

NUMERICAL ANALYSIS OF AUGMENTED PLANE WAVE METHODS FOR FULL-POTENTIAL ELECTRONIC STRUCTURE CALCULATIONS *

HUAJIE CHEN¹ AND REINHOLD SCHNEIDER¹

Abstract. This paper investigates the augmented plane wave methods which are widely used in full-potential electronic structure calculations. These methods introduce basis functions that describe different regions using different discretization schemes. We construct a nonconforming method based on this idea and present an *a priori* error analysis for both linear Schrödinger type equations and nonlinear Kohn–Sham equations. Some numerical experiments are presented to support our theory.

Mathematics Subject Classification. 65N15, 65N25, 35P30, 81Q05.

Received February 27, 2012. Revised June 11, 2014.
Published online 8 April 2015.

1. INTRODUCTION

Electronic structure theory describes the energies and distributions of electrons, which plays a fundamental role in many different fields: materials science, biochemistry, solid-state physics, and surface physics. The most widely used technique, which achieves so far the best compromise between accuracy and efficiency, is Kohn–Sham (KS) density functional theory (DFT) [28, 32]. For a system consisting of M nuclei (located at $\mathbf{R}_k \in \mathbb{R}^3$ with charge $Z_k \in \mathbb{N}$, $k = 1, \dots, M$) and N electrons, the KS-DFT model with local density approximation (LDA) gives rise to the following KS equations

$$H_{\Phi} \phi_i = \lambda_i \phi_i, \quad i = 1, \dots, N, \quad (1.1)$$

where the KS orbitals $\Phi = \{\phi_1, \dots, \phi_N\} \in (H^1(\mathbb{R}^3))^N$ satisfy the orthonormality conditions $\int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}$, and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ are the lowest N eigenvalues of H_{Φ} . The Hamiltonian H_{Φ} in (1.1) is given by

$$H_{\Phi} = -\frac{1}{2} \Delta + V_{\text{eff}}(\rho_{\Phi}(\mathbf{r})) \quad \text{with} \quad V_{\text{eff}}(\rho_{\Phi}(\mathbf{r})) = v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\rho_{\Phi}(\mathbf{r})) + v_{\text{xc}}(\rho_{\Phi}(\mathbf{r})), \quad (1.2)$$

Keywords and phrases. Kohn–Sham density functional theory, augmented plane wave methods, nonconforming, *a priori* error estimate.

* *Financial support from the Alexander von Humboldt Foundation under grant CHN 1138663 STP.*

¹ Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany.
hjchen@lsec.cc.ac.cn; schneidr@math.tu-berlin.de

where $\rho_{\Phi}(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2$ is the electron density, $v_{\text{eff}}(\rho_{\Phi}(\mathbf{r}))$ represents the effective potential, $v_{\text{ext}}(\mathbf{r}) = -\sum_{k=1}^M \frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|}$ is the electrostatic potential generated by the nuclei, $v_{\text{H}}(\rho_{\Phi}(\mathbf{r})) = \int_{\mathbb{R}^3} \frac{\rho_{\Phi}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$ is the Coulomb potential representing interactions between the electrons, and $v_{\text{xc}}(\rho_{\Phi}(\mathbf{r}))$ is the exchange-correlation potential. A self-consistent field (SCF) iteration algorithm is commonly resorted to for this nonlinear problem. In each iteration of the algorithm, a Hamiltonian $H_{\tilde{\Phi}}$ is constructed from a trial electronic state $\tilde{\Phi}$, and a linear eigenvalue problem is then solved to obtain the low-lying eigenvalues. In this paper, we will focus on the periodic versions of the KS models. In the periodic setting, the simulation domain is no longer \mathbb{R}^3 , but a unit cell Ω of some periodic lattice of \mathbb{R}^3 . In the periodic KS framework, periodic boundary conditions are imposed to the KS orbitals at the boundary of the simulation cell Ω , which is the standard method to compute condensed phase properties with a limited number of atoms in the simulation cell.

In studies of the electronic structure of periodic solids, it is nature to use plane wave basis functions which correspond to Bloch functions labeled by the k -vector of the first Brillouin zone. A pseudopotential approximation (see, *e.g.* [39]) is always combined with the plane wave method, since in full-potential calculations, an extremely large number of plane waves are needed to describe the cusps [24, 27] and the rapid oscillations of the electron density. The pseudopotential approximations remove the core electrons, replace them and the strong ionic potential by a smooth pseudopotential that acts on a set of pseudo wave functions rather than the true valence wave functions, and hence allow the wave functions to be expanded using a much smaller basis set. Although the pseudopotentials give satisfactory results in most cases, they sometimes fail, and to our best knowledge, a mathematical analysis of the pseudopotential approximations is still lacking. The full-potential calculations can be regarded as a benchmark for the pseudopotential calculations. Moreover, the core electrons have to be considered sometimes and are responsible for many properties. Therefore, the full-potential calculations are necessary in practice, where more efficient basis sets than plane waves shall be used.

