INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 8, Number 4, Pages 615–640

CONVERGENCE AND COMPLEXITY OF ADAPTIVE FINITE ELEMENT METHODS FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

LIANHUA HE AND AIHUI ZHOU

Abstract. In this paper, we study adaptive finite element approximations in a perturbation framework, which makes use of the existing adaptive finite element analysis of a linear symmetric elliptic problem. We analyze the convergence and complexity of adaptive finite element methods for a class of elliptic partial differential equations when the initial finite element mesh is sufficiently fine. For illustration, we apply the general approach to obtain the convergence and complexity of adaptive finite element methods for a nonsymmetric problem, a nonlinear problem as well as an unbounded coefficient eigenvalue problem.

Key Words. Adaptive finite element, convergence, complexity, eigenvalue, nonlinear, nonsymmetric, unbounded.

1. Introduction

The purpose of this paper is to study the convergence and complexity of adaptive finite element computations for a class of elliptic partial differential equations of second order and to apply our general approach to three problems: a nonsymmetric problem, a nonlinear problem, and an eigenvalue problem with an unbounded coefficient. One technical tool for motivating this work is the relationship between the general problem and a linear symmetric elliptic problem, which is derived from some perturbation arguments (see Theorem 3.1 and Lemma 3.1).

Since Babuška and Vogelius [3] gave an analysis of an adaptive finite element method (AFEM) for linear symmetric elliptic problems in one dimension, there has been much work on the convergence and complexity of adaptive finite element methods in the literature. For instance, Dörfler [10] presented the first multidimensional convergence result and Binev, Dehmen, and DeVore [5] showed the first complexity work, which have been improved and generalized in [5, 6, 9, 12, 13, 18, 19, 20, 21, 25], from convergence to convergent rate and complexity. For a nonsymmetric problem, in particular, Mekchay and Nochetto [18] imposed a quasi-orthogonality property instead of the Pythagoras equality to prove the convergence of AFEM while Morin, Siebrt, and Veeser [21] showed the convergence of error and estimator simultaneously with the strict error reduction and derived the convergence of the estimator by exploiting the (discrete) local lower but not the upper bound. In this paper, we can get the convergence and optimal complexity of nonsymmetric problems from our general approach directly. For a nonlinear problem, Chen, Holst and Xu [7] proved

Received by the editors October 28, 2010 and, in revised form, May 18, 2011.

2000 *Mathematics Subject Classification.* 65N15, 65N25, 65N30.

This work was partially supported by the National Science Foundation of China under grants 10871198 and 10971059, the National Basic Research Program of China under grant 2005CB321704, and the National High Technology Research and Development Program of China under grant 2009AA01A134.

the convergence of an adaptive finite element algorithm for Poisson-Boltzmann equation while we are able to obtain the convergence and optimal complexity of AFEM for a class of nonlinear problems now. For a smooth coefficient eigenvalue problem, Dai, Xu, and Zhou [9] gave the convergence and optimal complexity of AFEM for symmetric elliptic eigenvalue problems with piecewise smooth coefficients (see, also convergence analysis of a special case [12, 13]). In this paper, we will derive similar results for an unbounded coefficient eigenvalue problem from our general conclusions, too. We mention that a similar perturbation approach was used in [9].

This paper is organized as follows. In Section 2, we review some existing results on the convergence and complexity analysis of AFEM for the typical problem. In Section 3, we generalize results to a general model problem by using a perturbation argument when the initial finite element mesh is sufficiently fine. In Section 4 and Section 5, we provide three typical applications for illustration, including theory and numerics.

2. Adaptive FEM for a typical problem

In this section, we review some existing results on the convergence and complexity analysis of AFEM for a boundary value problem in the literature.

Let $\Omega \subset \mathbb{R}^d (d \geq 2)$ be a bounded polytopic domain. We shall use the standard notation for Sobolev spaces $W^{s,p}(\Omega)$ and their associated norms and seminorms, see, e.g., [1, 8]. For $p = 2$, we denote $H^s(\Omega) = W^{s,2}(\Omega)$ and $H_0^1(\Omega) = \{v \in H^1(\Omega) : v \mid_{\partial\Omega} = 0\}$, where $v \mid_{\partial\Omega} = 0$ is understood in the sense of trace, $\| \cdot \|_{s,\Omega} = \| \cdot \|_{s,2,\Omega}$. The space $H^{-1}(\Omega)$, the dual space of $H_0^1(\Omega)$, will also be used. Throughout this paper, we shall use C to denote a generic positive constant which may stand for different values at its different occurrences. We will also use $A \lesssim B$ to mean that $A \leq CB$ for some constant C that is independent of mesh parameters. All constants involved are independent of mesh sizes.

2.1. A boundary value problem. Consider a homogeneous boundary value problem:

$$\left\{ \begin{array}{rl} Lu := -\nabla \cdot (\mathbf{A}\nabla u) & = f \ \text{ in } \ \Omega, \\ u & = 0 \ \text{ on } \partial\Omega, \end{array} \right. \tag{1}$$

where $\mathbf{A} : \Omega \to \mathbb{R}^{d\times d}$ is piecewise Lipschitz over initial triangulation $\mathcal{T}_0$ and symmetric positive definite with smallest eigenvalue uniformly bounded away from 0 and $f \in L^2(\Omega)$.

Remark 2.1. *The choice of homogeneous boundary condition is made for ease of presentation, since similar results are valid for other boundary conditions* [6].

The weak form of (1) reads: find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \qquad \forall v \in H_0^1(\Omega), \tag{2}$$

where $a(\cdot, \cdot) = (\mathbf{A}\nabla\cdot, \nabla\cdot)$. It is seen that $a(\cdot, \cdot)$ is bounded and coercive on $H_0^1(\Omega)$, i.e., for any $w, v \in H^1(\Omega)$ there exist constants $0 < c_a \leq C_a < \infty$ such that

$$|a(w, v)| \leq C_a\|w\|_{1,\Omega}\|v\|_{1,\Omega} \quad \text{and} \quad c_a\|v\|_{1,\Omega}^2 \leq a(v, v) \ \forall w, v \in H_0^1(\Omega).$$

The energy norm $\| \cdot \|_{a,\Omega}$, which is equivalent to $\| \cdot \|_{1,\Omega}$, is defined by $\|w\|_{a,\Omega} = \sqrt{a(w, w)}$. It is known that (2) is well-posed, that is, there exists a unique solution for any $f \in H^{-1}(\Omega)$.

Let $\{\mathcal{T}_h\}$ be a shape regular family of nested conforming meshes over Ω: there exists a constant γ^* such that

$$\frac{h_\tau}{\rho_\tau} \le \gamma^* \qquad \forall \tau \in \bigcup_h \mathcal{T}_h,$$

where, for each $\tau \in \mathcal{T}_h$, h_τ is the diameter of τ, ρ_τ is the diameter of the biggest ball contained in τ, and $h = \max\{h_\tau : \tau \in \mathcal{T}_h\}$. Let $\mathcal{E}_h$ denote the set of interior sides (edges or faces) of $\mathcal{T}_h$. Let $S_0^h(\Omega) \subset H_0^1(\Omega)$ be a family of nested finite element spaces consisting of continuous piecewise polynomials over $\mathcal{T}_h$ of fixed degree $n \ge 1$, which vanish on $\partial\Omega$.

Define the Galerkin-projection $P_h : H_0^1(\Omega) \to S_0^h(\Omega)$ by

$$a(u - P_h u, v) = 0 \quad \forall v \in S_0^h(\Omega). \tag{3}$$

For any $u \in H_0^1(\Omega)$, there apparently hold:

$$\|P_h u\|_{a,\Omega} \lesssim \|u\|_{a,\Omega} \quad \text{and} \quad \lim_{h\to 0} \|u - P_h u\|_{a,\Omega} = 0.$$

Now we introduce the following quantity:

$$\rho_\Omega(h) = \sup_{f\in L^2(\Omega), \|f\|_{0,\Omega}=1} \inf_{v\in S_0^h(\Omega)} \|L^{-1}f - v\|_{a,\Omega},$$

then $\rho_\Omega(h) \to 0$ as $h \to 0$ (see, e.g., [2, 31]).

A standard finite element scheme for (2) is: find $u_h \in S_0^h(\Omega)$ such that

$$a(u_h, v) = (f, v) \qquad \forall v \in S_0^h(\Omega). \tag{4}$$

We see that $u_h = P_h u$.

By a contradiction argument, we have (c.f., e.g., [32])

Lemma 2.1. *As operators over $H_0^1(\Omega)$, there holds*

$$\lim_{h\to 0} \|\mathcal{K}(I - P_h)\| = 0$$

if $\mathcal{K}$ is a compact operator over $H_0^1(\Omega)$.

2.2. Adaptive algorithm. Given an initial triangulation $\mathcal{T}_0$, we shall generate a sequence of nested conforming triangulations $\mathcal{T}_k$ using the following loop:

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

More precisely, to get $\mathcal{T}_{k+1}$ from $\mathcal{T}_k$ we first solve the discrete equation to get u_k on $\mathcal{T}_k$. The error is then estimated using u_k and used to mark a set of elements that are to be refined. Elements are refined in such a way that the triangulation is still shape regular and conforming. We assume that we have the exact solutions of finite-dimensional problems.[1]

Now we review the residual type a posteriori error estimators for finite element solutions of (1). Let $\mathbb{T}$ denote the class of all conforming refinements by bisection of $\mathcal{T}_0$. For $\mathcal{T}_h \in \mathbb{T}$ and any $v \in S_0^h(\Omega)$ we define the element residual $\tilde{\mathcal{R}}_\tau(v)$ and the jump residual $\tilde{J}_e(v)$ by

$$\begin{aligned} \tilde{\mathcal{R}}_\tau(v) &= f - Lv = f + \nabla\cdot(\mathbf{A}\nabla v) \qquad \text{in } \tau \in \mathcal{T}_h, \\ \tilde{J}_e(v) &= -\mathbf{A}\nabla v^+ \cdot \nu^+ - \mathbf{A}\nabla v^- \cdot \nu^- = [[\mathbf{A}\nabla v]]_e \cdot \nu_e \text{ on } e \in \mathcal{E}_h, \end{aligned}$$

[1] Indeed we have ignored two important practical issues: the inexact solution of the resulting algebraic system and the numerical integration. By the similar perturbation argument, it would be seen that some approximations to the linear system will be sufficient.

where e is the common side of elements τ^+ and τ^- with unit outward normals ν^+ and ν^-, respectively, and $\nu_e = \nu^-$. Let ω_e be the union of elements which share the side e and ω_τ be the union of elements sharing a side with τ.

For $\tau \in \mathcal{T}_h$, we define the local error indicator $\tilde{\eta}_h(v, \tau)$ by

$$\tilde{\eta}_h^2(v,\tau) = h_\tau^2 \|\tilde{\mathcal{R}}_\tau(v)\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_h, e\subset\partial\tau} h_e \|\tilde{J}_e(v)\|_{0,e}^2$$

and the oscillation $\widetilde{osc}_h(v,\tau)$ by

$$\widetilde{osc}_h^2(v,\tau) = h_\tau^2 \|\tilde{\mathcal{R}}_\tau(v) - \overline{\tilde{\mathcal{R}}_\tau(v)}\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_h, e\subset\partial\tau} h_e \|\tilde{J}_e(v) - \overline{\tilde{J}_e(v)}\|_{0,e}^2,$$

where $\overline{w}$ is the L^2-projection of $w \in L^2(\Omega)$ to polynomials of some degree on τ or e.

Given a subset $\mathcal{T}' \subset \mathcal{T}_h$, we define the error estimator $\tilde{\eta}_h(v, \mathcal{T}')$ and the oscillation $\widetilde{osc}_h(v, \mathcal{T}')$ by

$$\tilde{\eta}_h^2(v,\mathcal{T}') = \sum_{\tau\in\mathcal{T}'} \tilde{\eta}_h^2(v,\tau) \quad \text{and} \quad \widetilde{osc}_h^2(v,\mathcal{T}') = \sum_{\tau\in\mathcal{T}'} \widetilde{osc}_h^2(v,\tau).$$

For $\tau \in \mathcal{T}_h$, we also need notations

$$\eta_h^2(\mathbf{A},\tau) = h_\tau^2(\|\mathrm{div}\mathbf{A}\|_{0,\infty,\tau}^2 + h_\tau^{-2}\|\mathbf{A}\|_{0,\infty,\omega_\tau}^2)$$

and

$$osc_h^2(\mathbf{A},\tau) = h_\tau^2(\|\mathrm{div}\mathbf{A} - \overline{\mathrm{div}\mathbf{A}}\|_{0,\infty,\tau}^2 + h_\tau^{-2}\|\mathbf{A} - \bar{\mathbf{A}}\|_{0,\infty,\omega_\tau}^2),$$

where $\overline{v}$ is the best L^∞-approximation in the space of discontinuous polynomials of some degree.

