

# ADAPTIVE FINITE ELEMENT APPROXIMATIONS FOR KOHNSHAM MODELS <sup>\*</sup>

HUAJIE CHEN<sup>†</sup>, XIAOYING DAI<sup>‡</sup>, XINGAO GONG<sup>§</sup>, LIANHUA HE<sup>¶</sup>, AND AIHUI ZHOU<sup>||</sup>

**Abstract.** The Kohn-Sham model is a powerful, widely used approach for computation of ground state electronic energies and densities in chemistry, materials science, biology, and nanoscience. In this paper, we study adaptive finite element approximations for the Kohn-Sham model. Based on the residual type a posteriori error estimators proposed in this paper, we introduce an adaptive finite element algorithm with a quite general marking strategy and prove the convergence of the adaptive finite element approximations. Using Dörfler's marking strategy, we then get the convergence rate and quasi-optimal complexity. We also carry out several typical numerical experiments that not only support our theory, but also show the robustness and efficiency of the adaptive finite element computations in electronic structure calculations.

**Keywords:** Kohn-Sham density functional theory, nonlinear eigenvalue problem, adaptive finite element approximation, convergence, complexity.

**AMS subject classifications:** 35Q55, 65N15, 65N25, 65N30, 81Q05.

**1. Introduction.** The Kohn-Sham density functional model is a powerful, widely used approach for computation of ground state electronic energies and densities in chemistry, materials science, biology, and nanosciences. Consider a molecular system consisting of  $M$  nuclei of charges  $\{Z_1, \dots, Z_M\}$  located at the positions  $\{\mathbf{R}_1, \dots, \mathbf{R}_M\}$  and  $N$  electrons in the non-relativistic and spin-unpolarized setting. By density functional theorem (DFT) [35, 36], the ground state solutions of the system may be obtained by solving the lowest  $N$  eigenpairs of the following Kohn-Sham equation

$$\begin{cases} \left( -\frac{1}{2}\Delta + V_{\text{ext}}(x) + \frac{1}{2} \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|} dy + V_{\text{xc}}(\rho) \right) \phi_i &= \mu_i \phi_i \quad \text{in } \mathbb{R}^3, \quad i = 1, 2, \dots, N, \\ \int_{\mathbb{R}^3} \phi_i \phi_j &= \delta_{ij}, \end{cases} \quad (1.1)$$

where  $V_{\text{ext}}(x) = -\sum_{k=1}^M \frac{Z_k}{|x - \mathbf{R}_k|}$  is the electrostatic potential generated by the nuclei,

$\rho(x) = \sum_{i=1}^N |\phi_i(x)|^2$  is the electron density, and  $V_{\text{xc}}(\rho)$  denotes the exchange-correlation potential.

---

<sup>\*</sup>This work was partially supported by the Funds for Creative Research Groups of China under grant 11021101, the National Basic Research Program of China under grant 2011CB309703, the National Science Foundation of China under grants 11101416 and 91330202, the National 863 Project of China under grant 2012AA01A309, and the National Center for Mathematics and Interdisciplinary Sciences of Chinese Academy of Sciences.

<sup>†</sup>Department of Mathematics, Technische Universität München, Germany (chenh@ma.tum.de).

<sup>‡</sup>LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China (daixy@lsec.cc.ac.cn).

<sup>§</sup>Department of Physics, Fudan University, Shanghai 200433, China (xggong@fudan.edu.cn).

<sup>¶</sup>LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China (helh@lsec.cc.ac.cn).

<sup>||</sup>LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China (azhou@lsec.cc.ac.cn).

Since the core electrons do not participate in the chemical binding and remain almost unchanged, a pseudopotential approximation is usually resorted to in practical computations of the Kohn-Sham equation, which is to replace the Coulomb potential of the nucleus and the effects of the core electrons by an effective ionic potential acting on the valence electrons. Therefore, under the pseudopotential framework, only valence electrons are involved. The pseudopotential consists of two terms: a local component  $V_{\text{loc}}$  (whose associated operator is the multiplication by the function  $V_{\text{loc}}$ ) and a nonlocal component  $V_{\text{nl}}$  (an operator whose expression is given in Section 2). The resulted equation is still (1.1) but  $V_{\text{ext}}(x) = V_{\text{loc}}(x) + V_{\text{nl}}(x)$ ,  $N$  now being the number of valence electrons, and  $\{\phi_i\}_{i=1}^N$  being the set of the pseudo-orbitals of the valence electrons.

We understand that the Kohn-Sham approach achieves so far the best balance between accuracy and efficiency among all the different formalisms of electronic structure theory, and simulations of large-scale material systems with Kohn-Sham DFT are still computationally very demanding (say, thousands of electrons or more). As a result, efficient numerical algorithms that can be scalable on parallel computing platforms are desirable to enable DFT calculations at larger scale and for more complex systems. We see that real-space techniques and methods for electronic structure calculations have been derived much attention from scientific and engineering computing communities and remarkably developed during the last two decades, among which the finite element method possesses several significant advantages [6, 26, 46, 47, 56, 57]. Although the finite element method employs more degrees of freedom than that of traditional methods like plane waves and Gaussians, it results in sparse algebraic eigenvalue problems and thus it is scalable on parallel computing platforms due to the strictly local basis functions, it is variational, and it is friendly to implement adaptive refinement approaches. Consequently, the computational accuracy and efficiency of the finite element approximations can be well controlled.

We observe that even in the pseudopotential setting, the eigenfunctions of (1.1) still vary rapidly around nuclei or chemical bonds [6, 18, 32]. Hence it is also natural to apply adaptive finite element (AFE) approaches to improve the approximation accuracy and reduce the computational cost. Indeed, we see that AFE computations have been quite successfully used in solving Kohn-Sham equations and electronic structure calculations. Tsuchida and Tsukada combined the finite element method with the adaptive curvilinear coordinate approach for electronic structure calculation of some molecules [58, 59]; Shen and Zhang introduced some adaptive tetrahedral finite element discretizations in their theses [51, 63] and calculated several typical molecular systems efficiently [32, 52, 64, 65]; Bylaska et.al used adaptive piecewise linear finite element method on completely unstructured simplex meshes to resolve the rapid variation electronic wave functions around atomic nuclei [10]; Dai et.al designed some parallel adaptive and localization based finite element algorithms for typical quantum chemistry and nanometer material computations containing more than one thousand atoms using tens of hundreds of processors on computer cluster [17, 18, 20, 22]; Gavini et.al constructed a finite element mesh using unstructured coarse-graining technique and computed materials systems [44, 55]; Yang successfully scaled their AFE simulations to over 6000 CPU cores on the Tianhe-1A supercomputer in his thesis [61]. The AFE simulations carried out in this paper also show the robustness and efficiency of the AFE computations in electronic structure calculations. We may refer to [27, 56] and references cited therein for other interesting discussions on adaptive finite element method (AFEM).

We see that it is significant to understand the mechanism of AFE computations, analyze the AFE approximations of Kohn-Sham equations, and give a mathematical justification of the AFE algorithm. We note that the AFE computations are based on some a posteriori error estimators and there are a little work concerning analysis of the a posteriori error estimators and convergence of AFE approximations for DFT. In [14, 15], the authors of this paper considered the nonlinear eigenvalue problems derived from the orbital-free DFT and obtained the convergence and optimal complexity of the AFE algorithm. We understand that the orbital-free DFT is viewed as a simplification of the Kohn-Sham DFT, in which only one eigenpair is involved. In this paper, we shall propose and analyze two AFE algorithms for Kohn-Sham DFT calculations and study the associated convergence and quasi-optimal complexity.

Let us now give an informal description of the main results of this paper. We propose and analyze two AFE algorithms: Algorithm 3.1 and Algorithm 4.1, which are based on the residual type a posteriori error estimators. We show the a posteriori error estimates (see Theorem 4.4) and prove that

- Under some reasonable assumptions, all limit points of the AFE approximations of the ground state solutions are ground state solutions (see Theorem 3.5).
- Under other reasonable assumptions, some eigenpairs (in particular, ground state solutions) can be well approximated by AFE approximations with some convergence rate (see Theorem 4.10).

In addition, we also study quasi-optimal complexity of AFE approximations (see Theorem 4.13).

We mention that Algorithm 3.1 and Algorithm 4.1 may be viewed as some extensions of associated existing algorithms for linear elliptic partial differential equations of second order and have been in fact used for years, for instance, in package RealSPACES (Real Space Parallel Adaptive Calculation of Electronic Structure) of the State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences. As we see, the numerical analysis for AFE approximation has been also derived much attention from the mathematical community. Since Babuška and Vogelius [4] gave an analysis of an AFEM for linear symmetric elliptic problems in one dimension, there has been much investigation on the convergence and complexity of AFEMs in literature (see, e.g., [9, 12, 21, 23, 30, 53] and the references cited therein). In the context of the finite element approximations of linear eigenvalue problems, in particular, we see that there are a number of works concerning a posteriori error estimates [8, 19, 24, 34, 37, 39, 60], AFEM convergence [21, 29, 30, 31, 33] and complexity [19, 21, 29, 33].

However, there are several crucial difficulties in numerical analysis of the Kohn-Sham equation: it is a nonlinear eigenvalue problem whose eigenvalues may be degenerate, and a number of eigenpairs must be involved; the associated energy functional is nonconvex with respect to density  $\rho$ , as a result, there is no uniqueness result for the ground state solutions; the energy functional is invariance under unitary transforms, which also induces redundancy of the ground state solutions. To handle these difficulties arising from the Kohn-Sham equations, we shall present some sophisticated arguments and consider the convergence under the distance between solution sets; investigate the convergence rate and optimal complexity under certain inf-sup assumption; and exploit the relationship between the finite element nonlinear eigenvalue approximations and the associated finite element boundary value approximations. Thanks to our previous works [13, 14, 15, 19, 21, 33, 66, 67] where the

perturbation argument was introduced for analyzing AFEM of eigenvalue problems and the compact approach was specialized for handling the nonlinear effects, combining the crucial technical results proposed also in this paper, we are then able to analyze our adaptive finite element algorithms for Kohn-Sham equations, prove the convergence and get the complexity.

The rest of this paper is organized as follows. In Section 2, we provide some preliminaries for Kohn-Sham DFT problem setting and residual type a posteriori error estimator based AFE methods. We prove the convergence of AFE approximations in Section 3 and analyze the convergence rate and optimal complexity of an AFE algorithm in Section 4. In Section 5, we present some numerical experiments that support the theory. Finally, we give some concluding remarks.

**2. Preliminaries.** Physically, the Kohn-Sham model is set in  $\mathbb{R}^3$ . However, due to the exponential decay of the ground state wavefunction of the Schrödinger equation (c.f., e.g., [2, 62]) and the fact that Kohn-Sham model is an approximation of Schrödinger equation,  $\mathbb{R}^3$  is usually replaced by some polyhedral domain  $\Omega \subset \mathbb{R}^3$  in practical computations for Kohn-Sham equation.

For  $\kappa \in \mathbb{R}^{N \times N}$ , we denote its Frobenius norm by  $|\kappa|$ . For  $p \geq 1$  and  $s \geq 0$ , we denote by  $W^{s,p}(\Omega)$  the standard Sobolev spaces with the induced norm  $\|\cdot\|_{s,p,\Omega}$  (see, e.g. [1, 16]). For  $p = 2$ , we denote by  $H^s(\Omega) = W^{s,2}(\Omega)$  with the norm  $\|\cdot\|_{s,\Omega} = \|\cdot\|_{s,2,\Omega}$ , and  $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$ , where  $v|_{\partial\Omega} = 0$  is understood in the sense of trace. The space  $H^{-1}(\Omega)$ , the dual of  $H_0^1(\Omega)$ , will also be used. Let  $\mathcal{H} = (H_0^1(\Omega))^N$  be the Hilbert space with  $H_1$  inner product

$$(\Phi, \Psi) = \sum_{i=1}^N \int_{\Omega} \phi_i \psi_i \quad \text{for } \Phi = (\phi_1, \dots, \phi_N), \Psi = (\psi_1, \dots, \psi_N) \in \mathcal{H}.$$

Let  $\mathbb{Q}$  be a subspace with orthonormality constraints:

$$\mathbb{Q} = \{\Phi \in \mathcal{H} : \Phi^T \Phi = I^{N \times N}\},$$

where  $\Phi^T \Psi = \left( \int_{\Omega} \phi_i \psi_j \right)_{ij} \in \mathbb{R}^{N \times N}$ . For  $\Phi \in \mathcal{H}$  and a subdomain  $\omega \subset \Omega$ , we shall

denote by  $\rho_{\Phi} = \sum_{i=1}^N |\phi_i|^2$  and (sometimes abuse the notation for simplicity) by

$$\|\Phi\|_{s,\omega} = \left( \sum_{i=1}^N \|\phi_i\|_{s,\omega}^2 \right)^{1/2}, \quad s = 0, 1; \quad \|\Phi\|_{0,p,\omega} = \left( \sum_{i=1}^N \|\phi_i\|_{0,p,\omega}^p \right)^{1/p}, \quad 1 \leq p \leq 6.$$

In our discussions, we shall use the following sets:

$$\mathcal{S}^{N \times N} = \{M \in \mathbb{R}^{N \times N} : M^T = M\}, \quad \mathcal{A}^{N \times N} = \{M \in \mathbb{R}^{N \times N} : M^T = -M\}.$$

For any  $\Phi \in \mathbb{Q}$ , we may decompose  $\mathcal{H}$  into a direct sum of three subspaces (see, e.g., [25]):

$$\mathcal{H} = \mathcal{S}_{\Phi} \oplus \mathcal{A}_{\Phi} \oplus \mathcal{T}_{\Phi},$$

where  $\mathcal{S}_{\Phi} = \Phi \mathcal{S}^{N \times N}$ ,  $\mathcal{A}_{\Phi} = \Phi \mathcal{A}^{N \times N}$ , and  $\mathcal{T}_{\Phi} = \{\Psi \in \mathcal{H} : \Psi^T \Phi = 0 \in \mathbb{R}^{N \times N}\}$ .

For convenience, the symbol  $\lesssim$  will be used throughout this paper, and  $A \lesssim B$  means that  $A \leq CB$  for some constant  $C$  that is independent of mesh parameters. We use  $\mathcal{P}(p, (c_1, c_2))$  to denote a class of functions satisfying some growth conditions:

$$\mathcal{P}(p, (c_1, c_2)) = \{f : \exists a_1, a_2 \in \mathbb{R} \text{ such that } c_1 t^p + a_1 \leq f(t) \leq c_2 t^p + a_2 \quad \forall t \geq 0\}$$

with  $c_1 \in \mathbb{R}$  and  $c_2, p \in [0, \infty)$ .

**2.1. Problem setting.** Consider the following general form of Kohn-Sham energy functional

$$E(\Phi) = \int_{\Omega} \left( \frac{1}{2} \sum_{i=1}^N |\nabla \phi_i|^2 + V_{\text{loc}} \rho_{\Phi} + \sum_{i=1}^N \phi_i V_{\text{nl}} \phi_i + e_{\text{xc}}(\rho_{\Phi}) \right) + \frac{1}{2} D(\rho_{\Phi}, \rho_{\Phi}) \quad (2.1)$$

for  $\Phi = (\phi_1, \phi_2, \dots, \phi_N) \in \mathcal{H}$ , which includes the cases of Coulomb potentials and pseudopotential approximations. For the Coulomb potential setting,  $V_{\text{loc}} = -\sum_{k=1}^M \frac{Z_k}{|x - \mathbf{R}_k|}$  and  $V_{\text{nl}} = 0$ . While for the pseudopotential approximations,  $V_{\text{loc}}$  is the local part of pseudopotential and  $V_{\text{nl}}$  is a nonlocal pseudopotential operator (see, e.g., [40]) given by

$$V_{\text{nl}} \phi = \sum_{j=1}^n (\phi, \zeta_j) \zeta_j$$

with  $\zeta_j \in L^2(\Omega)$  ( $j = 1, 2, \dots, n$ ),  $n \in \mathbb{N}$ .  $D(\rho_{\Phi}, \rho_{\Phi})$  is the electron-electron Coulomb defined by

$$D(f, g) = \int_{\Omega} f(g * r^{-1}) = \int_{\Omega} \int_{\Omega} f(x) g(y) \frac{1}{|x - y|} dx dy,$$

and  $e_{\text{xc}}(t)$  is some real function over  $[0, \infty)$ . In our analysis, we require  $V_{\text{loc}}$  belongs to  $L^2(\Omega)$ . We point out that  $V_{\text{loc}} \in L^2(\Omega)$  is a very mild condition, which is satisfied by both the Coulomb potential  $V_{\text{ext}}(x) = -\sum_{k=1}^M \frac{Z_k}{|x - \mathbf{R}_k|}$  and the local part of pseudopotential. Since  $e_{\text{xc}} : [0, \infty) \rightarrow \mathbb{R}$  does not have a simple analytical expression, we shall use some approximations and assume throughout this paper that

$$e_{\text{xc}}(t) \in \mathcal{P}(3, (c_1, c_2)) \text{ with } c_1 \geq 0 \quad \text{or} \quad e_{\text{xc}}(t) \in \mathcal{P}(4/3, (c_1, c_2)), \quad (2.2)$$

which is satisfied by almost all the LDAs.

The ground state of the system is obtained by solving the minimization problem

$$\inf \{E(\Phi) : \Phi \in \mathbb{Q}\}, \quad (2.3)$$

and we refer to [3, 11, 13] for the discussion of existence of a minimizer. Note that the energy functional (2.1) is invariant with respect to any unitary transform, i.e.

$$E(\Phi) = E(\Phi U) = E\left(\left(\sum_{j=1}^N u_{ij} \phi_j\right)_{i=1}^N\right) \quad \forall U = (u_{ij})_{i,j=1}^N \in \mathcal{O}^{N \times N}, \quad (2.4)$$

where  $\mathcal{O}^{N \times N}$  is the set of orthogonal matrices. It follows from (2.4) that if  $\Phi$  is a minimizer of (2.3), then  $\Phi U$  is also a minimizer for any orthogonal matrix  $U$ . For any  $\Psi \in \mathcal{H}$ , we define the equivalence class

$$[\Psi] = \{\Psi U, \forall U \in \mathcal{O}^{N \times N}\}.$$

We see that any minimizer  $\Phi = (\phi_1, \dots, \phi_N)$  of (2.3) satisfies the following weak form (i.e. the Euler-Lagrange equation associated with the minimization problem):

$$\begin{cases} (H_\Phi \phi_i, v) = \left( \sum_{j=1}^N \lambda_{ij} \phi_j, v \right) \quad \forall v \in H_0^1(\Omega), \quad i = 1, 2, \dots, N, \\ \int_{\Omega} \phi_i \phi_j = \delta_{ij}, \end{cases} \quad (2.5)$$

where  $H_\Phi$  is the Kohn-Sham Hamiltonian operator as

$$H_\Phi = -\frac{1}{2}\Delta + V_{\text{loc}} + V_{\text{nl}} + \int_{\Omega} \frac{\rho_\Phi(y)}{|\cdot - y|} dy + e'_{\text{xc}}(\rho_\Phi) \quad (2.6)$$

and

$$\Lambda = (\lambda_{ij})_{i,j=1}^N = \left( \int_{\Omega} \phi_j H_\Phi \phi_i \right)_{i,j=1}^N \quad (2.7)$$

is the Lagrange multiplier. Since the uniqueness of the ground state solution is unknown even up to a unitary transform, we define the set of ground states by

$$\Theta = \left\{ (\Lambda, \Phi) \in \mathbb{R}^{N \times N} \times \mathbb{Q} : E(\Phi) = \min_{\Psi \in \mathbb{Q}} E(\Psi) \text{ and } (\Lambda, \Phi) \text{ solves (2.5)} \right\}. \quad (2.8)$$

Note that the electron density  $\rho_\Phi$  and the operator  $H_\Phi$  are also invariant under any unitary transform, we may diagonalize the matrix of Lagrange multipliers  $\Lambda$ . More precisely, there exists a  $U \in \mathcal{O}^{N \times N}$ , such that the Lagrange multiplier is diagonal for  $\Psi = \Phi U = (\psi_1, \dots, \psi_N)$ , i.e.,

$$\int_{\Omega} \psi_j H_\Psi \psi_i = \mu_i \delta_{ij}.$$

Consequently, instead of (2.5), we may consider a form with diagonal multiplier as follows:

$$\begin{cases} (H_\Psi \psi_i, v) = (\mu_i \psi_i, v) \quad \forall v \in H_0^1(\Omega), \quad i = 1, 2, \dots, N, \\ \int_{\Omega} \psi_i \psi_j = \delta_{ij}, \end{cases} \quad (2.9)$$

which is the standard Kohn-Sham equation.

Note that any solution of (2.5) can be obtained from a unitary transform of some solution of (2.9). That is, once we get all solution of (2.9), we then obtain all solution of (2.5). Consequently, we also call (2.5) Kohn-Sham equation.

It is well known that the ground state has one electron in each of the  $N$  orbitals with the lowest  $N$  eigenvalues [40]. Therefore, the ground state solutions in (2.8) can be obtained by solving the lowest  $N$  eigenpairs of (2.9).