One of the most commonly used methods in full-potential calculations is the augmented plane wave (APW) method [39, 45], which is among the most accurate methods of performing electronic structure calculations for crystals. The APW method was originally proposed by Slater [47] in 1937. In spite of its demanding computational cost due to the energy dependency, it has been widely and successfully used (see *e.g.* [20]). Several improvements of the basis set were tried to get rid of the energy dependency, the first really successful one was the linearization scheme introduced by Andersen [3] in 1975, leading to the linearized augmented plane wave (LAPW) method [31]. It was further developed by including local atomic orbitals (LAPW+lo) to have enough variational flexibility in the radial basis functions [38, 44, 46]. Several widely used quantum chemistry and solid-state physics softwares are based on these methods such as *Exciting* [54], *FLEUR* [53], and *WIEN2k* [55].

In the APW method, the unit cell Ω is partitioned into two types of regions (the so-called ‘‘muffin-tin’’ division [39], see Fig. 1): (i) spheres \mathcal{C}_i centered at atomic sites \mathbf{R}_i with radius R_i , in which the effective potential is assumed to be symmetric, *i.e.* $V_{\text{eff}}(\rho_{\Phi}(\mathbf{r})) = V(|\mathbf{r} - \mathbf{R}_i|)$; (ii) the remaining interstitial region \mathcal{D} . The basis functions consist of augmentations of plane waves as follows

$$|\Omega|^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{r}} \rightarrow \begin{cases} |\Omega|^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{in } \mathcal{D}, \\ \sum_{lm} \alpha_{lm}^{\mathbf{k}} \chi_l(r_i, \varepsilon) Y_{lm}(\hat{\mathbf{r}}_i) & \text{in } \mathcal{C}_i, \end{cases} \quad (1.3)$$

where $\mathbf{r}_i = \mathbf{r} - \mathbf{R}_i$, $r = |\mathbf{r}|$, $\hat{\mathbf{r}} = \mathbf{r}/r$, $Y_{lm}(\hat{\mathbf{r}})$ denotes the spherical harmonic functions on S^2 , $\chi_l(r, \varepsilon)$ is the solution of the radial Schrödinger equation at energy parameter ε

$$-\frac{1}{2r^2} \frac{d}{dr} \left(r^2 \frac{d\chi_l}{dr} \right) + \left(\frac{l(l+1)}{2r^2} + V(r) - \varepsilon \right) \chi_l = 0, \quad (1.4)$$

and the coefficients $\alpha_{lm}^{\mathbf{k}}$ are determined such that each angular component of the basis function is matched through the spherical surface (see Sect. 3). The philosophy of the APW method is a procedure for solving the

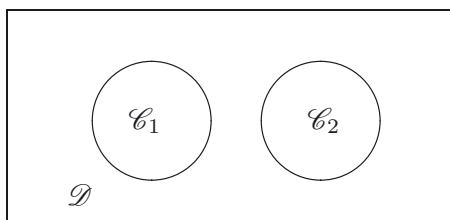


FIGURE 1. The “muffin-tin” division of the unit cell Ω into spheres \mathcal{C}_i centered at atoms and interstitial region \mathcal{D} .

KS equations of a many-electron system by introducing a basis set that is in some ways the “best of both worlds”. The smoothly varying parts of the wave functions between the atoms are represented by plane waves, and the rapidly varying parts near the nuclei are represented by radial atomic functions time spherical harmonics inside the sphere around each nucleus.

The purpose of this paper is to construct a discretization method based on the idea of the (L)APW method in order to provide a numerical analysis for full-potential electronic structure calculations. As far as we know, there is no numerical analysis concerning the APW method in literature. We shall first consider linear Schrödinger type eigenvalue problems with an effective potential, which appear in each step of the SCF iterations. Afterwards we extend the analysis to the nonlinear KS equations under certain reasonable coercivity assumptions. Our analysis shows that, in principle, the (L)APW approach can provide super algebraic and nearly exponential convergence for full-potential calculations. This explains why the (L)APW codes are frequently used for sufficiently accurate electronic structure calculations from a numerical point of view.

Standard plane wave and finite element methods for numerically solving (1.1) are based on variational principles, which consist in constructing a finite dimensional subspace of the Sobolev space $H^1(\Omega)$. Such conforming methods have been extensively studied for linear problems (see, e.g. [5, 14]). In contrast, the (L)APW basis functions are not continuous on the spherical surface due to the truncation of angular momentums in (1.3), so the finite dimensional approximation space no longer lies in $H^1(\Omega)$. Therefore, one gets a nonconforming method by using the (L)APW basis functions. There are a lot of existing results on analysis of the nonconforming finite element methods for solving linear elliptic problems, see e.g. [17, 30, 36, 48]. The advantages of the nonconforming idea lie in: (i) coupling different variational discretizations so as to take profit of the efficiency of each of them; (ii) more flexible and economical adaptive procedures as the refinements can be made on the subdomain where it is needed. The nonconforming idea in the APW method is also highly related to the mortar methods, which match incompatible grids with a suitable variational operator ensuring an optimal transmission of information between adjacent subdomains (see, e.g. [8–10]).