For $\mathcal{T}' \subset \mathcal{T}_h$, we finally set

$$\eta_h(\mathbf{A},\mathcal{T}') = \max_{\tau\in\mathcal{T}'} \eta_h(\mathbf{A},\tau) \quad \text{and} \quad osc_h(\mathbf{A},\mathcal{T}') = \max_{\tau\in\mathcal{T}'} osc_h(\mathbf{A},\tau).$$

We now recall the well-known upper and lower bounds for the energy error in terms of the residual type estimator (see, e.g., [18, 20, 28]).

Theorem 2.1. *Let $u \in H_0^1(\Omega)$ be the solution of (2) and $u_h \in S_0^h(\Omega)$ be the solution of (4). Then there exist constants $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3 > 0$ depending only on the shape regularity γ^*, C_a and c_a such that*

$$\|u - u_h\|_{a,\Omega}^2 \le \tilde{C}_1 \tilde{\eta}_h^2(u_h, \mathcal{T}_h) \tag{5}$$

and

$$\tilde{C}_2 \tilde{\eta}_h^2(u_h,\mathcal{T}_h) \le \|u - u_h\|_{a,\Omega}^2 + \tilde{C}_3 \widetilde{osc}_h^2(u_h,\mathcal{T}_h). \tag{6}$$

We replace the subscript h by an iteration counter called k and call the adaptive algorithm without oscillation marking as **Algorithm** D_0, which is defined by:

Choose a parameter $0 < \theta < 1$:

1. Pick an initial mesh $\mathcal{T}_0$, and let $k = 0$.
2. Solve the system on $\mathcal{T}_k$ for the discrete solution u_k.
3. Compute the local indicators $\{\tilde{\eta}_k(u_k, \tau) : \tau \in \mathcal{T}_k\}$.
4. Construct $\mathcal{M}_k \subset \mathcal{T}_k$ by **Marking Strategy** E_0 and parameter θ.
5. Refine $\mathcal{T}_k$ to get a new conforming mesh $\mathcal{T}_{k+1}$ by Procedure **REFINE**.
6. Solve the system on $\mathcal{T}_{k+1}$ for the discrete solution u_{k+1}.
7. Let $k = k + 1$ and go to Step 2.

Marking Strategy E_0, which is crucial for our adaptive methods, is stated as follows:

Given a parameter $0 < \theta < 1$:

1. Construct a minimal subset $\mathcal{M}_k$ of $\mathcal{T}_k$ by selecting some elements in $\mathcal{T}_k$ such that
$$\tilde{\eta}_k(u_k, \mathcal{M}_k) \geq \theta \tilde{\eta}_k(u_k, \mathcal{T}_k).$$
2. Mark all the elements in $\mathcal{M}_k$.

Due to [6], the procedure **REFINE** here is not required to satisfy the Interior Node Property of [18, 20].

Given a fixed number $b \geq 1$, for any $\mathcal{T}_k \in \mathbb{T}$ and a subset $\mathcal{M}_k \subset \mathcal{T}_k$ of marked elements,

$$\mathcal{T}_{k+1} = \textbf{REFINE}(\mathcal{T}_k, \mathcal{M}_k)$$

outputs a conforming triangulation $\mathcal{T}_{k+1} \in \mathbb{T}$, where at least all elements of $\mathcal{M}_k$ are bisected b times. We define $R_{\mathcal{T}_k \to \mathcal{T}_{k+1}} = \mathcal{T}_k \backslash (\mathcal{T}_k \cap \mathcal{T}_{k+1})$ as the set of refined elements, thus $\mathcal{M}_k \subset R_{\mathcal{T}_k \to \mathcal{T}_{k+1}}$.

Lemma 2.2. *([26]) Assume that $\mathcal{T}_0$ verifies condition (b) of section 4 in [26]. For $k \geq 0$ let $\{\mathcal{T}_k\}_{k \geq 0}$ be any sequence of refinements of $\mathcal{T}_0$ where $\mathcal{T}_{k+1}$ is generated from $\mathcal{T}_k$ by $\mathcal{T}_{k+1} = \textbf{REFINE}(\mathcal{T}_k, \mathcal{M}_k)$ with a subset $\mathcal{M}_k \subset \mathcal{T}_k$. Then*

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \lesssim \sum_{j=0}^{k-1} \#\mathcal{M}_j \quad \forall k \geq 1 \tag{7}$$

is valid, where the hidden constant depends on $\mathcal{T}_0$ and b.

The convergence of **Algorithm** D_0 is shown in [6] and stated as follows.

Theorem 2.2. *Let $\{u_k\}_{k \in \mathbb{N}_0}$ be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces $\{S_0^k(\Omega)\}_{k \in \mathbb{N}_0}$ produced by* **Algorithm** *D_0. Then there exist constants $\tilde{\gamma} > 0$ and $\tilde{\xi} \in (0,1)$ depending only on the shape regularity of $\mathcal{T}_0$, b and the marking parameter θ, such that for any two consecutive iterates we have*

$$\begin{aligned} & \|u - u_{k+1}\|_{a,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_{k+1}^2(u_{k+1}, \mathcal{T}_{k+1}) \\ \leq \quad & \tilde{\xi}^2 \big(\|u - u_k\|_{a,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_k^2(u_k, \mathcal{T}_k) \big). \end{aligned}$$

Indeed, the constant $\tilde{\gamma}$ has the following form

$$\tilde{\gamma} = \frac{1}{(1 + \delta^{-1}) \Lambda_1 \eta_0^2(\mathbf{A}, \mathcal{T}_0)}, \tag{8}$$

where $\eta_0^2(\mathbf{A}, \mathcal{T}_0) = \eta_{\mathcal{T}_0}^2(\mathbf{A}, \mathcal{T}_0)$, $\Lambda_1 = (d+1) C_0^2 / c_a$ with C_0 some positive constant and constant $\delta \in (0,1)$.

Following [6, 9, 25], we have

Lemma 2.3. *Let $u_H \in S_0^H(\Omega)$ and $u_h \in S_0^h(\Omega)$ be finite element solutions of (2) over a conforming mesh $\mathcal{T}_H$ and its any refinement $\mathcal{T}_h$ with marked element $\mathcal{M}_H$. Suppose that they satisfy the decrease property*

$$\begin{aligned} & \|u - u_h\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{osc}_h^2(u_h, \mathcal{T}_h) \\ \leq \quad & \tilde{\beta}_*^2 \big(\|u - u_H\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{osc}_H^2(u_H, \mathcal{T}_H) \big) \end{aligned}$$

with constants $\tilde{\gamma}_* > 0$ *and* $\tilde{\beta}_* \in (0, \sqrt{\frac{1}{2}})$. *Then the set* $\mathcal{R} = R_{\mathcal{T}_H \to \mathcal{T}_h}$ *satisfies the following inequality*

$$\tilde{\eta}_H(u_H, \mathcal{R}) \geq \hat{\theta}\tilde{\eta}_H(u_H, \mathcal{T}_H)$$

with $\hat{\theta}^2 = \frac{\tilde{C}_2(1-2\tilde{\beta}_*^2)}{\tilde{C}_0(\tilde{C}_1+(1+2C\tilde{C}_1)\tilde{\gamma}_*)}$, *where* $C = \Lambda_1 osc_0^2(\mathbf{A}, \mathcal{T}_0)$ and $\tilde{C}_0 = \max(1, \frac{\tilde{C}_3}{\tilde{\gamma}_*})$.

3. A general framework

Let $u \in H_0^1(\Omega)$ satisfy

$$a(u, v) + (Vu, v) = (\ell u, v) \qquad \forall v \in H_0^1(\Omega), \tag{9}$$

where $\ell : H_0^1(\Omega) \to L^2(\Omega)$ is an operator and $V : H_0^1(\Omega) \to L^2(\Omega)$ is a linear bounded operator. Some applications of ℓ and V will be shown in section 4.

Let $K : L^2(\Omega) \to H_0^1(\Omega)$ be the operator defined by

$$a(Kw, v) = (w, v) \ \forall v \in H_0^1(\Omega).$$

Then K is a compact operator from $L^2(\Omega)$ to $H_0^1(\Omega)$ and (9) becomes as

$$u + KVu = K\ell u.$$

We assume that for any $f \in H^{-1}(\Omega)$, there exists a unique solution $u \in H_0^1(\Omega)$ satisfying

$$a(u, v) + (Vu, v) = (f, v) \qquad \forall v \in H_0^1(\Omega),$$

which implies $(I + KV)^{-1}$ exists as an operator over $H_0^1(\Omega)$. An application of the open-mapping theorem yields that $(I + KV)^{-1}$ is bounded as an operator over $H_0^1(\Omega)$.

For $h \in (0, 1)$, let $u_h \in S_0^h(\Omega)$ be a solution of discretization

$$a(u_h, v) + (Vu_h, v) = (\ell_h u_h, v) \ \forall v \in S_0^h(\Omega), \tag{10}$$

where $\ell_h : S_0^h(\Omega) \to L^2(\Omega)$ is some operator. Note that we may view ℓ_h as a perturbation to ℓ, for which we assume that there exists $\kappa_1(h) \in (0, 1)$ such that

$$\|K(\ell u - \ell_h u_h)\|_{a,\Omega} = \mathcal{O}(\kappa_1(h))\|u - u_h\|_{a,\Omega}, \tag{11}$$

where $\kappa_1(h) \to 0$ as $h \to 0$.

Note that (10) can be written as

$$u_h + P_h KVu_h = P_h K\ell_h u_h,$$

where P_h is defined by (3). We have for $w^h = K\ell_h u_h - KVu_h$ that

$$u_h = P_h w^h. \tag{12}$$

Now we shall establish a relationship between the error estimates of finite element approximations of (9) and finite element approximations of (1), from which various a posteriori error estimators for (10) can be easily obtained since the a posteriori error estimators for (4) have been well-constructed.

Theorem 3.1. *There exists* $\kappa(h) \in (0, 1)$ *such that* $\kappa(h) \to 0$ *as* $h \to 0$ *and*

$$\|u - u_h\|_{a,\Omega} = \|w^h - P_h w^h\|_{a,\Omega} + \mathcal{O}(\kappa(h))\|u - u_h\|_{a,\Omega}. \tag{13}$$

Proof. By the definition of w^h, (12) and note that $P_h u_h = u_h$, we have

$$\begin{aligned} u - w^h &= K\ell u - KVu - (K\ell_h u_h - KVu_h) \\ &= K(\ell u - \ell_h u_h) + KVP_h(w^h - u) + KV(P_h - I)(u - u_h), \end{aligned}$$

hence

$$(I + KVP_h)(u - w^h) = K(\ell u - \ell_h u_h) + KV(P_h - I)(u - u_h). \tag{14}$$

Since $KV : H_0^1(\Omega) \to H_0^1(\Omega)$ is compact, we get from Lemma 2.1 that

$$\lim_{h\to 0} \|KV(I - P_h)\| = 0,$$

which together with the following equality

$$I + KVP_h = (I + KV) + KV(P_h - I)$$

leads to that $(I + KVP_h)^{-1}$ exists as an operator over $H_0^1(\Omega)$ when $h \ll 1$ and

$$\limsup_{h\to 0} \|(I + KVP_h)^{-1}\| < \infty. \tag{15}$$

Set

$$\kappa(h) = \|(I + KVP_h)^{-1}\|\big(\kappa_1(h) + \|KV(I - P_h)\|\big), \tag{16}$$

we have that $\kappa(h) \to 0$ as $h \to 0$ and

$$\|u - w^h\|_{a,\Omega} \le \tilde{C}\kappa(h)\|u - u_h\|_{a,\Omega}, \tag{17}$$

where (11), (14) and (15) are used.

Since (12) implies

$$u - u_h = w^h - P_h w^h + u - w^h,$$

we get (13) from (17). This completes the proof. □

Theorem 3.1 implies that the error of the general problem is equivalent to that of the typical problem with $\ell_h u_h - V u_h$ as a source term up to the high order term. However, the high order term can not be estimated easily in the analysis of convergence and optimal complexity of AFEM for the general problem, for instance, for a nonsymmetric problem, a nonlinear problem as well as an unbounded coefficient eigenvalue problem.

