mesh does not seem to be a problem in practice (see numerical experiments in §5).

REMARK 3.5.4 (Vanishing Convection). If $\mathbf{b} = 0$, then Theorem 1 has no restriction on the initial mesh. This thus extends the convergent result of Morin et al. [6, 7] to variable diffusion coefficient and zero order terms.

REMARK 3.5.5 (Optimal β). The choice of β can be optimized. In fact, we can easily see that

$$\alpha = \Lambda_0 - \beta\Lambda_1, \qquad \mu = \rho_1 + \frac{\beta}{1-\beta}\rho_2\Lambda_2$$

yields a unique value $0 < \beta_* < 1$ for which $\alpha = \mu$ and the contraction constant ξ of Theorem 1 is minimal. This β_* depends on geometric constant $\Lambda_0, \Lambda_1, \Lambda_2$ as well on $\theta, \hat{\theta}$ and h_0, but it is not computable.

4. Proofs of Lemmas. Let $\widehat{\mathcal{T}}_H \subset \mathcal{T}_H$ be a set of marked elements obtained from procedure MARK. Let $\mathcal{T}_h$ be a refined mesh obtained from procedure REFINE, and let $\mathbb{V}_H \subset \mathbb{V}_h$ be nested spaces corresponding to compatible meshes $\mathcal{T}_H$ and $\mathcal{T}_h$, respectively. For convenience, set

$$e_h := u - u_h, \qquad e_H := u - u_H, \qquad \varepsilon_H := u_h - u_H.$$

4.1. Proof of Lemma 2.1: Quasi-Orthogonality. In view of Galerkin orthogonality (3.1), namely $\mathcal{B}[e_h, v_h] = 0$, $v_h \in \mathbb{V}_h$, we have

$$|\!|\!| e_H |\!|\!|^2 = |\!|\!| e_h |\!|\!|^2 + |\!|\!| \varepsilon_H |\!|\!|^2 + \mathcal{B}[\varepsilon_H, e_h].$$

If $\mathbf{b} = 0$, then $\mathcal{B}$ is symmetric and $\mathcal{B}[\varepsilon_H, e_h] = \mathcal{B}[e_h, \varepsilon_H] = 0$. For $\mathbf{b} \neq 0$, instead, $\mathcal{B}[\varepsilon_H, e_h] \neq 0$, and we must account for this term. It is easy to see that $\nabla \cdot \mathbf{b} = 0$ and integration by parts yield

$$\mathcal{B}[\varepsilon_H, e_h] = \mathcal{B}[e_h, \varepsilon_H] + \langle \mathbf{b} \cdot \nabla \varepsilon_H, e_h \rangle - \langle \mathbf{b} \cdot \nabla e_h, \varepsilon_H \rangle = 2 \langle \mathbf{b} \cdot \nabla \varepsilon_H, e_h \rangle .$$

Hence

$$|\!|\!| e_h |\!|\!|^2 = |\!|\!| e_H |\!|\!|^2 - |\!|\!| \varepsilon_H |\!|\!|^2 - 2 \langle \mathbf{b} \cdot \nabla \varepsilon_H, e_h \rangle .$$

Using Cauchy-Schwarz inequality and replacing the $H^1(\Omega)$-norm by the energy norm we have, for any $\delta > 0$ to be chosen later,

$$-2 \langle \mathbf{b} \cdot \nabla \varepsilon_H, e_h \rangle \leq \delta \, \|e_h\|_{L^2}^2 + \frac{\|\mathbf{b}\|_{L^\infty}^2}{\delta c_B} \, |\!|\!| \varepsilon_H |\!|\!|^2 .$$

We then realize the need to relate $L^2(\Omega)$ and energy norms to replace $\|e_h\|_{L^2}$ by $|\!|\!| e_h |\!|\!|$. This requires a standard duality argument whose proof is reported in Ciarlet [3].

LEMMA 4.1 (Duality). *Let $f \in L^2(\Omega)$ and $u \in H^{1+s}(\Omega)$ for some $0 < s \leq 1$ be the solution, where s depends on the angles of $\partial\Omega$ ($s = 1$ if Ω is convex). Then, there exists a constant C_6, depending only on the shape regularity constant γ^* and the coercivity constant c_B, such that*

$$\|e_h\|_{L^2} \leq C_6 h^s \, |\!|\!| e_h |\!|\!| . \tag{4.1}$$

Inserting this estimate in the preceding two bounds, and using $h \leq h_0$, the meshsize of the initial mesh, we deduce

$$\left(1 - \delta C_6^2 h_0^{2s}\right) |\!|\!| e_h |\!|\!|^2 \leq |\!|\!| e_H |\!|\!|^2 - \left(1 - \frac{\|\mathbf{b}\|_{L^\infty}^2}{\delta c_B}\right) |\!|\!| \varepsilon_H |\!|\!|^2 .$$

We now choose $\delta = \frac{\|\mathbf{b}\|_{L^\infty}}{C_6 \sqrt{c_B} h_0^s}$ to equate both parenthesis, as well as the assumption that h_0 is sufficiently small for $\delta C_6^2 h_0^{2s} = C^* h_0^s \, \|\mathbf{b}\|_{L^\infty} < 1$ with $C^* := C_6/\sqrt{c_B}$. We end up with

$$|\!|\!| e_h |\!|\!|^2 \leq \frac{1}{1 - C^* \, \|\mathbf{b}\|_{L^\infty} \, h_0^s} \, |\!|\!| e_H |\!|\!|^2 - |\!|\!| \varepsilon_H |\!|\!|^2 .$$

This implies (2.6) and concludes the proof. □

4.2. Proof of Lemma 3.1 : Error Reduction. Upon restricting the test function v in (3.2) to $\mathbb{V}_h \supset \mathbb{V}_H$, we obtain the error representation

$$\mathcal{B}[\varepsilon_H, v_h] = \sum_{T \in \mathcal{T}_H} \int_T \overline{R_T} v_h + \int_T (R_T - \overline{R_T}) v_h + \sum_{S \in \mathcal{S}_H} \int_S J_S \, v_h \qquad \forall \, v_h \in \mathbb{V}_h, \tag{4.2}$$

where we use the abbreviations $R_T = R_T(u_H), J_S = J_S(u_H)$, and $\overline{R_T} = \Pi_T^{n-1} R_T$ denotes the L_2-projection of R_T onto the space of polynomials $\mathbb{P}_{n-1}(T)$ over the element $T \in \mathcal{T}_H$. Except for avoiding the oscillation terms of the jump residual J_S, the proof goes back to [4, 6, 7]. We proceed in three steps.

1. *Interior Residual.* Let $T \in \mathcal{T}_H$, and let x_T be an interior node of T generated by the procedure REFINE. Let $\psi_T \in \mathbb{V}_h$ be a bubble function which satisfies $\psi_T(x_T) = 1$, vanishes on ∂T, and $0 \le \psi_T \le 1$; hence $\mathsf{supp}\,(\psi_T) \subset T$. Since $\overline{R_T} \in \mathbb{P}_{n-1}(T)$ and $\psi_T > 0$ in a polyhedron of measure comparable with that of T, we have

$$C \left\|\overline{R_T}\right\|^2_{L^2(T)} \le \int_T \psi_T \, \overline{R_T}^2 = \int_T \overline{R_T}(\psi_T \, \overline{R_T}).$$