The KS equations form a nonlinear integro-differential eigenvalue problem with multiple eigenvalues to be considered, for which the numerical analysis is a difficult task. To our best knowledge, there are only a handful of very recent works concerning this problem, see, Cancès *et al.* [13], Chen *et al.* [16] and Suryanarayana *et al.* [49], and none of these numerical analysis can be applied to the APW method. In this paper, we shall establish the convergence of the eigenpair approximations and obtain an *a priori* error estimate, using the techniques that are related to the arguments in [13, 16, 34]. The results in this paper deal with the *a priori* error analysis, while the *a posteriori* error analysis is even more challenging and shall be investigated in our future works.

An outline of this paper is as follows. In Section 2, we state the model problem and some regularity results. In Section 3, we construct a nonconforming method analogous to the (L)APW methods with a complete basis set and obtain an *a priori* error estimate for linear eigenvalue problems. Further, the (L)APW methods are investigated under this framework and a numerical analysis is given. In Section 4, we give an *a priori* error estimate for ground state solutions of the nonlinear KS equations. In Section 5, we present several numerical experiments to support our theoretical results. Finally, some concluding remarks are made in Section 6.

2. PRELIMINARY

Throughout this paper, we shall use C to denote a generic positive constant which may stand for different values at its different occurrences and is independent of the discretization parameters. For convenience, the symbol \lesssim will be used. The notation $A \lesssim B$ means that $A \leq CB$ for some constant C that is independent of the discretization parameters. For $\mathbf{r} \in \mathbb{R}^3$, we shall denote by $r = |\mathbf{r}|$ and $\hat{\mathbf{r}} = \mathbf{r}/r$. Moreover, we shall denote $\sum_{l=0}^\infty \sum_{m=-l}^l$ by \sum_{lm} , and $\sum_{l=0}^L \sum_{m=-l}^l$ by \sum_{lm}^L for simplicity.

Let $\Omega \subset \mathbb{R}^3$ be the simulation domain with periodic boundary conditions. We denote by \mathcal{R} the periodic lattice, and by \mathcal{R}^* the dual lattice. For simplicity, we assume that $\Omega = [-\frac{\ell}{2}, \frac{\ell}{2}]^3$ ($\ell > 0$), $\mathcal{R} = \ell\mathbb{Z}^3$, and $\mathcal{R}^* = \frac{2\pi}{\ell}\mathbb{Z}^3$. For $\mathbf{k} \in \mathcal{R}^*$, we denote by $e_{\mathbf{k}}(\mathbf{r}) = |\Omega|^{-1/2}e^{i\mathbf{k}\cdot\mathbf{r}}$ the plane wave with wave vector \mathbf{k} . The family $\{e_{\mathbf{k}}\}_{\mathbf{k} \in \mathcal{R}^*}$ forms an orthonormal basis set of

$$L^2_{\#}(\Omega) = \{u \in L^2_{\text{loc}}(\mathbb{R}^3) : u \text{ is } \mathcal{R}\text{-periodic}\}.$$

For all $u \in L^2_{\#}(\Omega)$, we have

$$u(\mathbf{r}) = \sum_{\mathbf{k} \in \mathcal{R}^*} \hat{u}_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r}) \quad \text{with} \quad \hat{u}_{\mathbf{k}} = |\Omega|^{-1/2} \int_{\Omega} u(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}.$$

We introduce the Sobolev spaces of \mathcal{R} -periodic functions

$$H^s_{\#}(\Omega) = \left\{ u(\mathbf{r}) = \sum_{\mathbf{k} \in \mathcal{R}^*} \hat{u}_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r}) : \sum_{\mathbf{k} \in \mathcal{R}^*} (1 + |\mathbf{k}|^2)^s |\hat{u}_{\mathbf{k}}|^2 < \infty \right\},$$

where $s \in \mathbb{R}_+$ and $k = |\mathbf{k}|$. For $K \in \mathbb{N}$, we define

$$\mathcal{V}_K = \left\{ v_K \in L^2_{\#}(\Omega) : v_K(\mathbf{r}) = \sum_{\mathbf{k} \in \mathcal{R}^*, |\mathbf{k}| \leq \frac{2\pi}{\ell} K} c_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r}) \right\}.$$

For all $s \in \mathbb{R}_+$ and each $v \in H^s_{\#}(\Omega)$, the best approximation of v in \mathcal{V}_K for H^r -norm ($r \leq s$) is

$$\Pi_K v = \sum_{\mathbf{k} \in \mathcal{R}^*, |\mathbf{k}| \leq \frac{2\pi}{\ell} K} \hat{v}_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r}).$$