3.1. Adaptive algorithm. Following the element residual $\tilde{\mathcal{R}}_\tau(u_h)$ and the jump residual $\tilde{J}_e(u_h)$ for (4), we define the element residual $\mathcal{R}_\tau(u_h)$ and the jump residual $J_e(u_h)$ for (10) as follows:

$$\begin{aligned}
\mathcal{R}_\tau(u_h) &= \ell_h u_h - V u_h - L u_h = \ell_h u_h - V u_h + \nabla\cdot(\mathbf{A}\nabla u_h) \text{ in } \tau \in \mathcal{T}_h,\\
J_e(u_h) &= -\mathbf{A}\nabla u_h^+ \cdot \nu^+ - \mathbf{A}\nabla u_h^- \cdot \nu^- = [[\mathbf{A}\nabla u_h]]_e \cdot \nu_e \text{ on } e \in \mathcal{E}_h.
\end{aligned}$$

For $\tau \in \mathcal{T}_h$, we define the local error indicator $\eta_h(u_h, \tau)$ by

$$\eta_h^2(u_h, \tau) = h_\tau^2 \|\mathcal{R}_\tau(u_h)\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_h, e\subset\partial\tau} h_e \|J_e(u_h)\|_{0,e}^2$$

and the oscillation $osc_h(u_h, \tau)$ by

$$osc_h^2(u_h, \tau) = h_\tau^2 \|\mathcal{R}_\tau(u_h) - \overline{\mathcal{R}_\tau(u_h)}\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_h, e\subset\partial\tau} h_e \|J_e(u_h) - \overline{J_e(u_h)}\|_{0,e}^2,$$

where e , ν^+ and ν^- are defined as those in section 2.

Given a subset $\mathcal{T}' \subset \mathcal{T}_h$, we define the error estimator $\eta_h(u_h, \mathcal{T}')$ by

$$\eta_h^2(u_h, \mathcal{T}') = \sum_{\tau\in\mathcal{T}'} \eta_h^2(u_h, \tau) \tag{18}$$

and the oscillation $osc_h(u_h, \mathcal{T}')$ by

$$osc_h^2(u_h, \mathcal{T}') = \sum_{\tau\in\mathcal{T}'} osc_h^2(u_h, \tau). \tag{19}$$

Let $h_0 \in (0, 1)$ be the mesh size of the initial mesh $\mathcal{T}_0$ and define

$$\tilde{\kappa}(h_0) = \sup_{h\in(0,h_0]} \max\{h, \kappa(h)\}.$$

Obviously, $\tilde{\kappa}(h_0) \ll 1$ if $h_0 \ll 1$.

To analyze the convergence and complexity of finite element approximations, we need to establish some relationship between the two level approximations. We use $\mathcal{T}_H$ to denote a coarse mesh and $\mathcal{T}_h$ to denote a refined mesh of $\mathcal{T}_H$. Recall that $w^h = K(\ell_h u_h - V u_h)$ and $w^H = K(\ell_H u_H - V u_H)$.

Lemma 3.1. *If $h, H \in (0, h_0]$, then*

$$\|u - u_h\|_{a,\Omega} = \|w^H - P_h w^H\|_{a,\Omega} + \mathcal{O}(\tilde{\kappa}(h_0)) \left(\|u - u_h\|_{a,\Omega} + \|u - u_H\|_{a,\Omega}\right), \tag{20}$$

$$\eta_h(u_h, \mathcal{T}_h) = \tilde{\eta}_h(P_h w^H, \mathcal{T}_h) + \mathcal{O}(\tilde{\kappa}(h_0)) \left(\|u - u_h\|_{a,\Omega} + \|u - u_H\|_{a,\Omega}\right), \tag{21}$$

and

$$osc_h(u_h, \mathcal{T}_h) = \widetilde{osc}_h(P_h w^H, \mathcal{T}_h) + \mathcal{O}(\tilde{\kappa}(h_0)) \left(\|u - u_h\|_{a,\Omega} + \|u - u_H\|_{a,\Omega}\right). \tag{22}$$

Proof. First, we prove (20). It follows that

$$\begin{aligned}
\|P_h(w^h - w^H) + u - w^H\|_{a,\Omega} &\lesssim \|w^h - w^H\|_{a,\Omega} + \|u - w^H\|_{a,\Omega} \\
&\lesssim \|u - w^H\|_{a,\Omega} + \|u - w^h\|_{a,\Omega},
\end{aligned}$$

which together with (17) implies

$$\|P_h(w^h - w^H) + u - w^H\|_{a,\Omega} \lesssim \tilde{\kappa}(h_0)(\|u - u_H\|_{a,\Omega} + \|u - u_h\|_{a,\Omega}). \tag{23}$$

Observing that identity (12) leads to

$$u - u_h = w^H - P_h w^H + P_h(w^H - w^h) + u - w^H,$$

we then obtain (20) from (23).

Next, we turn to prove (22). Due to $Lw^h = \ell_h u_h - V u_h$ and $Lw^H = \ell_H u_H - V u_H$, we know that $w^h - w^H$ is the solution of the typical boundary value problem with $\ell_h u_h - \ell_H u_H + V u_H - V u_h$ as a source term. Set $E = P_h(w^h - w^H)$ and since

$$\tilde{\mathcal{R}}_\tau(P_h(w^h - w^H)) = \ell_h u_h - \ell_H u_H + V u_H - V u_h - L(P_h(w^h - w^H)),$$

we have

$$\begin{aligned}
&\widetilde{osc}_h^2(P_h(w^h - w^H), \mathcal{T}_h) = \sum_{\tau \in \mathcal{T}_h} \widetilde{osc}_h^2(E, \tau) \\
= &\sum_{\tau \in \mathcal{T}_h} \left(h_\tau^2 \|\tilde{\mathcal{R}}_\tau(E) - \overline{\tilde{\mathcal{R}}_\tau(E)}\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|\tilde{J}_e(E) - \overline{\tilde{J}_e(E)}\|_{0,e}^2\right) \\
\le &\sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\tilde{\mathcal{R}}_\tau(E) + LE - \overline{(\tilde{\mathcal{R}}_\tau(E) + LE)}\|_{0,\tau}^2 \\
&+ \sum_{\tau \in \mathcal{T}_h} \left(h_\tau^2 \|LE - \overline{LE}\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|\tilde{J}_e(E) - \overline{\tilde{J}_e(E)}\|_{0,e}^2\right).
\end{aligned} \tag{24}$$

Following the proof of Proposition 3.3 in [6], we see that

$$\sum_{\tau \in \mathcal{T}_h} \left(h_\tau^2 \|LE - \overline{LE}\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|\tilde{J}_e(E) - \overline{\tilde{J}_e(E)}\|_{0,e}^2\right)$$

can be bounded by

$$\sum_{\tau \in \mathcal{T}_h} C_0^2 osc_h^2(\mathbf{A}, \tau) \|P_h(w^h - w^H)\|_{1,\omega_\tau}^2 \lesssim osc_h^2(\mathbf{A}, \mathcal{T}_h) \|P_h(w^h - w^H)\|_{a,\Omega}^2.$$

Hence using the fact $osc_h(\mathbf{A},\mathcal{T}_h)\le osc_0(\mathbf{A},\mathcal{T}_0)$, we obtain

$$\begin{aligned}&\sum_{\tau\in\mathcal{T}_h}\big(h_\tau^2\|LE-\overline{LE}\|_{0,\tau}^2+\sum_{e\in\mathcal{E}_h,e\subset\partial\tau}h_e\|\tilde J_e(E)-\overline{\tilde J_e(E)}\|_{0,e}^2\big)\\ \lesssim\ & osc_0^2(\mathbf{A},\mathcal{T}_0)\|P_h(w^h-w^H)\|_{a,\Omega}^2. \qquad (25)\end{aligned}$$

Using the inverse inequality, the bounded property of V and (11), we get

$$\begin{aligned}&\big(\sum_{\tau\in\mathcal{T}_h}h_\tau^2\|\tilde{\mathcal{R}}_\tau(E)+LE-\overline{(\tilde{\mathcal{R}}_\tau(E)+LE)}\|_{0,\tau}^2\big)^{1/2}\\ \lesssim\ &\big(\sum_{\tau\in\mathcal{T}_h}\|h_\tau(\ell_hu_h-\ell_Hu_H+Vu_H-Vu_h)\|_{0,\tau}^2\big)^{1/2}\\ \lesssim\ &\|K(\ell_hu_h-\ell_Hu_H)\|_{a,\Omega}+h\|u_H-u_h\|_{a,\Omega}\\ \lesssim\ &\|K(\ell_hu_h-\ell u)\|_{a,\Omega}+\|K(\ell_Hu_H-\ell u)\|_{a,\Omega}\\ &+h\|u-u_H\|_{a,\Omega}+h\|u-u_h\|_{a,\Omega}\\ \lesssim\ &\tilde\kappa(h_0)\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right). \qquad (26)\end{aligned}$$

Note that

$$\begin{aligned}\|P_h(w^h-w^H)\|_{a,\Omega}&\lesssim\|w^h-w^H\|_{a,\Omega}\\ &\lesssim\|u-w^h\|_{a,\Omega}+\|u-w^H\|_{a,\Omega},\end{aligned}$$

which together with (17) implies

$$\|P_h(w^h-w^H)\|_{a,\Omega}\ \lesssim\ \tilde\kappa(h_0)\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right). \qquad (27)$$

Combining (24), (25), (26) and (27), we conclude that

$$\widetilde{osc}_h(P_h(w^h-w^H),\mathcal{T}_h)\ \lesssim\ \tilde\kappa(h_0)\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right). \qquad (28)$$

Due to $u_h=P_hw^H+P_h(w^h-w^H)$, $\widetilde{osc}_h(u_h,\mathcal{T}_h)=osc_h(u_h,\mathcal{T}_h)$, (28) and the definition of oscillation, we arrive at (22).

Finally, we prove (21). By (6) and (28), we have

$$\begin{aligned}&\tilde\eta_h(P_h(w^h-w^H),\mathcal{T}_h)\\ \lesssim\ &\|(w^h-w^H)-P_h(w^h-w^H)\|_{a,\Omega}+\widetilde{osc}_h(P_h(w^h-w^H),\mathcal{T}_h)\\ \lesssim\ &\|u-w^h\|_{a,\Omega}+\|u-w^H\|_{a,\Omega}+\tilde\kappa(h_0)\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right)\\ \lesssim\ &\tilde\kappa(h_0)\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right). \qquad (29)\end{aligned}$$

From (29) and the fact that

$$\tilde\eta_h(P_hw^h,\mathcal{T}_h)=\tilde\eta_h(P_hw^H+P_h(w^h-w^H),\mathcal{T}_h),$$

we obtain

$$\tilde\eta_h(P_hw^h,\mathcal{T}_h)=\tilde\eta_h(P_hw^H,\mathcal{T}_h)+\mathcal{O}(\tilde\kappa(h_0))\left(\|u-u_h\|_{a,\Omega}+\|u-u_H\|_{a,\Omega}\right),$$

which is nothing but (21) since $\tilde\eta_h(P_hw^h,\mathcal{T}_h)=\eta_h(u_h,\mathcal{T}_h)$. This completes the proof. □

Theorem 3.2. *Let h_0 be small enough and $h\in(0,h_0]$. There exist constants C_1,C_2 and C_3, which only depend on the shape regularity constant γ^*, C_a and c_a such that*

$$\|u-u_h\|_{a,\Omega}^2\le C_1\eta_h^2(u_h,\mathcal{T}_h) \qquad (30)$$

and

$$C_2\eta_h^2(u_h,\mathcal{T}_h)\le\|u-u_h\|_{a,\Omega}^2+C_3osc_h^2(u_h,\mathcal{T}_h). \qquad (31)$$

Proof. Recall that $Lw^h = \ell_h u_h - Vu_h$. From (5) and (6) we have

$$\|w^h - P_h w^h\|_{a,\Omega}^2 \leq \tilde{C}_1 \tilde{\eta}_h^2(P_h w^h, \mathcal{T}_h) \tag{32}$$

and

$$\tilde{C}_2 \tilde{\eta}_h^2(P_h w^h, \mathcal{T}_h) \leq \|w^h - P_h w^h\|_{a,\Omega}^2 + \tilde{C}_3 \widetilde{osc}_h^2(P_h w^h, \mathcal{T}_h). \tag{33}$$

Thus we obtain (30) and (31) from (12), (13), (32) and (33). In particular, we may choose C_1, C_2 and C_3 satisfying

$$C_1 = \tilde{C}_1(1 + \tilde{C}\tilde{\kappa}(h_0))^2, C_2 = \tilde{C}_2(1 - \tilde{C}\tilde{\kappa}(h_0))^2, C_3 = \tilde{C}_3(1 - \tilde{C}\tilde{\kappa}(h_0))^2. \tag{34}$$

This completes the proof. □

Remark 3.1. *Either to ensure that the discrete problem is well-posed or to provide a structure-preserving approximation, we shall require that h_0 is small enough for a finite element approximation to (9) (see, e.g.,* [15, 31]*). We refer to* [7, 18] *for the initial mesh size requirement in adaptive finite element computations for nonlinear and nonsymmetirc boundary value problems.*

Now we address step **MARK** of solving (10) in detail, which we call **Marking Strategy** E. Similar to **Marking Strategy** E_0 for (4), we define **Marking Strategy** E for (10) to enforce error reduction as follows:

Given a parameter $0 < \theta < 1$:

1. Construct a minimal subset $\mathcal{M}_k$ of $\mathcal{T}_k$ by selecting some elements in $\mathcal{T}_k$ such that
$$\eta_k(u_k, \mathcal{M}_k) \geq \theta \eta_k(u_k, \mathcal{T}_k).$$
2. Mark all the elements in $\mathcal{M}_k$.