For convenience, define  $\mathcal{F} : \mathbb{R}^{N \times N} \times \mathcal{H} \rightarrow \mathcal{H}^*$  by

$$\langle \mathcal{F}(\Lambda, \Phi), \Gamma \rangle = \sum_{i=1}^N (H_{\Phi} \phi_i - \sum_{j=1}^N \lambda_{ij} \phi_j, \gamma_i) \quad \forall \Gamma = (\gamma_i)_{i=1}^N \in \mathcal{H}.$$

The Fréchet derivative of  $\mathcal{F}$  with respect to  $\Phi$  at  $(\Lambda, \Phi)$  is denoted by  $\mathcal{F}'_{\Phi}(\Lambda, \Phi) : \mathcal{H} \rightarrow \mathcal{H}^*$  as follows

$$\begin{aligned} \langle \mathcal{F}'_{\Phi}(\Lambda, \Phi) \Psi, \Gamma \rangle &= \frac{1}{4} E''(\Phi)(\Psi, \Gamma) - \sum_{i,j=1}^N (\lambda_{ij} \psi_j, \gamma_i) \\ &= \sum_{i=1}^N (H_{\Phi} \psi_i - \sum_{j=1}^N \lambda_{ij} \psi_j, \gamma_i) + 4 \sum_{i,j=1}^N (e''_{xc}(\rho_{\Phi}) \phi_i \psi_i, \phi_j \gamma_j) + \sum_{i,j=1}^N 4D(\phi_i \psi_i, \phi_j \gamma_j). \end{aligned}$$

To study the convergence and complexity, we need the following assumptions [13]

- A1**  $|e'_{xc}(t)| + |te''_{xc}(t)| \in \mathcal{P}(p_1, (c_1, c_2))$  for some  $p_1 \in [0, 2]$ .
- A2** There exists a constant  $\alpha \in (0, 1]$  such that  $|e''_{xc}(t)| + |te'''_{xc}(t)| \lesssim 1 + t^{\alpha-1} \quad \forall t > 0$ .
- A3**  $(\Lambda, \Phi)$  is a solution of (2.5) and there exists a constant  $\beta > 0$  depending on  $(\Lambda, \Phi)$  such that

$$\inf_{\Gamma \in \mathcal{T}_{\Phi}} \sup_{\Psi \in \mathcal{T}_{\Phi}} \frac{\langle \mathcal{F}'_{\Phi}(\Lambda, \Phi) \Psi, \Gamma \rangle}{\|\Psi\|_{1,\Omega} \|\Gamma\|_{1,\Omega}} \geq \beta. \quad (2.10)$$

**REMARK 2.1.** We see that Assumption **A2** implies Assumption **A1** and the commonly used  $X_{\alpha}$  and LDA exchange-correlation energy functionals satisfy Assumption **A2**.

Assumption **A3** is equivalent to that  $\mathcal{F}'_{\Phi}(\Lambda, \Phi)$  is an isomorphism from  $\mathcal{T}_{\Phi}$  to  $\mathcal{T}_{\Phi}$ . We observe that if Assumption **A3** is satisfied for  $\Phi \in \mathbb{Q}$ , then Assumption **A3** is satisfied for any  $\tilde{\Phi} \in [\Phi]$  with the same constant  $\beta$ , too. We see that a stronger condition than (2.10) that

$$\langle \mathcal{F}'_{\Phi}(\Lambda, \Phi) \Gamma, \Gamma \rangle \geq \gamma \|\Gamma\|_{1,\Omega}^2 \quad \forall \Gamma \in \mathcal{T}_{\Phi}$$

is used in [11, 50], which is satisfied for a linear self-adjoint operator when there is a gap between the lowest  $N$ th eigenvalue and  $(N+1)$ th eigenvalue [50].

**2.2. Adaptive finite element approximations.** Let  $d_{\Omega}$  be the diameter of  $\Omega$  and  $\{\mathcal{T}_h\}$  be a shape regular family of nested conforming meshes over  $\Omega$  with size  $h \in (0, d_{\Omega})$ : there exists a constant  $\gamma^*$  such that

$$\frac{h_{\tau}}{\rho_{\tau}} \leq \gamma^* \quad \forall \tau \in \mathcal{T}_h, \quad (2.11)$$

where  $h_{\tau}$  is the diameter of  $\tau$  for each  $\tau \in \mathcal{T}_h$ ,  $\rho_{\tau}$  is the diameter of the biggest ball contained in  $\tau$ , and  $h = \max\{h_{\tau} : \tau \in \mathcal{T}_h\}$ . Let  $\mathcal{E}_h$  denote the set of interior faces (edges or sides) of  $\mathcal{T}_h$ .

Let  $S^{h,k}(\Omega)$  be a subspace of continuous functions on  $\Omega$  such that

$$S^{h,k}(\Omega) = \{v \in C(\bar{\Omega}) : v|_{\tau} \in P_{\tau}^k \quad \forall \tau \in \mathcal{T}_h\},$$

where  $P_\tau^k$  is the space of polynomials of degree no greater than  $k$  over  $\tau$ . Let  $S_0^{h,k}(\Omega) = S^{h,k}(\Omega) \cap H_0^1(\Omega)$ . We shall denote  $S_0^{h,k}(\Omega)$  by  $S_0^h(\Omega)$  for simplification of notation afterwards and let  $V_h = (S_0^h(\Omega))^N$ .

We consider the following finite element approximations of (2.3):

$$\inf\{E(\Phi_h) : \Phi_h \in V_h \cap \mathbb{Q}\}. \quad (2.12)$$

We see from [3, 13] that the minimizer of (2.12) exists under condition (2.2). Note that any minimizer  $\Phi_h = (\phi_{1,h}, \dots, \phi_{N,h})$  of (2.12) solves the Euler-Lagrange equation

$$\begin{cases} (H_{\Phi_h} \phi_{i,h}, v) = \left( \sum_{j=1}^N \lambda_{i,j,h} \phi_{j,h}, v \right) \quad \forall v \in S_0^h(\Omega), \quad i = 1, 2, \dots, N, \\ \int_{\Omega} \phi_{i,h} \phi_{j,h} = \delta_{ij} \end{cases} \quad (2.13)$$

with the Lagrange multiplier

$$\Lambda_h = (\lambda_{i,j,h})_{i,j=1}^N = \left( \int_{\Omega} \phi_{j,h} H_{\Phi_h} \phi_{i,h} \right)_{i,j=1}^N.$$

Define the set of finite dimensional ground state solutions:

$$\Theta_h = \left\{ (\Lambda_h, \Phi_h) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap V_h) : E(\Phi_h) = \min_{\Psi \in \mathbb{Q} \cap V_h} E(\Psi) \text{ and } (\Lambda_h, \Phi_h) \text{ solves (2.13)} \right\}.$$

We have from [13] that the finite dimensional approximations are uniformly bounded, i.e., there exists a constant  $C$  such that

$$\sup_{(\Lambda_h, \Phi_h) \in \Theta_h, h \in (0, d_\Omega)} (\|\Phi_h\|_{1,\Omega} + |\Lambda_h|) < C. \quad (2.14)$$

Using a unitary transform, we can diagonalize  $\Lambda_h$  and obtain a discrete Kohn-Sham equation

$$\begin{cases} (H_{\Psi_h} \psi_{i,h}, v) = (\mu_{i,h} \psi_{i,h}, v) \quad \forall v \in S_0^h(\Omega), \quad i = 1, 2, \dots, N, \\ \int_{\Omega} \psi_{i,h} \psi_{j,h} = \delta_{ij} \end{cases} \quad (2.15)$$

with  $\mu_{i,h} = (H_{\Psi_h} \psi_{i,h}, \psi_{i,h})$ .

Similar to the continuous case, we have that any solution of (2.13) can be obtained from a unitary transform of some solution of (2.15). That is,

$$\Theta_h = \{ (\Lambda_h, \Phi_h) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap V_h) : \Phi_h \in [\Psi_h] \text{ and } \Lambda_h = \Phi_h^T H_{\Phi_h} \Phi_h, \forall \Psi_h \text{ with } (\mu_h, \Psi_h) \in \Xi_h \},$$

where

$$\Xi_h = \left\{ (\mu_h, \Psi_h) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap V_h) : E(\Psi_h) = \min_{\Psi \in \mathbb{Q} \cap V_h} E(\Psi) \text{ and } (\mu_h, \Psi_h) \text{ solves (2.15)} \right\}.$$

Since (2.15) is solvable, to get  $\Theta_h$ , we always resort to solving (2.15) in practice.

An adaptive mesh-refining algorithm usually consists of the following loop [12, 21]:

**Solve**  $\rightarrow$  **Estimate**  $\rightarrow$  **Mark**  $\rightarrow$  **Refine**.



**Solve.** This step computes the piecewise polynomial finite element approximation with respect to a given mesh. To simplify the analysis and do as the most work on numerical study of convergence of AFE approximations, we shall assume throughout this paper that we have the exact solutions of discretized problems<sup>1</sup>.

**Estimate.** Given a partition  $\mathcal{T}_h$  and the corresponding output  $(\Lambda_h, \Phi_h)$  from the “Solve” step, “Estimate” computes the a posteriori error estimator  $\{\eta_h(\Phi_h, \tau)\}_{\tau \in \mathcal{T}_h}$ , which is defined as follows. Define the element residual  $\mathcal{R}_\tau(\Phi_h)$  and the jump  $J_e(\Phi_h)$  by

$$\mathcal{R}_\tau(\Phi_h) = \left( H_{\Phi_h} \phi_{i,h} - \sum_{j=1}^N \lambda_{ij,h} \phi_{j,h} \right)_{i=1}^N \quad \text{in } \tau \in \mathcal{T}_h,$$

$$J_e(\Phi_h) = \left( j_e(\phi_{i,h}) \right)_{i=1}^N, \quad j_e(\phi_{i,h}) = \frac{1}{2} \nabla \phi_{i,h}|_{\tau_1} \cdot \vec{n}_1 + \frac{1}{2} \nabla \phi_{i,h}|_{\tau_2} \cdot \vec{n}_2,$$

where  $e$  is the common face of elements  $\tau_1$  and  $\tau_2$  with unit outward normals  $\vec{n}_1$  and  $\vec{n}_2$ , respectively. Let  $\omega_h(e)$  be the union of elements that share the face  $e$ , and  $\omega_h(\tau)$  be the union of elements that share an edge with  $\tau$ . For  $\tau \in \mathcal{T}_h$ , we define local error indicator  $\eta_h(\Phi_h, \tau)$  and the oscillation  $\text{osc}_h(\Phi_h, \tau)$  by

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

$$\text{osc}_h(\Phi_h, \tau) = h_\tau \|\mathcal{R}_\tau(\Phi_h) - \overline{\mathcal{R}_\tau(\Phi_h)}\|_{0,\tau},$$

where  $\overline{w}$  is the  $L^2$ -projection of  $w \in L^2(\Omega)$  to polynomials of some degree on  $\tau$  or  $e$ . Given a subset  $\omega \subset \Omega$ , we define the error estimator  $\eta_h(\Phi_h, \omega)$  and the oscillation  $\text{osc}_h(\Phi_h, \omega)$  by

$$\eta_h^2(\Phi_h, \omega) = \sum_{\tau \in \mathcal{T}_h, \tau \subset \omega} \eta_h^2(\Phi_h, \tau) \quad \text{and} \quad \text{osc}_h^2(\Phi_h, \omega) = \sum_{\tau \in \mathcal{T}_h, \tau \subset \omega} \text{osc}_h^2(\Phi_h, \tau).$$

**Mark.** We shall replace the subscript  $h$  (or  $h_k$ ) by an iteration counter  $k$  whenever convenient afterwards. Based on the a posteriori error indicators  $\{\eta_k(\Phi_k, \tau)\}_{\tau \in \mathcal{T}_k}$ , “Mark” gives a strategy to choose a subset of elements  $\mathcal{M}_k$  of  $\mathcal{T}_k$  for refinement. One of the most widely used marking strategy to enforce error reduction is the so-called Dörfler strategy.

Dörfler Strategy. Given a parameter  $0 < \theta < 1$  :

1. Construct a subset  $\mathcal{M}_k$  of  $\mathcal{T}_k$  by selecting some elements in  $\mathcal{T}_k$  such that

$$\sum_{\tau \in \mathcal{M}_k} \eta_k^2(\Phi_k, \tau) \geq \theta \sum_{\tau \in \mathcal{T}_k} \eta_k^2(\Phi_k, \tau). \quad (2.16)$$

2. Mark all the elements in  $\mathcal{M}_k$ .

A weaker strategy, which is called “Maximum Strategy”, only requires that the set of marked elements  $\mathcal{M}_k$  contains at least one element of  $\mathcal{T}_k$  holding the largest value estimator [29, 30]. Namely, there exists at least one element  $\tau_k^{\max} \in \mathcal{M}_k$  such that

$$\eta_k(\Phi_k, \tau_k^{\max}) = \max_{\tau \in \mathcal{T}_k} \eta_k(\Phi_k, \tau). \quad (2.17)$$

<sup>1</sup> Similar conclusion can be expected for the case where the errors of numerical integrations and nonlinear algebraic solvers are included (see Section 6). And we understand that the assumption is indeed a very important practical issue.

It is easy to check that the most commonly used marking strategies, e.g., Dörfler's strategy and Equidistribution strategy, fulfill this condition.

**Refine.** Given the partition  $\mathcal{T}_k$  and the set of marked elements  $\mathcal{M}_k$ , "Refine" produces a new partition  $\mathcal{T}_{k+1}$  by refining all elements in  $\mathcal{M}_k$  at least one time. We restrict ourself to a shape-regular bisection for the refinement. Define

$$\mathcal{R}_{\mathcal{T}_k \rightarrow \mathcal{T}_{k+1}} = \mathcal{T}_k \setminus (\mathcal{T}_k \cap \mathcal{T}_{k+1})$$

as the set of refined elements, we have  $\mathcal{M}_k \subset \mathcal{R}_{\mathcal{T}_k \rightarrow \mathcal{T}_{k+1}}$ . Note that usually more than the marked elements in  $\mathcal{M}_k$  are refined in order to keep the mesh conforming.

**3. Convergence of adaptive finite element approximations.** In this section, we propose and investigate an AFE algorithm with Maximum Strategy for Kohn-Sham equations as follows:

ALGORITHM 3.1. **AFE algorithm with Maximum Strategy**

1. Pick an initial mesh  $\mathcal{T}_0$ , and let  $k = 0$ .
2. Solve (2.15) on  $\mathcal{T}_k$  to get discrete solutions  $(\mu_{i,k}, \psi_{i,k})(i = 1, \dots, N)$  and then  $\Theta_k$ .
3. Compute local error indicators  $\eta_k(\Psi_k, \tau)$  for all  $\tau \in \mathcal{T}_k$ .
4. Construct  $\mathcal{M}_k \subset \mathcal{T}_k$  by Maximum Strategy.
5. Refine  $\mathcal{T}_k$  to get a new conforming mesh  $\mathcal{T}_{k+1}$ .
6. Let  $k = k + 1$  and go to 2.

We shall prove that all the limit points of the AFE approximations generated by Algorithm 3.1 are ground state solutions of (2.5), for which we shall use the similar arguments in [14, 30, 66, 67]. Given an initial mesh  $\mathcal{T}_0$ , Algorithm 3.1 generates a sequence of meshes  $\mathcal{T}_1, \mathcal{T}_2, \dots$ , and associated discrete subspaces

$$S_0^{h_0}(\Omega) \subsetneq S_0^{h_1}(\Omega) \subsetneq \dots \subsetneq S_0^{h_n}(\Omega) \subsetneq S_0^{h_{n+1}}(\Omega) \subsetneq \dots \subsetneq S_\infty(\Omega) \subseteq H_0^1(\Omega),$$

where  $S_\infty(\Omega) = \overline{\cup_{k=1}^\infty S_0^{h_k}(\Omega)}^{H_0^1(\Omega)}$ . Similar to the definition for  $V_h$ , we set  $V_\infty = (S_\infty(\Omega))^N$ . We have that  $V_\infty$  is a Hilbert space with the inner product inherited from  $\mathcal{H}$  and

$$\lim_{k \rightarrow \infty} \inf_{\Psi_k \in V_{h_k}} \|\Psi_k - \Psi_\infty\|_{1,\Omega} = 0 \quad \forall \Psi_\infty \in V_\infty. \quad (3.1)$$

Using a direct calculation (see [13]), we derive that

$$\inf_{\tilde{\Psi}_k \in V_{h_k} \cap \mathbb{Q}} \|\tilde{\Psi}_k - \Psi_\infty\|_{1,\Omega} \lesssim \inf_{\Psi_k \in V_{h_k}} \|\Psi_k - \Psi_\infty\|_{1,\Omega} \quad \forall \Psi_\infty \in V_\infty \cap \mathbb{Q}$$

for any  $k \in \mathbb{N}$ , and hence

$$\lim_{k \rightarrow \infty} \inf_{\tilde{\Psi}_k \in V_{h_k} \cap \mathbb{Q}} \|\tilde{\Psi}_k - \Psi_\infty\|_{1,\Omega} = 0 \quad \forall \Psi_\infty \in V_\infty \cap \mathbb{Q}. \quad (3.2)$$

From [3, 13], we know that if Assumption **A2** is satisfied, then the minimizer of energy functional (2.1) in  $V_\infty \cap \mathbb{Q}$  exists.

We see that any minimizer  $\Phi_\infty = (\phi_{1,\infty}, \dots, \phi_{N,\infty}) \in V_\infty \cap \mathbb{Q}$  solves the following Euler-Lagrange equation

$$\begin{cases} (H_{\Phi_\infty} \phi_{i,\infty}, v) = \left( \sum_{j=1}^N \lambda_{ij,\infty} \phi_{j,\infty}, v \right) \quad \forall v \in S_\infty(\Omega), \quad i = 1, 2, \dots, N, \\ \int_\Omega \phi_{i,\infty} \phi_{j,\infty} = \delta_{ij} \end{cases} \quad (3.3)$$

with the Lagrange multiplier

$$\Lambda_\infty = (\lambda_{ij,\infty})_{i,j=1}^N = \left( \int_{\Omega} \phi_{j,\infty} H_{\Phi_\infty} \phi_{i,\infty} \right)_{i,j=1}^N. \quad (3.4)$$

Define

$$\Theta_\infty = \{(\Lambda_\infty, \Phi_\infty) \in \mathbb{R}^{N \times N} \times (V_\infty \cap \mathbb{Q}) : E(\Phi_\infty) = \min_{\Psi \in V_\infty \cap \mathbb{Q}} E(\Psi) \text{ and } (\Lambda_\infty, \Phi_\infty) \text{ solves (3.3)}\}.$$

Using similar arguments to those in the proof of Theorem 4.1 in [14], we can prove that the AFE approximations for the Kohn-Shan equation converge to some limiting pair in  $\Theta_\infty$ .

LEMMA 3.1. *Let  $\{\Theta_k\}_{k \in \mathbb{N}}$  be the sequence obtained by Algorithm 3.1. We have*

$$\begin{aligned} \lim_{k \rightarrow \infty} E_k &= \min_{\Psi \in V_\infty \cap \mathbb{Q}} E(\Psi), \\ \lim_{k \rightarrow \infty} d_{\mathcal{H}}(\Theta_k, \Theta_\infty) &= 0, \end{aligned}$$

where  $E_k = E(\Phi_k)((\Lambda, \Phi) \in \Theta_k)$  and the distance between sets  $X, Y \subset \mathbb{R}^{N \times N} \times \mathcal{H}$  is defined by

$$d_{\mathcal{H}}(X, Y) = \sup_{(\Lambda, \Phi) \in X} \inf_{(\mu, \Psi) \in Y} (|\Lambda - \mu| + \|\Phi - \Psi\|_{1,\Omega}).$$

*Proof.* Let  $(\Lambda_k, \Phi_k) \in \Theta_k$  for  $k = 1, 2, \dots$ , and  $\{(\Lambda_{k_m}, \Phi_{k_m})\}_{m \in \mathbb{N}}$  be any subsequence of  $\{(\Lambda_k, \Phi_k)\}_{k \in \mathbb{N}}$  with  $1 \leq k_1 < k_2 < \dots < k_m < \dots$ .