The more regular v , the faster the convergence of this truncated series to v : for real numbers r and s with $0 < r \leq s$, we have

$$\|v - \Pi_K v\|_{H^r_{\#}(\Omega)} = \min_{v_K \in \mathcal{V}_K} \|v - v_K\|_{H^r_{\#}(\Omega)} \lesssim K^{r-s} \|v\|_{H^s_{\#}(\Omega)} \quad \forall v \in H^s_{\#}(\Omega). \tag{2.1}$$

As a model problem, we consider the following Schrödinger type linear eigenvalue problem, which can be viewed as a linearization of (1.1): find $\lambda \in \mathbb{R}$ and $u \in H^1_{\#}(\Omega)$ such that $\|u\|_{L^2(\Omega)} = 1$ and

$$a(u, v) = \lambda(u, v) \quad \forall v \in H^1_{\#}(\Omega), \tag{2.2}$$

where the bilinear form $a : H^1_{\#}(\Omega) \times H^1_{\#}(\Omega) \rightarrow \mathbb{R}$ is defined by:

$$a(u, v) = \frac{1}{2} \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} V_{\text{eff}} uv \tag{2.3}$$

with the effective potential V_{eff} being a smooth potential except at the positions of nuclei (see (2.7)).

It is sufficient for us to consider the system with a single nucleus located at the origin, the algorithms and analysis of which can be easily generalized to multi-nuclei systems. Throughout this paper, we shall denote by \mathcal{C} the sphere centered at the origin, by R the radius of the sphere, by Γ the spherical surface, and by \mathcal{D} the interstitial region.

Note that the plane wave approximations of (2.2) can not achieve as good convergence rate as (2.1) due to the cusps of electron density at the nuclear positions (see, e.g. [22–24]). In our analysis, we rely on the regularity result in weighted Sobolev space for Schrödinger type eigenvalue problems, which is developed in [21]. This type of analysis has been introduced to investigate singularities of boundary value problems in conical domains with corners and edges, we refer to [6, 19, 26] for more details. In our case the geometry is fairly simple, while the singular electrostatic potential generated by the nuclei fits perfectly in this treatment.

Let $\varrho \in L^2_{\#}(\Omega)$ be a continuous function such that $\varrho(\mathbf{r}) = r$ in the neighborhood of 0, $\varrho \in C^{\infty}_{\text{loc}}(\mathbb{R}^3 \setminus S)$ with $S = \ell\mathbb{Z}^3$ being the set of singular points of ϱ , and $\varrho(\mathbf{r}) > 0$ for $\mathbf{r} \in \mathbb{R}^3 \setminus S$. We define the k th weighted Sobolev space with index γ by

$$\mathcal{K}^{k,\gamma}(\Omega) = \left\{ u \in L^2_{\#}(\Omega) : \varrho^{|\alpha|-\gamma} \partial^{\alpha} u \in L^2_{\#}(\Omega) \quad \forall |\alpha| \leq k \right\}, \tag{2.4}$$

where $k \in \mathbb{N}$ and $\gamma \in \mathbb{R}$. Note that the difference between the Sobolev space $H^k(\Omega)$ and the weighted Sobolev space $\mathcal{K}^{k,\gamma}(\Omega)$ only lays in the introduction of weight function $\varrho^{|\alpha|-\gamma}$. Consider a subspace of $\mathcal{K}^{s,\gamma}(\Omega)$ consisting of functions with the asymptotic expansions

$$u(\mathbf{r}) \sim \sum_{j \in \mathbb{N}_0} c_j(\hat{\mathbf{r}}) r^j \quad \text{as } r \rightarrow 0, \tag{2.5}$$

where c_j belongs to the finite dimensional subspace $M_j = \text{span}\{Y_{lm}, 0 \leq l \leq j, |m| \leq l\} \subset C^{\infty}(S^2)$. We define the weighted Sobolev spaces with asymptotic type (2.5) as

$$\mathcal{K}^{k,\gamma}(\Omega) = \left\{ u \in \mathcal{K}^{k,\gamma}(\Omega) : u(\mathbf{r}) - \omega(r) \sum_{j=0}^N c_j(\hat{\mathbf{r}}) r^j \in \mathcal{K}^{k,\gamma+N}(\Omega), \quad \forall N \in \mathbb{N} \right\}, \tag{2.6}$$

where $\omega(r)$ is a smooth cut-off function, i.e. $\omega = 1$ near 0 and $\omega = 0$ outside some neighborhood of 0. More precisely, $u(\mathbf{r}) - \omega(r) \sum_{j=0}^N c_j(\hat{\mathbf{r}}) r^j$ in Definition (2.6) should be the periodic extension of the function defined on Ω by this expression.

Definition 2.1. A function u is called asymptotically well behaved if $u \in \mathcal{K}^{\infty,\gamma}(\Omega)$ for $\gamma < 3/2$.