The adaptive algorithm of solving (10), which we call **Algorithm** D, is nothing but **Algorithm** D_0 when $\tilde{\eta}_k$ are replaced by η_k and **Marking Strategy** E_0 is replaced by **Marking Strategy** E.

3.2. Convergence. We now prove that **Algorithm** D of (10) is a contraction with respect to the sum of the energy error plus the scaled error estimator.

Theorem 3.3. *Let $\theta \in (0,1)$ and $\{u_k\}_{k\in\mathbb{N}_0}$ be a sequence of finite element solutions of (10) corresponding to a sequence of nested finite element spaces $\{S_0^k(\Omega)\}_{k\in\mathbb{N}_0}$ produced by* **Algorithm** D*. If $h_0 \ll 1$, then there exist constants $\gamma > 0$ and $\xi \in (0,1)$ depending only on the shape regularity constant γ^*, C_a, c_a and the marking parameter θ such that*

$$\begin{aligned} &\|u - u_{k+1}\|_{a,\Omega}^2 + \gamma \eta_{k+1}^2(u_{k+1}, \mathcal{T}_{k+1}) \\ \leq\ &\xi^2\big(\|u - u_k\|_{a,\Omega}^2 + \gamma \eta_k^2(u_k, \mathcal{T}_k)\big). \end{aligned} \tag{35}$$

Here,

$$\gamma = \frac{\tilde{\gamma}}{1 - C_4 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \tag{36}$$

with C_4 a positive constant.

Proof. For convenience, we use u_h, u_H to denote u_{k+1} and u_k, respectively.

We conclude from Theorem 2.2, $w^h = K(\ell_h u_h - Vu_h)$ and $w^H = K(\ell_H u_H - Vu_H)$ that there exist constants $\tilde{\gamma} > 0$ and $\tilde{\xi} \in (0,1)$ satisfying

$$\begin{aligned} &\|w^H - P_h w^H\|_{a,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_h^2(P_h w^H, \mathcal{T}_h) \\ \leq\ &\tilde{\xi}^2\big(\|w^H - P_H w^H\|_{a,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_H^2(P_H w^H, \mathcal{T}_H)\big). \end{aligned}$$

Hence using the fact that $u_H = P_H w^H$, we obtain

$$\begin{aligned} & \|w^H - P_h w^H\|_{a,\Omega}^2 + \tilde{\gamma}\tilde{\eta}_h^2(P_h w^H, \mathcal{T}_h) \\ \le \ & \tilde{\xi}^2\big(\|w^H - u_H\|_{a,\Omega}^2 + \tilde{\gamma}\eta_H^2(u_H, \mathcal{T}_H)\big). \end{aligned} \tag{37}$$

By (20) and (21), there exists a constant $\hat{C} > 0$ such that

$$\begin{aligned} & \|u - u_h\|_{a,\Omega}^2 + \tilde{\gamma}\eta_h^2(u_h, \mathcal{T}_h) \\ \le \ & (1+\delta_1)\|w^H - P_h w^H\|_{a,\Omega}^2 + (1+\delta_1)\tilde{\gamma}\tilde{\eta}_h^2(P_h w^H, \mathcal{T}_h) \\ & + \hat{C}(1+\delta_1^{-1})\tilde{\kappa}^2(h_0)(\|u - u_h\|_{a,\Omega}^2 + \|u - u_H\|_{a,\Omega}^2) \\ & + \hat{C}(1+\delta_1^{-1})\tilde{\kappa}^2(h_0)\tilde{\gamma}(\|u - u_h\|_{a,\Omega}^2 + \|u - u_H\|_{a,\Omega}^2), \end{aligned}$$

where the Young's inequality is used and $\delta_1 \in (0,1)$ satisfies

$$(1+\delta_1)\tilde{\xi}^2 < 1. \tag{38}$$

It thus follows from (17), (37), and identity $\tilde{\eta}_H(P_H w^H, \mathcal{T}_H) = \eta_H(u_H, \mathcal{T}_H)$ that there exists a positive constant C^* depending on $\hat{C}$ and $\tilde{\gamma}$ such that

$$\begin{aligned} & \|u - u_h\|_{a,\Omega}^2 + \tilde{\gamma}\eta_h^2(u_h, \mathcal{T}_h) \\ \le \ & (1+\delta_1)\tilde{\xi}^2\big(\|w^H - u_H\|_{a,\Omega}^2 + \tilde{\gamma}\eta_H^2(u_H, \mathcal{T}_H)\big) \\ & + C^*\delta_1^{-1}\tilde{\kappa}^2(h_0)(\|u - u_h\|_{a,\Omega}^2 + \|u - u_H\|_{a,\Omega}^2) \\ \le \ & (1+\delta_1)\tilde{\xi}^2\left(\big(1 + \tilde{C}\tilde{\kappa}(h_0)\big)^2\|u - u_H\|_{a,\Omega}^2 + \tilde{\gamma}\eta_H^2(u_H, \mathcal{T}_H)\right) \\ & + C^*\delta_1^{-1}\tilde{\kappa}^2(h_0)\left(\|u - u_h\|_{a,\Omega}^2 + \|u - u_H\|_{a,\Omega}^2\right). \end{aligned}$$

Hence, if $h_0 \ll 1$, then there exists a positive constant C_4 depending on C^* and $\tilde{C}$ such that

$$\begin{aligned} & \|u - u_h\|_{a,\Omega}^2 + \tilde{\gamma}\eta_h^2(u_h, \mathcal{T}_h) \\ \le \ & (1+\delta_1)\tilde{\xi}^2\left(\|u - u_H\|_{a,\Omega}^2 + \tilde{\gamma}\eta_H^2(u_H, \mathcal{T}_H)\right) \\ & + C_4\tilde{\kappa}(h_0)\|u - u_H\|_{a,\Omega}^2 + C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)\|u - u_h\|_{a,\Omega}^2. \end{aligned}$$

Consequently,

$$\begin{aligned} & \|u - u_h\|_{a,\Omega}^2 + \frac{\tilde{\gamma}}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}\eta_h^2(u_h, \mathcal{T}_h) \\ \le \ & \frac{(1+\delta_1)\tilde{\xi}^2 + C_4\tilde{\kappa}(h_0)}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}\|u - u_H\|_{a,\Omega}^2 + \frac{(1+\delta_1)\tilde{\xi}^2\tilde{\gamma}}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}\eta_H^2(u_H, \mathcal{T}_H). \end{aligned}$$

Since $h_0 \ll 1$ implies $\tilde{\kappa}(h_0) \ll 1$, we have that the constant ξ defined by

$$\xi = \left(\frac{(1+\delta_1)\tilde{\xi}^2 + C_4\tilde{\kappa}(h_0)}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}\right)^{1/2}$$

satisfies $\xi \in (0,1)$. Therefore,

$$\begin{aligned} & \|u - u_h\|_a^2 + \frac{\tilde{\gamma}}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}\eta_h^2(u_h, \mathcal{T}_h) \\ \le \ & \xi^2\left(\|u - u_H\|_{a,\Omega}^2 + \frac{(1+\delta_1)\tilde{\xi}^2\tilde{\gamma}}{(1+\delta_1)\tilde{\xi}^2 + C_4\tilde{\kappa}(h_0)}\eta_H^2(u_H, \mathcal{T}_H)\right). \end{aligned}$$

Finally, we arrive at (35) by using the fact that

$$\frac{(1+\delta_1)\tilde{\xi}^2\tilde{\gamma}}{(1+\delta_1)\tilde{\xi}^2 + C_4\tilde{\kappa}(h_0)} < \gamma.$$

This completes the proof. □

3.3. Complexity. We shall study the complexity in a class of functions defined by

$$\mathcal{A}^s_\gamma = \{v \in H^1_0(\Omega) : |v|_{s,\gamma} < \infty\},$$

where $\gamma > 0$ is some constant,

$$|v|_{s,\gamma} = \sup_{\varepsilon>0} \varepsilon \inf_{\{\mathcal{T}\subset\mathcal{T}_0 : \inf(\|v-v'\|^2_{a,\Omega}+(\gamma+1)osc^2_{\mathcal{T}}(v',\mathcal{T}))^{1/2}\le\varepsilon : v'\in S^{\mathcal{T}}_0(\Omega)\}} \big(\#\mathcal{T} - \#\mathcal{T}_0\big)^s$$

and $\mathcal{T} \subset \mathcal{T}_0$ means $\mathcal{T}$ is a refinement of $\mathcal{T}_0$ and $S^{\mathcal{T}}_0(\Omega)$ is the associated finite element space. It is seen from the definition that, for all $\gamma > 0$, $\mathcal{A}^s_\gamma = \mathcal{A}^s_1$. For simplicity, here and hereafter, we use $\mathcal{A}^s$ to stand for $\mathcal{A}^s_1$, and use $|v|_s$ to denote $|v|_{s,\gamma}$. So $\mathcal{A}^s$ is the class of functions that can be approximated within a given tolerance ε by continuous piecewise polynomial functions of degree n over a partition $\mathcal{T}$ with number of degrees of freedom $\#\mathcal{T} - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s}|v|^{1/s}_s$.

In order to study the complexity of **Algorithm D** for solving (10), we need some preparations. Recall that associated with u_H, the solution of (10) in each mesh $\mathcal{T}_H$, $w^H = K(\ell_H u_H - V u_H)$ satisfies

$$a(w^H, v) = (\ell_H u_H - V u_H, v)\ \forall v \in H^1_0(\Omega). \tag{39}$$

Using the similar procedure as in the proof of Theorem 3.3, we have

Lemma 3.2. *Let u_H and u_h be discrete solutions of (10) over a conforming mesh $\mathcal{T}_H$ and its any refinement $\mathcal{T}_h$ with marked set $\mathcal{M}_H$. Suppose that they satisfy the decrease property*

$$\begin{aligned} &\|u-u_h\|^2_{a,\Omega} + \gamma_* osc^2_h(u_h,\mathcal{T}_h) \\ \le\ & \beta^2_*\big(\|u-u_H\|^2_{a,\Omega} + \gamma_* osc^2_H(u_H,\mathcal{T}_H)\big) \end{aligned}$$

with constants $\gamma_* > 0$ and $\beta_* \in (0, \sqrt{\frac{1}{2}})$. If $h_0 \ll 1$, then the set $\mathcal{R} = R_{\mathcal{T}_H\to\mathcal{T}_h}$ satisfies the following inequality

$$\eta_H(u_H,\mathcal{R}) \ge \hat{\theta}\eta_H(u_H,\mathcal{T}_H)$$

with $\hat{\theta}^2 = \frac{\tilde{C}_2(1-2\tilde{\beta}^2_*)}{\tilde{C}_0(\tilde{C}_1+(1+2C\tilde{C}_1)\tilde{\gamma}_*)}$, $C = \Lambda_1 osc^2_0(\mathbf{A},\mathcal{T}_0)$ and $\tilde{C}_0 = \max(1, \frac{\tilde{C}_3}{\tilde{\gamma}_*})$, where $\tilde{\beta}_*$ and $\tilde{\gamma}_*$ are defined in (41) with δ_1 being chosen such that $\tilde{\beta}_* \in (0, \sqrt{\frac{1}{2}})$.