First, following [66, 67] (see also [14]), we have from (2.14) and the Eberlein-Smulian Theorem that there exists a weakly convergent subsequence  $\{\Phi_{k_{m_j}}\}_{j \in \mathbb{N}}$  and  $\Phi_\infty \in V_\infty$  satisfying

$$\Phi_{k_{m_j}} \rightharpoonup \Phi_\infty \quad \text{in } \mathcal{H}, \quad (3.5)$$

thus it is sufficient to prove

$$E(\Phi_\infty) = \min_{\Psi \in V_\infty \cap \mathbb{Q}} E(\Psi), \quad (3.6)$$

$$\lim_{j \rightarrow \infty} (\|\Phi_{k_{m_j}} - \Phi_\infty\|_{1,\Omega} + |\Lambda_{k_{m_j}} - \Lambda_\infty|) = 0. \quad (3.7)$$

Since  $H_0^1(\Omega)$  is compactly imbedded in  $L^p(\Omega)$  for  $p \in [2, 6)$ , we have that  $\Phi_{k_{m_j}} \rightarrow \Phi_\infty$  strongly in  $(L^p(\Omega))^N$  as  $j \rightarrow \infty$ . Hence, we obtain that

$$\begin{aligned} \lim_{j \rightarrow \infty} \int_{\Omega} V_{\text{loc}}(x) \rho_{\Phi_{k_{m_j}}} &= \int_{\Omega} V_{\text{loc}}(x) \rho_{\Phi_\infty}, \\ \lim_{j \rightarrow \infty} \int_{\Omega} \sum_{i=1}^N \phi_{i,k_{m_j}} V_{\text{nl}} \phi_{i,k_{m_j}} &= \int_{\Omega} \sum_{i=1}^N \phi_{i,\infty} V_{\text{nl}} \phi_{i,\infty}, \\ \lim_{j \rightarrow \infty} \int_{\Omega} e_{\text{xc}}(\rho_{\Phi_{k_{m_j}}}) &= \int_{\Omega} e_{\text{xc}}(\rho_{\Phi_\infty}), \\ \lim_{j \rightarrow \infty} D(\rho_{\Phi_{k_{m_j}}}, \rho_{\Phi_{k_{m_j}}}) &= D(\rho_{\Phi_\infty}, \rho_{\Phi_\infty}), \end{aligned}$$

where (2.2) is used for the third equality. Besides, from (3.5) we have

$$\liminf_{j \rightarrow \infty} \|\nabla \Phi_{k_{m_j}}\|_{0,\Omega} \geq \|\nabla \Phi_\infty\|_{0,\Omega}.$$

Thus,

$$\liminf_{j \rightarrow \infty} E(\Phi_{k_{m_j}}) \geq E(\Phi_\infty). \quad (3.8)$$

Let  $\Psi_\infty$  be a minimizer of the energy functional in  $V_\infty \cap \mathbb{Q}$ . (3.2) implies that there exists a sequence  $\{\Psi_j\}_{j \in \mathbb{N}}$  such that  $\Psi_j \in V_{k_{m_j}} \cap \mathbb{Q}$  and  $\Psi_j \rightarrow \Psi_\infty$  in  $\mathcal{H}$ . Therefore,

$$E(\Psi_\infty) = \lim_{j \rightarrow \infty} E(\Psi_j). \quad (3.9)$$

Note that  $\{\Phi_{k_{m_j}}\}$  converge to  $\Phi_\infty$  strongly in  $(L^2(\Omega))^N$  leads to  $\Phi_\infty \in V_\infty \cap \mathbb{Q}$ , we have

$$E(\Phi_\infty) \geq E(\Psi_\infty). \quad (3.10)$$

Since  $\Phi_{k_{m_j}}$  is a minimizer of the energy functional in  $V_{k_{m_j}} \cap \mathbb{Q}$ , we obtain

$$E(\Psi_j) \geq E(\Phi_{k_{m_j}}),$$

which together with (3.8), (3.9) and (3.10) leads to

$$\liminf_{j \rightarrow \infty} E(\Phi_{k_{m_j}}) \geq E(\Phi_\infty) \geq E(\Psi_\infty) = \lim_{j \rightarrow \infty} E(\Psi_j) \geq \liminf_{j \rightarrow \infty} E(\Phi_{k_{m_j}}).$$

This implies

$$\lim_{j \rightarrow \infty} E(\Phi_{k_{m_j}}) = E(\Phi_\infty) = \min_{\Psi \in V_\infty \cap \mathbb{Q}} E(\Psi)$$

and thus  $(\Lambda_\infty, \Phi_\infty) \in \Theta_\infty$ .

Therefore, we get that each term of  $E(\Phi)$  converges and in particular

$$\lim_{j \rightarrow \infty} \|\nabla \Phi_{k_{m_j}}\|_{0,\Omega} = \|\nabla \Phi_\infty\|_{0,\Omega}. \quad (3.11)$$

Since  $(H_0^1(\Omega))^N$  is a Hilbert space under norm  $\|\nabla \cdot\|_{0,\Omega}$ , we conclude from (3.5) and (3.11) that

$$\lim_{j \rightarrow \infty} \|\nabla(\Phi_{k_{m_j}} - \Phi_\infty)\|_{0,\Omega} = 0,$$

which together with (2.7), (3.4) and (3.6) implies (3.7). This completes the proof.  $\square$

To show that the limit in  $V_\infty \cap \mathbb{Q}$  is indeed a ground state solution, we turn to the convergence of the a posteriori error estimators. Following the ideas in [14, 29, 30, 43], we split the partition  $\mathcal{T}_k$  into two sets  $\mathcal{T}_k^+$  and  $\mathcal{T}_k^0$ , where

$$\mathcal{T}_k^+ = \{\tau \in \mathcal{T}_k : \tau \in \mathcal{T}_l, \forall l \geq k\} \quad \text{and} \quad \mathcal{T}_k^0 = \mathcal{T}_k \setminus \mathcal{T}_k^+.$$

Actually,  $\mathcal{T}_k^+$  is the set of elements that are not refined any more, and  $\mathcal{T}_k^0$  consists of those elements that will eventually be refined. We denote by

$$\Omega_k^+ = \cup_{\tau \in \mathcal{T}_k^+} \omega_k(\tau) \quad \text{and} \quad \Omega_k^0 = \cup_{\tau \in \mathcal{T}_k^0} \omega_k(\tau).$$

Since the mesh size function  $h_k \equiv h_k(x)$  associated with  $\mathcal{T}_k$  is monotonically decreasing and bounded from below by 0, we have that

$$h_\infty(x) = \lim_{k \rightarrow \infty} h_k(x)$$

is well-defined for almost all  $x \in \Omega$  and hence defines a function in  $L^\infty(\Omega)$ . Moreover, the convergence is uniform (see [43]), more precisely, if  $\{h_k\}_{k \in \mathbb{N}}$  is the sequence of mesh size functions generated by Algorithm 3.1, then

$$\lim_{k \rightarrow \infty} \|h_k - h_\infty\|_{0, \infty, \Omega} = 0 \quad (3.12)$$

and

$$\lim_{k \rightarrow \infty} \|h_k \chi_{\Omega_k^0}\|_{0, \infty, \Omega} = 0, \quad (3.13)$$

where  $\chi_{\Omega_k^0}$  is the characteristic function of  $\Omega_k^0$ .

**LEMMA 3.2.** *Let  $(\Lambda_h, \Phi_h) \in \Theta_h$ . If Assumption **A1** is satisfied, then there exists a constant  $C_\eta > 0$  depending only on the mesh regularity, such that  $\eta_h(\Phi_h, \Omega) \leq C_\eta$  and*

$$\eta_h(\Phi_h, \tau) \lesssim \|\Phi_h\|_{0,6,\omega_h(\tau)} + \|\Phi_h\|_{1,\omega_h(\tau)} \quad \forall \tau \in \mathcal{T}_h.$$

*Proof.* Using (2.14), the inverse inequality, the Hölder inequality, the trace inequality and Assumption **A1**, we have

$$\begin{aligned} h_\tau \|\mathcal{R}_\tau(\Phi_h)\|_{0,\tau} &= h_\tau \left( \sum_{i=1}^N \left\| - \sum_{j=1}^N \lambda_{ij,h} \phi_{j,h} - \frac{1}{2} \Delta \phi_{i,h} + V_{\text{loc}} \phi_{i,h} + V_{\text{nl}} \phi_{i,h} \right. \right. \\ &\quad \left. \left. + e'_{\text{xc}}(\rho_{\Phi_h}) \phi_{i,h} + (r^{-1} * \rho_{\Phi_h}) \phi_{i,h} \right\|_{0,\tau}^2 \right)^{1/2} \\ &\lesssim \sum_{i=1}^N h_\tau \left( \|\phi_{i,h}\|_{0,\tau} + \|\Delta \phi_{i,h}\|_{0,\tau} + \|V_{\text{loc}} \phi_{i,h}\|_{0,\tau} + \sum_{j=1}^n \|\zeta_j\|_{0,\tau}^2 \|\phi_{i,h}\|_{0,\tau} \right. \\ &\quad \left. + \|e'_{\text{xc}}(\rho_{\Phi_h}) \phi_{i,h}\|_{0,\tau} + \|(r^{-1} * \rho_{\Phi_h}) \phi_{i,h}\|_{0,\tau} \right) \\ &\lesssim \|\Phi_h\|_{0,6,\omega_h(\tau)} + \|\Phi_h\|_{1,\omega_h(\tau)} \end{aligned}$$

and

$$\begin{aligned} h_e^{1/2} \|J_e(\Phi_h)\|_{0,e} &= h_e^{1/2} \left( \sum_{i=1}^N \left\| \frac{1}{2} \nabla \phi_{i,h} |_{\tau_1} \cdot \vec{n}_1 + \frac{1}{2} \nabla \phi_{i,h} |_{\tau_2} \cdot \vec{n}_2 \right\|_{0,e}^2 \right)^{1/2} \\ &\lesssim h_e^{1/2} \left( \sum_{i=1}^N (\|\nabla \phi_{i,h} |_{\tau_1}\|_{0,e}^2 + \|\nabla \phi_{i,h} |_{\tau_2}\|_{0,e}^2) \right)^{1/2} \\ &\lesssim h_e^{1/2} \left( h_e^{-1} \sum_{i=1}^N \|\nabla \phi_{i,h}\|_{0,\omega_h(\tau)}^2 \right)^{1/2} \\ &\lesssim \|\Phi_h\|_{1,\omega_h(\tau)}. \end{aligned}$$

Hence we obtain

$$\eta_h(\Phi_h, \tau) \lesssim \|\Phi_h\|_{0,6,\omega_h(\tau)} + \|\Phi_h\|_{1,\omega_h(\tau)} \quad \forall \tau \in \mathcal{T}_h,$$

which together with the Sobolev inequality implies  $\eta_h(\Phi_h, \Omega) \leq C_\eta$ , where the constant  $C_\eta > 0$  depends only on the data and the mesh regularity. This completes the proof.  $\square$

Using similar procedure as in [14, 30], we can prove that the maximal error indicator  $\max_{\tau \in \mathcal{M}_k} \eta_k(\Phi_k, \tau)$  tends to zero.

**LEMMA 3.3.** *Let  $\{\Phi_k\}_{k \in \mathbb{N}}$  be the sequence produced by Algorithm 3.1. If Assumption **A1** is satisfied, then*

$$\lim_{k \rightarrow \infty} \max_{\tau \in \mathcal{M}_k} \eta_k(\Phi_k, \tau) = 0.$$

*Proof.* We see from Lemma 3.1 that for any subsequence  $\{\Phi_{k_m}\}$  of  $\{\Phi_k\}$ , there exist a convergent subsequence  $\{\Phi_{k_{m_j}}\}$  and  $\Phi_\infty$  satisfying  $(\Lambda_\infty, \Phi_\infty) \in \Theta_\infty$  such that

$$\Phi_{k_{m_j}} \rightarrow \Phi_\infty \quad \text{in } \mathcal{H}. \quad (3.14)$$

Hence it is only necessary for us to prove that

$$\lim_{j \rightarrow \infty} \max_{\tau \in \mathcal{M}_{k_{m_j}}} \eta_{k_{m_j}}(\Phi_{k_{m_j}}, \tau) = 0.$$

For simplicity, we denote the subsequence  $\{\Phi_{k_{m_j}}\}_{j \in \mathbb{N}}$  by  $\{\Phi_k\}_{k \in \mathbb{N}}$ , and  $\{\mathcal{T}_{k_{m_j}}\}_{j \in \mathbb{N}}$  by  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$ . We obtain from Lemma 3.2 that

$$\begin{aligned} \eta_k(\Phi_k, \tau_k) &\lesssim \|\Phi_k\|_{0,6,\omega_k(\tau_k)} + \|\Phi_k\|_{1,\omega_k(\tau_k)} \\ &\lesssim \|\Phi_k - \Phi_\infty\|_{1,\Omega} + \|\Phi_\infty\|_{0,6,\omega_k(\tau_k)} + \|\Phi_k\|_{1,\omega_k(\tau_k)}, \end{aligned} \quad (3.15)$$

where  $\tau_k \in \mathcal{M}_k$  be such that

$$\eta_k(\Phi_k, \tau_k) = \max_{\tau \in \mathcal{M}_k} \eta_k(\Phi_k, \tau).$$

Note that (3.14) implies that the first term on the right-hand side of (3.15) goes to zero. Since  $\tau_k \in \mathcal{M}_k \subset \mathcal{T}_k^0$ , we have from (3.13) that

$$|\omega_k(\tau_k)| \lesssim h_{\tau_k}^3 \leq \|h_k \chi_{\Omega_k^0}\|_{0,\infty,\Omega}^3 \rightarrow 0 \quad \text{as } k \rightarrow \infty,$$

which implies that the other two terms on the right-hand side of (3.15) go to zero, too. This completes the proof.  $\square$

Define a global residual  $\mathbf{R}_h(\Phi_h) \in \mathcal{H}^*$  by

$$\langle \mathbf{R}_h(\Phi_h), \Gamma \rangle = \sum_{i=1}^N (H_{\Phi_h} \phi_{i,h} - \sum_{j=1}^N \lambda_{ij,h} \phi_{j,h}, \gamma_i) \quad \forall \Gamma = (\gamma_i)_{i=1}^N \in \mathcal{H}. \quad (3.16)$$

We see that

$$\langle \mathbf{R}_h(\Phi_h), \Gamma \rangle = \sum_{\tau \in \mathcal{T}_h} \left( (\mathcal{R}_\tau(\Phi_h), \Gamma)_\tau + \sum_{e \in \mathcal{E}_h, e \subset \partial \tau} (J_e(\Phi_h), \Gamma)_e \right) \quad \forall \Gamma \in \mathcal{H}. \quad (3.17)$$

Thus

$$|\langle \mathbf{R}_h(\Phi_h), \Gamma \rangle| \lesssim \sum_{\tau \in \mathcal{T}_h} \eta_h(\Phi_h, \tau) \|\Gamma\|_{1, \omega_h(\tau)} \quad \forall \Gamma \in \mathcal{H}. \quad (3.18)$$

Thanks to Lemma 3.2 and Lemma 3.3, by carrying out the similar procedure as the proof for Lemma 4.3 of [14], we can obtain a weak convergence of  $\mathbf{R}_k(\Phi_k)$  as follows.

LEMMA 3.4. *Let  $\{\Phi_k\}_{k \in \mathbb{N}}$  be the sequence produced by Algorithm 3.1. If Assumption **A1** is satisfied, then*

$$\lim_{k \rightarrow \infty} \langle \mathbf{R}_k(\Phi_k), \Gamma \rangle = 0 \quad \forall \Gamma \in \mathcal{H}. \quad (3.19)$$

Now we turn to prove the main result of this section, that is, the limit of the AFE approximations for the Kohn-Shan equation is a ground state solution.

THEOREM 3.5. (convergence) *Let  $\{\Theta_k\}_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 3.1. If the initial mesh  $\mathcal{T}_0$  is sufficiently fine and Assumption **A1** is satisfied, then*

$$\lim_{k \rightarrow \infty} E_k = \min_{\Psi \in \mathbb{Q}} E(\Psi), \quad (3.20)$$

$$\lim_{k \rightarrow \infty} d_{\mathcal{H}}(\Theta_k, \Theta) = 0. \quad (3.21)$$

*Proof.* Let  $\{(\Lambda_k, \Phi_k)\}_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 3.1. We know from Lemma 3.1 that for any subsequence  $\{(\Lambda_{k_m}, \Phi_{k_m})\}_{m \in \mathbb{N}}$ , there exists a convergent subsequence  $\{(\Lambda_{k_{m_j}}, \Phi_{k_{m_j}})\}_{j \in \mathbb{N}}$  and  $(\Lambda_\infty, \Phi_\infty) \in \Theta_\infty$  such that

$$\begin{aligned} \Phi_{k_{m_j}} &\rightarrow \Phi_\infty \quad \text{in } \mathcal{H}, \\ \Lambda_{k_{m_j}} &\rightarrow \Lambda_\infty \quad \text{in } \mathbb{R}^{N \times N}. \end{aligned}$$

Consequently, it is only necessary for us to prove  $(\Lambda_\infty, \Phi_\infty) \in \Theta$ , which implies (3.20) and (3.21) directly. For simplicity, we denote by  $\{(\Lambda_k, \Phi_k)\}_{k \in \mathbb{N}}$  the convergent subsequence  $\{(\Lambda_{k_{m_j}}, \Phi_{k_{m_j}})\}_{j \in \mathbb{N}}$ , and by  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  the corresponding subsequence  $\{\mathcal{T}_{k_{m_j}}\}_{j \in \mathbb{N}}$ .

We first show that the limiting eigenpair  $(\Lambda_\infty, \Phi_\infty)$  is also an eigenpair of (2.5). We have from (3.16) that for any  $\Gamma \in \mathcal{H}$

$$\begin{aligned} (H_{\Phi_\infty} \Phi_\infty - \Lambda_\infty \Phi_\infty, \Gamma) &= (H_{\Phi_\infty} \Phi_\infty - \Lambda_\infty \Phi_\infty, \Gamma) - \langle \mathbf{R}_k(\Phi_k), \Gamma \rangle + \langle \mathbf{R}_k(\Phi_k), \Gamma \rangle \\ &= (H_{\Phi_\infty} \Phi_\infty - H_{\Phi_k} \Phi_k, \Gamma) - (\Lambda_\infty \Phi_\infty - \Lambda_k \Phi_k, \Gamma) \\ &\quad + \langle \mathbf{R}_k(\Phi_k), \Gamma \rangle. \end{aligned} \quad (3.22)$$

By a direct calculation using Assumption **A1**, we get

$$(H_{\Phi_\infty} \Phi_\infty - H_{\Phi_k} \Phi_k, \Gamma) \lesssim \|\Phi_\infty - \Phi_k\|_{1, \Omega} \|\Gamma\|_{1, \Omega},$$

which together with (3.22) leads to

$$(H_{\Phi_\infty} \Phi_\infty - \Lambda_\infty \Phi_\infty, \Gamma) \lesssim (\|\Phi_\infty - \Phi_k\|_{1, \Omega} + |\Lambda_\infty - \Lambda_k|) \|\Gamma\|_{1, \Omega} + \langle \mathbf{R}_k(\Phi_k), \Gamma \rangle. \quad (3.23)$$

We get from  $\Lambda_k \rightarrow \Lambda_\infty$  and  $\Phi_k \rightarrow \Phi_\infty$  in  $\mathcal{H}$  that the first term on the right-hand side of (3.23) goes to zero when  $k$  goes to infinity. We obtain from Lemma 3.4 that the other term on the right-hand side of (3.23) goes to zero, and hence

$$(H_{\Phi_\infty} \Phi_\infty, \Gamma) = (\Lambda_\infty \Phi_\infty, \Gamma) \quad \forall \Gamma \in \mathcal{H}.$$

Then we shall show that for a sufficiently fine initial mesh, the limiting eigenpair  $(\Lambda_\infty, \Phi_\infty)$  is a ground state solution in  $\Theta$ . Similar to [14], we set

$$\mathcal{W} = \{(\Lambda, \Phi) \in \mathbb{R}^{N \times N} \times \mathcal{H} : (\Lambda, \Phi) \text{ solves (2.5)}\}.$$

Note that  $\Theta \subsetneq \mathcal{W}$ . Using the fact

$$\lim_{h \rightarrow 0} \inf_{\Psi \in V_h} \|\Psi - \Phi\|_{1, \Omega} = 0 \quad \forall \Phi \in \mathcal{H},$$

we can choose an initial mesh  $\mathcal{T}_0$  such that

$$E_0 \equiv \min_{\Phi_{h_0} \in V_{h_0} \cap \mathcal{Q}} E(\Phi_{h_0}) < \min_{(M, \Psi) \in \mathcal{W} \setminus \Theta} E(\Psi),$$

Due to  $\mathcal{T}_0 \subset \mathcal{T}_k$ , we have  $E_k \leq E_0$  and hence  $(\Lambda_\infty, \Phi_\infty) \in \Theta$ . This completes the proof.  $\square$

**4. Quasi-optimality of adaptive finite element methods.** In this section we propose and analyze the following AFE algorithm using Dörfler's marking strategy.

**ALGORITHM 4.1. AFE algorithm with Dörfler Strategy**

1. Pick a given mesh  $\mathcal{T}_0$ , and let  $k = 0$ .
2. Solve (2.15) on  $\mathcal{T}_k$  to get discrete solutions  $(\mu_{i,k}, \psi_{i,k})(i = 1, \dots, N)$ , and then  $\Theta_k$ .
3. Compute local error indicators  $\eta_k(\Psi_k, \tau)$  for all  $\tau \in \mathcal{T}_k$ .
4. Construct  $\mathcal{M}_k \subset \mathcal{T}_k$  by Dörfler Strategy and parameter  $\theta$ .
5. Refine  $\mathcal{T}_k$  to get a new conforming mesh  $\mathcal{T}_{k+1}$ .
6. Let  $k = k + 1$  and go to 2.

We shall study the convergence rate and quasi-optimal complexity of Algorithm 4.1, for which we shall apply the perturbation arguments (c.f., e.g., [15, 21, 33]) and certain relationship between nonlinear problem (2.5) and its associated linear boundary value problem (see (A.1)).