We shall make the assumption on the effective potential that

$$V_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\rho(\mathbf{r})) + v_{\text{s}}(\mathbf{r}) \quad \text{with } v_{\text{s}} \in C^{\infty}_{\#}(\Omega), \tag{2.7}$$

where $v_{\text{ext}}(\mathbf{r})$ is a periodic function that equals $-Z/|\mathbf{r}|$ ($Z \in \mathbb{N}$) in the neighborhood of 0, and belongs to $C^{\infty}_{\text{loc}}(\mathbb{R}^3 \setminus S)$, ρ is an asymptotically well behaved function, and

$$C^{\infty}_{\#}(\Omega) = \{v \in C^{\infty}_{\text{loc}}(\mathbb{R}^3) : v \text{ is } \mathcal{R}\text{-periodic}\}.$$

Here, $v_{\text{H}}(\rho(\mathbf{r}))$ denotes the \mathcal{R} -periodic Coulomb potential generated by the \mathcal{R} -periodic electron density ρ :

$$v_{\text{H}}(\rho(\mathbf{r})) = 4\pi \sum_{\mathbf{k} \in \mathcal{R}^* \setminus \{0\}} |\mathbf{k}|^{-2} \hat{\rho}_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r}). \tag{2.8}$$

The following lemma concerning the regularity of the eigenfunctions of (2.2) is heavily used in our analysis, the proof of which is given in ([21], Thms. 1, 4 and Prop. 1).

Lemma 2.2. *If u is an eigenfunction of (2.2) with V_{eff} satisfying (2.7), then u is asymptotically well behaved.*

Lemma 2.3. *For any $s > 0$ and $k \in \mathbb{N}$, $k > s + \frac{3}{2}$, if $u \in \mathcal{K}^{k,\gamma}(\Omega)$ with $\gamma < 3/2$, then the restriction of u to \mathcal{C} satisfies $u|_{\mathcal{C}} \in H^s([0, R] \times S^2)$.*

Proof. For any $s > 0$, we take $k > s + \frac{3}{2}$, and express the function $u \in \mathcal{K}^{k,\gamma}(\Omega)$ as

$$u(\mathbf{r}) = \omega(r) \sum_{j=0}^m c_j(\hat{\mathbf{r}})r^j + \Phi_{m+1}(\mathbf{r}), \tag{2.9}$$

where $\Phi_{m+1} \in \mathcal{K}^{k,\beta}(\Omega)$ with $\beta < \frac{3}{2} + m$. For sufficiently large m , $\Phi_{m+1} \in \mathcal{K}^{k,k}(\Omega)$. Therefore, we have from the Definition (2.4) that

$$\sum_{|\alpha| \leq k} \int_{\mathcal{C}} \wp^{2|\alpha|-2k} |\partial^\alpha \Phi_{m+1}|^2 < \infty.$$

This implies $\Phi_{m+1} \in H^k(\mathcal{C})$. According to Sobolev’s lemma, we have $H^k(\mathcal{C}) \subset C^s(\mathcal{C})$ for $s < k - \frac{3}{2}$, and hence $\Phi_{m+1} \in H^s([0, R] \times S^2)$.

Note that the first part of (2.9) already belongs to $C^\infty([0, R] \times S^2)$, we obtain $u|_{\mathcal{C}} \in H^s([0, R] \times S^2)$, which completes the proof. \square

The following lemma will be used in our analysis, which states the relationship between two Sobolev norms.

Lemma 2.4. *If $v \in H^1(\mathcal{C}) \cap H^1([0, R] \times S^2)$, then there exists a constant C such that*

$$\|v\|_{H^1(\mathcal{C})} \leq C \|v\|_{H^1([0,R] \times S^2)}.$$

Proof. Since in spherical coordinates

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}, \tag{2.10}$$

where the last two terms multiplied by r^2 is Δ_{S^2} , i.e. the Laplace–Beltrami operator on S^2 . We have

$$\begin{aligned} \|v\|_{H^1(\mathcal{C})}^2 &= - \int_{\mathcal{C}} v \Delta v + \int_{\Gamma} v \frac{\partial v}{\partial r} \Big|_{r=R} + \int_{\mathcal{C}} v^2 \\ &= \int_0^R r^2 dr \int_{S^2} \left(v^2 + \left(\frac{\partial v}{\partial r} \right)^2 \right) - \int_0^R dr \int_{S^2} (v \Delta_{S^2} v) \\ &\leq R^2 \int_{S^2} \|v\|_{H^1([0,R])}^2 + \int_0^R \|v\|_{H^1(S^2)}^2 dr \\ &\leq C \|v\|_{H^1([0,R] \times S^2)}^2, \end{aligned}$$

where Green’s formula is used for the second equality. This completes the proof. \square