Since $\psi_T \overline{R_T}$ is a piecewise polynomial of degree $\le n$ over $\mathcal{T}_h$, it is thus an admissible test function in (4.2) which vanishes outside T (and in particular on all $S \in \mathcal{S}_H$). Therefore

$$\begin{aligned} C \left\|\overline{R_T}\right\|^2_{L^2(T)} &\le \mathcal{B}[\varepsilon_H, \psi_T \overline{R_T}] + \int_T (\overline{R_T} - R_T)\psi_T \overline{R_T} \\ &\le C \left(H_T^{-1} \left\|\varepsilon_H\right\|_{H^1(T)} + \left\|R_T - \overline{R_T}\right\|_{L^2(T)} \right) \left\|\overline{R_T}\right\|_{L^2(T)}, \end{aligned}$$

because of an inverse inequality for $\psi_T \overline{R_T}$. This, together with the triangle inequality, yields the desired estimate for $H_T^2 \left\|R_T\right\|^2_{L^2(T)}$:

$$H_T^2 \left\|R_T\right\|^2_{L^2(T)} \le C \left(\left\|\varepsilon_H\right\|^2_{H^1(T)} + H_T^2 \left\|R_T - \overline{R_T}\right\|^2_{L^2(T)} \right). \tag{4.3}$$

2. *Jump Residual.* Let $S \in \mathcal{S}_H$ be an interior side of $T_1 \in \widehat{\mathcal{T}}_H$, and let $T_2 \in \mathcal{T}_H$ be the other element sharing S. Let x_S be an interior node of S created by the procedure REFINE. Let $\psi_S \in \mathbb{V}_h$ be a bubble function in $\omega_S := T_1 \cup T_2$ such that $\psi_S(x_S) = 1$, ψ_S vanishes on $\partial\omega_S$, and $0 \le \psi_S \le 1$; hence $\mathsf{supp}\,(\psi_S) \subset \omega_S$.

Since u_H is continuous, then $[\![\nabla u_H]\!]_S$ is parallel to ν_S, i.e. $[\![\nabla u_H]\!]_S = j_S \, \nu_S$. Moreover, the coefficient matrix $\mathbf{A}(x)$ being continuous implies

$$J_S = \mathbf{A}(x) \, [\![\nabla u_H]\!]_S \cdot \nu_S = j_S \, \mathbf{A}(x)\nu_S \cdot \nu_S = a(x) \, j_S,$$

where $a(x) := \mathbf{A}(x)\nu_S \cdot \nu_S$ satisfies $0 < \underline{a}_S \le a(x) \le \overline{a}_S$ with $\underline{a}_S, \overline{a}_S$ the smallest and largest eigenvalues of $\mathbf{A}(x)$ on S. Consequently,

$$\left\|J_S\right\|^2_{L^2(S)} \le \overline{a}_S^2 \int_S j_S^2 \le C \overline{a}_S^2 \int_S j_S^2 \psi_S \le C \frac{\overline{a}_S^2}{\underline{a}_S} \int_S (j_S \, \psi_S) J_S, \tag{4.4}$$

where the second inequality follows from j_S being a polynomial and $\psi_S > 0$ in a polygon of measure comparable with that of S.

We now extend j_S to ω_S by first mapping to the reference element, next extending constantly along the normal to $\hat{S}$ and finally mapping back to ω_S. The resulting extension $\mathsf{E}_h(j_S)$ is a piecewise polynomial of degree $\le n-1$ in ω_S so that $\psi_S \mathsf{E}_h(j_S) \in \mathbb{V}_h$, and satisfies $\left\|\psi_S \mathsf{E}_h(j_S)\right\|_{L^2(\omega_S)} \le C H_S^{1/2} \left\|j_S\right\|_{L^2(S)}$. Since $v_h = \psi_S \mathsf{E}_h(j_S)$ is an admissible test function in (4.2) which vanishes on all sides of $\mathcal{S}_H$ but S, we arrive at

$$\begin{aligned} \int_S J_S (j_S \psi_S) &= \mathcal{B}[\varepsilon_H, v_h] - \int_{T_1} R_{T_1} \, v_h - \int_{T_2} R_{T_2} \, v_h \\ &\le C \left(H_S^{-1/2} \left\|\varepsilon_H\right\|_{H^1_S(\omega_S)} + H_S^{1/2} \sum_{i=1}^{2} \left\|R_{T_i}\right\|_{L^2(T_i)} \right) \left\|j_S\right\|_{L^2(S)}. \end{aligned} \tag{4.5}$$

Therefore

$$H_S \|J_S\|^2_{L^2(S)} \le C \left(\|\varepsilon_H\|^2_{H^1(\omega_S)} + \sum_{i=1}^{2} H^2_{T_i} \|R_{T_i}\|^2_{L^2(T_i)} \right). \tag{4.6}$$

3. *Final Estimate.* To remove the interior residual from the right hand side of (4.6) we resort to (4.3) since T_1 and T_2 contain an interior node according to procedure REFINE. Hence

$$H_S \|J_S\|^2_{L^2(S)} \le C \left(\|\varepsilon_H\|^2_{H^1(\omega_S)} + \sum_{i=1}^{2} H^2_{T_i} \left\|R_{T_i} - \overline{R_{T_i}}\right\|^2_{L^2(T_i)} \right). \tag{4.7}$$

The asserted estimate for $\eta_H(T)^2$ is thus obtained by adding this bound to (4.3). The constant C depends on the shape regularity constant γ^* and the ratio $\overline{a}_S^2/\underline{a}_S$ of eigenvalues of $\mathbf{A}(x)$ on S. □

REMARK 4.2.1 (Positivity). The use of $\mathbf{A}(x)$ being positive definite in (4.4) avoids having oscillation terms on S. This comes at the expense of a constant depending on $\overline{a}_S^2/\underline{a}_S$. If we were to proceed in the usual manner, as in [1, 8, 12], we would end up with oscillation of the form

$$\begin{aligned} H_S^{1/2} \|(\mathbf{A}(x) - \mathbf{A}(x_S)) [\![\nabla u_H]\!]_S \cdot \nu_S\|_{L^2(S)} &= H_S^{1/2} \|(a(x) - a(x_S)) j_S\|_{L^2(S)} \\ &\le C H_S^{3/2} \|\mathbf{A}\|_{W^1_\infty(S)} \|j_S\|_{L^2(S)} \\ &\le C H_S \left\| H_S^{1/2} J_S \right\|_{L^2(S)}, \end{aligned}$$

where $C > 0$ also depends on the ratio $\overline{a}_S/\underline{a}_S$ dictated by the variation of $a(x)$ on S. This oscillation can be absorbed into the term $H_S^{1/2} \|J_S\|_{L^2(S)}$ provided that the meshsize H_S is sufficiently small; see [8]. We do not need this assumption in our present discussion.

REMARK 4.2.2 (Continuity of $\mathbf{A}$). The continuity of $\mathbf{A}$ is instrumental in avoiding jump oscillation which in turn makes computations simpler. However, jump oscillation cannot be avoid when $\mathbf{A}$ exhibits discontinuities across inter-element boundaries of the initial mesh. We get instead of (4.7)

$$C H_S \|J_S\|^2_{L^2(S)} \le \|\varepsilon_H\|^2_{H^1(\omega_S)} + \sum_{i=1}^{2} H^2_{T_i} \left\|R_{T_i} - \overline{R_{T_i}}\right\|^2_{L^2(T_i)} + H_S \left\|J_S - \overline{J_S}\right\|^2_{L^2(S)}, \tag{4.8}$$

where $\overline{J_S}$ is L_2-projection of J_S onto $\mathbb{P}_{n-1}(S)$. To obtain estimate (4.8) we proceed as follows. Starting from a polynomial $\overline{J_S}$, we get an estimate similar to that of (4.4)

$$C \left\|\overline{J_S}\right\|^2_{L^2(S)} \le \int_S \psi_S \overline{J_S}^2 = \int_S J_S(\psi_S \overline{J_S}) + \int_S (\overline{J_S} - J_S)(\psi_S \overline{J_S}). \tag{4.9}$$