To establish the relationship, we define

$$a(\Phi, \Gamma) = \sum_{i=1}^N \frac{1}{2} (\nabla \phi_i, \nabla \gamma_i) \quad \forall \Phi = (\phi_i)_{i=1}^N, \Gamma = (\gamma_i)_{i=1}^N \in \mathcal{H}.$$

One sees that there exists a constant  $c_a > 0$  such that

$$a(\Gamma, \Gamma) \geq c_a \|\Gamma\|_{1, \Omega}^2 \quad \forall \Gamma \in \mathcal{H}. \quad (4.1)$$

Let  $\mathcal{L} : \mathcal{H} \rightarrow \mathcal{H}^*$  be the operator defined by

$$\langle \mathcal{L}(\Phi), \Gamma \rangle = a(\Phi, \Gamma) \quad \forall \Gamma \in \mathcal{H},$$

and  $K : \mathcal{H}^* \rightarrow \mathcal{H}$  be the inverse operator of  $\mathcal{L}$  such that

$$a(K\Phi, \Gamma) = \langle \Phi, \Gamma \rangle \quad \forall \Gamma \in \mathcal{H}.$$



Note that (4.1) implies that  $K$  is well defined and there holds

$$\|K\Phi\|_{1,\Omega} \lesssim \|\Phi\|_{-1,\Omega} \quad \forall \Phi \in \mathcal{H}^*. \quad (4.2)$$

Let  $P_h : \mathcal{H} \rightarrow V_h$  be the  $H^1$ -projection defined by

$$a(\Phi - P_h\Phi, \Gamma) = 0 \quad \forall \Phi \in \mathcal{H}, \Gamma \in V_h. \quad (4.3)$$

For any  $\Phi \in \mathcal{H}$ , there hold

$$\|P_h\Phi\|_{1,\Omega} \lesssim \|\Phi\|_{1,\Omega} \quad \text{and} \quad \lim_{h \rightarrow 0} \|\Phi - P_h\Phi\|_{1,\Omega} = 0. \quad (4.4)$$

**4.1. Basic estimate.** First we recall an a priori error estimate, whose proof is referred to [13]. Define

$$X_{\Phi,h} = \mathcal{S}^{N \times N} \times (V_h \cap (\mathcal{S}_\Phi \oplus \mathcal{T}_\Phi)).$$

**THEOREM 4.1.** *Let  $(\Lambda, \Phi)$  be a solution of (2.5). If Assumptions **A2** and **A3** are satisfied, then there exists  $\delta > 0$  such that for sufficiently small  $h$ , (2.13) has a unique local solution  $(\Lambda_h, \Phi_h) \in X_{\Phi,h} \cap B_\delta((\Lambda, \Phi))$ . Moreover, there hold*

$$\|\Phi - \Phi_h\|_{1,\Omega} \lesssim \inf_{\Psi \in V_h} \|\Phi - \Psi\|_{1,\Omega}, \quad (4.5)$$

$$|\Lambda_h - \Lambda| \lesssim \|\Phi_h - \Phi\|_{1,\Omega}^2 + \|\Phi_h - \Phi\|_{0,\Omega}, \quad (4.6)$$

$$\|\Phi - \Phi_h\|_{0,\Omega} \lesssim r(h) \|\Phi - \Phi_h\|_{1,\Omega} \quad (4.7)$$

with  $r(h) \rightarrow 0$  as  $h \rightarrow 0$ .

Using Theorem 4.1, we can denote afterwards by  $(\Lambda_h, \Phi_h) \in X_{\Phi,h} \cap B_\delta((\Lambda, \Phi))$  the unique local discrete approximation of  $(\Lambda, \Phi) \in \Theta$ .

For simplicity, we denote by  $V = V_{\text{loc}} + V_{\text{nl}}$  and  $\mathcal{N}(\rho_\Phi) = \int_\Omega \frac{\rho_\Phi(y)}{|\cdot - y|} dy + e'_{xc}(\rho_\Phi)$ .

**LEMMA 4.2.** *Let  $(\Lambda, \Phi)$  be a solution of (2.5) and  $h_0 \in (0, 1)$  be the mesh size of the initial mesh  $\mathcal{T}_0$ . If Assumptions **A2** and **A3** are satisfied, then there exists  $\hat{\kappa}(h)$  such that  $\hat{\kappa}(h) \rightarrow 0$  as  $h \rightarrow 0$  and*

$$\|V(\Phi_h - \Phi)\|_{-1,\Omega} + \|\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_\Phi)\Phi\|_{-1,\Omega} \lesssim \hat{\kappa}(h) \|\Phi - \Phi_h\|_{1,\Omega}. \quad (4.8)$$

*Proof.* For any  $\Psi \in \mathcal{H}$ , by using the Hölder inequality and the Young's inequality, we have that for any  $\varepsilon > 0$ , there holds

$$\begin{aligned} \|\Psi\|_{0,3,\Omega} &\leq \|\Psi\|_{0,\Omega}^{1/3} \|\Psi\|_{0,4,\Omega}^{2/3} = (\varepsilon^{-2/3} \|\Psi\|_{0,\Omega}^{1/3}) (\varepsilon^{2/3} \|\Psi\|_{0,4,\Omega}^{2/3}) \\ &\lesssim \frac{\varepsilon^{-2}}{3} \|\Psi\|_{0,\Omega} + \frac{2\varepsilon}{3} \|\Psi\|_{1,\Omega}, \end{aligned}$$

which together with (4.7) implies that there exists a positive constant  $C$  independent of  $h$  and  $\varepsilon$  such that

$$\|\Phi - \Phi_h\|_{0,3,\Omega} \leq C (\varepsilon^{-2} r(h) + \varepsilon) \|\Phi - \Phi_h\|_{1,\Omega} \quad \forall h \in (0, h_0].$$

Therefore, by the Hölder inequality, we get

$$\begin{aligned} \|V_{\text{loc}}(\Phi - \Phi_h)\|_{-1,\Omega} &= \sup_{\Gamma \in \mathcal{H}} \frac{(V_{\text{loc}}(\Phi_h - \Phi), \Gamma)}{\|\Gamma\|_{1,\Omega}} \leq \|V_{\text{loc}}\|_{0,\Omega} \|\Phi - \Phi_h\|_{0,3,\Omega} \\ &\lesssim (\varepsilon^{-2}r(h) + \varepsilon) \|\Phi - \Phi_h\|_{1,\Omega}. \end{aligned} \quad (4.9)$$

For the nonlocal pseudopotential operator, we derive

$$(V_{\text{nl}}(\Phi_h - \Phi), \Gamma) \lesssim \|\Phi - \Phi_h\|_{0,\Omega} \|\Gamma\|_{0,\Omega} \quad \forall \Gamma \in \mathcal{H}$$

from the fact that

$$\left( \sum_{j=1}^n (\zeta_j, \phi_{i,h} - \phi_i) \zeta_j, v \right) \lesssim \|\phi_{i,h} - \phi_i\|_{0,\Omega} \|v\|_{0,\Omega} \quad \forall v \in H_0^1(\Omega), \quad i = 1, \dots, N.$$

Therefore, we have

$$\|V_{\text{nl}}(\Phi_h - \Phi)\|_{-1,\Omega} = \sup_{\Gamma \in \mathcal{H}} \frac{(V_{\text{nl}}(\Phi_h - \Phi), \Gamma)}{\|\Gamma\|_{1,\Omega}} \lesssim \|\Phi - \Phi_h\|_{0,\Omega} \lesssim r(h) \|\Phi - \Phi_h\|_{1,\Omega}. \quad (4.10)$$

For the exchange-correlation part, we have that there exists  $\xi = (\xi_1, \dots, \xi_N)$  with  $\xi_i = \delta_i \phi_{i,h} + (1 - \delta_i) \phi_i$  and  $\delta_i \in [0, 1]$  ( $i = 1, \dots, N$ ), such that

$$(e'_{\text{xc}}(\rho_{\Phi_h})\Phi_h - e'_{\text{xc}}(\rho_{\Phi})\Phi, \Gamma) = \sum_{i=1}^N \int_{\Omega} (e'_{\text{xc}}(\rho_{\xi}) + 2\xi_i^2 e''_{\text{xc}}(\rho_{\xi})) (\phi_{i,h} - \phi_i) \gamma_i.$$

This together with Assumption **A2** leads to

$$\begin{aligned} (e'_{\text{xc}}(\rho_{\Phi_h})\Phi_h - e'_{\text{xc}}(\rho_{\Phi})\Phi, \Gamma) &\lesssim \sum_{i=1}^N \int_{\Omega} (\rho_{\xi} + \rho_{\xi}^{\alpha}) |\phi_{i,h} - \phi_i| \cdot |\gamma_i| \\ &\lesssim \sum_{i=1}^N (\|\rho_{\xi}^{\alpha}\|_{0,3/\alpha,\Omega} \|\phi_{i,h} - \phi_i\|_{0,\Omega} \|\gamma_i\|_{0,6/(3-2\alpha),\Omega} + \|\rho_{\xi}\|_{0,3,\Omega} \|\phi_{i,h} - \phi_i\|_{0,\Omega} \|\gamma_i\|_{0,6,\Omega}) \\ &\lesssim \|\Phi_h - \Phi\|_{0,\Omega} \|\Gamma\|_{1,\Omega} \quad \forall \Gamma \in \mathcal{H}, \end{aligned} \quad (4.11)$$

where the Hölder inequality and the fact

$$\|\rho_{\xi}\|_{0,3,\Omega} \leq \|\xi\|_{0,6,\Omega}^2 \leq \|\Phi\|_{0,6,\Omega}^2 + \|\Phi_h\|_{0,6,\Omega}^2 \leq \bar{C}$$

are used. For the Coulomb potential, we obtain from the Young's inequality and the Uncertainty Principle [49] that

$$\|r^{-1} * (\rho_{\Phi} - \rho_{\Phi_h})\|_{0,\infty,\Omega} \lesssim \sum_{i=1}^N \|\nabla(\phi_i + \phi_{i,h})\|_{0,\Omega} \|\phi_i - \phi_{i,h}\|_{0,\Omega} \lesssim \|\Phi - \Phi_h\|_{0,\Omega}.$$

Therefore, we have that for any  $v \in H_0^1(\Omega)$  and  $1 \leq i \leq N$ , there holds

$$\begin{aligned} &\int_{\Omega} ((r^{-1} * \rho_{\Phi_h})\phi_{i,h} - (r^{-1} * \rho_{\Phi})\phi_i) v \\ &= \int_{\Omega} (r^{-1} * \rho_{\Phi_h})(\phi_{i,h} - \phi_i) v + \int_{\Omega} r^{-1} * (\rho_{\Phi_h} - \rho_{\Phi}) \phi_i v \\ &\lesssim \|r^{-1} * \rho_{\Phi_h}\|_{0,\infty,\Omega} \|\phi_{i,h} - \phi_i\|_{0,\Omega} \|v\|_{0,\Omega} + \|r^{-1} * (\rho_{\Phi_h} - \rho_{\Phi})\|_{0,\infty,\Omega} \|\phi_i\|_{0,\Omega} \|v\|_{0,\Omega} \\ &\lesssim \|\phi_i - \phi_{i,h}\|_{0,\Omega} \|v\|_{0,\Omega} + \|\Phi - \Phi_h\|_{0,\Omega} \|v\|_{0,\Omega}, \end{aligned}$$

which implies

$$((r^{-1} * \rho_{\Phi_h})\Phi_h - (r^{-1} * \rho_{\Phi})\Phi, \Gamma) \lesssim \|\Phi - \Phi_h\|_{0,\Omega} \|\Gamma\|_{0,\Omega} \quad \forall \Gamma \in \mathcal{H}. \quad (4.12)$$

Consequently, we obtain from (4.11), (4.12) and the definition of  $\mathcal{N}$  that

$$\|\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_{\Phi})\Phi\|_{-1,\Omega} = \sup_{\Gamma \in \mathcal{H}} \frac{(\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_{\Phi})\Phi, \Gamma)}{\|\Gamma\|_{1,\Omega}} \lesssim \|\Phi - \Phi_h\|_{0,\Omega}. \quad (4.13)$$

Taking,  $\varepsilon = r(h)^{1/3}$  and setting  $\hat{\kappa}(h) = r(h)^{1/3}$ , we have that  $\hat{\kappa}(h) \rightarrow 0$  as  $h \rightarrow 0$ . Combining (4.7), (4.9), (4.10) and (4.13), we complete the proof of (4.8).  $\square$

We now exploit the relationship between the nonlinear eigenvalue problem and its associated linear boundary value problem, which will be employed in our analysis. We rewrite (2.5) and (2.13) as

$$\Phi = K(\Phi\Lambda - V\Phi - \mathcal{N}(\rho_{\Phi})\Phi),$$

$$\Phi_h = P_h K(\Phi_h\Lambda_h - V\Phi_h - \mathcal{N}(\rho_{\Phi_h})\Phi_h), \quad (4.14)$$

respectively. Set  $W^h = K(\Phi_h\Lambda_h - V\Phi_h - \mathcal{N}(\rho_{\Phi_h})\Phi_h)$ , we have  $\Phi_h = P_h W^h$ .

**THEOREM 4.3.** *Let  $(\Lambda, \Phi)$  be a solution of (2.5). If Assumptions **A2** and **A3** are satisfied, then there exists  $\kappa(h) \in (0, 1)$  such that  $\kappa(h) \rightarrow 0$  as  $h \rightarrow 0$  and*

$$\|\Phi - \Phi_h\|_{1,\Omega} = \|W^h - P_h W^h\|_{1,\Omega} + \mathcal{O}(\kappa(h)) \|\Phi - \Phi_h\|_{1,\Omega}. \quad (4.15)$$

*Proof.* By the definition of  $W^h$ , we have

$$\Phi - W^h = K(\Phi\Lambda - \Phi_h\Lambda_h) + KV(\Phi_h - \Phi) + K(\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_{\Phi})\Phi). \quad (4.16)$$

For the first term on the right-hand side of (4.16), we obtain from (4.2) and (4.7) that

$$\begin{aligned} \|K(\Phi\Lambda - \Phi_h\Lambda_h)\|_{1,\Omega} &\leq \|\Phi\Lambda - \Phi_h\Lambda_h\|_{0,\Omega} \lesssim \|(\Phi - \Phi_h)\Lambda\|_{0,\Omega} + \|\Phi_h(\Lambda - \Lambda_h)\|_{0,\Omega} \\ &\lesssim \|\Phi - \Phi_h\|_{0,\Omega} |\Lambda| + |\Lambda - \Lambda_h| \lesssim r(h) \|\Phi - \Phi_h\|_{1,\Omega}. \end{aligned} \quad (4.17)$$

Using Lemma 4.2, we can estimate the second term on the right-hand side of (4.16) as follows

$$\|KV(\Phi - \Phi_h)\|_{1,\Omega} \lesssim \|V(\Phi - \Phi_h)\|_{-1,\Omega} \lesssim \hat{\kappa}(h) \|\Phi - \Phi_h\|_{1,\Omega}. \quad (4.18)$$

Using (4.2), (4.7) and (4.8), we obtain for the last term of (4.16) that

$$\|K(\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_{\Phi})\Phi)\|_{1,\Omega} \lesssim \|\mathcal{N}(\rho_{\Phi_h})\Phi_h - \mathcal{N}(\rho_{\Phi})\Phi\|_{-1,\Omega} \lesssim r(h) \|\Phi - \Phi_h\|_{1,\Omega}.$$

Set  $\kappa(h) = r(h) + \hat{\kappa}(h)$ , we derive from (4.16), (4.17), and (4.18) that

$$\|\Phi - W^h\|_{1,\Omega} \leq \hat{C}\kappa(h) \|\Phi - \Phi_h\|_{1,\Omega}, \quad (4.19)$$

with  $\hat{C}$  being some constant. Note that (4.14) implies

$$\Phi - \Phi_h = W^h - P_h W^h + \Phi - W^h,$$

which together with (4.19) leads to (4.15). This completes the proof.  $\square$

**4.2. A posteriori error estimates.** Define

$$\tilde{\kappa}(h_0) = \sup_{h \in (0, h_0]} \kappa(h) \quad (4.20)$$

and note that  $\tilde{\kappa}(h_0) \ll 1$  if  $h_0 \ll 1$ . Based on the relevant results for linear boundary value problems (see Appendix), we have the following estimates for AFE approximations.

**THEOREM 4.4.** *Let  $(\Lambda, \Phi)$  be a solution of (2.5),  $h_0 \ll 1$  and  $h \in (0, h_0]$ . If Assumptions **A2** and **A3** are satisfied, then there exist positive constants  $C_1, C_2$  and  $C_3$  depending on the coercivity constant  $c_a$  (in (4.1)) and the shape regularity constant  $\gamma^*$  (in (2.11)), such that*

$$\|\Phi - \Phi_h\|_{1, \Omega}^2 \leq C_1 \eta_h^2(\Phi_h, \Omega), \quad (4.21)$$

$$C_2 \eta_h^2(\Phi_h, \Omega) \leq \|\Phi - \Phi_h\|_{1, \Omega}^2 + C_3 \text{osc}_h^2(\Phi_h, \Omega). \quad (4.22)$$

*Proof.* Due to  $\mathcal{L}W^h = \Phi_h \Lambda_h - V \Phi_h - \mathcal{N}(\rho_{\Phi_h}) \Phi_h$ , we obtain from (A.7) and (A.8) that

$$\|W^h - P_h W^h\|_{1, \Omega}^2 \leq \tilde{C}_1 \tilde{\eta}_h^2(P_h W^h, \Omega), \quad (4.23)$$

$$\tilde{C}_2 \tilde{\eta}_h^2(P_h W^h, \Omega) \leq \|W^h - P_h W^h\|_{1, \Omega}^2 + \tilde{C}_3 \tilde{\text{osc}}_h^2(P_h W^h, \Omega), \quad (4.24)$$

where the constants  $\tilde{C}_1, \tilde{C}_2$  and  $\tilde{C}_3$  are given in Theorem A.1,  $\tilde{\eta}_h^2(P_h W^h, \Omega)$  and  $\tilde{\text{osc}}_h^2(P_h W^h, \Omega)$  are defined by (A.5) and (A.6) with  $\Gamma$  being replaced by  $P_h W^h$ . It is easy to see that  $\tilde{\eta}_h(P_h W^h, \Omega) = \eta_h(\Phi_h, \Omega)$  and  $\tilde{\text{osc}}_h(P_h W^h, \Omega) = \text{osc}_h(\Phi_h, \Omega)$  from their definitions and the fact that  $\Phi_h = P_h W^h$ .

We have from (4.15) and (4.20) that

$$\|\Phi - \Phi_h\|_{1, \Omega} \leq (1 + \hat{C} \tilde{\kappa}(h_0)) \|W^h - P_h W^h\|_{1, \Omega},$$

which together with (4.23) leads to (4.21) by taking the constant

$$C_1 = \tilde{C}_1 (1 + \hat{C} \tilde{\kappa}(h_0))^2. \quad (4.25)$$

Similarly, we get (4.22) from (4.14), (4.15) and (4.24). In particular, we may choose  $C_2$  and  $C_3$  by

$$C_2 = \tilde{C}_2 (1 - \hat{C} \tilde{\kappa}(h_0))^2, \quad C_3 = \tilde{C}_3 (1 - \hat{C} \tilde{\kappa}(h_0))^2. \quad (4.26)$$

This completes the proof.  $\square$

We shall now present the following property that will be used in our analysis.

**LEMMA 4.5.** *Let  $(\Lambda_h, \Psi_h)$  be solution of (2.13). For any  $\Psi'_h = \Psi_h U$  with  $U$  being some orthogonal matrix, there hold*

$$\frac{1}{N} \eta_h^2(\Psi'_h, \tau) \leq \eta_h^2(\Psi_h, \tau) \leq N \eta_h^2(\Psi'_h, \tau), \quad \forall \tau \in \mathcal{T}_h, \quad (4.27)$$

and

$$\frac{1}{N} \text{osc}_h^2(\Psi'_h, \tau) \leq \text{osc}_h^2(\Psi_h, \tau) \leq N \text{osc}_h^2(\Psi'_h, \tau), \quad \forall \tau \in \mathcal{T}_h. \quad (4.28)$$

*Proof.* We write  $U = (\alpha_{i,j})_{i,j=1}^N$ . Since  $U$  is orthogonal, we have  $\sum_{l=1}^N \alpha_{i,l} \alpha_{j,l} = \sum_{l=1}^N \alpha_{l,i} \alpha_{l,j} = \delta_{ij}$  for  $i, j = 1, \dots, N$ .