3. ERROR ESTIMATES FOR LINEAR SCHRÖDINGER TYPE EQUATIONS

In this section, we shall consider a nonconforming approximation of (2.2) and extend the analysis to the (L)APW methods. We introduce the broken Sobolev space

$$H_{\#\delta}(\Omega) = \{v \in L^2_{\#}(\Omega) : v|_{\mathcal{C}} \in H^1(\mathcal{C}), v|_{\mathcal{D}} \in H^1(\mathcal{D})\} \tag{3.1}$$

equipped with the broken Sobolev norm

$$\|v\|_{\delta} = \|v\|_{H^1(\mathcal{C})} + \|v\|_{H^1(\mathcal{D})}.$$

Define the subspace $H_{\#L}(\Omega) \subset H_{\#\delta}(\Omega)$ by

$$H_{\#L}(\Omega) = \{v \in H_{\#\delta}(\Omega) : b(v, \psi) = 0, \forall \psi \in M_L\}, \tag{3.2}$$

where $M_L = \text{span}\{Y_{lm}, 0 \leq l \leq L, |m| \leq l\}$ and the bilinear form $b : L^2_{\#}(\Omega) \times L^2(S^2) \rightarrow \mathbb{R}$ is given by

$$b(v, \psi) = \int_{\Gamma} \psi(v^+ - v^-)d\Gamma$$

with v^{\pm} being the traces of v taken from outside and inside the sphere, respectively.

Note that Definition 2.1 and Lemma 2.3 imply that an asymptotically well behaved function belongs to $H^s([0, R] \times S^2)$ in the sphere and $H^s(\mathcal{D})$ in the interstitial region for any $s > 0$. For simplicity of notation, we introduce the space

$$\tilde{H}_L^s(\Omega) = \{v \in H_{\#L}(\Omega) : v|_{\mathcal{D}} \in H^s(\mathcal{D}), v|_{\mathcal{C}} \in H^s([0, R] \times S^2)\}$$

with the induced norm

$$\|v\|_{\tilde{H}_L^s(\Omega)} = \|v\|_{H^s(\mathcal{D})} + \|v\|_{H^s([0, R] \times S^2)}.$$

3.1. A nonconforming method

Define the space of functions on \mathcal{D} expanded in plane waves

$$\mathcal{P}_K(\mathcal{D}) = \left\{ u \in H^1(\mathcal{D}) : u(\mathbf{r}) = \sum_{|\mathbf{k}| \leq \frac{2\pi}{L}K} c_{\mathbf{k}} e_{\mathbf{k}}(\mathbf{r})|_{\mathcal{D}} \right\}$$

and the space of functions on \mathcal{C} expanded in polynomial times spherical harmonics

$$\mathcal{B}_{NL}(\mathcal{C}) = \left\{ u \in H^1(\mathcal{C}) : u(\mathbf{r}) = \sum_{lm}^L \sum_{n=0}^N c_{nlm} \chi_n(r) Y_{lm}(\hat{\mathbf{r}}) \right\},$$

where $\{\chi_n\}_{n=0}^N$ forms a basis set of the space of polynomials on $[0, R]$ with degree at most N . Let

$$\mathcal{S}_{NL}^K(\Omega) = \{u \in L^2_{\#}(\Omega) : u|_{\mathcal{C}} \in \mathcal{B}_{NL}(\mathcal{C}) \text{ and } u|_{\mathcal{D}} \in \mathcal{P}_K(\mathcal{D})\} \tag{3.3}$$

and $\varrho = \min\{K, N, L\}$. We may assume throughout this paper that $\varrho \geq C \max\{K, N, L\}$.

Based on the above constructions, we have the nonconforming approximation space

$$V_{\varrho}(\Omega) = H_{\#L}(\Omega) \cap \mathcal{S}_{NL}^K(\Omega), \tag{3.4}$$

which does not lie in $H_{\#}^1(\Omega)$. The nonconforming approximation of (2.2) is: find $\lambda_{\varrho}(= \lambda_{NL}^K) \in \mathbb{R}$ and $u_{\varrho}(= u_{NL}^K) \in V_{\varrho}$ such that $\|u_{\varrho}\|_{L^2(\Omega)} = 1$ and

$$a_{\delta}(u_{\varrho}, v) = \lambda_{\varrho}(u_{\varrho}, v) \quad \forall v \in V_{\varrho}, \tag{3.5}$$

where the bilinear form $a_{\delta}(\cdot, \cdot) : H_{\#\delta}(\Omega) \times H_{\#\delta}(\Omega) \rightarrow \mathbb{R}$ is given by

$$a_{\delta}(u, v) = \frac{1}{2} \int_{\mathcal{C}} \nabla u \cdot \nabla v + \frac{1}{2} \int_{\mathcal{D}} \nabla u \cdot \nabla v + \int_{\Omega} V_{\text{eff}} uv. \tag{3.6}$$