In contrast to (4.4), we see that the oscillation term $(\overline{J_S} - J_S)$ cannot be avoided when $\mathbf{A}$ has a discontinuity across S. We estimate the first term on the right hand side of (4.9) exactly as we have argued with (4.5) and thereby arrive at

$$\int_S J_S(\overline{J_S}\psi_S) \le C \left(H_S^{-1/2} \|\varepsilon_H\|_{H^1_S(\omega_S)} + H_S^{1/2} \sum_{i=1}^{2} \|R_{T_i}\|_{L^2(T_i)} \right) \left\|\overline{J_S}\right\|_{L^2(S)}.$$

This and a further estimate of the second term on the right hand side of (4.9), yield

$$H_S \left\|\overline{J_S}\right\|^2_{L^2(S)} \leq C \left(\|\varepsilon_H\|^2_{H^1(\omega_S)} + \sum_{i=1}^{2} H^2_{T_i} \|R_{T_i}\|^2_{L^2(T_i)} + H_S \left\|J_S - \overline{J_S}\right\|^2_{L^2(S)} \right),$$

whence the assertion (4.8) follows using triangle inequality for $\|J_S\|_{L^2(S)}$. Combining with (4.3), we deduce an estimate for $\eta_H(T)$ similar to (3.11), namely,

$$\eta_H(T)^2 \leq C \left(\|\varepsilon_H\|^2_{H^1(\omega_T)} + \mathsf{osc}_H(\omega_T)^2 \right),$$

with the new oscillation term involving jumps on interior sides

$$\mathsf{osc}_H(T)^2 := H^2_T \left\|R_T - \overline{R_T}\right\|^2_{L^2(T)} + \sum_{S \subset \partial T} H_S \left\|J_S - \overline{J_S}\right\|^2_{L^2(S)}. \tag{4.10}$$

In §6.1 we discuss the case of a discontinuous $\mathbf{A}$. We show an oscillation reduction property of $\mathsf{osc}_H(T)$, defined by (4.10), similar to Lemma 3.2.

4.3. Proof of Lemma 3.2 : Oscillation Reduction. The proof hinges on the Marking Strategy O and the Interior Node Property. We point out that if $T \in \mathcal{T}_h$ is contained in $T' \in \widehat{\mathcal{T}}_H$, then REFINE gives a reduction factor $\gamma_0 < 1$ of element size:

$$h_T \leq \gamma_0 H_{T'}. \tag{4.11}$$

The proof proceeds in three steps as follows.

1. *Relation between Oscillations.* We would like to relate $\mathsf{osc}_h(T')$ and $\mathsf{osc}_H(T')$ for any $T' \in \mathcal{T}_H$. To this end, we note that for all $T \in \mathcal{T}_h$ contained in T', we can write

$$R_T(u_h) = R_T(u_H) - \mathcal{L}_T(\varepsilon_H) \qquad \text{in } T,$$

where $\varepsilon_H = u_h - u_H$ as before and

$$\mathcal{L}_T(\varepsilon_H) := -\nabla\cdot(\mathbf{A}\nabla\varepsilon_H) + \mathbf{b}\cdot\nabla\varepsilon_H + c\,\varepsilon_H \qquad \text{in } T.$$

By Young's inequality, we have for all $\delta > 0$

$$\begin{aligned}\mathsf{osc}_h(T)^2 &= h^2_T \left\| R_T(u_h) - \overline{R_T(u_h)} \right\|^2_{L^2(T)} \\ &\leq (1+\delta) h^2_T \left\| R_T(u_H) - \overline{R_T(u_H)} \right\|^2_{L^2(T)} + (1+\delta^{-1}) h^2_T \left\| \mathcal{L}_T(\varepsilon_H) - \overline{\mathcal{L}_T(\varepsilon_H)} \right\|^2_{L^2(T)},\end{aligned}$$

where $\overline{R_T(u_h)}$, $\overline{R_T(u_H)}$, and $\overline{\mathcal{L}_T(\varepsilon_H)}$ are L^2-projections of $R_T(u_h)$, $R_T(u_H)$, and $\mathcal{L}_T(\varepsilon_H)$ onto polynomials of degree $\leq n-1$ on T. We next observe that

$$\left\| \mathcal{L}_T(\varepsilon_H) - \overline{\mathcal{L}_T(\varepsilon_H)} \right\|_{L^2(T)} \leq \|\mathcal{L}_T(\varepsilon_H)\|_{L^2(T)}$$

and that, according to (4.11),

$$h_T \leq \gamma_{T'} H_{T'}$$

provided $\gamma_{T'} = \gamma_0$ if $T' \in \widehat{\mathcal{T}}_H$ and $\gamma_{T'} = 1$ otherwise. Therefore, if $\mathcal{T}_h(T')$ denotes all $T \in \mathcal{T}_h$ contained in T',

$$\begin{aligned} \mathsf{osc}_h(T')^2 &= \sum_{T\in\mathcal{T}_h(T')} \mathsf{osc}_h(T)^2 \\ &\le (1+\delta)\gamma_{T'}^2 \mathsf{osc}_H(T')^2 + (1+\delta^{-1}) \sum_{T\in\mathcal{T}_h(T')} h_T^2 \left\|\mathcal{L}_T(\varepsilon_H)\right\|_{L^2(T)}^2, \end{aligned} \tag{4.12}$$

because $R_T(u_H) = R_{T'}(u_H)$ and $\overline{R_T(u_H)}$ is the best approximation of $R_{T'}(u_H)$ in T.

2. *Estimate of* $\mathcal{L}_T(\varepsilon_H)$. In order to estimate $\|\mathcal{L}_T(\varepsilon_H)\|_{L^2(T)}$ in terms of $\|\varepsilon_H\|_{H^1(T)}$, we first split it as follows

$$\|\mathcal{L}_T(\varepsilon_H)\|_{L^2(T)} \le \|\nabla\cdot(\mathbf{A}\nabla\varepsilon_H)\|_{L^2(T)} + \|\mathbf{b}\cdot\nabla\varepsilon_H\|_{L^2(T)} + \|c\,\varepsilon_H\|_{L^2(T)}$$

and denote these terms N_A, N_B, and N_C, respectively. Since

$$N_A \le \|(\nabla\cdot\mathbf{A})\cdot\nabla\varepsilon_H\|_{L^2(T)} + \|\mathbf{A} : H(\varepsilon_H)\|_{L^2(T)}$$

where $H(\varepsilon_H)$ is the Hessian of ε_H in T, invoking the Lipschitz continuity of $\mathbf{A}$ together with an inverse estimate in T, we infer that

$$N_A \le C_A \left(\|\nabla\varepsilon_H\|_{L^2(T)} + h_T^{-1} \|\nabla\varepsilon_H\|_{L^2(T)} \right),$$

where C_A depends on $\mathbf{A}$ and the shape regularity constant γ^*. Besides, we readily have

$$N_B \le C_B \|\nabla\varepsilon_H\|_{L^2(T)}, \qquad N_C \le C_C \|\varepsilon_H\|_{L^2(T)},$$

where C_B, C_C depend on $\mathbf{b}$, c. Combining these estimates, we arrive at

$$h_T^2 \|\mathcal{L}_T(\varepsilon_H)\|_{L^2(T)}^2 \le C_* \|\varepsilon_H\|_{H^1(T)}^2. \tag{4.13}$$