Proof. Recall that $w^h = K(\ell_h u_h - V u_h)$ and $w^H = K(\ell_H u_H - V u_H)$. Due to (20) and (22), we have

$$\|w^H - P_h w^H\|_{a,\Omega} = \|u-u_h\|_{a,\Omega} + \mathcal{O}(\tilde{\kappa}(h_0))\left(\|w^H - P_H w^H\|_{a,\Omega} + \|w^H - P_h w^H\|_{a,\Omega}\right)$$

and

$$\widetilde{osc}_h(P_h w^H,\mathcal{T}_h) = osc_h(u_h,\mathcal{T}_h) + \mathcal{O}(\tilde{\kappa}(h_0))\left(\|w^H - P_H w^H\|_{a,\Omega} + \|w^H - P_h w^H\|_{a,\Omega}\right).$$

Proceed the same procedure as in the proof of Theorem 3.3, then for problem (39), we have

$$\begin{aligned} &\|w^H - P_h w^H\|^2_{a,\Omega} + \tilde{\gamma}_*\widetilde{osc}^2_h(P_h w^H,\mathcal{T}_h) \\ \le\ & \tilde{\beta}^2_*\big(\|w^H - P_H w^H\|^2_{a,\Omega} + \tilde{\gamma}_*\widetilde{osc}^2_H(P_H w^H,\mathcal{T}_H)\big) \end{aligned} \tag{40}$$

with

$$\tilde{\beta}_* = \left(\frac{(1+\delta_1)\beta_*^2 + C_5 \tilde{\kappa}(h_0)}{1 - C_5 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \right)^{1/2}, \quad \tilde{\gamma}_* = \frac{\gamma_*}{1 - C_5 \delta_1^{-1} \tilde{\kappa}^2(h_0)}, \tag{41}$$

where C_5 is some positive constant and $\delta_1 \in (0,1)$ is some constant as shown in the proof of Theorem 3.3.

Combining $u_H = P_H w^H$ with Lemma 2.3 and (40), we get the desired result. This completes the proof. □

The key to relate the best mesh with AFEM triangulations is the fact that procedure **MARK** selects the marked set $\mathcal{M}_k$ with minimal cardinality.

Lemma 3.3. *Let $u \in \mathcal{A}^s$, $\mathcal{T}_k$ be a conforming partition obtained from $\mathcal{T}_0$ produced by* **Algorithm** *D, and $\theta \in (0, \sqrt{\frac{C_2\gamma}{C_3(C_1+(1+2CC_1)\gamma)}})$. If $h_0 \ll 1$, then the following estimate is valid:*

$$\#\mathcal{M}_k \lesssim \left(\|u - u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k, \mathcal{T}_k) \right)^{-1/2s} |u|_s^{1/s}, \tag{42}$$

where the hidden constant depends on the discrepancy between $\sqrt{\frac{C_2\gamma}{C_3(C_1+(1+2CC_1)\gamma)}}$ and θ with C defined in Lemma 2.3.

Proof. Let $\alpha, \alpha_1 \in (0,1)$ satisfy $\alpha_1 \in (0, \alpha)$ and

$$\theta^2 < \frac{C_2\gamma}{C_3(C_1 + (1+2CC_1)\gamma)}(1-\alpha^2).$$

Choose $\delta_1 \in (0,1)$ to satisfy (38) and

$$(1+\delta_1)^2 \alpha_1^2 \le \alpha^2, \tag{43}$$

which implies

$$(1+\delta_1)\alpha_1^2 < 1. \tag{44}$$

Set

$$\varepsilon = \frac{1}{\sqrt{2}} \alpha_1 \left(\|u - u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k, \mathcal{T}_k) \right)^{1/2}$$

and let $\mathcal{T}_\varepsilon$ be a refinement of $\mathcal{T}_0$ with minimal degrees of freedom satisfying

$$\|u - u_\varepsilon\|_{a,\Omega}^2 + (\gamma+1) osc_\varepsilon^2(u_\varepsilon, \mathcal{T}_\varepsilon) \le \varepsilon^2. \tag{45}$$

It follows from the definition of $\mathcal{A}^s$ that

$$\#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s} |u|_s^{1/s}.$$

Let $\mathcal{T}_* = \mathcal{T}_\varepsilon \oplus \mathcal{T}_k$ be the smallest common refinement of $\mathcal{T}_k$ and $\mathcal{T}_\varepsilon$. Note that $w^\varepsilon = K(\ell_\varepsilon u_\varepsilon - V u_\varepsilon)$ satisfies

$$L w^\varepsilon = \ell_\varepsilon u_\varepsilon - V u_\varepsilon,$$

we get from the definition of oscillation and the Young's inequality that

$$\widetilde{osc}_*^2(P_* w^\varepsilon, \tau) \le 2\widetilde{osc}_*^2(P_\varepsilon w^\varepsilon, \tau) + 2C_0^2 osc_*^2(\mathbf{A}, \tau) \|P_\varepsilon w^\varepsilon - P_* w^\varepsilon\|_{1,\omega_\tau}^2 \ \forall \tau \in \mathcal{T}_*,$$

which together with the monotonicity property $osc_*(\mathbf{A}, \mathcal{T}_*) \le osc_0(\mathbf{A}, \mathcal{T}_0)$ yields

$$\widetilde{osc}_*^2(P_* w^\varepsilon, \mathcal{T}_*) \le 2\widetilde{osc}_*^2(P_\varepsilon w^\varepsilon, \mathcal{T}_*) + 2C\|P_\varepsilon w^\varepsilon - P_* w^\varepsilon\|_{a,\Omega}^2,$$

where $C = \Lambda_1 osc_0^2(\mathbf{A}, \mathcal{T}_0)$. Due to the orthogonality

$$\|w^\varepsilon - P_* w^\varepsilon\|_{a,\Omega}^2 = \|w^\varepsilon - P_\varepsilon w^\varepsilon\|_{a,\Omega}^2 - \|P_* w^\varepsilon - P_\varepsilon w^\varepsilon\|_{a,\Omega}^2,$$

we arrive at

$$\begin{aligned} & \|w^\varepsilon - P_* w^\varepsilon\|_{a,\Omega}^2 + \frac{1}{2C}\widetilde{osc}_*^2(P_* w^\varepsilon, \mathcal{T}_*) \\ \le\ & \|w^\varepsilon - P_\varepsilon w^\varepsilon\|_{a,\Omega}^2 + \frac{1}{C} osc_\varepsilon^2(P_\varepsilon w^\varepsilon, \mathcal{T}_\varepsilon). \end{aligned}$$

Since (8) implies $\tilde{\gamma} \le \frac{1}{2C}$, we obtain that

$$\begin{aligned} & \|w^\varepsilon - P_* w^\varepsilon\|_{a,\Omega}^2 + \tilde{\gamma}\widetilde{osc}_*^2(P_* w^\varepsilon, \mathcal{T}_*) \\ \le\ & \|w^\varepsilon - P_\varepsilon w^\varepsilon\|_{a,\Omega}^2 + \frac{1}{C} osc_\varepsilon^2(P_\varepsilon w^\varepsilon, \mathcal{T}_\varepsilon) \\ \le\ & \|w^\varepsilon - P_\varepsilon w^\varepsilon\|_{a,\Omega}^2 + (\tilde{\gamma} + \sigma) osc_\varepsilon^2(P_\varepsilon w^\varepsilon, \mathcal{T}_\varepsilon) \end{aligned}$$

with $\sigma = \frac{1}{C} - \tilde{\gamma} \in (0,1)$. Applying the similar argument in the proof of Theorem 3.3 when (21) is replaced by (22), we then get that

$$\begin{aligned} & \|u - u_*\|_{a,\Omega}^2 + \gamma osc_*^2(u_*, \mathcal{T}_*) \\ \le\ & \alpha_0^2 \left(\|u - u_\varepsilon\|_{a,\Omega}^2 + (\gamma + \sigma) osc_\varepsilon^2(P_\varepsilon w^\varepsilon, \mathcal{T}_\varepsilon)\right) \\ \le\ & \alpha_0^2 \left(\|u - u_\varepsilon\|_{a,\Omega}^2 + (\gamma + 1) osc_\varepsilon^2(P_\varepsilon w^\varepsilon, \mathcal{T}_\varepsilon)\right), \end{aligned} \tag{46}$$

where

$$\alpha_0^2 = \frac{(1+\delta_1) + C_4\tilde{\kappa}(h_0)}{1 - C_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}$$

and C_4 is the constant appearing in the proof of Theorem 3.3. Thus, by (45) and (46), it follows

$$\|u - u_*\|_{a,\Omega}^2 + \gamma osc_*^2(u_*, \mathcal{T}_*) \le \breve{\alpha}^2\big(\|u - u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k, \mathcal{T}_k)\big)$$

with $\breve{\alpha} = \frac{1}{\sqrt{2}}\alpha_0\alpha_1$. In view of (44), we have $\breve{\alpha}^2 \in (0, \frac{1}{2})$ when $h_0 \ll 1$. Let $\mathcal{R} = R_{\mathcal{T}_k \to \mathcal{T}_*}$, by Lemma 3.2, we have that $\mathcal{T}_*$ satisfies

$$\eta_k(u_k, \mathcal{R}) \ge \breve{\theta}\eta_k(u_k, \mathcal{T}_k),$$

where $\breve{\theta}^2 = \frac{\tilde{C}_2(1-2\hat{\alpha}^2)}{\tilde{C}_0(\tilde{C}_1 + (1+2C\tilde{C}_1)\hat{\gamma})}$, $\hat{\gamma} = \frac{\gamma}{1 - C_5\delta_1^{-1}\tilde{\kappa}^2(h_0)}$, $\tilde{C}_0 = \max(1, \frac{\tilde{C}_3}{\hat{\gamma}})$, and

$$\hat{\alpha}^2 = \frac{(1+\delta_1)\breve{\alpha}^2 + C_5\tilde{\kappa}(h_0)}{1 - C_5\delta_1^{-1}\tilde{\kappa}^2(h_0)}.$$

It follows from the definition of γ (see (36)) and $\tilde{\gamma}$ (see (8)) that $\hat{\gamma} < 1$ and hence $\tilde{C}_0 = \frac{\tilde{C}_3}{\hat{\gamma}}$. Since $h_0 \ll 1$, we obtain that $\hat{\gamma} > \gamma$ and $\hat{\alpha} \in (0, \frac{1}{\sqrt{2}}\alpha)$ from (43). It is easy to see from (34) and $\hat{\gamma} > \gamma$ that

$$\begin{aligned} \breve{\theta}^2 &= \frac{\tilde{C}_2(1-2\hat{\alpha}^2)}{\frac{\tilde{C}_3}{\hat{\gamma}}(\tilde{C}_1 + (1+2C\tilde{C}_1)\hat{\gamma})} \ge \frac{\tilde{C}_2}{\tilde{C}_3(\frac{\tilde{C}_1}{\hat{\gamma}} + 1 + 2C\tilde{C}_1)}(1-\alpha^2) \\ &= \frac{\frac{C_2}{(1-\tilde{C}\tilde{\kappa}(h_0))^2}}{\frac{C_3}{(1-\tilde{C}\tilde{\kappa}(h_0))^2}\left(\frac{C_1}{\hat{\gamma}((1+\tilde{C}\tilde{\kappa}(h_0))^2)} + 1 + 2C\frac{C_1}{(1+\tilde{C}\tilde{\kappa}(h_0))^2}\right)}(1-\alpha^2) \\ &\ge \frac{C_2}{C_3(\frac{C_1}{\gamma} + (1+2CC_1))}(1-\alpha^2) = \frac{C_2\gamma}{C_3(C_1 + (1+2CC_1)\gamma)}(1-\alpha^2) > \theta^2 \end{aligned}$$

when $h_0 \ll 1$. Thus

$$\begin{aligned} \#\mathcal{M}_k &\le \#\mathcal{R} \le \#\mathcal{T}_* - \#\mathcal{T}_k \le \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \\ &\le \left(\frac{1}{\sqrt{2}}\alpha_1\right)^{-1/s}\left(\|u - u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k, \mathcal{T}_k)\right)^{-1/2s}|u|_s^{1/s}, \end{aligned}$$

which is the desired estimate (42) with an explicit dependence on the discrepancy between θ and $\sqrt{\frac{C_2\gamma}{C_3(C_1+(1+2CC_1)\gamma)}}$ via α_1. This completes the proof. $\square$

As a consequence, we obtain the optimal complexity as follows.