On the one hand, we obtain from  $\Psi'_h = \Psi_h U$  that

$$\psi'_{i,h} = \sum_{j=1}^N \alpha_{j,i} \psi_{j,h}, \quad i = 1, \dots, N.$$

Denote the Lagrange multiplier corresponding to  $\Psi'_h$  by  $\Lambda'_h$ . Since  $\Psi'_h = \Psi_h U$  implies  $H_{\Psi'_h} = H_{\Psi_h}$ , we get

$$\Lambda'_h = (\Psi'_h)^T H_{\Psi'_h} \Psi'_h = (\Psi_h U)^T H_{\Psi_h} \Psi_h U = U^T \Psi_h^T H_{\Psi_h} \Psi_h U = U^T \Lambda_h U.$$

Therefore,

$$\Psi'_h \Lambda'_h = \Psi_h U U^T \Lambda_h U = \Psi_h \Lambda_h U,$$

that is,

$$\sum_{j=1}^N \lambda'_{j,h} \psi'_{j,h} = \sum_{l,j=1}^N \alpha_{l,i} \lambda_{l,j} \psi_{j,h}, \quad i = 1, \dots, N.$$

Consequently, for any  $\tau \in \mathcal{T}_h$

$$\begin{aligned} \eta_h^2(\Psi'_h, \tau) &= h_\tau^2 \|\mathcal{R}_\tau(\Psi'_h)\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|J_e(\Psi'_h)\|_{0,e}^2 \\ &= \sum_{i=1}^N \left( h_\tau^2 \|H_{\Psi'_h} \psi'_{i,h}\|_{0,\tau}^2 - \sum_{j=1}^N \lambda'_{j,h} \psi'_{j,h} \|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|j_e(\psi'_{i,h})\|_{0,e}^2 \right) \\ &= \sum_{i=1}^N \left( h_\tau^2 \|H_{\Psi_h} \sum_{l=1}^N \alpha_{l,i} \psi_{l,h} - \sum_{l,j=1}^N \alpha_{l,i} \lambda_{l,j} \psi_{j,h}\|_{0,\tau}^2 \right. \\ &\quad \left. + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|j_e(\sum_{l=1}^N \alpha_{l,i} \psi_{l,h})\|_{0,e}^2 \right). \end{aligned}$$

Thus, by triangle inequality and Hölder inequality, we may estimate as follows

$$\begin{aligned} \eta_h^2(\Psi'_h, \tau) &\leq \sum_{i=1}^N \left( h_\tau^2 \left( \sum_{l=1}^N \alpha_{l,i} \|H_{\Psi_h} \psi_{l,h} - \sum_{j=1}^N \lambda_{l,j} \psi_{j,h}\|_{0,\tau} \right)^2 \right. \\ &\quad \left. + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \left( \sum_{l=1}^N \alpha_{l,i} \|j_e(\psi_{l,h})\|_{0,e} \right)^2 \right) \\ &\leq \sum_{i=1}^N \left( \left( \sum_{l=1}^N \alpha_{l,i}^2 \right) \left( h_\tau^2 \sum_{l=1}^N \|H_{\Psi_h} \psi_{l,h} - \sum_{j=1}^N \lambda_{l,j} \psi_{j,h}\|_{0,\tau}^2 \right. \right. \\ &\quad \left. \left. + \sum_{l=1}^N \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|j_e(\psi_{l,h})\|_{0,e}^2 \right) \right) \\ &= \sum_{i=1}^N \left( \sum_{l=1}^N h_\tau^2 \|H_{\Psi_h} \psi_{l,h} - \sum_{j=1}^N \lambda_{l,j} \psi_{j,h}\|_{0,\tau}^2 + \sum_{l=1}^N \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|j_e(\psi_{l,h})\|_{0,e}^2 \right) \\ &= N \eta_h^2(\Psi_h, \tau), \quad \forall \tau \in \mathcal{T}_h, \end{aligned}$$

where the fact  $\sum_{i=1}^N \alpha_{i,i}^2 = 1$  is used. That is,

$$\eta_h^2(\Psi'_h, \tau) \leq N\eta_h^2(\Psi_h, \tau), \quad \forall \tau \in \mathcal{T}_h. \quad (4.29)$$

On the other hand,  $\Psi'_h = \Psi_h U$  implies  $\Psi_h = \Psi'_h U^T$ . Hence,

$$\psi_{i,h} = \sum_{j=1}^N \alpha_{i,j} \psi'_{j,h}, \quad i = 1, \dots, N.$$

By the similar process we obtain that

$$\eta_h^2(\Psi_h, \tau) \leq N\eta_h^2(\Psi'_h, \tau), \quad \forall \tau \in \mathcal{T}_h. \quad (4.30)$$

Similarly, there have

$$\text{osc}_h^2(\Psi'_h, \tau) \leq N\text{osc}_h^2(\Psi_h, \tau), \quad \forall \tau \in \mathcal{T}_h \quad (4.31)$$

and

$$\text{osc}_h^2(\Psi_h, \tau) \leq N\text{osc}_h^2(\Psi'_h, \tau), \quad \forall \tau \in \mathcal{T}_h. \quad (4.32)$$

We obtain (4.27) from (4.29) and (4.30), and get (4.28) from (4.31) and (4.32). This completes the proof.  $\square$

Thanks to Lemma 4.5, we can get the bounds of  $\|\Phi - \Phi_h\|_{1,\Omega}$  by computable terms  $\eta_h^2(\Psi_h, \Omega)$  and  $\text{osc}_h^2(\Psi_h, \Omega)$ , other than the uncomputable term  $\eta_h^2(\Phi_h, \Omega)$  and  $\text{osc}_h^2(\Phi_h, \Omega)$  as in Theorem 4.4, and then get the a posteriori error estimate for distance between the ground states and its approximation as follows.

**THEOREM 4.6.** (a posteriori error estimate) *Suppose  $h_0 \ll 1$  and  $h \in (0, h_0]$ . Let  $(\mathbb{P}_h, \Psi_h)$  be solution of (2.15), if Assumptions **A2** and **A3** are satisfied, then there hold*

$$d_{\mathcal{H}}^2(\overline{\Theta}_h, \Theta) \lesssim \eta_h^2(\Psi_h, \Omega), \quad (4.33)$$

$$\eta_h^2(\Psi_h, \Omega) \lesssim d_{\mathcal{H}}^2(\overline{\Theta}_h, \Theta) + \text{osc}_h^2(\Psi_h, \Omega), \quad (4.34)$$

here  $\overline{\Theta}_h = \{(\Lambda_h, \Phi_h) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap V_h) : \Phi_h \in [\Psi_h], \text{ and } \Lambda_h = \Phi_h^T H_{\Phi_h} \Phi_h\} \subseteq \Theta_h$ .

Our analysis is based on the following crucial technical result, which can be obtain directly from Lemma 4.5.

**LEMMA 4.7.** *Let  $(\Lambda_h, \Phi_h)$  be any solution of (2.13). If there exists constant  $\theta \in (0, 1)$  satisfying*

$$\sum_{\tau \in \mathcal{M}_h} \eta_h^2(\Phi_h, \tau) \geq \theta \eta_h^2(\Phi_h, \Omega), \quad (4.35)$$

then for any  $\Phi'_h = \Phi_h U$  with  $U$  being some orthogonal matrix, there exists a constant  $\theta' \in (0, 1)$ , such that

$$\sum_{\tau \in \mathcal{M}_h} \eta_h^2(\Phi'_h, \tau) \geq \theta' \eta_h^2(\Phi'_h, \Omega). \quad (4.36)$$

In further, we have  $\theta' = \frac{\theta}{N^2}$ .

**4.3. Convergence rate.** Now we turn to analyze the convergence rate of Algorithm 4.1. Similar to [15, 21], we shall first establish some relationships between two level finite element approximations. We use  $\mathcal{T}_H$  to denote a coarse mesh and  $\mathcal{T}_h$  to denote a refined mesh of  $\mathcal{T}_H$ .

LEMMA 4.8. *Let  $h, H \in (0, h_0]$  and  $(\Lambda, \Phi)$  be a solution of (2.5). If Assumptions **A2** and **A3** are satisfied, then*

$$\|\Phi - \Phi_h\|_{1,\Omega} = \|W^H - P_h W^H\|_{1,\Omega} + \mathcal{O}(\tilde{\kappa}(h_0)) (\|\Phi - \Phi_H\|_{1,\Omega} + \|\Phi - \Phi_h\|_{1,\Omega}), \quad (4.37)$$

$$\text{osc}_h(\Phi_h, \Omega) = \widetilde{\text{osc}}_h(P_h W^H, \Omega) + \mathcal{O}(\tilde{\kappa}(h_0)) (\|\Phi - \Phi_H\|_{1,\Omega} + \|\Phi - \Phi_h\|_{1,\Omega}), \quad (4.38)$$

and

$$\eta_h(\Phi_h, \Omega) = \tilde{\eta}_h(P_h W^H, \Omega) + \mathcal{O}(\tilde{\kappa}(h_0)) (\|\Phi - \Phi_H\|_{1,\Omega} + \|\Phi - \Phi_h\|_{1,\Omega}). \quad (4.39)$$

*Proof.* First, we obtain (4.37) from (4.4), (4.19) and the identity

$$\Phi - \Phi_h = W^H - P_h W^H + P_h(W^H - W^h) + \Phi - W^H.$$

For the estimate of (4.38), we get from  $\Phi_h = P_h W^H + P_h(W^h - W^H)$  that

$$\widetilde{\text{osc}}_h(P_h W^h, \Omega) \leq \widetilde{\text{osc}}_h(P_h W^H, \Omega) + \widetilde{\text{osc}}_h(P_h(W^h - W^H), \Omega), \quad (4.40)$$

where  $\widetilde{\text{osc}}$  is given in Appendix. Using (4.14) and the fact  $\widetilde{\text{osc}}_h(\Phi_h, \Omega) = \text{osc}_h(\Phi_h, \Omega)$ , we know that it is only necessary to estimate  $\widetilde{\text{osc}}_h(P_h(W^h - W^H), \Omega)$ .

Since  $\mathcal{L}W^h = \Phi_h \Lambda_h - V \Phi_h - \mathcal{N}(\rho_{\Phi_h}) \Phi_h$  and  $\mathcal{L}W^H = \Phi_H \Lambda_H - V \Phi_H - \mathcal{N}(\rho_{\Phi_H}) \Phi_H$ , we obtain

$$\mathcal{L}(W^h - W^H) = \Phi_h \Lambda_h - \Phi_H \Lambda_H + V(\Phi_H - \Phi_h) + \mathcal{N}(\rho_{\Phi_H}) \Phi_H - \mathcal{N}(\rho_{\Phi_h}) \Phi_h.$$

Let  $G = P_h(W^h - W^H)$  and  $\tilde{\mathcal{R}}_\tau(G)$  be defined by (A.4) with  $\Gamma$  being replaced by  $G$ . We have

$$\tilde{\mathcal{R}}_\tau(G) = \Phi_h \Lambda_h - \Phi_H \Lambda_H + V(\Phi_H - \Phi_h) + \mathcal{N}(\rho_{\Phi_H}) \Phi_H - \mathcal{N}(\rho_{\Phi_h}) \Phi_h - \mathcal{L}G$$

and

$$\begin{aligned} \widetilde{\text{osc}}_h^2(P_h(W^h - W^H), \Omega) &= \sum_{\tau \in \mathcal{T}_h} \widetilde{\text{osc}}_h^2(G, \tau) = \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\tilde{\mathcal{R}}_\tau(G) - \overline{\tilde{\mathcal{R}}_\tau(G)}\|_{0,\tau}^2 \\ &\leq \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\tilde{\mathcal{R}}_\tau(G) + \mathcal{L}G - \overline{(\tilde{\mathcal{R}}_\tau(G) + \mathcal{L}G)}\|_{0,\tau}^2 + \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\mathcal{L}G - \overline{\mathcal{L}G}\|_{0,\tau}^2. \end{aligned} \quad (4.41)$$

Using the inverse inequality, and the fact that  $\Phi_h \Lambda_h$  and  $\Phi_H \Lambda_H$  are piecewise polynomials vectors over  $\mathcal{T}_h$  and  $\mathcal{T}_H$  respectively, (4.7), and (4.8), we may estimate as follows

$$\begin{aligned} & \left( \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\tilde{\mathcal{R}}_\tau(G) + \mathcal{L}G - \overline{(\tilde{\mathcal{R}}_\tau(G) + \mathcal{L}G)}\|_{0,\tau}^2 \right)^{1/2} \\ & \lesssim \sum_{\tau \in \mathcal{T}_h} h_\tau (\|V(\Phi_H - \Phi_h)\|_{0,\tau} + \|\mathcal{N}(\rho_{\Phi_H}) \Phi_H - \mathcal{N}(\rho_{\Phi_h}) \Phi_h\|_{0,\tau}) \\ & \lesssim \tilde{\kappa}(h_0) (\|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi - \Phi_H\|_{1,\Omega}). \end{aligned} \quad (4.42)$$

Combining the inverse inequality, (4.4) and (4.19), we arrive at

$$\begin{aligned} & \left( \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\mathcal{L}G - \overline{\mathcal{L}G}\|_{0,\tau}^2 \right)^{1/2} \lesssim \left( \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|\mathcal{L}G\|_{0,\tau}^2 \right)^{1/2} \lesssim \|G\|_{1,\Omega} \\ & = \|P_h(W^h - W^H)\|_{1,\Omega} \lesssim \tilde{\kappa}(h_0) (\|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi - \Phi_H\|_{1,\Omega}). \end{aligned} \quad (4.43)$$

Taking (4.41), (4.42) and (4.43) into account, we have

$$\widetilde{\text{osc}}_h(P_h(W^h - W^H), \Omega) \lesssim \tilde{\kappa}(h_0) (\|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi - \Phi_H\|_{1,\Omega}), \quad (4.44)$$

which together with (4.40) leads to (4.38).

Finally, we shall prove (4.39). We obtain from (A.8), (4.19) and (4.44) that

$$\begin{aligned} \tilde{\eta}_h(P_h(W^h - W^H), \Omega) & \lesssim \|(W^h - W^H) - P_h(W^h - W^H)\|_{1,\Omega} \\ & \quad + \widetilde{\text{osc}}_h(P_h(W^h - W^H), \Omega) \\ & \lesssim \tilde{\kappa}(h_0) (\|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi - \Phi_H\|_{1,\Omega}). \end{aligned}$$

This together with the fact

$$\tilde{\eta}_h(P_h W^h, \Omega) = \tilde{\eta}_h(P_h W^H + P_h(W^h - W^H), \Omega)$$

leads to

$$\tilde{\eta}_h(P_h W^h, \Omega) = \tilde{\eta}_h(P_h W^H, \Omega) + \mathcal{O}(\tilde{\kappa}(h_0)) (\|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi - \Phi_H\|_{1,\Omega}),$$

which is nothing but (4.39). This completes the proof.  $\square$

For the convenience of the statement of the following results, we need some definition. For  $(\Lambda, \Phi) \in \Theta$  and  $\Phi_h \in V_h$ , we say the equivalence class  $[\Phi_h]$  approximate the equivalence class  $[\Phi]$  if

$$D_{\mathcal{H}}([\Phi_h], [\Phi]) < D_{\mathcal{H}}([\Phi_h], [\tilde{\Phi}]), \quad \forall (\tilde{\Lambda}, \tilde{\Phi}) \in \Theta \text{ and } [\Phi] \neq [\tilde{\Phi}],$$

the distance between sets  $X, Y \subset \mathcal{H}$  is defined by

$$D_{\mathcal{H}}(X, Y) = \sup_{\Phi \in X} \inf_{\Psi \in Y} \|\Phi - \Psi\|_{1,\Omega}.$$

Thanks to Theorem A.2, Lemma 4.7, and Lemma 4.8, by using the similar argument in [15, 19, 21], we get the following theorem.

**THEOREM 4.9.** (error reduction) *Let  $\theta \in (0, 1)$  and  $h_0 \ll 1$ . Let  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces  $\{V_k\}_{k \in \mathbb{N}_0}$  produced by Algorithm 4.1. Assume  $[\Psi_{k_i}]$  is an approximation of some  $[\Phi]$  with  $\Phi$  being one solution of (2.5), denote  $k_{i+1} (> k_i)$  the minimal index among all indexes  $k (> k_i)$  which satisfy that  $[\Psi_k]$  approximates  $[\Phi]$ . If Assumption **A2** is true and  $(\Lambda, \Phi)$  satisfies Assumption **A3**, then*

$$\|\Phi - \Phi_{k_{i+1}}\|_{1,\Omega}^2 + \gamma \eta_{k_{i+1}}^2(\Phi_{k_{i+1}}, \mathcal{T}_{k_{i+1}}) \leq \xi^2 (\|\Phi - \Phi_{k_i}\|_{1,\Omega}^2 + \gamma \eta_{k_i}^2(\Phi_{k_i}, \mathcal{T}_{k_i})) \quad (4.45)$$

with  $\Phi_{k_{i+1}} \in X_{\Phi, k_{i+1}}$  and  $\Phi_{k_i} \in X_{\Phi, k_i}$  satisfying the a priori error estimates (4.5) and (4.7) when  $h$  is replaced by  $h_{k_{i+1}}$  and  $h_{k_i}$ , respectively,  $\gamma > 0$  and  $\xi \in (0, 1)$  some constants depending only on the coercivity constant  $c_a$ , the shape regularity constant  $\gamma^*$ , and the marking parameter  $\theta$ .



*Proof.* For convenience, we use  $\Phi_h, \Phi_H$  to denote  $\Phi_{k_{i+1}}$  and  $\Phi_{k_i}$ , respectively. Then it is sufficient to prove that for  $\Phi_h$  and  $\Phi_H$ , there holds,

$$\|\Phi - \Phi_h\|_{1,\Omega}^2 + \gamma\eta_h^2(\Phi_h, \Omega) \leq \xi^2(\|\Phi - \Phi_H\|_{1,\Omega}^2 + \gamma\eta_H^2(\Phi_H, \Omega)).$$

Note that  $\Phi_H$  and  $\Psi_H$  are solutions of (2.13) and (2.15), respectively. From the relationship of (2.13) and (2.15), we have that if  $[\Phi_H]$  and  $[\Psi_H]$  approximate the same  $[\Phi]$ , then  $\Phi_H = \Psi_H U_H$  with  $U_H$  being some unitary transform. Therefore, we obtain from Lemma 4.7 that Dörfler Marking strategy in Algorithm 4.1 implies that there exists a constant  $\theta' = \frac{\theta}{N^2} \in (0, 1)$ , such that

$$\sum_{\tau \in \mathcal{M}_H} \eta_H^2(\Phi_H, \tau) \geq \theta' \eta_H^2(\Phi_H, \Omega).$$

Thus, from  $W^H = K(\Phi_H \Lambda_H - V \Phi_H - \mathcal{N}(\rho_{\Phi_H}) \Phi_H)$  and  $\Phi_H = P_H W^H$ , we have that Dörfler strategy is satisfied for  $W^H$  with  $\theta' = \frac{\theta}{N^2}$ . So we conclude from Theorem A.2 that there exist constants  $\tilde{\gamma} > 0$  and  $\tilde{\xi} \in (0, 1)$  satisfying

$$\|W^H - P_h W^H\|_{1,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_h^2(P_h W^H, \Omega) \leq \tilde{\xi}^2(\|W^H - \Phi_H\|_{1,\Omega}^2 + \tilde{\gamma} \eta_H^2(\Phi_H, \Omega)), \quad (4.46)$$

where the fact  $\tilde{\eta}_H(P_h W^H, \Omega) = \eta_H(\Phi_H, \Omega)$  is used.

From (4.19), we get that there exists constant  $\hat{C}_1 > 0$  such that

$$(1 + \hat{C}_1 \tilde{\kappa}(h_0)) \|\Phi - \Phi_H\|_{1,\Omega}^2 + \tilde{\gamma} \eta_H^2(\Phi_H, \Omega) \geq \|W^H - P_h W^H\|_{1,\Omega}^2 + \tilde{\gamma} \eta_H^2(\Phi_H, \Omega) \quad (4.47)$$

We obtain from Lemma 4.8 and the Young's inequality that there exists constant  $\hat{C}_2 > 0$  such that

$$\begin{aligned} \|\Phi - \Phi_h\|_{1,\Omega}^2 + \tilde{\gamma} \eta_h^2(\Phi_h, \Omega) &\leq (1 + \delta_1) \|W^H - P_h W^H\|_{1,\Omega}^2 + (1 + \delta_1) \tilde{\gamma} \tilde{\eta}_h^2(P_h W^H, \Omega) \\ &\quad + \hat{C}_2 (1 + \delta_1^{-1}) \tilde{\kappa}^2(h_0) (\|\Phi - \Phi_h\|_{1,\Omega}^2 + \|\Phi - \Phi_H\|_{1,\Omega}^2). \end{aligned} \quad (4.48)$$

where  $\delta_1 \in (0, 1)$  satisfies  $(1 + \delta_1)\xi < 1$ .

Combining (4.46), (4.47) with (4.48), we have that

$$\begin{aligned} &\left(1 - \hat{C}_2 (1 + \delta_1^{-1}) \tilde{\kappa}^2(h_0)\right) \|\Phi - \Phi_h\|_{1,\Omega}^2 + \tilde{\gamma} \eta_h^2(\Phi_h, \Omega) \\ &\leq \left((1 + \delta_1) \tilde{\xi}^2 + (1 + \delta_1) \tilde{\xi}^2 \hat{C}_1 \tilde{\kappa}(h_0) + \hat{C}_2 (1 + \delta_1^{-1}) \tilde{\kappa}^2(h_0)\right) \|\Phi - \Phi_H\|_{1,\Omega}^2 \\ &\quad + (1 + \delta_1) \tilde{\xi}^2 \tilde{\gamma} \eta_H^2(\Phi_H, \Omega). \end{aligned}$$

Since  $h_0 \ll 1$  implies  $\tilde{\kappa}(h_0) \ll 1$ , there holds

$$\begin{aligned} &\|\Phi - \Phi_h\|_{1,\Omega}^2 + \frac{\tilde{\gamma}}{1 - \hat{C}_3 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \eta_h^2(\Phi_h, \Omega) \\ &\leq \frac{(1 + \delta_1) \tilde{\xi}^2 + \hat{C}_3 \tilde{\kappa}(h_0)}{1 - \hat{C}_3 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \left( \|\Phi - \Phi_H\|_{1,\Omega}^2 + \frac{\tilde{\xi}^2 \tilde{\gamma}}{(1 + \delta_1) \tilde{\xi}^2 + \hat{C}_3 \tilde{\kappa}(h_0)} \eta_H^2(\Phi_H, \Omega) \right), \end{aligned}$$

with  $\hat{C}_3$  some constant depending on  $\hat{C}_1$  and  $\hat{C}_2$ . Note that  $h_0 \ll 1$  implies  $\tilde{\kappa}(h_0) \ll 1$ , we see that the constant  $\xi$  defined by

$$\xi = \left( \frac{(1 + \delta_1) \tilde{\xi}^2 + \hat{C}_3 \tilde{\kappa}(h_0)}{1 - \hat{C}_3 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \right)^{1/2}$$

satisfies  $\xi \in (0, 1)$  when  $h_0 \ll 1$ .