3. *Choice of* δ. We insert (4.13) into (4.12) and add over $T' \in \mathcal{T}_H$. Recalling the definition of $\gamma_{T'}$ and utilizing (3.10), we deduce

$$\begin{aligned} \sum_{T'\in\mathcal{T}_H} \gamma_{T'}^2 \mathsf{osc}_H(T')^2 &= \gamma_0^2 \sum_{T'\in\widehat{\mathcal{T}}_H} \mathsf{osc}_H(T')^2 + \sum_{T'\in\mathcal{T}_H\setminus\widehat{\mathcal{T}}_H} \mathsf{osc}_H(T')^2 \\ &= \mathsf{osc}_H(\Omega)^2 - (1-\gamma_0^2) \sum_{T'\in\widehat{\mathcal{T}}_H} \mathsf{osc}_H(T')^2 \\ &\le (1-(1-\gamma_0^2)\hat\theta^2))\mathsf{osc}_H(\Omega)^2, \end{aligned}$$

where $\hat\theta$ is the user's parameter in (3.10). Moreover, since $C_* \|\varepsilon_H\|_{H^1}^2 \le C_o |\!|\!| \varepsilon_H |\!|\!|^2$ with $C_o = C_* c_B^{-1}$ in light of (2.3), we end up with

$$\mathsf{osc}_h(\Omega)^2 \le (1+\delta)(1-(1-\gamma_0^2)\hat\theta^2)\mathsf{osc}_H(\Omega)^2 + (1+\delta^{-1})C_o |\!|\!| \varepsilon_H |\!|\!|^2.$$

To complete the proof, we finally choose δ sufficiently small so that

$$\rho_1 = (1+\delta)(1-(1-\gamma_0^2)\hat\theta^2) < 1, \qquad \rho_2 = (1+\delta^{-1})C_o. \qquad \square$$

5. Numerical Experiments. We test performance of the adaptive algorithm AFEM with several examples. We are thus able to study how meshes adapt to various effects from lack of regularity of solutions and convexity of domains to data smoothness, boundary layers, changing boundary conditions, etc. For simplicity, we strict our experiments to the case of piecewise linear finite element solutions with polygonal domains in 2 dimensions. The implementation is done within the finite element toolbox ALBERT of Schmidt and Siebert [10, 11] which provides tools for adaptivity.

5.1. Implementation. We employ the four main procedures as given by Morin et al. [6, 7]: SOLVE, ESTIMATE, MARK, and REFINE. We slightly modified the built-in adaptive solver for elliptic problems of ALBERT toolbox [10] to make it work for the general PDE (1.1) and mixed boundary conditions, as follows:
• SOLVE. We used built-in solvers provided by ALBERT toolbox, such as GMRES, BICG, or CG.
• ESTIMATE. We modified ALBERT for computing the estimator so that it works for (1.1), and added procedures for computing oscillations which are not provided.
• MARK. We utilized the same algorithm introduced by Morin et al [6, 7] for finding a marked set $\widehat{\mathcal{T}}_H$.
• REFINE. We employed 3 newest bisections for each refinement step to enforce the Interior Node Property.

For simplicity and convenience of presentation, we introduce the following notation:
• DOF := number of elements in a given mesh, which is comparable with number of degree of freedoms;
• EOC := $\frac{\log(e(k-1)/e(k))}{\log(\mathsf{DOF}(k)/\mathsf{DOF}(k-1))}$, experimental order of convergence, $e(k) := \|u - u_k\|_{H^1}$;
• EOC(η) := $\frac{\log(\eta_{k-1}/\eta_k)}{\log(\mathsf{DOF}(k)/\mathsf{DOF}(k-1))}$, experimental order of convergence of estimator, $\eta_k := \eta_k(\Omega)$;
• $Z_e := e(k)/e(k-1)$ and $Z_o := \mathsf{osc}_k/\mathsf{osc}_{k-1}$, reduction factors of error and oscillation;
• Eff := η_k/e_k, effectivity index, i.e. the ratio between the estimator and error.

5.2. Experiment 1 : Oscillatory Coefficients and Nonconvex Domain. We consider the PDE (1.1) with Dirichlet boundary condition $u = g$ on the nonconvex L-shape domain $\Omega := [-1,1]^2 \setminus [0,1] \times [-1,0]$. We also take the exact solution

$$u(r) = r^{\frac{2}{3}} \sin(\frac{2}{3}\theta),$$

where $r^2 := x^2 + y^2$ and $\theta := \tan^{-1}(y/x) \in [0, 2\pi)$. We deal with variable coefficients $\mathbf{A}(x,y) = a(x,y)\mathbf{I}$, $\mathbf{b}(x,y)$, $c(x,y)$ defined by

$$a(x,y) = \frac{1}{4 + P(\sin(\frac{2\pi x}{\epsilon}) + \sin(\frac{2\pi y}{\epsilon}))}, \tag{5.1}$$

$$\mathbf{b}(x,y) = (-y, x). \tag{5.2}$$

$$c(x,y) = A_c(\cos^2(k_1 x) + \cos^2(k_2 x)), \tag{5.3}$$

where P, ϵ, A_c, k_1 and k_2 are parameters. The functions f in (1.1) and g are defined accordingly. To see how AFEM performs comparing to standard uniform refinement, results of AFEM and standard uniform refinement are provided in Tables 5.1 and 5.2. Some examples of adapted refined meshes from AFEM are also displayed in Figure 5.1. Observations and conclusions about AFEM performance are as follows:

DOF	$\|u-u_k\|_{H^1}$	EOC	Z_e	Z_o	Eff
153	1.9800e-01	0.7766	0.6963	0.6537	0.5829
323	1.5470e-01	0.3303	0.7813	0.4862	0.7699
478	1.0859e-01	0.9028	0.7020	0.5570	0.7945
869	7.0835e-02	0.7148	0.6523	0.3853	0.8197
1404	5.6474e-02	0.4723	0.7973	0.4895	0.9184
2266	4.1940e-02	0.6216	0.7426	0.4038	0.9625
3689	3.1117e-02	0.6125	0.7419	0.6052	1.0033
7103	2.1326e-02	0.5767	0.6854	0.4379	1.0024
13729	1.4849e-02	0.5494	0.6963	0.5205	0.9849

TABLE 5.1

Experiment 1 (Oscillatory coefficients and nonconvex domain): The parameters of AFEM are $\theta = \hat{\theta} = 0.6$*, and those controlling the oscillatory coefficients are* $P = 1.8, \epsilon = 0.4, A_c = 1.0, k_1 = k_2 = 4.0$*, as described in (5.1)-(5.3). The experimental order of convergence* EOC *is close to the optimal value 0.5, which indicates quasi-optimal meshes. The oscillation reduction factor* Z_o *is small than the error reduction factor, which confirms that oscillation decreases faster than error.*

DOF	$\|u-u_k\|_{H^1}$	EOC	Z_e	Z_o
384	1.5833e-01	0.4224	0.5568	0.3431
1536	8.3427e-02	0.4622	0.5269	0.1772
6144	5.0608e-02	0.3606	0.6066	0.1851
24576	3.1809e-02	0.3350	0.6285	0.2458

TABLE 5.2

Experiment 1 (Oscillatory coefficients and nonconvex domain): Standard uniform refinement is performed using the same values for parameters P, ϵ, A_c, k_1*, and* k_2 *as that of AFEM given by Table 5.1 above.* EOC *is now suboptimal and close to the expected value* $1/3$*.*