Since $v_{\text{ext}} \in L_{\#}^2(\Omega)$ and $\int_{\Omega} v_{\text{H}}(\rho) uv \leq C_H \|\rho\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}$ (C_H is a constant, see [13], (3.17)), we have from (2.7) that for all $v \in H_{\#\delta}(\Omega)$,

$$\begin{aligned} a_{\delta}(v, v) &\geq \frac{1}{2} \|\nabla v\|_{L^2(\mathcal{C})}^2 + \frac{1}{2} \|\nabla v\|_{L^2(\mathcal{D})}^2 - \|v_{\text{ext}}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}^{1/2} \|v\|_{L^6(\Omega)}^{3/2} \\ &\quad - C_H \|\rho\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}^2 - \|v_s\|_{L^{\infty}(\Omega)} \|v\|_{L^2(\Omega)}^2 \\ &\geq C \|v\|_{\delta}^2 - \nu \|v\|_{L^2(\Omega)}^2, \end{aligned} \tag{3.7}$$

where C and ν are positive constants. Using a shift of the bilinear form $a_{\delta}^{\nu}(u, v) = a_{\delta}(u, v) + \nu(u, v)$, (3.5) is equivalent to

$$a_{\delta}^{\nu}(u_{\varrho}, v) = (\lambda_{\varrho} + \nu)(u_{\varrho}, v) \quad \forall v \in V_{\varrho}$$

with the same $(\lambda_{\varrho}, u_{\varrho})$. Therefore, we may assume afterwards without loss of generality that there exists $C > 0$, such that

$$a_{\delta}(v, v) \geq C \|v\|_{\delta}^2 \quad \forall v \in H_{\#\delta}(\Omega). \tag{3.8}$$

Note that (3.8) together with the fact $H_{\#}^1(\Omega) \subset H_{\#\delta}(\Omega)$ implies $a(v, v) \geq C \|v\|_{H_{\#}^1(\Omega)}$ for all $v \in H_{\#}^1(\Omega)$.

To obtain the *a priori* error estimate of the nonconforming approximations, we shall first construct a basis set of V_{ϱ} . Let $\{\chi_n\}_{n=0}^N$ be a basis set that spans the space of polynomials on $[0, R]$ with degree at most N and satisfies (see, e.g. Fig. 2)

$$\chi_n(R) = 0 \quad \text{for } n = 0, \dots, N-1, \quad \text{and} \quad \chi_N(r) = \frac{r}{R}. \tag{3.9}$$

Note that the subspace $\tilde{\mathcal{B}}_{NL} \subset \mathcal{B}_{NL}$ defined by

$$\tilde{\mathcal{B}}_{NL} = \left\{ u \in H_0^1(\mathcal{C}) : u(\mathbf{r}) = \sum_{lm}^L \sum_{n=0}^{N-1} c_{nlm} \chi_n(r) Y_{lm}(\hat{\mathbf{r}}) \right\}$$

is a finite dimensional subspace of $H_0^1(\mathcal{C})$.

Define

$$\omega_{\mathbf{k}}(\mathbf{r}) = \begin{cases} |\Omega|^{-\frac{1}{2}} e^{i\mathbf{k} \cdot \mathbf{r}} & \text{in } \mathcal{D}, \\ \sum_{lm}^L \beta_{lm}^{\mathbf{k}} \chi_N(r) Y_{lm}(\hat{\mathbf{r}}) & \text{in } \mathcal{C}, \end{cases}$$

where the coefficients

$$\beta_{lm}^{\mathbf{k}} = |\Omega|^{-1/2} 4\pi i^l j_l(kR) Y_{lm}^*(\hat{\mathbf{k}}) / \chi_N(R) \tag{3.10}$$

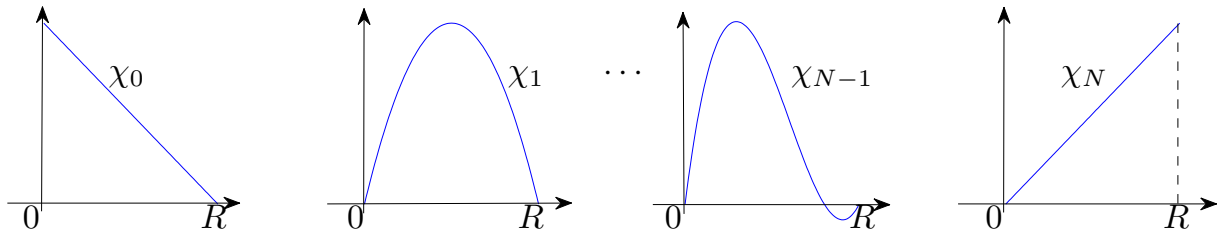


FIGURE 2. Schematic plots of radial basis functions $\chi_n(r)$ $n = 0, \dots, N$ satisfying (3.9).

are determined by the continuity constraint in (3.2) together with the scattering expansion [40]

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{lm} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}). \tag{3.11}$$

Set $\tilde{V}_\varrho = \tilde{\mathcal{B}}_{NL} \oplus \text{span}\{\omega_{\mathbf{k}}(\mathbf{r}) : |\mathbf{k}| \leq K\}$. It is obvious from the definition of $\omega_{\mathbf{k}}$ that $\tilde{V}_\varrho \subset V_\varrho$, which together with the fact

$$\dim(V_\varrho) = \dim(\mathcal{P}_K) + \dim(\mathcal{B}_{NL}) - \dim(M_L) = \dim(\mathcal{P}_K) + \dim(\tilde{\mathcal{B}}_{NL}) = \dim(\tilde{V}_\varrho)$$

implies $V_\varrho = \tilde{V}_\varrho$. Having a set of basis functions that span V_ϱ , we can derive the error for the best approximation in this finite dimensional space.