Theorem 3.4. *Let $u \in \mathcal{A}^s$ and $\{u_k\}_{k\in\mathbb{N}_0}$ be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces $\{S_0^k(\Omega)\}_{k\in\mathbb{N}_0}$ produced by* **Algorithm** *D. If $h_0 \ll 1$, then*

$$\|u-u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k,\mathcal{T}_k) \lesssim (\#\mathcal{T}_k - \#\mathcal{T}_0)^{-2s}|u|_s^2,$$

where the hidden constant depends on the discrepancy between $\sqrt{\frac{C_2\gamma}{C_3(C_1+(1+2CC_1)\gamma)}}$ and θ.

Proof. It follows from (7) and (42) that

$$\begin{aligned}\#\mathcal{T}_k - \#\mathcal{T}_0 &\lesssim \sum_{j=0}^{k-1}\#\mathcal{M}_j \\ &\lesssim \sum_{j=0}^{k-1}\left(\|u-u_j\|_{a,\Omega}^2 + \gamma osc_j^2(u_j,\mathcal{T}_j)\right)^{-1/2s}|u|_s^{1/s}.\end{aligned}$$

Note that (31) implies

$$\|u-u_j\|_{a,\Omega}^2 + \gamma\eta_j^2(u_j,\mathcal{T}_j) \le \check{C}\left(\|u-u_j\|_{a,\Omega}^2 + \gamma osc_j^2(u_j,\mathcal{T}_j)\right),$$

where $\check{C} = \max(1+\frac{\gamma}{C_2}, \frac{C_3}{C_2})$. It then turns out

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \lesssim \sum_{j=0}^{k-1}\left(\|u-u_j\|_{a,\Omega}^2 + \gamma\eta_j^2(u_j,\mathcal{T}_j)\right)^{-1/2s}|u|_s^{1/s}.$$

Due to (35), we obtain for $0 \le j < k$ that

$$\|u-u_k\|_{a,\Omega}^2 + \gamma\eta_k^2(u_k,\mathcal{T}_k) \le \xi^{2(k-j)}\left(\|u-u_j\|_{a,\Omega}^2 + \gamma\eta_j^2(u_j,\mathcal{T}_j)\right).$$

Consequently,

$$\begin{aligned}\#\mathcal{T}_k - \#\mathcal{T}_0 &\lesssim |u|_s^{1/s}\left(\|u-u_k\|_{a,\Omega}^2 + \gamma\eta_k^2(u_k,\mathcal{T}_k)\right)^{-1/2s}\sum_{j=0}^{k-1}\xi^{\frac{k-j}{s}} \\ &\lesssim |u|_s^{1/s}\left(\|u-u_k\|_{a,\Omega}^2 + \gamma\eta_k^2(u_k,\mathcal{T}_k)\right)^{-1/2s},\end{aligned}$$

the last inequality holds because of the fact $\xi < 1$.

Since $osc_k(u_k,\mathcal{T}_k) \le \eta_k(u_k,\mathcal{T}_k)$, we arrive at

$$\#\mathcal{T}_k - \#\mathcal{T}_0 \lesssim \left(\|u-u_k\|_{a,\Omega}^2 + \gamma osc_k^2(u_k,\mathcal{T}_k)\right)^{-1/2s}|u|_s^{1/s}.$$

This completes the proof. $\square$

4. Applications

In this section, we provide three typical examples to show that our general theory is quite useful.

4.1. A nonsymmetric problem. The first example is a nonsymmetric elliptic partial differential equation of second order. We consider the following problem: find $u \in H_0^1(\Omega)$ such that

$$\left\{ \begin{array}{rl} -\nabla \cdot (\mathbf{A}\nabla u) + \mathbf{b} \cdot \nabla u + cu & = f \ \text{ in } \ \Omega, \\ u & = 0 \ \text{ on } \partial\Omega, \end{array} \right. \tag{47}$$

where $\Omega \subset \mathbb{R}^d (d \geq 2)$ is a ploytopic domain, $\mathbf{A} : \Omega \to \mathbb{R}^{d\times d}$ is piecewise Lipschitz over initial triangulation $\mathcal{T}_0$ and symmetric positive definite with smallest eigenvalue uniformly bounded away from 0, $\mathbf{b} \in [L^\infty(\Omega)]^d$ is divergence free, $c \in L^\infty(\Omega)$, and $f \in L^2(\Omega)$.

The weak form of (47) is as follows: find $u \in H_0^1(\Omega)$ such that

$$(\mathbf{A}\nabla u, \nabla v) + (\mathbf{b} \cdot \nabla u, v) + (cu, v) = (f, v) \quad \forall v \in H_0^1(\Omega). \tag{48}$$

We assume that (48) is well-posed, namely (48) is uniquely solvable for any $f \in H^{-1}(\Omega)$. (A simple sufficient condition for this assumption to be satisfied is that $c \geq 0$.)

A finite element discretization of (48) reads: find $u_h \in S_0^h(\Omega)$ such that

$$(\mathbf{A}\nabla u_h, \nabla v) + (\mathbf{b} \cdot \nabla u_h, v) + (cu_h, v) = (f, v) \ \forall v \in S_0^h(\Omega). \tag{49}$$

It is seen that (49) has a unique solution u_h if $h \ll 1$ (see, e.g., [31]) and (49) is a special case of (10), in which $Vw = \mathbf{b} \cdot \nabla w + cw$ and $\ell w = \ell_h w = f \ \forall w \in H_0^1(\Omega)$. Consequently, $\kappa_1(h) = 0$ and $w^h = K(f - Vu_h)$.

Obviously, $V : H_0^1(\Omega) \to L^2(\Omega)$ is a linear bounded operator and KV is a compact operator over $H_0^1(\Omega)$.

Set

$$\kappa(h) = \|(I + KVP_h)^{-1}\| \|KV(I - P_h)\|,$$

we have the conclusion of Theorem 3.1.

In this application, the element residual and jump residual become

$$\begin{array}{rcl} \mathcal{R}_\tau(u_h) & = & f - \mathbf{b} \cdot \nabla u_h - cu_h + \nabla \cdot (\mathbf{A}\nabla u_h) \text{ in } \tau \in \mathcal{T}_h, \\ J_e(u_h) & = & [[\mathbf{A}\nabla u_h]]_e \cdot \nu_e \text{ on } e \in \mathcal{E}_h \end{array}$$

while the corresponding error estimator $\eta_h(u_h, \mathcal{T}_h)$ and the oscillation $osc_h(u_h, \mathcal{T}_h)$ are defined by (18) and (19), respectively. Thus Theorem 3.3 and Theorem 3.4 ensure the convergence and optimal complexity of AFEM for the nonsymmetric problem (47).

4.2. A nonlinear problem. In this subsection, we derive the convergence and optimal complexity of AFEM for a nonlinear problem from our general theory.

Consider the following nonlinear problem: find $u \in H_0^1(\Omega)$ such that

$$\left\{ \begin{array}{rl} \mathcal{L}u \equiv -\Delta u + f(x, u) & = 0 \ \text{ in } \ \Omega, \\ u & = 0 \ \text{ on } \partial\Omega, \end{array} \right. \tag{50}$$

where $\Omega \subset \mathbb{R}^d (d = 1, 2, 3)$ is a polytopic domain and $f(x, y)$ is a smooth function on $\mathbb{R}^d \times \mathbb{R}^1$.

For convenience, we shall drop the dependence of variable x in $f(x, u)$ in the following exposition. We assume that (50) has a solution $u \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$ for some $s \in (1/2, 1]$. Setting

$$b(w, v) = (\nabla w, \nabla v) + (f(w), v),$$

then

$$b(u, v) = 0 \quad \forall v \in H_0^1(\Omega).$$

For any $w \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$, the linearized operator $\mathcal{L}'_w$ at w (namely, the Fréchet derivative of $\mathcal{L}$ at w) is then given by

$$\mathcal{L}'_w = -\Delta + f'(w).$$

We assume that $\mathcal{L}'_u : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism in the neighborhood of u.

As a result, $u \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$ must be an isolated solution of (50).

A finite element discretization of (50) reads: find $u_h \in S_0^h(\Omega)$ such that

$$b(u_h, v) = 0 \quad \forall v \in S_0^h(\Omega). \tag{51}$$

It is seen that (51) has a unique solution u_h in the neighbour of u if $h \ll 1$ (see, e.g., [30]). Let $a(\cdot,\cdot) = (\nabla\cdot, \nabla\cdot)$, $K = (-\Delta)^{-1} : L^2(\Omega) \to H_0^1(\Omega)$, $V = 0$ and $\ell_h w = -f(w)$ for any $w \in S_0^h(\Omega)$, then (51) becomes (10).

Let $P'_h : H_0^1(\Omega) \to S_0^h(\Omega)$ be defined by

$$b'(u; w - P'_h w, v) = 0 \quad \forall v \in S_0^h(\Omega), \tag{52}$$

where $b'(u; \phi, v) \equiv (\mathcal{L}'_u \phi, v) = (\nabla\phi, \nabla v) + (f'(u)\phi, v)$. It is seen that as operators over $H_0^1(\Omega)$

$$\lim_{h\to 0} \|K(I - P'_h)\| = 0.$$

Moreover, using Aubin-Nitsche duality argument we have

$$\|u - P'_h u\|_{0,\Omega} \lesssim \tilde{r}(h)\|u - P'_h u\|_{a,\Omega}, \tag{53}$$

where $\tilde{r}(h) \to 0$ as $h \to 0$.

Lemma 4.1. *Assume that* $\|u_h\|_{0,\infty,\Omega} \lesssim 1$ *and* $\|u - u_h\|_{a,\Omega} \to 0$ *as* $h \to 0$. *If* $h \ll 1$, *then*

$$\|P'_h u - u_h\|_{a,\Omega} \lesssim \|u - u_h\|_{0,3/2,\Omega}\|u - u_h\|_{a,\Omega}, \tag{54}$$

$$\|u - u_h\|_{0,\Omega} = \mathcal{O}(r(h))\|u - u_h\|_{a,\Omega}, \tag{55}$$

where $r(h) \to 0$ *as* $h \to 0$.

Proof. For any $w, \chi, v \in H_0^1(\Omega)$, set $\eta(t) = b(w + t(\chi - w), v)$. From identity

$$\eta(1) = \eta(0) + \eta'(0) + \int_0^1 \eta''(t)(1-t)dt,$$

we obtain that

$$b(\chi, v) = b(w, v) + b'(w; \chi - w, v) + R(w, \chi, v),$$

where

$$R(w, \chi, v) = \int_0^1 \eta''(t)(1-t)dt.$$

Thus $u_h \in S_0^h(\Omega)$ solves (51) if and only if

$$b'(u; u - u_h, v) = R(u, u_h, v) \ \ \forall v \in S_0^h(\Omega). \tag{56}$$

A straightforward calculation shows that

$$\eta''(t) = (f''(w + t(\chi - w))(\chi - w)^2, v).$$

Since $f(x, y)$ is smooth, there exists a constant $C_\zeta > 0$ such that

$$|R(w, \chi, v)| \le C_\zeta \|w - \chi\|_{0,3/2,\Omega}\|w - \chi\|_{1,\Omega}\|v\|_{1,\Omega} \tag{57}$$

when $\max(\|w\|_{0,\infty,\Omega}, \|\chi\|_{0,\infty,\Omega}) \le \zeta$.

Combining (52) with (56), we have

$$b'(u; P_h'u - u_h, v) = R(u, u_h, v).$$

Note that Sobolev imbedding theorem implies $u \in L^\infty(\Omega)$, we then obtain from $\|u_h\|_{0,\infty,\Omega} \lesssim 1$ and (57) that

$$\|u_h - P_h'u\|_{a,\Omega} \lesssim \|u - u_h\|_{0,3/2,\Omega} \|u - u_h\|_{a,\Omega},$$

which is nothing but (54).