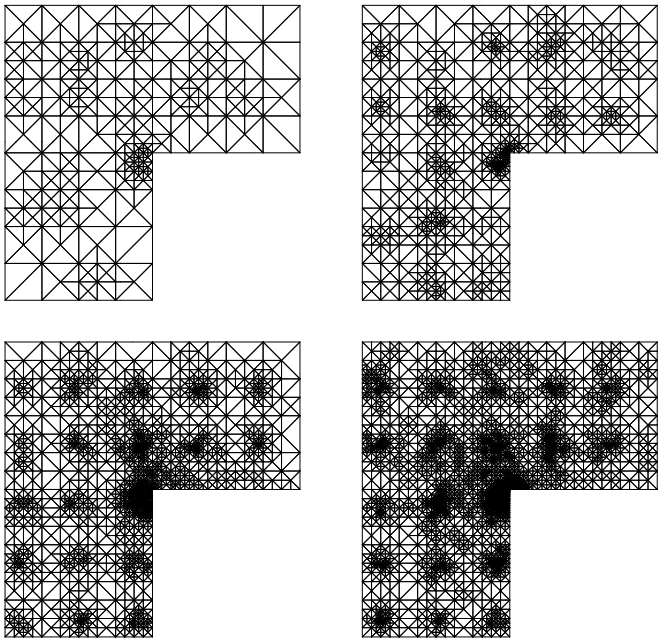

FIG. 5.1. *Experiment 1 (Oscillatory coefficients and nonconvex domain): Parameters of AFEM are* $\theta = \hat{\theta} = 0.6$*, those of oscillatory coefficient* $\mathbf{A}$ *are* $P = 1.8, \epsilon = 0.4$*, and both* $\mathbf{b}$ *and* c *vanish. This mesh sequence shows that mesh refinement is dictated by geometric (corner) singularities as well as periodic variations of the diffusion coefficient.*

• AFEM gives optimal rate order of convergence ≈ 0.5 while standard uniform refinement achieves the suboptimal rate 0.33 expected from theory; see Tables 5.1 and 5.2.
• AFEM performs with effectivity index close to 1.0 and reduction factors of error and oscillation close to 0.7 and 0.5 as DOF increases (see Table 5.1). The oscillation thus decreases faster than the error and becomes insignificant asymptotically for k large.
• Figure 5.1 depicts the effect of a corner singularity and periodic variation of diffusion in mesh grading; here both $\mathbf{b}$ and c vanish.

5.3. Experiment 2 : Convection Dominated-Diffusion. We consider the convection dominated-diffusion elliptic model problem (1.1) with Dirichlet boundary condition $u = g$ on convex domain $\Omega := [0,1]^2$, with isotropic diffusion coefficient $\mathbf{A} = \epsilon\mathbf{I}$, $\epsilon = 10^{-3}$, convection velocity $\mathbf{b} = (y, \frac{1}{2} - x)$, and without the zero order term $c = 0$. The Dirichlet boundary condition $g(x,y)$ on $\partial\Omega$ is the continuous piecewise linear function given by

$$g(x,y) = \begin{cases} 1 & \{.2 + \xi \le x \le .5 - \xi;\ y = 0\}, \\ 0 & \partial\Omega \setminus \{.2 \le x \le .5;\ y = 0\}, \\ \text{linear} & \{(.2 \le x \le .2 + \xi) \text{ or } (.5 - \xi \le x \le .5); y = 0\}, \end{cases}$$

where $\xi = 5 \cdot 10^{-3}$. The parameters of AFEM are $\theta = 0.65$, $\hat{\theta} = 0.6$. Results are reported in Tables 5.3, 5.4 and Figures 5.2, 5.3. Conclusions and observations follow:
• We observe from Tables 5.3, 5.4 that at about the same level of DOF, AFEM gives much smaller error estimators, which implies that AFEM reduces error estimator much faster than standard uniform refinement. Since the computational decay of estimator $\eta(\Omega)$ is close to the optimal value 0.5, the resulting meshes are quasi-optimal.
• Figure 5.2 depicts graded meshes which capture the nature of transport of a pulse from the boundary inside the domain. Figure 5.3 displays solutions without oscillations even though AFEM is not stabilized. Mesh grading is more noticeable at the boundary and internal layers, whereas the rest of the mesh barely changes.

DOF	$\eta(\Omega)$	EOC(η)	Z_o
1084	1.47435e-02	9.39685	0.32533
1350	1.05868e-02	1.50924	0.58414
1777	6.96381e-03	1.52417	0.37741
2487	4.89377e-03	1.04943	0.35142
3672	3.44244e-03	0.90280	0.40186
5785	2.53812e-03	0.67048	0.47686
10232	1.78286e-03	0.61939	0.41112
19708	1.23101e-03	0.56504	0.47391
42564	8.09329e-04	0.54466	0.36558

TABLE 5.3

Experiment 2 (Convection Dominated-Diffusion): Parameters of AFEM are $\theta = 0.65$, $\hat{\theta} = 0.6$, and model parameters $\epsilon = 0.001$ and $\xi = 0.005$. The optimal decay ≈ 0.5 of estimator $\eta(\Omega)$ is computational evidence of optimal meshes.

5.4. Experiment 3 : Drift-Diffusion Model. We consider a model problem that comes from a mathematical model in semi-conductors and chemotaxis.

DOF	$\eta(\Omega)$	EOC(η)	Z_o
2048	2.42501e-02	0.05403	0.21985
4096	1.64387e-02	0.56090	0.31395
8192	1.31271e-02	0.32454	0.21355
16384	9.79366e-03	0.42263	0.27699
32768	7.20200e-03	0.44345	0.23857
65536	5.57928e-03	0.36832	0.25634

TABLE 5.4

Experiment 2 (Convection Dominated-Diffusion): Standard uniform refinement is performed using the same values for parameters ϵ and ξ as that of AFEM given in Table 5.3. The resolution with about 6.5×10^4 elements is comparable with adapted meshes of about 2×10^3 elements.

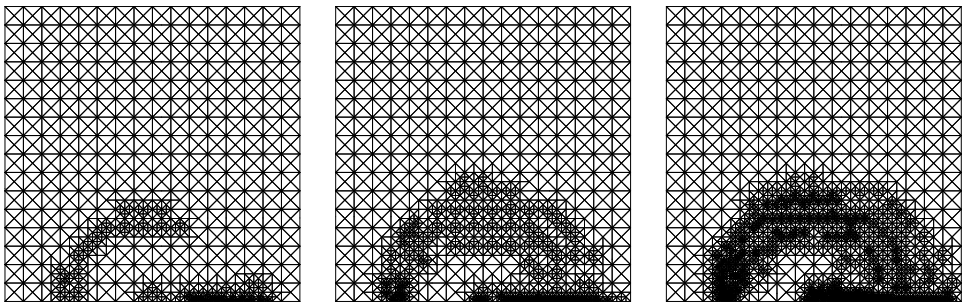

FIG. 5.2. *Experiment 2: Adaptively refined meshes after 3, 5, and 7 iterations. AFEM defects the region of rapid variation (circular transport of a pulse) and boundary layer in the outflow. The rest of the mesh remains unchanged*

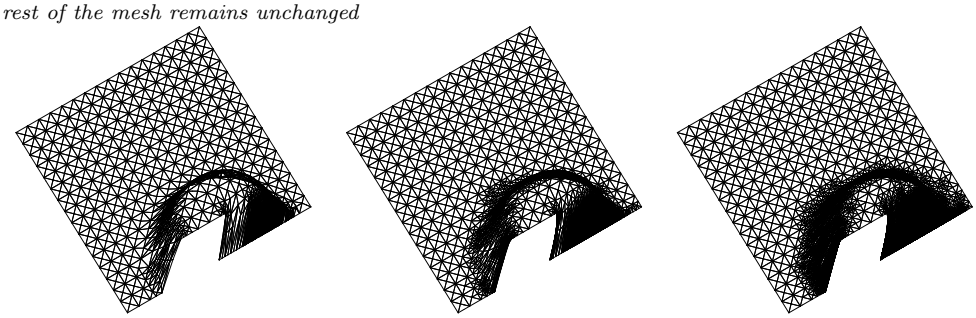

FIG. 5.3. *Experiment 2: The corresponding plots of solutions after 3, 5, and 7 iterations. No oscillations are detected even though AFEM is not stabilized. Adaptivity incorporates a nonlinear stabilization mechanism.*

$$\begin{aligned} -\nabla\cdot(\nabla u + \chi u \nabla\psi) &= 0 \qquad && \text{in } \Omega := [0,1]^2, \\ u &= g \qquad && \text{on } \Gamma \subset \partial\Omega, \\ \partial_\nu u &= 0 \qquad && \text{on } \partial\Omega \setminus \Gamma, \end{aligned}$$

where χ is a constant. The radial function ψ is defined on Ω by

$$\psi(x,y) := \begin{cases} 1 & \{\sqrt{x^2+y^2} \le r_1\}, \\ \alpha & \{\sqrt{x^2+y^2} \ge r_1 + \alpha\}, \\ \text{linear} & \{r_1 < \sqrt{x^2+y^2} < r_1 + \alpha\}, \end{cases}$$