Finally, we arrive at (4.45) by using the fact that

$$\frac{\xi^2 \tilde{\gamma}}{(1 + \delta_1) \tilde{\xi}^2 + \hat{C}_3 \tilde{\kappa}(h_0)} < \gamma \quad \text{with} \quad \gamma = \frac{\tilde{\gamma}}{1 - \hat{C}_3 \delta_1^{-1} \tilde{\kappa}^2(h_0)}. \quad (4.49)$$

This completes the proof.  $\square$

We have from Theorem 3.5 that if  $\{\Psi_k\}$  is obtained by Algorithm 4.1, then there exists a subsequence  $\{[\Psi_{k_i}]\}$  that converge to some equivalent class  $[\Phi]$ , where  $\Phi$  is a solution of (2.5). Here, a sequence  $\{[\Psi_{k_i}]\}$  converges to a equivalent class  $[\Phi]$  means that there exist unitary matrices  $U_{k_i} \in \mathcal{O}^{N \times N}$ , such that

$$\lim_{i \rightarrow \infty} \Psi_{k_i} U_{k_i} = \Phi.$$

Therefore, combining Theorem 4.9, we have the following theorem.

**THEOREM 4.10.** (convergence rate) *Let  $\theta \in (0, 1)$  and  $h_0 \ll 1$ . Let  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  be a sequence of finite element approximations obtained by Algorithm 4.1 and  $\{[\Psi_{k_i}]\}$  be the subsequence that converges to some  $[\Phi]$ , where  $\Phi$  is a solution of (2.5). If Assumptions **A2** and **A3** are satisfied, then there holds*

$$\|\Phi - \Phi_{k_{i+1}}\|_{1, \Omega}^2 + \gamma \eta_{k_{i+1}}^2(\Phi_{k_{i+1}}, \mathcal{T}_{k_{i+1}}) \leq \xi^2 (\|\Phi - \Phi_{k_i}\|_{1, \Omega}^2 + \gamma \eta_{k_i}^2(\Phi_{k_i}, \mathcal{T}_{k_i})), \quad (4.50)$$

where  $\Phi_{k_{i+1}} \in X_{\Phi, k_{i+1}}$  and  $\Phi_{k_i} \in X_{\Phi, k_i}$  satisfy the a priori error estimates (4.5) and (4.7) with  $h$  being replaced by  $h_{k_{i+1}}$  and  $h_{k_i}$ , respectively,  $\gamma > 0$  and  $\xi \in (0, 1)$  are constants depending only on the coercivity constant  $c_a$ , the shape regularity constant  $\gamma^*$  and the marking parameter  $\theta$ . Therefore, the  $k_m$ -th iteration solution of Algorithm 4.1 satisfies

$$\|\Phi - \Phi_{k_m}\|_{1, \Omega}^2 + \gamma \eta_{k_m}^2(\Phi_{k_m}, \mathcal{T}_{k_m}) \leq \xi^{2m} (\|\Phi - \Phi_{k_0}\|_{1, \Omega}^2 + \gamma \eta_{k_0}^2(\Phi_{k_0}, \mathcal{T}_{k_0})), \quad (4.51)$$

and

$$|\Lambda - \Lambda_{k_m}| \lesssim \xi^{2m}. \quad (4.52)$$

In further, we have

$$d_{\mathcal{H}}(\Theta_{k_m}, \Theta) \lesssim \xi^{2m}. \quad (4.53)$$

**4.4. Complexity.** Finally, we study the complexity of Algorithm 4.1 in a class of functions. Following [12, 21], define

$$\mathcal{A}_\gamma^s = \{\Psi \in \mathcal{H} : |\Psi|_{s, \gamma} < \infty\},$$

where  $\gamma > 0$  is some constant and

$$|\Psi|_{s, \gamma} = \sup_{\varepsilon > 0} \varepsilon \inf_{\{\mathcal{T} \subset \mathcal{T}_0 : \inf_{\Psi_{\mathcal{T}} \in V_{\mathcal{T}}} (\|\Psi - \Psi_{\mathcal{T}}\|_{1, \Omega}^2 + (\gamma+1) \text{osc}_{\mathcal{T}}^2(\Psi_{\mathcal{T}}, \mathcal{T}))^{1/2} \leq \varepsilon\}} \inf (\#\mathcal{T} - \#\mathcal{T}_0)^s$$

and  $\mathcal{T} \subset \mathcal{T}_0$  means  $\mathcal{T}$  is a refinement of  $\mathcal{T}_0$ . We see that, for all  $\gamma > 0$ ,  $\mathcal{A}_\gamma^s = \mathcal{A}_1^s$ . For simplicity, we use  $\mathcal{A}^s$  to stand for  $\mathcal{A}_1^s$ , and use  $|\Psi|_s$  to denote  $|\Psi|_{s, \gamma}$ . So  $\mathcal{A}^s$  is the class of functions that can be approximated within a given tolerance  $\varepsilon$  by continuous

piecewise polynomial functions over a partition  $\mathcal{T}_k$  with number of degrees of freedom satisfying  $\#\mathcal{T}_k - \#\mathcal{T}_0 \lesssim \varepsilon^{-1/s} |\Psi|_s^{1/s}$ .

LEMMA 4.11. *Suppose  $\theta \in (0, 1)$  and  $h_0 \ll 1$ . Let  $\Psi_H$  and  $\Psi_h$  be the solutions of (2.15) over a conforming mesh  $\mathcal{T}_H$  and its refinement  $\mathcal{T}_h$ , and  $[\Psi_H]$  and  $[\Psi_h]$  approximate the same solution class  $[\Phi]$ , where  $\Phi$  is a solution of (2.5). Suppose Assumption **A2** is true. If for some  $\Phi \in [\Phi]$  satisfying (2.10), we have*

$$\|\Phi - \Phi_h\|_{a,\Omega}^2 + \gamma_* \text{osc}_h^2(\Phi_h, \Omega) \leq \beta_*^2 (\|\Phi - \Phi_H\|_{a,\Omega}^2 + \gamma_* \text{osc}_H^2(\Phi_H, \Omega)) \quad (4.54)$$

with  $\Phi_h \in X_{\Phi,h}$  and  $\Phi_H \in X_{\Phi,H}$  satisfying the a priori error estimates (4.5) and (4.7) when  $h$  is replaced by  $h$  and  $H$ , respectively,  $\gamma_* > 0$  and  $\beta_* \in (0, \sqrt{\frac{1}{2}})$ . Then, the set  $\mathcal{R} = \mathcal{R}_{\mathcal{T}_H \rightarrow \mathcal{T}_h}$  satisfies the following inequality

$$\sum_{\tau \in \mathcal{R}} \eta_H^2(\Phi_H, \tau) \geq \hat{\theta} \sum_{\tau \in \mathcal{T}_H} \eta_H^2(\Phi_H, \tau),$$

here  $\hat{\theta} = \frac{\tilde{C}_2(1-2\tilde{\beta}_*^2)}{\tilde{C}_0(\tilde{C}_1+(1+2\tilde{C}_*^2\tilde{C}_1)\tilde{\gamma}_*)}$ , with  $\tilde{C}_0, \tilde{\beta}_*, \tilde{C}_*$  and  $\tilde{\gamma}_*$  being constants defined in the proof.

*Proof.* For  $W^H = K(\Phi_H \Lambda_H - V \Phi_H - \mathcal{N}(\rho_{\Phi_H}) \Phi_H)$ , we observe from Lemma 4.8 that

$$\begin{aligned} \|\Phi - \Phi_h\|_{1,\Omega} &= \|W^H - P_h W^H\|_{1,\Omega} \\ &\quad + \mathcal{O}(\tilde{\kappa}(h_0)) (\|W^H - P_H W^H\|_{1,\Omega} + \|W^H - P_h W^H\|_{1,\Omega}), \\ \text{osc}_h(\Phi_h, \Omega) &= \widetilde{\text{osc}}_h(P_h W^H, \Omega) + \mathcal{O}(\tilde{\kappa}(h_0)) (\|W^H - P_H W^H\|_{1,\Omega} + \|W^H - P_h W^H\|_{1,\Omega}). \end{aligned}$$

Proceeding the similar procedure as in the proof of Theorem 4.9, we have

$$\begin{aligned} &\|W^H - P_h W^H\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{\text{osc}}_h^2(P_h W^H, \Omega) \\ &\leq \tilde{\beta}_*^2 (\|W^H - P_H W^H\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{\text{osc}}_H^2(P_H W^H, \Omega)) \end{aligned} \quad (4.55)$$

with

$$\tilde{\beta}_* = \left( \frac{\beta_*^2(1 + \delta_1) + \hat{C}_4 \tilde{\kappa}(h_0)}{1 - \hat{C}_4 \delta_1^{-1} \tilde{\kappa}^2(h_0)} \right)^{1/2}, \quad \tilde{\gamma}_* = \frac{\gamma_*}{1 - \hat{C}_4 \delta_1^{-1} \tilde{\kappa}^2(h_0)}, \quad (4.56)$$

where  $\hat{C}_4$  is some positive constant and  $\delta_1 \in (0, 1)$  is some constant as shown in the proof of Theorem 4.9.

Set  $\tilde{C}_0 = \max\{1, \frac{\tilde{C}_3}{\tilde{\gamma}_*}\}$ , we get from (A.8) that

$$\begin{aligned} (1 - 2\tilde{\beta}_*^2) \tilde{C}_2 \tilde{\eta}_H^2(P_H W^H, \Omega) &\leq (1 - 2\tilde{\beta}_*^2) (\|W^H - P_H W^H\|_{a,\Omega}^2 + \tilde{C}_3 \widetilde{\text{osc}}_H^2(P_H W^H, \Omega)) \\ &\leq \tilde{C}_0 (1 - 2\tilde{\beta}_*^2) (\|W^H - P_H W^H\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{\text{osc}}_H^2(P_H W^H, \Omega)), \end{aligned}$$

which together with (4.55) produces

$$\begin{aligned} \frac{\tilde{C}_2}{\tilde{C}_0} (1 - 2\tilde{\beta}_*^2) \sum_{\tau \in \mathcal{T}_H} \tilde{\eta}_H^2(P_H W^H, \tau) &\leq (\|W^H - P_H W^H\|_{a,\Omega}^2 + \tilde{\gamma}_* \widetilde{\text{osc}}_H^2(P_H W^H, \Omega) \\ &\quad - \|W^H - P_h W^H\|_{a,\Omega}^2 - 2\tilde{\gamma}_* \widetilde{\text{osc}}_h^2(P_h W^H, \Omega)). \end{aligned} \quad (4.57)$$

Thus using equality

$$\|W^H - P_H W^H\|_{a,\Omega}^2 - \|W^H - P_h W^H\|_{a,\Omega}^2 = \|P_H W^H - P_h W^H\|_{a,\Omega}^2$$

and Theorem A.3, we obtain that

$$\|W^H - P_H W^H\|_{a,\Omega}^2 - \|W^H - P_h W^H\|_{a,\Omega}^2 \leq \tilde{C}_1 \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(P_H W^H, \tau). \quad (4.58)$$

By the triangle inequality, the inverse inequality, and the Young's inequality, we get

$$\sum_{\tau \in \mathcal{T}_H \cap \mathcal{T}_h} \widetilde{\text{osc}}_H^2(P_H W^H, \tau) \leq 2 \sum_{\tau \in \mathcal{T}_H \cap \mathcal{T}_h} \widetilde{\text{osc}}_h^2(P_h W^H, \tau) + 2C_*^2 \|P_H W^H - P_h W^H\|_{a,\Omega}^2$$

where  $C_*$  is a positive constant depending on the shape regularity constant  $\gamma^*$ . Hence, using the fact

$$\widetilde{\text{osc}}_H^2(P_H W^H, \tau) \leq \tilde{\eta}_H^2(P_H W^H, \tau) \quad \forall \tau \in \mathcal{T}_H,$$

we may estimate as follows

$$\begin{aligned} & \widetilde{\text{osc}}_H^2(P_H W^H, \Omega) - 2\widetilde{\text{osc}}_h^2(P_h W^H, \Omega) \\ & \leq \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(P_H W^H, \tau) + \sum_{\tau \in \mathcal{T}_H \cap \mathcal{T}_h} \widetilde{\text{osc}}_H^2(P_H W^H, \tau) - 2 \sum_{\tau \in \mathcal{T}_H \cap \mathcal{T}_h} \widetilde{\text{osc}}_h^2(P_h W^H, \tau) \\ & \leq \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(P_H W^H, \tau) + 2C_*^2 \|P_H W^H - P_h W^H\|_{a,\Omega}^2 \\ & \leq (1 + 2C_*^2 \tilde{C}_1) \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(P_H W^H, \tau). \end{aligned} \quad (4.59)$$

Combining (4.57), (4.58) and (4.59), we then arrive at

$$\frac{\tilde{C}_2}{\tilde{C}_0} (1 - 2\tilde{\beta}_*^2) \sum_{\tau \in \mathcal{T}_H} \tilde{\eta}_H^2(P_H W^H, \tau) \leq (\tilde{C}_1 + (1 + 2C_*^2 \tilde{C}_1) \tilde{\gamma}_*) \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(P_H W^H, \tau),$$

that is,

$$\sum_{\tau \in \mathcal{R}} \eta_H^2(\Phi_H, \tau) \geq \hat{\theta} \sum_{\tau \in \mathcal{T}_H} \eta_H^2(\Phi_H, \tau)$$

with

$$\hat{\theta} = \frac{\tilde{C}_2 (1 - 2\tilde{\beta}_*^2)}{\tilde{C}_0 (\tilde{C}_1 + (1 + 2C_*^2 \tilde{C}_1) \tilde{\gamma}_*)}.$$

This completes the proof.  $\square$

Similar for the boundary value problem [12] and the linear eigenvalue problems [19], to analyze the complexity of Algorithm 3.1, we need more requirements than for the convergence rate.

ASSUMPTION 4.1.

1. The marking parameter  $\theta$  satisfies  $\theta \in (0, \theta_*)$ , with

$$\theta_* = \frac{1}{N^2} \frac{C_2 \gamma}{C_3 (C_1 + (1 + 2C_*^2 C_1) \gamma)}.$$

2. The marked  $\mathcal{M}_{h_k}$  satisfy (2.16) with minimal cardinality.
3. The distribution of refinement edges on  $\mathcal{T}_{h_0}$  satisfies condition (b) of section 4 in [54].

We mention that Dörfler Marking Strategy selects the marked set  $\mathcal{M}_k$  with minimal cardinality.

LEMMA 4.12. *Let  $\theta \in (0, 1)$  and  $h_0 \ll 1$ ,  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces  $\{V_k\}_{k \in \mathbb{N}_0}$  produced by Algorithm 4.1. Suppose Assumption **A2** is true. If  $[\Psi_k]$  approximates the solution class  $[\Phi]$ , where  $\Phi$  is a solution of (2.5), then for any  $\Phi \in [\Phi] \cap \mathcal{A}^s$  satisfying (2.10), we have*

$$\#\mathcal{M}_k \lesssim (\|\Phi - \Phi_k\|_{a,\Omega}^2 + \gamma \text{osc}_k^2(\Phi_k, \Omega))^{-1/2s} |\Phi|_s^{1/s}, \quad (4.60)$$

where  $\Phi_k \in X_{\Phi,k}$  satisfies the a priori error estimates (4.5) and (4.7) with  $h$  being replaced by  $h_{k_i}$ , and the hidden constant depends on the discrepancy between the marking parameter  $\frac{1}{N^2} \frac{C_2\gamma}{C_3(C_1 + (1+2C_*^2C_1)\gamma)}$  and  $\theta$ .

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

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

We choose  $\delta_1 \in (0, 1)$  to satisfy  $(1 + \delta_1)\tilde{\xi}^2 < 1$  and

$$(1 + \delta_1)^2 \alpha_1^2 \leq \alpha^2, \quad (4.61)$$

which implies

$$(1 + \delta_1)\alpha_1^2 < 1. \quad (4.62)$$

Define

$$\varepsilon = \frac{1}{\sqrt{2}} \alpha_1 (\|\Phi - \Phi_k\|_{a,\Omega}^2 + \gamma \text{osc}_k^2(\Phi_k, \Omega))^{1/2}$$

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

$$\|\Phi - \Phi_\varepsilon\|_{a,\Omega}^2 + (\gamma + 1) \text{osc}_\varepsilon^2(\Phi_\varepsilon, \Omega) \leq \varepsilon^2. \quad (4.63)$$

We get from  $\Phi \in \mathcal{A}^s$  that

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

Let  $\mathcal{T}_*$  be the smallest common refinement of  $\mathcal{T}_k$  and  $\mathcal{T}_\varepsilon$ . Since  $W^\varepsilon = K(\Phi_\varepsilon \Lambda_\varepsilon - V\Phi_\varepsilon - \mathcal{N}(\rho_{\Phi_\varepsilon})\Phi_\varepsilon)$ , we obtain from the triangle inequality, the inverse inequality, and the Young's inequality that

$$\widetilde{\text{osc}}_*^2(P_*W^\varepsilon, \Omega) \leq 2\widetilde{\text{osc}}_*^2(P_\varepsilon W^\varepsilon, \Omega) + 2C_*^2 \|P_\varepsilon W^\varepsilon - P_*W^\varepsilon\|_{a,\Omega}^2,$$

where  $P_\varepsilon$  and  $P_*$  are Galerkin projections on  $\mathcal{T}_\varepsilon$  and  $\mathcal{T}_*$  defined by (4.3). Note that

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

we have

$$\|W^\varepsilon - P_*W^\varepsilon\|_{a,\Omega}^2 + \frac{1}{2C_*^2} \widetilde{\text{osc}}_*^2(P_*W^\varepsilon, \Omega) \leq \|W^\varepsilon - P_\varepsilon W^\varepsilon\|_{a,\Omega}^2 + \frac{1}{C_*^2} \text{osc}_\varepsilon^2(P_\varepsilon W^\varepsilon, \Omega).$$

Since (A.9) implies  $\tilde{\gamma} \leq \frac{1}{2C_*^2}$ , we get that

$$\begin{aligned} \|W^\varepsilon - P_*W^\varepsilon\|_{a,\Omega}^2 + \tilde{\gamma}\tilde{\text{osc}}_*^2(P_*W^\varepsilon, \Omega) &\leq \|W^\varepsilon - P_\varepsilon W^\varepsilon\|_{a,\Omega}^2 + \frac{1}{C_*^2}\text{osc}_\varepsilon^2(P_\varepsilon W^\varepsilon, \Omega) \\ &\leq \|W^\varepsilon - P_\varepsilon W^\varepsilon\|_{a,\Omega}^2 + (\tilde{\gamma} + \sigma)\text{osc}_\varepsilon^2(P_\varepsilon W^\varepsilon, \Omega), \end{aligned}$$

where  $\sigma = \frac{1}{C_*^2} - \tilde{\gamma} \in (0, 1)$ . We may conclude from using the similar argument as that in proof of Theorem 4.9 that

$$\begin{aligned} \|\Phi - \Phi_*\|_{a,\Omega}^2 + \gamma\text{osc}_*^2(\Phi_*, \Omega) &\leq \alpha_0^2 (\|\Phi - \Phi_\varepsilon\|_{a,\Omega}^2 + (\gamma + \sigma)\text{osc}_\varepsilon^2(P_\varepsilon W^\varepsilon, \Omega)) \\ &\leq \alpha_0^2 (\|\Phi - \Phi_\varepsilon\|_{a,\Omega}^2 + (\gamma + 1)\text{osc}_\varepsilon^2(P_\varepsilon W^\varepsilon, \Omega)), \end{aligned} \quad (4.64)$$

where

$$\alpha_0^2 = \frac{(1 + \delta_1) + \hat{C}_3\tilde{\kappa}(h_0)}{1 - \hat{C}_3\delta_1^{-1}\tilde{\kappa}^2(h_0)}$$

and  $\hat{C}_3$  is the constant appearing in the proof of Theorem 4.9. We derive from (4.63) and (4.64) that

$$\|\Phi - \Phi_*\|_{a,\Omega}^2 + \gamma\text{osc}_*^2(\Phi_*, \mathcal{T}_*) \leq \check{\alpha}^2 (\|\Phi - \Phi_k\|_{a,\Omega}^2 + \gamma\text{osc}_k^2(\Phi_k, \mathcal{T}_k))$$

with  $\check{\alpha} = \frac{1}{\sqrt{2}}\alpha_0\alpha_1$ . Using (4.62), we obtain  $\check{\alpha}^2 \in (0, \frac{1}{2})$  when  $h_0 \ll 1$ . Set  $\check{\theta} = \frac{\tilde{C}_2(1-2\hat{\alpha}^2)}{\tilde{C}_0(\tilde{C}_1+(1+2C_*^2\tilde{C}_1)\hat{\gamma})}$ ,  $\hat{\gamma} = \frac{\gamma}{1-\hat{C}_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}$ ,  $\tilde{C}_0 = \max(1, \frac{\tilde{C}_3}{\tilde{\gamma}})$ , and  $\hat{\alpha}^2 = \frac{(1+\delta_1)\check{\alpha}^2 + \hat{C}_4\tilde{\kappa}(h_0)}{1-\hat{C}_4\delta_1^{-1}\tilde{\kappa}^2(h_0)}$ . Denote  $\mathcal{R} = \mathcal{R}_{\mathcal{T}_k \rightarrow \mathcal{T}_*}$  the refined elements from  $\mathcal{T}_k$  to  $\mathcal{T}_*$ , we obtain from Lemma 4.11 that  $\mathcal{T}_*$  satisfies

$$\sum_{\tau \in \mathcal{R}} \eta_k^2(\Phi_k, \tau) \geq \check{\theta} \sum_{\tau \in \mathcal{T}_k} \eta_k^2(\Phi_k, \tau).$$