Obviously, (54) implies

$$\|u_h - P_h'u\|_{0,\Omega} \lesssim \|u - u_h\|_{0,3/2,\Omega} \|u - u_h\|_{a,\Omega}. \tag{58}$$

Since Sobolev imbedding theorem leads to $\|u - u_h\|_{0,3/2,\Omega} \to 0$ as $h \to 0$, we get

$$\|u - P_h'u\|_{a,\Omega} \lesssim \|u - u_h\|_{a,\Omega}$$

if $h \ll 1$. Due to (53), we have

$$\|u - P_h'u\|_{0,\Omega} \lesssim \tilde{r}(h) \|u - u_h\|_{a,\Omega}. \tag{59}$$

We then arrive at (55) from (58), (59), the triangle inequality

$$\|u - u_h\|_{0,\Omega} \le \|u - P_h'u\|_{0,\Omega} + \|P_h'u - u_h\|_{0,\Omega},$$

and setting $r(h) = \tilde{r}(h) + \|u - u_h\|_{0,3/2,\Omega}$. This completes the proof. □

Note that $\|u - u_h\|_{a,\Omega} \to 0$ as $h \to 0$ implies that $\mathcal{L}'_{u_h} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism and (51) has a locally unique solution u_h in an neighborhood of u when $h \ll 1$. Now we shall show that Theorem 3.1 is applicable for (50). Let $\ell w = -f(w)\ \forall w \in H_0^1(\Omega)$. Since K is monotone and $f(x, y)$ is smooth, we have from Lemma 4.1 that

$$\begin{aligned} & \|K(f(u) - f(u_h))\|_{a,\Omega} \lesssim \|K(|u - u_h|)\|_{a,\Omega} \\ \lesssim \quad & \|u - u_h\|_{0,\Omega} \lesssim r(h) \|u - u_h\|_{a,\Omega}. \end{aligned}$$

Therefore we have (13) when we choose $\kappa(h) = r(h)$.

In this application, the element residual and jump residual become:

$$\begin{aligned} \mathcal{R}_\tau(u_h) &= -f(u_h) + \Delta u_h \qquad \text{in } \tau \in \mathcal{T}_h, \\ J_e(u_h) &= -\nabla u_h^+ \cdot \nu^+ - \nabla u_h^- \cdot \nu^- = [[\nabla u_h]]_e \cdot \nu_e \text{ on } e \in \mathcal{E}_h \end{aligned}$$

and the corresponding error estimator $\eta_h(u_h, \mathcal{T}_h)$ and the oscillation $osc_h(u_h, \mathcal{T}_h)$ are defined by (18) and (19), respectively. Note that the requirements $\|u_h\|_{0,\infty,\Omega} \lesssim 1$ and $\|u - u_h\|_{a,\Omega} \to 0$ as $h \to 0$ may be satisfied in adaptive finite element approximations.[2] Thus Theorem 3.3 and Theorem 3.4 may ensure the convergence and optimal complexity of AFEM for nonlinear problem (50).

[2] To meet the requirements is relatively easy for one and two dimensional cases. We refer to [7, 22] for the relevant discussions in the three dimensional setting.

4.3. An unbounded coefficient problem. Finally, we investigate a nonlinear eigenvalue problem, of which a coefficient is unbounded. It is known that electronic structure computations require solving the following Kohn-Sham equations [4, 14, 17]

$$\left(-\frac{1}{2}\Delta - \sum_{j=1}^{N_{atom}} \frac{Z_j}{|x-r_j|} + \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|}dy + V_{xc}(\rho)\right) u_i = \lambda_i u_i \ in \quad \mathbb{R}^3,$$

where N_{atom} is the total number of atoms in the system, Z_j is the valance charge of this ion (nucleus plus core electrons), r_j is the position of the j-th atom ($j = 1, \cdots, N_{atom}$),

$$\rho = \sum_{i=1}^{N_{occ}} c_i |u_i|^2$$

with u_i the i-th smallest eigenfunction, c_i the number of electrons on the i-th orbit, and N_{occ} the total number of the occupied orbits. The central computation in solving the Kohn-Sham equation is the repeated solution of the following eigenvalue problem: find $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$\left\{ \begin{array}{ll} -\frac{1}{2}\Delta u + Vu & = \lambda u \quad \text{in } \Omega, \\ \|u\|_{0,\Omega} & = 1, \end{array} \right. \tag{60}$$

where Ω is a polytopic domain in $\mathbb{R}^3$ and $V = V_{ne} + V_0$ is the so-called effective potential. Here, $V_0 \in L^\infty(\Omega)$ and

$$V_{ne}(x) = -\sum_{j=1}^{N_{atom}} \frac{Z_j}{|x-r_j|}.$$

The weak form of (60) is: find $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that $\|u\|_{0,\Omega} = 1$ and

$$\frac{1}{2}(\nabla u, \nabla v) + (Vu, v) = \lambda(u, v) \ \forall v \in H_0^1(\Omega). \tag{61}$$

Note that (61) has a countable sequence of real eigenvalues $\lambda_1 < \lambda_2 \le \lambda_3 \le \cdots$, and the corresponding eigenfunctions in $H_0^1(\Omega)$, $u_1, u_2, u_3, \cdots$, which can be assumed to satisfy $(u_i, u_j) = \delta_{ij}, i, j = 1, 2, \cdots$ (see, e.g., [14]).

A finite element discretization of (60) reads: find $(\lambda_h, u_h) \in \mathbb{R} \times S_0^h(\Omega)$ such that $\|u_h\|_{0,\Omega} = 1$ and

$$\frac{1}{2}(\nabla u_h, \nabla v) + (Vu_h, v) = \lambda_h(u_h, v) \quad \forall v \in S_0^h(\Omega). \tag{62}$$

Let $\ell_h : S_0^h(\Omega) \to L^2(\Omega)$ be defined by

$$\ell_h w = \lambda_h w \quad \forall w \in S_0^h(\Omega),$$

then (62) is a special case of (10) when $a(\cdot,\cdot) = \frac{1}{2}(\nabla\cdot, \nabla\cdot)$ and $K = (-\frac{1}{2}\Delta)^{-1} : L^2(\Omega) \to H_0^1(\Omega)$.

Using the uncertainty principle lemma (see, e.g., [27])

$$\int_{\mathbb{R}^3} \frac{w^2(x)}{|x|^2} \le 4 \int_{\mathbb{R}^3} |\nabla w|^2 \ \forall w \in C_0^\infty(\mathbb{R}^3)$$

and the fact that $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, we obtain

$$\int_\Omega \frac{w^2(x)}{|x|^2} \le 4 \int_\Omega |\nabla w|^2 \ \forall w \in H_0^1(\Omega).$$

Then for any $w \in H_0^1(\Omega)$, we have

$$\|V_{ne}w + V_0 w\|_{0,\Omega} \leq C\|w\|_{1,\Omega},$$

namely, V is a bounded operator from $H_0^1(\Omega)$ to $L^2(\Omega)$. Thus KV is a compact operator over $H_0^1(\Omega)$.

We consider the case of that $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ is some simple eigenpair of (60) with $\|u\|_{0,\Omega} = 1$. We see that (62) has an associated finite element eigenpair $(\lambda_h, u_h) \in \mathbb{R} \times S_0^h(\Omega)$ that satisfies $\|u_h\|_{0,\Omega} = 1$ and (c.f. [2])

$$\|u - u_h\|_{0,\Omega} \lesssim \rho_\Omega(h)\|u - u_h\|_{a,\Omega} \tag{63}$$

and

$$|\lambda - \lambda_h| \lesssim \|u - u_h\|_{a,\Omega}^2, \tag{64}$$

where

$$\rho_\Omega(h) = \sup_{f \in L^2(\Omega), \|f\|_{0,\Omega}=1} \inf_{v \in S_0^h(\Omega)} \|(-\frac{1}{2}\Delta + V)^{-1} f - v\|_{a,\Omega}.$$

Note that for $\ell w = \lambda w \; \forall w \in H_0^1(\Omega)$, there holds

$$K(\ell u - \ell_h u_h) = \lambda K(u - u_h) + (\lambda - \lambda_h)K u_h, \tag{65}$$

which together with (64) yields

$$\|K(\ell u - \ell_h u_h)\|_{a,\Omega} = O(\kappa_1(h))\|u - u_h\|_{a,\Omega},$$

where $\kappa_1(h) = \rho_\Omega(h) + \|u - u_h\|_{a,\Omega}$ satisfying $\kappa_1(h) \to 0$ as $h \to 0$.

In this application, the element residual and jump residual become:

$$\begin{aligned} \mathcal{R}_\tau(u_h) &= \lambda_h u_h - V u_h + \frac{1}{2}\Delta u_h \quad \text{in } \tau \in \mathcal{T}_h, \\ J_e(u_h) &= [[\frac{1}{2}\nabla u_h]]_e \cdot \nu_e \quad \text{on } e \in \mathcal{E}_h \end{aligned}$$

and the corresponding error estimator $\eta_h(u_h, \mathcal{T}_h)$ and the oscillation $osc_h(u_h, \mathcal{T}_h)$ are defined by (18) and (19), respectively. Then Theorem 3.3 and Theorem 3.4 ensure the convergence and optimal complexity of adaptive finite element eigenfunction approximations for the unbounded coefficient problem (60) (c.f. [9]). Besides, we can also get the convergence and optimal complexity of adaptive finite element eigenvalue approximations from (64).

5. Numerical examples

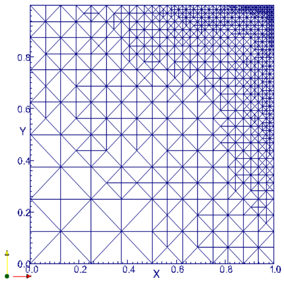
In this section, we will report some numerical results to illustrate our theory. Our numerical results were carried out on LSSC-II in the State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences, and our codes were based on the toolbox PHG of the Laboratory.

Example 1. We consider (47) when the homogenous Dirichlet boundary condition is replaced by $u = g$ on $\partial\Omega$ and $\Omega = (0,1)^3$ with the isotropic diffusion coefficient $\mathbf{A} = \epsilon I$, $\epsilon = 10^{-2}$, convection velocity $\mathbf{b} = (2,3,4)$, and c = 0 (c.f. [16] for a 2D case and Remark 2.1). The exact solution is given by

$$u = \left(x^3 - \exp\left(\frac{2(x-1)}{\epsilon}\right)\right)\left(y^2 - \exp\left(\frac{3(y-1)}{\epsilon}\right)\right)\left(z - \exp\left(\frac{4(z-1)}{\epsilon}\right)\right).$$

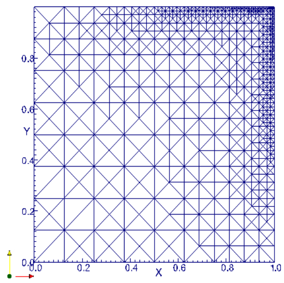
For small $\epsilon > 0$, the solution has the typical layer behavior in the neighbourhood of $x = 1$, $y = 1$, $z = 1$, respectively. The Dirichlet boundary condition is given by

$$g(x,y,z) = \begin{cases} 0 & x = 1 \; or \; y = 1 \; or \; z = 1, \\ u(x,y,z) & x = 0 \; or \; y = 0 \; or \; z = 0. \end{cases}$$



Z=0.0

FIGURE 1. The cross-section of an adaptive mesh of **Example 1** using linear finite elements



Z=0.0

FIGURE 2. The cross-section of an adaptive mesh of **Example 1** using quadratic finite elements

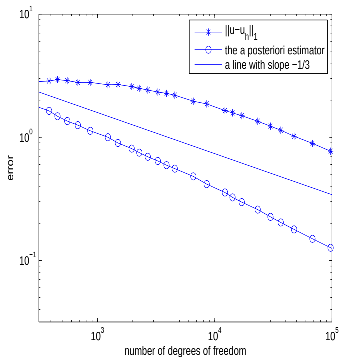


FIGURE 3. The convergence curves of **Example 1** using linear finite elements

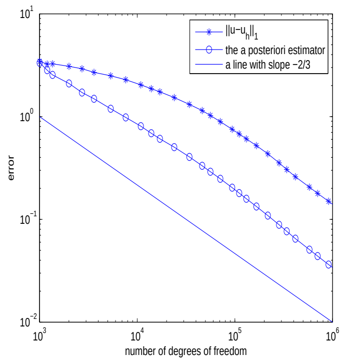


FIGURE 4. The convergence curves of **Example 1** using quadratic finite elements

Some adaptively refined meshes are displayed in Fig. 1 and Fig. 2. Our numerical results are presented in Fig. 3 and Fig. 4. It is shown from Fig. 4 that $\|u - u_h\|_1$ is proportional to the a posteriori error estimators, which indicates the efficiency and reliability of the a posteriori error estimators given in section 4.1. Besides, it is also seen from Fig. 3 and Fig. 4 that, by using linear finite elements and quadratic finite elements, the convergence curves of errors are approximately parallel to the line with slope $-1/3$ and the line with slope $-2/3$, respectively. These mean that the approximation error of the exact solution has optimal convergence rate, which coincides with our theory in section 3.2.