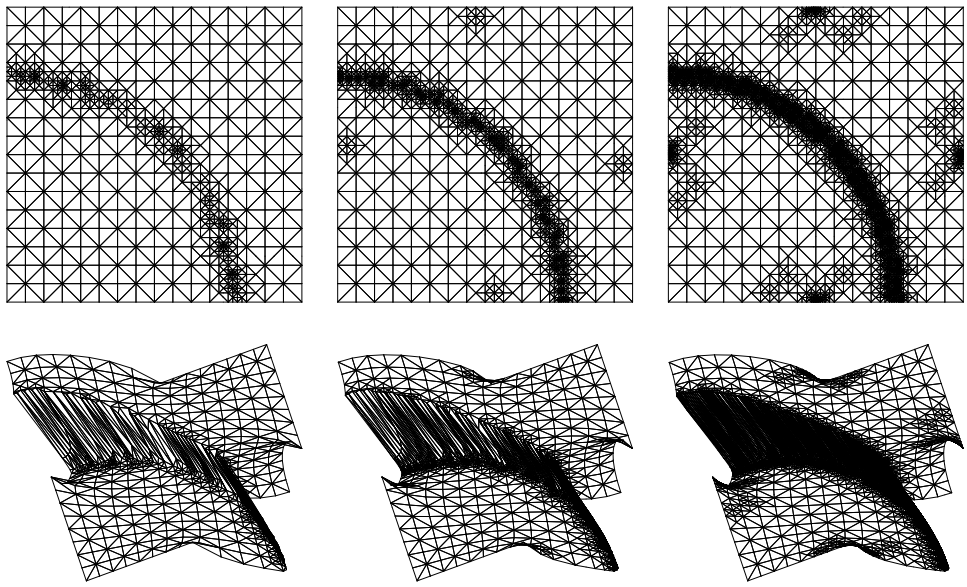

FIG. 5.4. *Experiment 3: Refined meshes after $k = 6, 8, 10$ iterations and corresponding solutions u_k. Mesh grading is quite pronounced in the internal layer where $\nabla\psi$ does not vanish, and at the midpoints of the boundary sides, where boundary conditions change.*

where α is a small parameter and $r_1 < 1$ is a constant. The Dirichlet boundary condition on Γ is assumed to be

$$g(x,y) = \begin{cases} 1 & \{x = 0; 0 \le y \le 0.5\} \bigcup \{y = 0; 0 \le x \le 0.5\}, \\ 0 & \{x = 1; 0.5 \le y \le 1\} \bigcup \{y = 1; 0.5 \le x \le 1\}. \end{cases}$$

We resort to the following transformation (exponential fitting) to symmetrize the problem

$$\rho := \exp(\chi\psi)u \implies -\nabla\cdot(\exp(-\chi\psi)\nabla\rho) = 0,$$

which gives a simpler form of model problem with variable scalar coefficient $a = \exp(-\chi\psi)$. We can now apply AFEM to solve for ρ and next compute $u = \exp(-\chi\psi)\rho$ at the nodes. Again, the experiment is performed by comparing the performance of AFEM with standard uniform refinement using parameters $\chi = 10.0, r_1 = 0.75$, and $\alpha = 0.04$ for the model problem, and parameters $\theta = 0.6$, $\hat{\theta} = 0.75$ for AFEM. Results reported in Tables 5.5, 5.6 and Figure 5.4. Conclusions and observations follow:

• From Tables 5.5, 5.6 we see again that AFEM outperforms the standard uniform refinement. Since the decay of estimator $\eta(\Omega)$ is optimal, we have computational evidence of optimal meshes.

• Figure 5.4 displays meshes and corresponding solutions u_k for iterations $k = 6, 8, 10$. Meshes adapt well to lack of smoothness, namely refinement concentrates in the transition layer, where $\nabla\psi$ does not vanish, and at the midpoints of boundary sides, where boundary conditions change.

6. Extensions. We extend the model problem (1.1) by considering now $\mathbf{A}$ with discontinuities aligned with the initial mesh and a velocity field $\mathbf{b}$ no longer divergence free. Note that if $\nabla\cdot\mathbf{b} \neq 0$, then the bilinear form $\mathcal{B}$ is non-coercive.

DOF	$\eta(\Omega)$	EOC(η)	Z_o
586	792.8314	5.3687	0.2105
630	45.6188	39.4377	0.0035
718	40.2721	0.9534	0.5993
770	21.6921	8.8487	0.3665
846	11.9151	6.3651	0.2232
1154	6.6450	1.8808	0.2673
1546	3.8249	1.8888	0.2527
2448	2.1441	1.2593	0.2069
4032	1.4556	0.7762	0.2859
6790	1.0867	0.5608	0.3408
12188	0.7370	0.6639	0.2532
23386	0.5180	0.5409	0.2870
45728	0.3632	0.5294	0.2611

TABLE 5.5

Experiment 3 (Drift-Diffusion Model): Parameters of AFEM are $\theta = 0.6$, $\hat{\theta} = 0.75$, and model parameters are $\chi = 10$, $r_1 = 0.75$ and $\alpha = 0.04$. The optimal decay ≈ 0.5 of estimator $\eta(\Omega)$ is computational evidence of quasi-optimal meshes. AFEM outperforms uniform refinement (compare with Table 5.6).

DOF	$\eta(\Omega)$	EOC(η)	Z_o
1024	179.8310	3.1860	0.0094
2048	30.7691	2.5471	0.0260
4096	11.0316	1.4798	0.0968
8192	3.9838	1.4694	0.1068
16384	2.1732	0.8744	0.1887
32768	1.2960	0.7457	0.2163
65536	0.8746	0.5675	0.2509

TABLE 5.6

Experiment 3 (Drift-Diffusion Model): Standard uniform refinement is performed using the same values for parameters χ, r_1 and α as for AFEM given in Table 5.5.

6.1. Discontinuous A. The assumption of continuity of $\mathbf{A}$ is used for obtaining error and oscillation reduction estimates, Lemma 3.1 and Lemma 3.2, in that the element oscillation $\mathsf{osc}_H(T)$ does not involve oscillation of jump residual on ∂T. Remark 4.2.2 shows that when $\mathbf{A}$ has discontinuities across element faces, we still obtain error reduction estimate (3.11) of Lemma 3.1, but this time the oscillation is defined by (4.10) and involves oscillation of jump residual. To prove convergence it suffices to show the oscillation reduction estimate (3.12), for the new concept of element oscillation, namely $\mathsf{osc}_H(T)^2 = \mathsf{osc}_{R,H}(T)^2 + \sum_{S\subset\partial T}\mathsf{osc}_{J,H}(S)^2$ with

$$\mathsf{osc}_{R,H}(T)^2 := H_T^2 \left\| R_T(u_H) - \overline{R_T(u_H)} \right\|^2_{L^2(T)} \qquad \forall\, T \in \mathcal{T}_H,$$

$$\mathsf{osc}_{J,H}(S)^2 := H_S \left\| J_S(u_H) - \overline{J_S(u_H)} \right\|^2_{L^2(S)} \qquad \forall\, S \in \mathcal{S}_H.$$

We proceed in three steps as follows.