Similar to the illustration in proof of 4.9, from the relationship of (2.13) and (2.15), we also have that  $\Psi_k = \Phi_k U_k$  with  $U_k$  being some unitary matrix. Therefore, from Lemma 4.7, we have that there exists  $\check{\theta}' = \frac{\check{\theta}}{N^2}$ , such that

$$\sum_{\tau \in \mathcal{R}} \eta_k^2(\Psi_k, \tau) \geq \check{\theta}' \sum_{\tau \in \mathcal{T}_k} \eta_k^2(\Psi_k, \tau). \quad (4.65)$$

We obtain from the definition of  $\gamma$  (see (4.49)) and  $\tilde{\gamma}$  (see (A.9)) that  $\frac{\tilde{C}_3}{\tilde{\gamma}} \geq \tilde{C}_3 C_*^2$ . Note that  $\tilde{C}_3$  and  $C_*$  are constants appeared in upper bound, without loss of generality, we can assume  $\tilde{C}_3 \geq 1$  and  $C_* \geq 1$ . Hence we have  $\tilde{C}_0 = \frac{\tilde{C}_3}{\tilde{\gamma}}$ . Since  $h_0 \ll 1$ , we get that  $\hat{\gamma} > \gamma$  and  $\hat{\alpha} \in (0, \frac{1}{\sqrt{2}}\alpha)$  from (4.61). We observe from (4.25), (4.26) and  $\hat{\gamma} > \gamma$  that

$$\begin{aligned} \check{\theta}' &= \frac{1}{N^2} \frac{\tilde{C}_2(1-2\hat{\alpha}^2)}{\frac{\tilde{C}_3}{\tilde{\gamma}}(\tilde{C}_1 + (1+2C_*^2\tilde{C}_1)\hat{\gamma})} \geq \frac{1}{N^2} \frac{\tilde{C}_2}{\tilde{C}_3(\frac{\tilde{C}_1}{\tilde{\gamma}} + 1 + 2C_*^2\tilde{C}_1)} (1 - \alpha^2) \\ &= \frac{1}{N^2} \frac{\frac{C_2}{(1-\tilde{C}\tilde{\kappa}(h_0))^2}}{\frac{C_3}{(1-\tilde{C}\tilde{\kappa}(h_0))^2}(\frac{C_1}{\tilde{\gamma}(1+\tilde{C}\tilde{\kappa}(h_0))^2} + 1 + 2C_*^2\frac{C_1}{(1+\tilde{C}\tilde{\kappa}(h_0))^2})} (1 - \alpha^2) \\ &\geq \frac{1}{N^2} \frac{C_2}{C_3(\frac{C_1}{\tilde{\gamma}} + (1+2C_*^2C_1))} (1 - \alpha^2) = \frac{1}{N^2} \frac{C_2\gamma}{C_3(C_1 + (1+2C_*^2C_1)\gamma)} (1 - \alpha^2) > \theta \end{aligned}$$

when  $h_0 \ll 1$ .

Therefore, from (4.65), we deduce

$$\sum_{\tau \in \mathcal{R}} \eta_k^2(\Psi_k, \tau) \geq \theta \sum_{\tau \in \mathcal{T}_k} \eta_k^2(\Psi_k, \tau). \quad (4.66)$$

Since  $\mathcal{M}_k$  satisfies (4.66) with minimal cardinality, we arrive at

$$\begin{aligned} \#\mathcal{M}_k &\leq \#\mathcal{R}_{\mathcal{T}_* \rightarrow \mathcal{T}_k} \leq \#\mathcal{T}_* - \#\mathcal{T}_k \leq \#\mathcal{T}_\varepsilon - \#\mathcal{T}_0 \\ &\lesssim \left(\frac{1}{\sqrt{2}}\alpha_1\right)^{-1/s} \left(\|\Phi - \Phi_k\|_{a,\Omega}^2 + \gamma \text{osc}_k^2(\Phi_k, \mathcal{T}_k)\right)^{-1/2s} |\Phi|_s^{1/s}, \end{aligned}$$

which is nothing but (4.60) with an explicit dependence on the discrepancy between  $\theta$  and  $\frac{C_2\gamma}{C_3(C_1+(1+2C_*^2C_1)\gamma)}$  via  $\alpha_1$ . This completes the proof.  $\square$

**THEOREM 4.13.** (optimal complexity) *Let  $\theta \in (0, 1)$  and  $h_0 \ll 1$ . Assume that Assumption **A2** is satisfied and (2.5) has  $m$  solutions (up to the invariance of unitary transform), which are denoted as  $[\Phi^{(l)}](l = 1, \dots, m)$  where  $m$  can be chosen to be  $\infty$ . Let  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  be a sequence of finite element solutions corresponding to a sequence of nested finite element spaces  $\{V_k\}_{k \in \mathbb{N}_0}$  produced by Algorithm 4.1. Then the following quasi-optimal bound is valid*

$$\#\mathcal{T}_n - \#\mathcal{T}_0 \lesssim \sum_{l=1}^m \left( \|\Phi^l - \Phi_{k_{n_l}}^l\|_{1,\Omega}^2 + \gamma \text{osc}_{k_{n_l}}^2(\Phi_{k_{n_l}}^l, \Omega) \right)^{-1/2s}, \quad (4.67)$$

where  $\Phi^l \in [\Phi^{(l)}] \cap \mathcal{A}^s$  satisfies (2.10),  $\Phi_{k_{n_l}}^l \in X_{\Phi^l, k_{n_l}}$  satisfies the a priori error estimates (4.5) and (4.7) with  $h$  being replaced by  $h_{k_{n_l}}$ , and the hidden constant depends on the exact solution  $\Phi^l$  and the discrepancy between  $\theta$  and  $\frac{1}{N^2} \frac{C_2\gamma}{C_3(C_1+(1+2C_*^2C_1)\gamma)}$ . Here,  $n_l$  and  $k_{n_l}$  are the total number and the maximal index of iteration which approximate  $[\Phi^{(l)}](l = 1, \dots, m)$  among the  $n$  iteration, respectively.

*Proof.* Assume that among the iterate solution spaces  $\{[\Psi_i]\}_{i=1}^n$ , there are  $n_l$  approximations for  $[\Phi^{(l)}](l = 1, \dots, m)$ , which are denoted by  $[\Psi_{k_i}^l](i = 1, \dots, n_l)$ . Here,  $\sum_{l=1}^m n_l = n$ , and  $n_l$  can be 0. Recall that (see Theorem 6.1 in [54])

$$\#\mathcal{T}_n - \#\mathcal{T}_0 \lesssim \sum_{l=1}^m \sum_{i=1}^{n_l} \#\mathcal{M}_{k_i^l},$$

we obtain from (4.60) that

$$\#\mathcal{T}_n - \#\mathcal{T}_0 \lesssim \sum_{l=1}^m \sum_{i=1}^{n_l} \left( \|\Phi^l - \Phi_{k_i^l}^l\|_{1,\Omega}^2 + \gamma \text{osc}_{k_i^l}^2(\Phi_{k_i^l}^l, \Omega) \right)^{-1/2s} (|\Phi^l|_s^{1/s}).$$

Note that (4.22) implies

$$\|\Phi^l - \Phi_{k_i^l}^l\|_{1,\Omega}^2 + \gamma \eta_{k_i^l}^2(\Phi_{k_i^l}^l, \Omega) \leq \check{C} \left( \|\Phi^l - \Phi_{k_i^l}^l\|_{1,\Omega}^2 + \gamma \text{osc}_{k_i^l}^2(\Phi_{k_i^l}^l, \Omega) \right),$$

where  $\check{C} = \max(1 + \frac{\gamma}{C_2}, \frac{C_3}{C_2})$ , we conclude

$$\#\mathcal{T}_n - \#\mathcal{T}_0 \lesssim \sum_{l=1}^m \sum_{i=1}^{n_l} \left( \|\Phi^l - \Phi_{k_i^l}^l\|_{1,\Omega}^2 + \gamma \eta_{k_i^l}^2(\Phi_{k_i^l}^l, \Omega) \right)^{-1/2s} (|\Phi^l|_s^{1/s}).$$

Since (4.50) yields

$$\|\Phi^l - \Phi_{k_{n_l}}^l\|_{1,\Omega}^2 + \gamma\eta_{k_{n_l}}^2(\Phi_{k_{n_l}}^l, \Omega) \leq \xi^{2(n_l-i)} (\|\Phi^l - \Phi_{k_i}^l\|_{1,\Omega}^2 + \gamma\eta_{k_i}^2(\Phi_{k_i}^l, \Omega)),$$

we arrive at

$$\begin{aligned} \#\mathcal{T}_n - \#\mathcal{T}_0 &\lesssim \sum_{l=1}^m \left( |\Phi^l|_s^{1/s} (\|\Phi^l - \Phi_{k_{n_l}}^l\|_{1,\Omega}^2 + \gamma\eta_{k_{n_l}}^2(\Phi_{k_{n_l}}^l, \Omega))^{-1/2s} \sum_{i=1}^{n_l} \xi^{\frac{n_l-i}{s}} \right) \\ &\lesssim \sum_{l=1}^m \left( |\Phi^l|_s^{1/s} (\|\Phi^l - \Phi_{k_{n_l}}^l\|_{1,\Omega}^2 + \gamma\eta_{k_{n_l}}^2(\Phi_{k_{n_l}}^l, \Omega))^{-1/2s} \right), \end{aligned}$$

where the fact  $\xi < 1$  is used.

Thus we obtain from  $\text{osc}_k(\Phi_k^l, \Omega) \leq \eta_k(\Phi_k^l, \Omega)$  that

$$\#\mathcal{T}_n - \#\mathcal{T}_0 \lesssim \sum_{l=1}^m \left( \|\Phi^l - \Phi_{k_{n_l}}^l\|_{1,\Omega}^2 + \gamma\text{osc}_{k_{n_l}}^2(\Phi_{k_{n_l}}^l, \Omega) \right)^{-1/2s}.$$

This completes the proof.  $\square$

**5. Numerical examples.** In this section, we shall present some numerical simulations for three typical molecular systems:  $C_9H_8O_4$  (Aspirin),  $C_5H_9O_2N$  ( $\alpha$  amino acid), and  $C_{60}$  (fullerene), which support our theory. Due to the length limitation for the paper, we only show the results for pseudopotential approximations for illustration.

Our numerical experiments are carried out on LSSC-III in the State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences, and our package RealSPACES (Real Space Parallel Adaptive Calculation of Electronic Structure) that are based on the toolbox PHG [68] of the State Key Laboratory of Scientific and Engineering Computing, Chinese Academy of Sciences.

In our computations, we use the norm-conserving pseudopotential obtained by fhi98PP software and the LDA exchange-correlation potential. We use Algorithm 4.1 and apply the standard quadratic finite element discretizations. Since the analytic solutions are not known even for the simplest systems, we only show the convergence curve of the a posteriori error estimator  $\eta_k^2(\Psi_k, \Omega)$  in our figures. The mesh and density illustrations are drawn using ParaView.

**Example 1:** Aspirin  $C_9H_8O_4$ .

The ground state energy obtained by SIESTA is  $-119.621$  *a.u.*. In our computations, we choose the computational domain to be  $\Omega = [-20.0, 20.0]^3$ .

The atomic configuration, the calculated ground state charge density and the associated computational mesh are shown in Figure 5.1. First, comparing the configuration figure (the left one of Figure 5.1) and the charge density figure (the middle one of Figure 5.1), we can see qualitatively that our calculations are correct, the carbon-hydrogen bonds, carbon-oxygen bonds, and the oxygen-hydrogen bonds are preserved very well. If we take a detailed look at the charge density figure, we can further see that the charge is more concentrative around the oxygen than around the carbon. We also see from the mesh figure (the right one of Figure 5.1) and the charge density figure that our error estimator can catch the oscillations of the charge density very well, which qualitatively confirms that our error estimator is efficient.

We now turn to analyze some quantitative behavior of our calculations. The convergence curve of the ground state energy is shown in the left of Figure 5.2. We



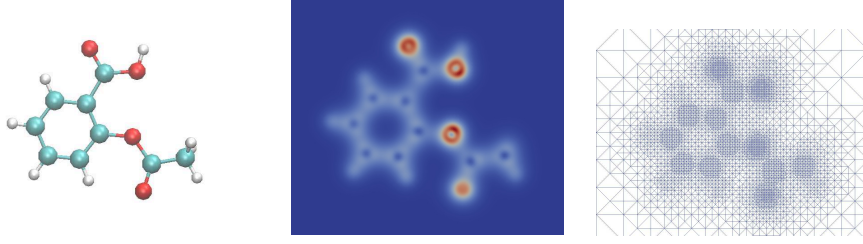


Fig. 5.1:  $C_9H_8O_4$ : configuration, charge density and mesh on plane  $z = 0$ .

observe that the ground state energy approximations converge to  $-119.918 a.u.$ , which is very close to the value given by SIESTA. This result validates our calculations quantitatively. We see from the right of Figure 5.2 that the convergence curve of the a posteriori error estimator is parallel to the line with slope  $-\frac{2}{3}$ , which means that it reaches the optimal convergence rate. From the analysis result for the a posteriori error estimator (Theorem 4.3) the optimal convergence of the a posteriori error estimator also indicates that the approximation of the eigenfunction space have reached the optimal convergence rate, which coincides with our theory in Section 4.

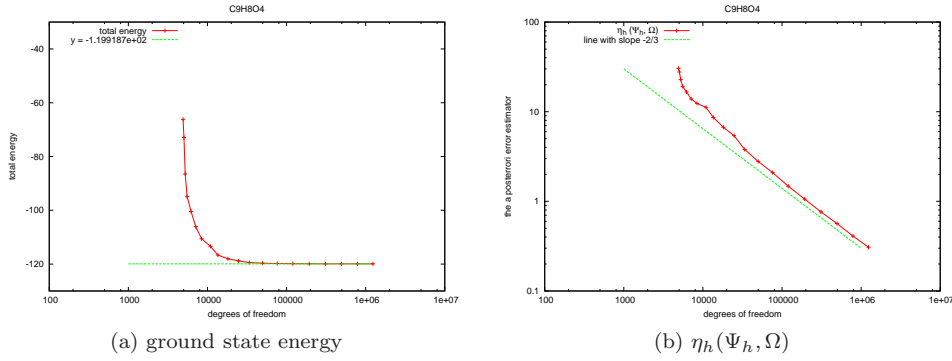


Fig. 5.2: The convergence curves of the ground state energy and  $\eta_h(\Psi_h, \Omega)$ .

**Example 2:**  $\alpha$  amino acid  $C_5H_9O_2N$ .

The ground state energy obtained by SIESTA is  $-75.494 a.u.$ . In our computations, we choose the computational domain to be  $\Omega = [-10.0, 10.0]^3$ .

The atomic configuration, the calculated ground state charge density and the associated computational mesh are shown in Figure 5.3. We have to point out that for  $C_5H_9O_2N$ , not more than 2 atoms stay in the same plane. Therefore, it is very difficult to find a plane where the configuration and the charge density coincide very well with each other as Example 1. Similar to Example 1, we also choose the plane  $z = 0$  as our viewpoint. Anyway, we can see from the figure for charge density and the figure for the adaptive mesh that our error indicator is very efficient. These results can validate our computations.

The convergence curves of the ground state energy and the a posteriori error estimator  $\eta_k(\Psi_k, \Omega)$  obtained by the quadratic finite elements are shown in Figure 5.4, from which we observe that the ground state energy approximations converge to

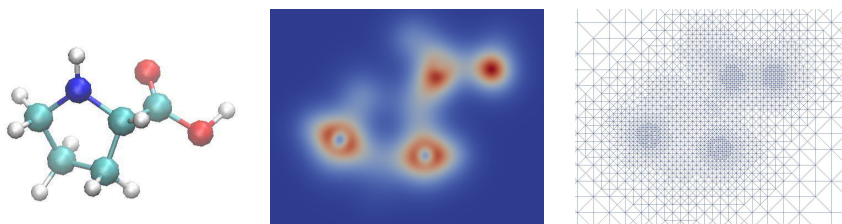


Fig. 5.3:  $C_5H_9O_2N$ : configuration, charge density and mesh on plane  $z = 0$ .

$-75.494 a.u.$ , and the a posteriori error estimator decays with a rate  $-\frac{2}{3}$ . This implies the similar conclusions as those for Example 1.

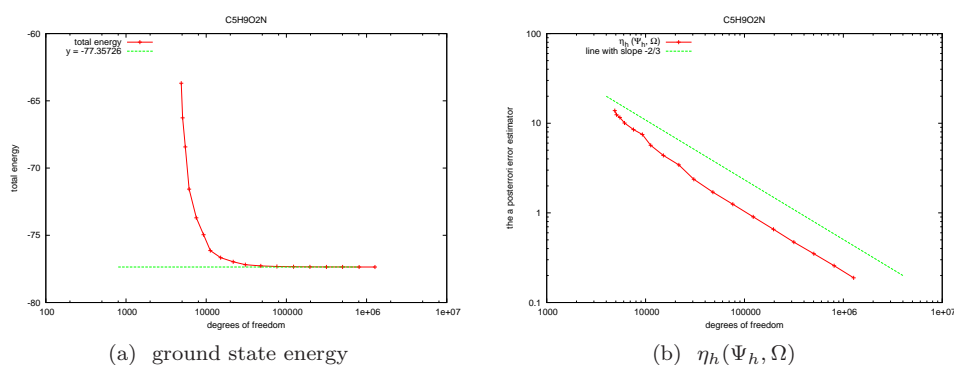


Fig. 5.4: The convergence curves of the ground state energy and  $\eta_h(\Psi_h, \Omega)$ .

**Example 3:** Fullerene  $C_{60}$ .

The ground state energy obtained by SIESTA is  $-341.340 a.u.$ . In our computations, we choose  $\Omega = [-30.0, 16.0] \times [-23.0, 22.0] \times [-24.0, 21.0]$  to be the computational domain.

We can see the preservation of carbon-hydrogen bonds in Figure 5.5, which validates our calculations. Figure 5.5, Figure 5.6 and Figure 5.7 show that more mesh points are placed around the atoms.

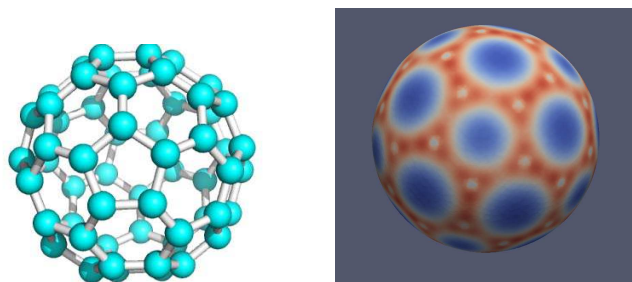


Fig. 5.5:  $C_{60}$ : configuration and charge density on a sphere.

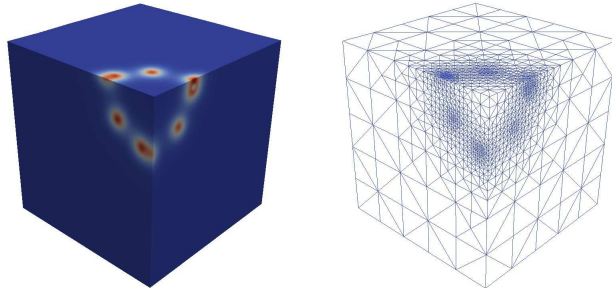


Fig. 5.6:  $C_{60}$ : charge density and mesh on an interior cross-section.

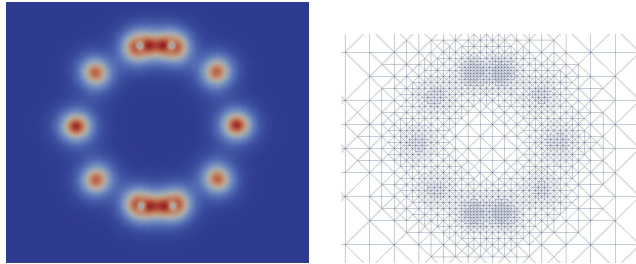


Fig. 5.7:  $C_{60}$ : charge density and mesh on plane  $z = 0$ .