Example 2. Consider the following nonlinear problem:

$$\begin{cases} -\Delta u + u^3 &= f \text{ in } \quad \Omega, \\ \qquad\quad u &= 0 \text{ on } \partial\Omega, \end{cases}$$

where $\Omega = (0,1)^3$. The exact solution is given by

$$u = \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)/(x_1^2 + x_2^2 + x_3^2)^{1/2}.$$

Since $-\Delta + 2u^2$ is nonsigular, the conditions required in section 4.2 are fulfilled.

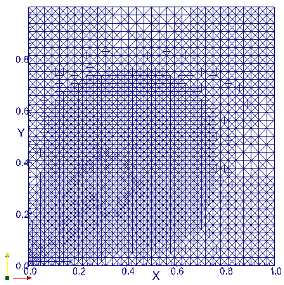


FIGURE 5. The cross-section of an adaptive mesh of **Example 2** using linear finite elements

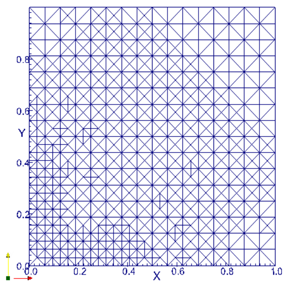


FIGURE 6. The cross-section of an adaptive mesh of **Example 2** using quadratic finite elements

Fig. 5 and Fig. 6 are two adaptively refined meshes, which show that the error indicator is good. It is shown from Fig. 7 and Fig. 8 that $\|u - u_h\|_1$ is proportional to the a posteriori error estimators, which implies that the a posteriori error estimators given in section 4.2 are efficient. Besides, similar conclusions to that of Example 1 can be obtained from Fig. 7 and Fig. 8, too.

Example 3. Consider the Kohn-Sham equation for helium atoms:

$$\left(-\frac{1}{2}\Delta - \frac{2}{|x|} + \int \frac{\rho(y)}{|x-y|}dy + V_{xc}\right) u = \lambda u \text{ in } \mathbb{R}^3$$

with $\int_{\mathbb{R}^3} |u|^2 = 1$, where $\rho = 2|u|^2$. In our computation of the ground state energy, we solve the following nonlinear eigenvalue problem: find $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that $\int_\Omega |u|^2 dx = 1$ and

$$\begin{cases} \left(-\frac{1}{2}\Delta - \frac{2}{|x|} + \int \frac{\rho(y)}{|x-y|}dy + V_{xc}\right) u &= \lambda u \text{ in } \Omega, \\ \qquad\qquad\qquad\qquad\qquad\qquad u &= 0 \text{ on } \partial\Omega, \end{cases} \tag{66}$$

where $\Omega = (-10.0, 10.0)^3$, and $V_{xc}(\rho) = -\frac{3}{2}\alpha(\frac{3}{\pi}\rho)^{\frac{1}{3}}$ with $\alpha = 0.77298$. Since (66) is a nonlinear eigenvalue problem, we need to linearize and solve them iteratively, which is called the self-consistent approach [4, 14, 17, 23]. In our computation, a Broyden-type quasi-Newton method [24] is used.

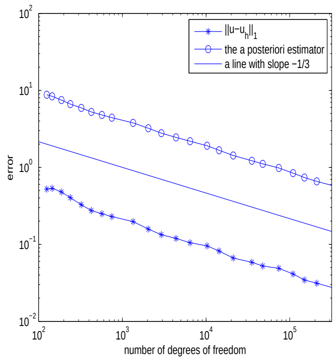


FIGURE 7. The convergence curves of **Example 2** using linear finite elements

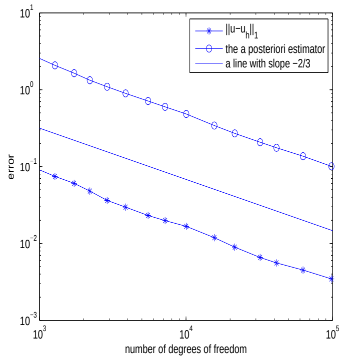


FIGURE 8. The convergence curves of **Example 2** using quadratic finite elements

In 1989, White [29] computed helium atoms over uniform cubic grids and obtained ground state energy -2.8522 a.u. by using 500,000 finite element bases. While the ground state energy of helium atoms in Software package fhi98PP [11] is -2.8346 a.u., which we take as a reference.

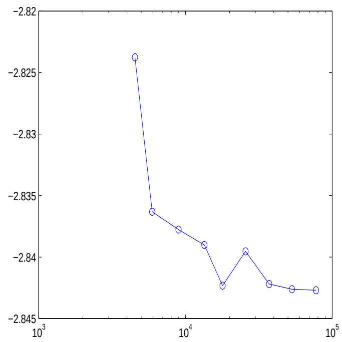


FIGURE 9. The ground state energy using linear finite elements

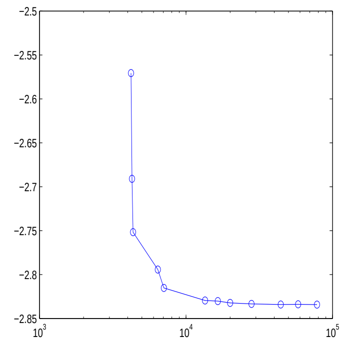
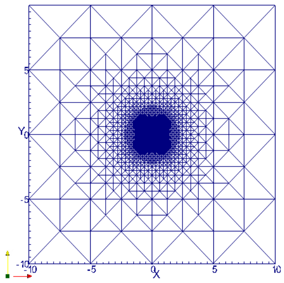


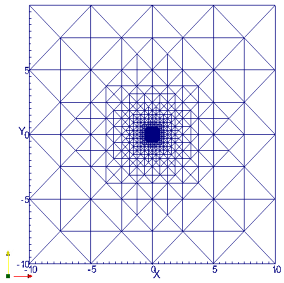
FIGURE 10. The ground state energy using quadratic finite elements

Our results are displayed in Fig. 9– Fig. 14. It is seen from Fig. 10 that the ground state energy in our computation is close to the reference with less 100,000 degrees of freedom when the quadratic finite element discretization is used. Some cross-sections of the adaptively refined meshes are displayed in Fig. 11 and Fig. 12. Since we do not have the exact solution, we list the convergence curves of the a posteriori error estimators in Fig. 13 and Fig. 14 only. It is shown from these



Z=0.0

FIGURE 11. The cross-section of an adaptive mesh of **Example 3** using linear finite elements



Z=0.0

FIGURE 12. The cross-section of an adaptive mesh of **Example 3** using quadratic finite elements

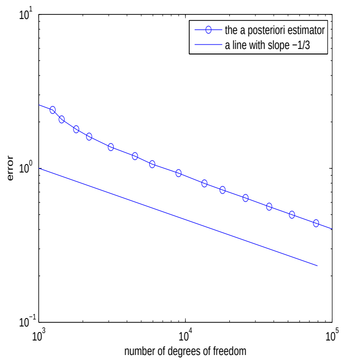


FIGURE 13. The convergence curve of **Example 3** using linear finite elements

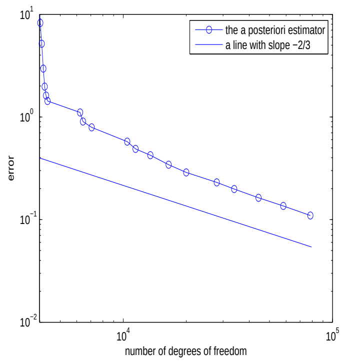


FIGURE 14. The convergence curve of **Example 3** using quadratic finite elements

figures that the a posteriori error estimators given in section 4.3 are convergent as predicted by our theory.

Acknowledgments

The authors would like to thank Dr. Huajie Chen, Dr. Xiaoying Dai, Prof. Xingao Gong, Prof. Lihua Shen, and Prof. Jinchao Xu for their stimulating discussions and fruitful cooperations that have motivated this work.

References

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] I. Babuska and J. E. Osborn, *Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems*, Math. Comput., **52** (1989), pp. 275-297.

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

[4] S. L. Beck, *Real-space mesh techniques in density-function theory*, Rev. Mod. Phys., **72** (2000), pp. 1041-1080.

[5] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., **97** (2004), pp. 219-268.

[6] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., **46** (2008), pp. 2524-2550.

[7] L. Chen, M. J. Holst, and J. Xu, *The finite element approximation of the nonlinear Poisson-Boltzmann equation*, SIAM J. Numer. Anal., **45** (2007), pp. 2298-2320.

[8] P. G. Ciarlet and J. L. Lions, eds., *Finite Element Methods, Volume II of Handbook of Numerical Analysis*, North–Holland, Amsterdam, 1991.

[9] X. Dai, J. Xu, and A. Zhou, *Convergence and optimal complexity of adaptive finite element eigenvalue computations*, Numer. Math., **110** (2008), pp. 313-355.

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

[11] M. Fuchs and M. Scheffler, *Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory*, Comput. Phys. Commun., **119** (1999), pp. 67-98.

[12] E. M. Garau, P. Morin, and C. Zuppa, *Convergence of adaptive finite element methods for eigenvalue problems*, M^3AS, **19** (2009), pp. 721-747.

[13] S. Giani and I. G. Graham, *A convergent adaptive method for elliptic eigenvalue problems*, SIAM J. Numer. Anal., **47** (2009), pp. 1067-1091.

[14] X. Gong, L. Shen, D. Zhang, and A. Zhou, *Finite element approximations for Schrödinger equations with applications to electronic structure computations*, J. Comput. Math., **26** (2008), pp. 310-323.

[15] T. Kerkhoven and J. W. Jerome, *L_∞-stability of finite element approximations to elliptic gradient equations*, Numer. Math., **57** (1990), pp. 561-575.

[16] P. Knobloch and L. Tobiska, *The P^1_{mod} element: A new nonconforming finite element for convection-diffusion problems*, SIAM J. Numer. Anal., **41** (2003), pp. 436-456.

[17] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. A, **140** (1965), pp. 4743-4754.

[18] K. Mekchay and R. H. Nochetto, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal., **43** (2005), pp. 1803-1827.

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

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

[21] P. Morin, K. G. Siebert, and A. Veeser, *A basic convergence result for conforming adaptive finite elements*, M^3AS, **18** (2008), pp. 707-737.

[22] R. H. Nochetto, A. Schmidt, K. G. Siebert, and A. Veeser, *Pointwise a posteriori error estimates for monotone semi-linear equations*, Numer. Math., **104** (2006), pp. 515-538.

[23] J. P. Perdew and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron*, Phys. Rev. B, **23** (1981), pp. 5048-5079.

[24] G. P. Srivastava, *Broyden's method for self-consistent field convergence acceleration*, J. Phys. A, **17** (1984), pp. 317–321.

[25] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., **7** (2007), pp. 245-269.

[26] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comput., **77** (2008), pp. 227-241.

[27] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, II: Fourier Analysis, Self-Adjointness*, Academic Press, San Diego, 1975.

[28] R. Verfürth, *A Riview of a Posteriori Error Estimates and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, New York, 1996.

[29] S. R. White, J. W. Wilkins, and M. P. Teter, *Finite-element method for electronic structure*, Phys. Rev. B, **39** (1989), pp. 5819-5833.

[30] J. Xu, *A novel two-grid method for semilinear elliptic equations*, SIAM J. Sci. Comput., **15** (1994), pp. 231-237.

[31] J. Xu and A. Zhou, *Local and parallel finite element algorithms based on two-grid discretizations*, Math. Comput., **69** (2000), pp. 881-909.

[32] A. Zhou, *Multi-level adaptive corrections in finite dimensional approximations*, J. Comput. Math., **28** (2010), pp. 45-54.

LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China and, Graduate University of Chinese Academy of Sciences, Beijing 100190, China
E-mail: helh@lsec.cc.ac.cn

LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
E-mail: azhou@lsec.cc.ac.cn