1. *Oscillation of Interior Residual.* Invoking the same arguments as in the proof of Lemma 3.2 in §4.3, we obtain an oscillation reduction estimate for interior residual

$$\mathsf{osc}_{R,h}(T')^2 \le (1+\delta)\gamma_{T'}^2 \mathsf{osc}_{R,H}(T')^2 + C_*(1+\delta^{-1})\,\|\varepsilon_H\|^2_{H^1(T')} \qquad \forall\, T' \in \mathcal{T}_H,$$

where $\mathsf{osc}_{R,h}(T')$ is defined to be $\mathsf{osc}_h(T')$ in (4.12).

2. *Oscillation of Jump Residual.* To obtain estimate for $\mathsf{osc}_{J,h}(S)$ we write

$$J_S(u_h) = \gamma_S \left[\!\left[\mathbf{A}\nabla u_H\right]\!\right]_S \cdot \nu_S + \left[\!\left[\mathbf{A}\nabla \varepsilon_H\right]\!\right]_S \cdot \nu_s = \gamma_S J_S(u_H) + J_S(\varepsilon_H),$$

where $\gamma_S = 1$ if $S \subset S' \in \mathcal{S}_H$ and $\gamma_S = 0$ otherwise, since $\mathbf{A}\nabla u_H$ is continuous on S in the second case. Using Young's inequality, we have for all $\delta > 0$

$$\begin{aligned}\mathsf{osc}_{J,h}(S)^2 \le (1+\delta)\gamma_S h_S \left\| J_S(u_H) - \overline{J_S(u_H)}\right\|^2_{L^2(S)} \\ + (1+\delta^{-1}) h_S \left\| J_S(\varepsilon_H) - \overline{J_S(\varepsilon_H)}\right\|^2_{L^2(S)},\end{aligned}$$

where $\overline{J_S(u_H)}$ and $\overline{J_S(\varepsilon_H)}$ are L_2-projections of $J_S(u_H)$ and $J_S(\varepsilon_H)$ onto $\mathbb{P}_{n-1}(S)$. For the second term we observe that

$$\begin{aligned}\left\| J_S(\varepsilon_H) - \overline{J_S(\varepsilon_H)}\right\|_{L^2(S)} &\le \|J_S(\varepsilon_H)\|_{L^2(S)} = \left\|\left[\!\left[\mathbf{A}\nabla\varepsilon_H\right]\!\right]_S \cdot \nu_S\right\|_{L^2(S)} \\ &\le \left\|\mathbf{A}^+\nabla\varepsilon_H^+ \cdot \nu_S\right\|_{L^2(S)} + \left\|\mathbf{A}^-\nabla\varepsilon_H^- \cdot \nu_S\right\|_{L^2(S)} \\ &\le \|\mathbf{A}\|_{L^\infty(\omega_S)} \left(\left\|\nabla\varepsilon_H^+\right\|_{L^2(S)} + \left\|\nabla\varepsilon_H^-\right\|_{L^2(S)}\right) \\ &\le C_A h_S^{-1/2} \|\varepsilon_H\|_{H^1(\omega_S)},\end{aligned}$$

where C_A depends on $\mathbf{A}$ and shape regularity constant γ^*. For simplicity, let $\mathcal{S}_h(T')$ denote all $S \in \mathcal{S}_h$ contained in $T' \in \mathcal{T}_H$; hence

$$\begin{aligned}\mathsf{osc}_{J,h}(T')^2 &= \sum_{S\in\mathcal{S}_h(T')} \mathsf{osc}_{J,h}(S)^2 \\ &\le (1+\delta) \sum_{S\in\mathcal{S}_h(T')} \gamma_S h_S \left\| J_S(u_H) - \overline{J_S(u_H)}\right\|^2_{L^2(S)} + (1+\delta^{-1}) C_A \|\varepsilon_H\|^2_{H^1(\omega_{T'})}.\end{aligned}$$

In light of reduction factor of element size $h_S \le \gamma_{T'} H_{S'}$, and definitions of γ_S and $\gamma_{T'}$, we obtain

$$\begin{aligned}\sum_{S\in\mathcal{S}_h(T')} \gamma_S h_S \left\| J_S(u_H) - \overline{J_S(u_H)}\right\|^2_{L^2(S)} &\le \gamma_{T'} \sum_{S'\in\mathcal{S}_H(T')} H_{S'} \left\| J_{S'}(u_H) - \overline{J_{S'}(u_H)}\right\|^2_{L^2(S')} \\ &= \gamma_{T'} \mathsf{osc}_{J,H}(T')^2,\end{aligned}$$

because for $S \subset S' \subset \partial T'$, we have $J_S(u_H) = J_{S'}(u_H)$ and $\overline{J_S(u_H)}$ is the best L_2 estimate for $J_S(u_H)$ on S. Therefore

$$\mathsf{osc}_{J,h}(T')^2 \le (1+\delta)\gamma_{T'} \mathsf{osc}_{J,H}(T')^2 + (1+\delta^{-1}) C_A \|\varepsilon_H\|^2_{H^1(\omega_{T'})} \qquad \forall\, T' \in \mathcal{T}_H.$$

3. *Choice of δ.* Combining results from steps 1 and 2 above using $\gamma_{T'} \le 1$, $C_{**} = \max\{C_*, C_A\}$ and definition of $\mathsf{osc}_h(T)$, we arrive at

$$\mathsf{osc}_h(T')^2 \le (1+\delta)\gamma_{T'} \mathsf{osc}_H(T')^2 + C_{**}(1+\delta^{-1}) \|\varepsilon_H\|^2_{H^1(\omega_{T'})}.$$

Proceeding as in step 3 of the proof of Lemma 3.2, this time with Marking Strategy O performed according to the new definition of $\mathsf{osc}_H(T)$, we arrive at

$$\mathsf{osc}_h(\Omega)^2 \le (1+\delta)(1-(1-\gamma_0)\hat{\theta}^2)\mathsf{osc}_H(\Omega)^2 + C_o(1+\delta^{-1}) |\!|\!| \varepsilon_H |\!|\!|^2,$$

with $C_o = C_{**} c_B^{-1}$. The assertion thus follows by choosing δ sufficiently small so that

$$\rho_1 := (1+\delta)(1-(1-\gamma_0)\hat{\theta}^2) < 1, \qquad \rho_2 := C_o(1+\delta^{-1}).$$

6.2. Non-coercive $\mathcal{B}$. In this section we prove convergence of AFEM for non-coercive bilinear forms $\mathcal{B}$. According to what we have so far, the assumption of $\mathcal{B}$ being coercive is used for proving quasi-orthogonality and for having equivalence between energy norm $\|\!|v|\!\|^2 := \mathcal{B}[v,v]$ and H^1-norm as in (2.3). Since now $\mathcal{B}$ is no longer coercive, we cannot define energy norm in this manner. We instead define energy norm by $\|\!|v|\!\|^2 := \int_\Omega \mathbf{A}\nabla v \cdot \nabla v + c\, v^2$, and we have equivalence of norms

$$c_E \|v\|^2_{H^1(\Omega)} \le \|\!|v|\!\|^2 \le C_E \|v\|^2_{H^1(\Omega)}, \tag{6.1}$$

where constants c_E and C_E depend only on data $\mathbf{A}, c$ and Ω. The lack of coercivity is now replaced by Gårding's inequality

$$\|\!|v|\!\|^2 - \gamma_G \|v\|^2_{L^2(\Omega)} \le \mathcal{B}[v,v] \qquad \forall\, v \in H^1_0(\Omega), \tag{6.2}$$

where $\gamma_G = \|\nabla\cdot\mathbf{b}\|_\infty/2$. To see this we integrate by parts the middle term of $\mathcal{B}[v,v]$,

$$\int_\Omega \mathbf{b}\cdot\nabla v\; v = \frac{1}{2}\int_\Omega \mathbf{b}\cdot\nabla(v^2) = -\int_\Omega \frac{\nabla\cdot\mathbf{b}}{2} v^2 \qquad \forall\, v \in H^1_0(\Omega).$$

The same calculation leads to the sharp upper bound for $\mathcal{B}[v,v]$:

$$\mathcal{B}[v,v] \le \|\!|v|\!\|^2 + \gamma_G \|v\|^2_{L^2(\Omega)} \qquad \forall\, v \in H^1_0(\Omega). \tag{6.3}$$

Existence and uniqueness of weak solutions follows from the maximum principle for $c \ge 0$ [5]. Schatz showed in [9] that the discrete problem (2.5) has a unique solution if the meshsize h is sufficiently small, i.e. $h \le h^*$ for some constant h^* depending on shape regularity and data but not computable; the results in [9] are valid also for graded meshes. Assuming $h_0 \le h^*$, to prove convergence of AFEM it thus suffices to prove quasi-orthogonality. We follow the steps of Lemma 2.1.