The convergence curve of the ground state energy approximations is shown in the right of Figure 5.8, from which we observe a convergence to  $-342.722$  *a.u.*, which is very close to the reference energy. The convergence curve of the a posteriori error estimator obtained by the quadratic finite element is shown in the left of Figure 5.8, from which we see that it reaches the optimal convergence rate.

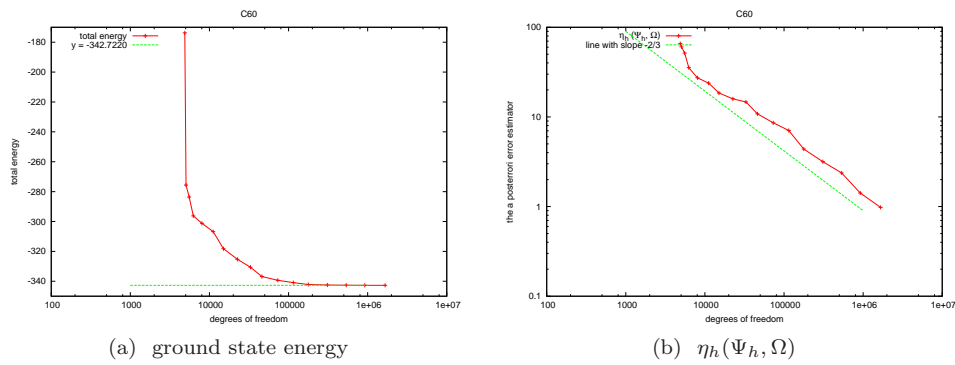


Fig. 5.8: The convergence curves of the ground state energy and  $\eta_h(\Psi_h, \Omega)$ .

**6. Concluding remarks.** In this paper, we have studied the AFE approximations of Kohn-Sham models. We have obtained the convergence and quasi-optimal complexity of the AFE approximations. We have also carried out some typical numer-

ical simulations that not only support our theory, but also show the robustness and efficiency of the adaptive finite element method in electronic structure calculations.

In our analysis of convergence rate and complexity of AFE approximations, for convenience, we have assumed that the numerical integration was exact and the nonlinear algebraic eigenvalue problem was exactly solved. Indeed, the same conclusion can be expected when the error resulting from the inexact solving of the nonlinear algebraic eigenvalue problem and the error coming from the inexact numerical integration are taken into account.

Suppose that  $(\Lambda, \Phi) \in \Theta$ , the associated exact solution over mesh  $\mathcal{T}_h$  is  $(\Lambda_h, \Phi_h)$ , and the inexact numerical solution is  $(\hat{\Lambda}_h, \hat{\Phi}_h)$ . If the numerical errors resulting from the solution of (nonlinear) algebraic system and the numerical integration are small enough, say, satisfy

$$\|\Phi_h - \hat{\Phi}_h\|_{1,\Omega}^2 + |\Lambda_h - \hat{\Lambda}_h| \lesssim \tilde{r}(h_0)\eta_h^2(\hat{\Phi}_h, \Omega)$$

with  $\tilde{r}(h_0) \ll 1$  for  $h_0 \ll 1$ , then we have from the following triangle inequality

$$\begin{aligned} \|\Phi - \hat{\Phi}_h\|_{1,\Omega} &\leq \|\Phi - \Phi_h\|_{1,\Omega} + \|\Phi_h - \hat{\Phi}_h\|_{1,\Omega}, \\ |\Lambda - \hat{\Lambda}_h| &\leq |\Lambda - \Lambda_h| + |\Lambda_h - \hat{\Lambda}_h|, \end{aligned}$$

and the similar perturbation arguments that the same convergence rate and quasi-optimal complexity can be derived.

Finally, we point out that, in this paper, we have not given the convergence rate and complexity for the AFE approximations for the Lagrange multipliers  $\Lambda$ . Indeed, the related optimal results for Lagrange multipliers are not so obvious, and we need do some more detailed analysis, which increase the length of this paper. We will report elsewhere.

**Appendix: A boundary value problem.** In this appendix, we shall provide some basic results for the AFE approximations of a model problem that was used in our previous analysis. Consider a homogeneous boundary value problem:

$$\begin{cases} \mathcal{L}\Phi = F & \text{in } \Omega, \\ \Phi = 0 & \text{on } \partial\Omega, \end{cases} \quad (\text{A.1})$$

where  $F = (f_i)_{i=1}^N \in (L^2(\Omega))^N$ . Note that (A.1) is equal to: Find  $\Phi \in \mathcal{H}$  such that

$$a(\Phi, \Gamma) = (F, \Gamma) \quad \forall \Gamma \in \mathcal{H}. \quad (\text{A.2})$$

A standard finite element scheme for (A.2) is: Find  $\Phi_h \in V_h$  satisfying

$$a(\Phi_h, \Gamma) = (F, \Gamma) \quad \forall \Gamma \in V_h. \quad (\text{A.3})$$

Let  $\mathbb{T}$  denote the class of all conforming refinements by bisections of  $\mathcal{T}_0$ . For  $\mathcal{T}_h \in \mathbb{T}$  and any  $\Gamma = (\gamma_i)_{i=1}^N \in V_h$ , we define the element residual  $\tilde{\mathcal{R}}_\tau(\Gamma)$  and the jump  $\tilde{J}_e(\Gamma)$  by

$$\begin{aligned} \tilde{\mathcal{R}}_\tau(\Gamma) &= \left( f_i + \frac{1}{2} \Delta \gamma_i \right)_{i=1}^N & \text{in } \tau \in \mathcal{T}_h, \\ \tilde{J}_e(\Gamma) &= \left( \frac{1}{2} \nabla \gamma_i|_{\tau_1} \cdot \vec{n}_1 + \frac{1}{2} \nabla \gamma_i|_{\tau_2} \cdot \vec{n}_2 \right)_{i=1}^N & \text{on } e \in \mathcal{E}_h, \end{aligned} \quad (\text{A.4})$$

where  $e$  is the common face of elements  $\tau_1$  and  $\tau_2$  with unit outward normals  $\vec{n}_1$  and  $\vec{n}_2$ , respectively. For  $\tau \in \mathcal{T}_h$ , we define the local error indicator  $\tilde{\eta}_h(\Gamma, \tau)$  by

$$\tilde{\eta}_h^2(\Gamma, \tau) = h_\tau^2 \|\tilde{\mathcal{R}}_\tau(\Gamma)\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|\tilde{J}_e(\Gamma)\|_{0,e}^2 \quad (\text{A.5})$$

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

$$\widetilde{\text{osc}}_h(\Gamma, \tau) = h_\tau \|\tilde{\mathcal{R}}_\tau(\Gamma) - \overline{\tilde{\mathcal{R}}_\tau(\Gamma)}\|_{0,\tau}. \quad (\text{A.6})$$

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

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

respectively. We see that a similar a posteriori error estimate to that for Poisson equation can be expected for (A.1) (c.f. [41, 42, 60]).

**THEOREM A.1.** *Let  $\Phi \in \mathcal{H}$  be the solution of (A.2) and  $\Phi_h \in V_h$  be the solution of (A.3). Then there exist constants  $\tilde{C}_1, \tilde{C}_2$  and  $\tilde{C}_3 > 0$  depending only on  $c_a$  in (4.1) and  $\gamma^*$  in (2.11) such that*

$$\|\Phi - \Phi_h\|_{1,\Omega}^2 \leq \tilde{C}_1 \tilde{\eta}_h^2(\Phi_h, \Omega), \quad (\text{A.7})$$

$$\tilde{C}_2 \tilde{\eta}_h^2(\Phi_h, \Omega) \leq \|\Phi - \Phi_h\|_{1,\Omega}^2 + \tilde{C}_3 \widetilde{\text{osc}}_h^2(\Phi_h, \Omega). \quad (\text{A.8})$$

An AFE algorithm for (A.2) is designed as follows (c.f. [12]):

ALGORITHM A.1.

1. Pick a given mesh  $\mathcal{T}_0$ , and let  $k = 0$ .
2. Solve (A.3) on  $\mathcal{T}_k$  to get discrete solution  $\Phi_k$ .
3. Compute local error indicators  $\tilde{\eta}_k(\Phi_k, \tau)$  for all  $\tau \in \mathcal{T}_k$ .
4. Construct  $\mathcal{M}_k \subset \mathcal{T}_k$  by Dörfler Strategy and parameter  $\theta$ .
5. Refine  $\mathcal{T}_k$  to get a new conforming mesh  $\mathcal{T}_{k+1}$ .
6. Let  $k = k + 1$  and go to 2.

Using the similar arguments to those for scalar linear elliptic boundary value problem (see, e.g. [12]), we have the following result for Algorithm A.1.

**THEOREM A.2.** *If  $\{\Phi_k\}_{k \in \mathbb{N}_0}$  is a sequence of finite element solutions produced by Algorithm A.1, then there exist constants  $\tilde{\gamma} > 0$  and  $\tilde{\xi} \in (0, 1)$  depending only on the shape regularity  $\gamma^*$  and the marking parameter  $\theta$ , such that for any two consecutive iterations*

$$\|\Phi - \Phi_{k+1}\|_{1,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_{k+1}^2(\Phi_{k+1}, \Omega) \leq \tilde{\xi}^2 (\|\Phi - \Phi_k\|_{1,\Omega}^2 + \tilde{\gamma} \tilde{\eta}_k^2(\Phi_k, \Omega)).$$

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

$$\tilde{\gamma} = \frac{1}{(1 + \delta^{-1}) \tilde{C}_*^2} \quad (\text{A.9})$$

with  $\tilde{C}_* > 0$  depending on the regularity constant  $\gamma^*$  and  $\delta \in (0, 1)$ .

For the distance between two nested solutions of (A.3), we have (c.f. [12])

THEOREM A.3. Let  $\Phi_H \in V_H$  and  $\Phi_h \in V_h$  be solutions of (A.3) respectively. If  $\mathcal{T}_h$  is a refinement of  $\mathcal{T}_H$  by marked element  $\mathcal{M}_H$  and refined elements  $\mathcal{R} = \mathcal{R}_{\mathcal{T}_H \rightarrow \mathcal{T}_h}$ , then

$$\|\Phi_H - \Phi_h\|_{1,\Omega}^2 \leq \tilde{C}_1 \sum_{\tau \in \mathcal{R}} \tilde{\eta}_H^2(\Phi_H, \tau).$$

#### REFERENCES

- [1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [2] S. Agmon, *Lectures on the Exponential Decay of Solutions of Second-Order Elliptic Operators*, Princeton University Press, Princeton, 1981.
- [3] A. Anantharaman and E. Cancès, *Existence of minimizers for Kohn-Sham models in quantum chemistry*, Ann. I. H. Poincaré-AN, **26** (2009), pp. 2425-2455.
- [4] I. Babuska and M. Vogelius, *Feedback and adaptive finite element solution of one-dimensional boundary value problems*, Numer. Math., **44** (1984), pp. 75-102.
- [5] E. Bänsch and K. Siebert, *A Posteriori Error Estimation for Nonlinear Problems by Duality Techniques*, Albert-Ludwigs-Univ., Math. Fak., 1995.
- [6] T.L. Beck, *Real-space mesh techniques in density-functional theory*, Rev. Mod. Phys., **72** (2000), pp. 1041-1080.
- [7] A.D. Becke, *A new inhomogeneity parameter in density-functional theory*, J. Phys. Chem., **109** (1998), pp. 2092-2098.
- [8] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica, **10** (2001), pp. 1-102.
- [9] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., **97** (2004), pp. 219-268.
- [10] E.J. Bylaska, M. Holst, and J.H. Weare, *Adaptive finite element method for solving the exact Kohn-Sham equation of density functional theory*, J. Chem. Theory Comput., **5** (2009), pp. 937-948.
- [11] E. Cancès, R. Chakir, and Y. Maday, *Numerical analysis of the planewave discretization of some orbital-free and Kohn-Sham models*, M2AN, **46** (2012), pp. 341-388.
- [12] 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.
- [13] H. Chen, X. Gong, L. He, Z. Yang, and A. Zhou, *Numerical analysis of finite dimensional approximations of Kohn-Sham equations*, Adv. Comput. Math., **38** (2013), pp. 225-256.
- [14] H. Chen, X. Gong, L. He, and A. Zhou, *Adaptive finite element approximations for a class of nonlinear eigenvalue problems in quantum physics*, Adv., Appl., Math., Mech., **3** (2011), pp. 493-518.
- [15] H. Chen, L. He, and A. Zhou, *Finite element approximations of nonlinear eigenvalue problems in quantum physics*, Comput. Methods Appl. Mech. Engrg., **200** (2011), pp. 1846-1865.
- [16] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.
- [17] X. Dai, *Adaptive and Localization Based Finite Element Discretizations for the First-Principles Electronic Structure Calculations*, Ph.D. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 2008.
- [18] X. Dai, X. Gong, Z. Yang, D. Zhang, and A. Zhou, *Finite volume discretizations for eigenvalue problems with applications to electronic structure calculations*, Multiscale Model. Simul., **9** (2011), pp. 208-240.
- [19] X. Dai, L. He, and A. Zhou, *Convergence rate and quasi-optimal complexity of adaptive finite element computations for multiple eigenvalues*, arXiv:1210.1846, 2012.
- [20] X. Dai, L. Shen, and A. Zhou, *A local computational scheme for higher order finite element eigenvalue approximations*, Inter. J. Numer. Anal. Model., **5** (2008), pp. 570-589.
- [21] X. Dai, J. Xu, and A. Zhou, *Convergence and optimal complexity of adaptive finite element eigenvalue computations*, Numer. Math., **110** (2008), pp. 313-355.
- [22] X. Dai and A. Zhou, *Three-scale finite element discretizations for quantum eigenvalue problems*, SIAM J. Numer. Anal., **46** (2008), pp. 295-324.

- [23] W. Dörfler, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., **33** (1996), pp. 1106-1124.
- [24] R.G. Durán, C. Padra, and R. Rodríguez, *A posteriori error estimates for the finite element approximation of eigenvalue problems*, Math. Mod. Meth. Appl. Sci., **13** (2003), pp. 1219-1229.
- [25] A. Edelman, T.A. Arias, and S.T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. appl., **20** (1998), pp. 303-353.
- [26] J. Fang, X. Gao, and A. Zhou, *A Kohn-Sham equation solver based on hexahedral finite elements*, J. Comput. Phys., **231** (2012), pp. 3166-3180.
- [27] J.L. Fattebert, R.D. Hornung, and A.M. Wissink, *Finite element approach for density functional theory calculations on locally refined meshes*, J. Comput. Phys., **223** (2007), pp. 759-773.
- [28] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and T. Ø. Sørensen, *Analytic structure of many-body Coulombic wave functions*, Comm. Math. Phys., **289** (2009), pp. 291-310.
- [29] E.M. Garau and P. Morin, *Convergence and quasi-optimality of adaptive FEM for Steklov eigenvalue problems*, IMA J. Numer. Anal., **31** (2011), pp. 914-946.
- [30] E.M. Garau, P. Morin, and C. Zuppa, *Convergence of adaptive finite element methods for eigenvalue problems*, M<sup>3</sup>AS, **19** (2009), pp. 721-747.
- [31] S. Giani and I. G. Graham, *A convergent adaptive method for elliptic eigenvalue problems*, SIAM J. Numer. Anal., **47** (2009), pp. 1067-1091.
- [32] X. Gong, L. Shen, D. Zhang, and A. Zhou, *Finite element approximations for Schrödinger equations with applications to electronic structure computations*, J. Comput. Math., **23** (2008), pp. 310-327.
- [33] L. He and A. Zhou, *Convergence and complexity of adaptive finite element methods for elliptic partial differential equations*, Inter. J. Numer. Anal. Model., **8** (2011), pp. 615-640.
- [34] V. Heuveline and R. Rannacher, *A posteriori error control for finite element approximations of elliptic eigenvalue problems*, Adv. Comput. Math., **15** (2001), pp. 107-138.
- [35] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev. B, **136** (1964), pp. 864-871.
- [36] W. Kohn and L.J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. A, **140** (1965), pp. 1133-1138.
- [37] M.G. Larson, *A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems*, SIAM J. Numer. Anal., **38** (2000), pp. 608-625.
- [38] C. Lee, W. Yang, and R.G. Parr, *Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, **37** (1988), pp. 785-789.
- [39] D. Mao, L. Shen, and A. Zhou, *Adaptive finite element algorithms for eigenvalue problems based on local averaging type a posteriori error estimates*, Adv. Comput. Math., **25** (2006), pp. 135-160.
- [40] R.M. Martin, *Electronic Structure: Basic Theory and Practical Method*, Cambridge University Press, Cambridge, 2004.
- [41] 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.
- [42] P. Morin, R.H. Nochetto, and K. Siebert, *Convergence of adaptive finite element methods*, SIAM Review, **44** (2002), pp. 631-658.
- [43] P. Morin, K.G. Siebert, and A. Veiser, *A basic convergence result for conforming adaptive finite elements*, Math. Models Methods Appl. Sci., **18** (2008), pp. 707-737.
- [44] P. Motamarri, M.R. Nowak, K. Leiter, J. Knap, and V. Gavini, *Higher-order adaptive finite-element methods for Kohn-Sham density functional theory*, J. Comput. Phys., **253** (2013), pp. 308-343.
- [45] R.G. Parr and W.T. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, Clarendon Press, Oxford, 1994.
- [46] J.E. Pask, B.M. Klein, P.A. Sterne, and C.Y. Fong, *Finite-element methods in electronic-structure theory*, Comput. Phys. Commun., **135** (2001), pp. 1-34.
- [47] J. Pask and P. Sterne, *Finite element methods in ab initio electronic structure calculations*, Modelling Simul. Mater. Sci. Eng., **13** (2005), pp. R71-R96.
- [48] J.P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, Phys. Rev. Lett., **77** (1996), pp. 3865-3868.

- [49] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. 2, Academic Press, New York, 1975.
- [50] R. Schneider, T. Rohwedder, A. Neelov, and J. Blauert, *Direct minimization for calculating invariant subspaces in density functional computations of the electronic structure*, J. Comput. Math., **27** (2009), pp. 360-387.
- [51] L. Shen, *Parallel Adaptive Finite Element Algorithms for Electronic Structure Computing based on Density Functional Theory*, Ph.D. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 2005.
- [52] L. Shen and A. Zhou, *A defect correction scheme for finite element eigenvalues with applications to quantum chemistry*, SIAM J. Sci. Comput., **28** (2006), pp. 321-338.
- [53] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., **7** (2007), pp. 245-269.
- [54] R. Stevenson, *The completion of locally refined simplicial partitions created by bisection*, Math. Comput., **77** (2008), pp. 227-241.
- [55] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, and M. Ortiz, *Non-periodic finite-element formulation of Kohn-Sham density functional theory*, J. Mech. Phys. Solids, **58** (2010), pp. 256-280.
- [56] T. Torsti, T. Eirola, J. Enkovaara, T. Hakala, P. Havu, V. Havu, T. Hoynalanmaa, J. Ignatius, M. Lyly, I. Makkonen, T.T. Rantala, J. Ruokolainen, K. Ruotsalainen, E. Rasanen, H. Saarikoski, and M.J. Puska, *Three real-space discretization techniques in electronic structure calculations*, Physica Status Solidi B, **243** (2006), pp. 1016-1053.
- [57] E. Tsuchida and M. Tsukada, *Electronic-structure calculations based on the finite-element method*, Phys. Rev. B, **52** (1995), pp. 5573-5578.
- [58] E. Tsuchida and M. Tsukada, *Adaptive finite-element method for electronic-structure calculations*, Phys. Rev. B, **54** (1996), pp. 7602-7605.
- [59] E. Tsuchida and M. Tsukada, *Large-scale electronic-structure calculations based on the adaptive finite-element method*, J. Phys. Soc. Jpn., **67** (1998), pp. 3844-3858.
- [60] R. Verfürth, *A Review of A Posteriori Error Estimates and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, New York, 1996.
- [61] Z. Yang, *Finite Volume Discretization Based First-Principles Electronic Structure Calculations*, Ph.D. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 2011.
- [62] H. Yserentant, *Regularity and Approximability of Electronic Wave Functions*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2010.
- [63] D. Zhang, *Applications of Finite Element Methods in Electronic Structure Calculations*, Ph.D. Thesis, Fudan University, 2007.
- [64] D. Zhang, L. Shen, A. Zhou, and X. Gong, *Finite element method for solving Kohn-Sham equations based on self-adaptive tetrahedral mesh*, Phy. Lett. A, **372** (2008), pp. 5071-5076.
- [65] D. Zhang, A. Zhou, and X. Gong, *Parallel mesh refinement of higher order finite elements for electronic structure calculations*, Commun. Comput. Phys., **4** (2008), pp. 1086-1105.
- [66] A. Zhou, *An analysis for finite dimensional approximations for the ground state solution of Bose-Einstein condensates*, Nonlinearity, **17** (2004), pp. 541-550.
- [67] A. Zhou, *Finite dimensional approximations for the electronic ground state solution of a molecular system*, Math. Meth. Appl. Sci., **30** (2007), pp. 429-447.
- [68] PHG, <http://lsec.cc.ac.cn/phg/>.