Using the same notation as in §4 for e_h, e_H and ε_H, expanding $\mathcal{B}[e_H, e_H]$, and noticing that $e_H = e_h + \varepsilon_H$ and $\mathcal{B}[e_h, \varepsilon_H] = 0$, we arrive at

$$\mathcal{B}[e_h,e_h] = \mathcal{B}[e_H,e_H] - \mathcal{B}[\varepsilon_H,\varepsilon_H] - \mathcal{B}[\varepsilon_H,e_h], \tag{6.4}$$

where this time integration by parts yields

$$\begin{aligned}\mathcal{B}[\varepsilon_H,e_h] &= \mathcal{B}[e_h,\varepsilon_H] + \langle \mathbf{b}\cdot\nabla\varepsilon_H, e_h\rangle - \langle \mathbf{b}\cdot\nabla e_h, \varepsilon_H\rangle \\ &= 2\,\langle \mathbf{b}\cdot\nabla\varepsilon_H, e_h\rangle + \langle \nabla\cdot\mathbf{b}\; e_h, \varepsilon_H\rangle\,.\end{aligned}$$

Consequently, using Cauchy-Schwarz inequality and (6.1), we have for all $\delta > 0$

$$|\mathcal{B}[\varepsilon_H,e_h]| \le (2\|\mathbf{b}\|_\infty \|\nabla\varepsilon_H\|_{L^2} + \|\nabla\cdot\mathbf{b}\|_\infty \|\varepsilon_H\|_{L^2})\,\|e_h\|_{L^2} \le C_b^2\delta\,\|\!|\varepsilon_H|\!\|^2 + \delta^{-1}\|e_h\|^2_{L^2},$$

where constant $C_b = \max\{2\|\mathbf{b}\|_\infty, \|\nabla\cdot\mathbf{b}\|_\infty\}\, c_E^{-1}/2$.

Using (6.2) and (6.3) to estimate terms $\mathcal{B}[e_h,e_h], \mathcal{B}[e_H,e_H], \mathcal{B}[\varepsilon_H,\varepsilon_H]$ in (6.4), and combining with the previous estimate, we infer that

$$\|\!|e_h|\!\|^2 - (\gamma_G + \delta^{-1})\|e_h\|^2_{L^2} \le \|\!|e_H|\!\|^2 + \gamma_G\|e_H\|^2_{L^2} - (1 - C_b^2\delta)\,\|\!|\varepsilon_H|\!\|^2 + \gamma_G\|\varepsilon_H\|^2_{L^2}.$$

Since $\|\varepsilon_H\|^2_{L^2} \le 2\|e_h\|^2_{L^2} + 2\|e_H\|^2_{L^2}$, estimates for $\|e_h\|_{L^2}$ and $\|e_H\|_{L^2}$ of the form (4.1), obtained via duality, imply

$$\Lambda_h\; \|\!|e_h|\!\|^2 \le \Lambda_H\,\|\!|e_H|\!\|^2 - \Lambda_\varepsilon\,\|\!|\varepsilon_H|\!\|^2, \tag{6.5}$$

where

$$\Lambda_h = 1 - C_6^2 h_0^{2s}(3\gamma_G + \delta^{-1}), \quad \Lambda_H = 1 + 3\gamma_G C_6^2 h_0^{2s}, \quad \Lambda_\varepsilon = 1 - C_b^2\delta.$$

Consequently, to get $\Lambda_h = \Lambda_\varepsilon$, it is easy to see that we need to choose δ depending on h_0,

$$\delta(h_0) = \frac{C_G h_0^{2s} + \sqrt{C_G^2 h_0^{4s} + 4C_b^2 C_6^2 h_0^{2s}}}{2C_b^2} > 0,$$

where $C_G = 3\gamma_G C_6^2$. The assertion thus follows provided the meshsize h_0 is sufficiently small so that $C_b^2\delta(h_0) < 1$. This can be achieved for $h_0^s \le \min\left\{C_6 C_b C_G^{-1}, (3C_6C_b)^{-1}\right\}$ because

$$\begin{aligned} C_b^2\delta(h_0) &= \frac{C_G}{2}h_0^{2s} + C_b C_6 h_0^s\sqrt{1 + h_0^{2s}C_G^2(4C_b^2C_6^2)^{-1}} \\ &\le 2C_bC_6h_0^s\big(1 + h_0^s C_G(4C_bC_6)^{-1}\big) < 3C_bC_6h_0^s \le 1. \end{aligned}$$

We conclude that if the meshsize h_0 of the initial mesh satisfies

$$h_0^s \le \min\left\{C_6C_bC_G^{-1}, (3C_6C_b)^{-1}, (h^*)^s\right\},$$

then quasi-orthogonality holds, i.e. for $\Lambda_0 := \Lambda_H/\Lambda_h$,

$$\left|\!\left|\!\left| e_h \right|\!\right|\!\right|^2 \le \Lambda_0 \left|\!\left|\!\left| e_H \right|\!\right|\!\right|^2 - \left|\!\left|\!\left| \varepsilon_H \right|\!\right|\!\right|^2, \tag{6.6}$$

and Λ_0 can be made arbitrarily close to 1 by decreasing h_0. Convergence of AFEM finally follows as in Theorem 1.

REFERENCES

[1] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, John Wiley & Sons, Inc., 2000.

[2] Z. Chen and F. Jia, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*, Math. Comp. (to appear)

[3] Ph. Ciarlet, *The Finite Element Method for Elliptic Problems.*, North-Holland, Amsterdam, 1978.

[4] W. Dörfler. *A convergent adaptive algorithm for poisson's equation*, SIAM J. Numer. Anal., 33(1996), pp.1106-1124.

[5] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Germany, 1983.

[6] P. Morin, R.H. Nochetto, and K.G. Siebert, *Data oscillation and convergence of adaptive FEM*, SIAM J. Numer. Anal., 38, 2 (2000), pp.466-488.

[7] P. Morin, R.H. Nochetto, and K.G. Siebert, *Convergence of adaptive finite element methods*, SIAM Review, 44 (2002), pp.631-658.

[8] R.H. Nochetto, *Removing the saturation assumption in a posteriori error analysis*,Istit. Lombardo Sci. Lett. Rend. A, 127 (1993), 67-82.

[9] A.H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. 28 (1974), pp.959-962.

[10] A. Schmidt and K.G. Siebert, *ALBERT: an adaptive hierarchical finite element toolbox*, Documentation, Preprint 06/2000, Universität Freiburg.

[11] A. Schmidt and K.G. Siebert, *ALBERT - software for scientific computations and applications*, Acta Mathematica Universitatis Comenianae 70 (2001), 105-122.

[12] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Technique*, Wiley-Teubner, Chichester, 1996.