Lemma 3.1. *If $u \in \tilde{H}_L^s(\Omega)$ for $s > 3/2$, then there exists a constant C such that*

$$\inf_{v_\varrho \in V_\varrho} \|u - v_\varrho\|_\delta \leq C \varrho^{-(s-3/2)} \|u\|_{\tilde{H}_L^s(\Omega)}. \tag{3.12}$$

Proof. Split u into two parts $u = u_1 + u_2$ such that

$$u_1(\mathbf{r}) = \begin{cases} u(\hat{\mathbf{r}}R) \cdot \frac{r}{R} & \text{in } \mathcal{C} \\ u & \text{in } \mathcal{D} \end{cases} \quad \text{and} \quad u_2 = u - u_1 \in H_0^1(\mathcal{C}),$$

where $u(\hat{\mathbf{r}}R)$ is the trace of $u|_{\mathcal{C}}$ on Γ . We approximate these two parts separately.

To approximate u_1 , we shall first extend $u|_{\mathcal{D}}$ smoothly into the sphere \mathcal{C} . Note that u can be represented by

$$u(\mathbf{r}) = \sum_{lm} u_{lm}(r) Y_{lm}(\hat{\mathbf{r}})$$

around the sphere surface with $u_{lm}(r) = \int_{S^2} u(\mathbf{r}) Y_{lm}(\hat{\mathbf{r}})$. We can define

$$\tilde{u}(\mathbf{r}) = \begin{cases} u(\mathbf{r}) & \text{in } \mathcal{D}, \\ \sum_{lm} \varphi_{lm}(r) Y_{lm}(\hat{\mathbf{r}}) & \text{in } \mathcal{C}, \end{cases} \tag{3.13}$$

where $\varphi_{lm}(r) = \tau(r) \sum_{n=1}^{\bar{s}+1} c_n u_{lm}(R + \frac{1}{n}(R-r))$ with \bar{s} being the ceil integer of s , c_n being the coefficients such

that $\sum_{n=1}^{\bar{s}+1} (-\frac{1}{n})^k c_n = 1$ ($k = 0, 1, \dots, \bar{s}$), and $\tau \in C^\infty([0, R])$ satisfying $\tau = 0$ in $[0, \frac{R}{3}]$ and $\tau = 1$ in $[\frac{2R}{3}, R]$. We

observe by a direct calculation that $u \in H^s(\mathcal{D})$ leads to $\tilde{u} \in H^s(\Omega)$, and

$$\|\tilde{u}\|_{H^s(\Omega)} \leq \beta_s \|u\|_{H^s(\mathcal{D})}, \tag{3.14}$$

where the constant β_s is only related to s , R and $\|\tau\|_{H^s[0,R]}$. Let

$$\tilde{c}_{\mathbf{k}} = |\Omega|^{-1/2} \int_{\Omega} \tilde{u}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \quad \text{and} \quad v_1 = \sum_{|\mathbf{k}| \leq K} \tilde{c}_{\mathbf{k}} \omega_{\mathbf{k}}. \tag{3.15}$$

We have from (3.14) that

$$\begin{aligned} \|u_1 - v_1\|_{H^1(\mathcal{D})} &= \|\tilde{u} - \sum_{|\mathbf{k}| \leq K} \tilde{c}_{\mathbf{k}} e_{\mathbf{k}}\|_{H^1(\mathcal{D})} \leq \|\tilde{u} - \sum_{|\mathbf{k}| \leq K} \tilde{c}_{\mathbf{k}} e_{\mathbf{k}}\|_{H^1(\Omega)} \\ &\leq CK^{-(s-1)} \|\tilde{u}\|_{H^s(\Omega)} \leq C\beta_s K^{-(s-1)} \|u\|_{H^s(\mathcal{D})}. \end{aligned} \tag{3.16}$$

Using Lemma 2.4, the definition of $\omega_{\mathbf{k}}$, and the trace theorem, we have

$$\begin{aligned} \|u_1 - v_1\|_{H^1(\mathcal{C})} &\lesssim \|u_1 - v_1\|_{H^1([0,R] \times S^2)} \\ &= \|\chi_N(r)(u(\hat{\mathbf{r}}R) - v_1(\hat{\mathbf{r}}R))|_{\mathcal{C}}\|_{H^1([0,R] \times S^2)} \\ &\lesssim \|u - v_1|_{\mathcal{C}}\|_{H^1(\Gamma)} \\ &\lesssim \|v_1|_{\mathcal{C}} - v_1|_{\mathcal{D}}\|_{H^1(\Gamma)} + \|u - v_1|_{\mathcal{D}}\|_{H^1(\Gamma)} \\ &\lesssim L^{-(s-3/2)} \|v_1\|_{H^s(\mathcal{D})} + \|u - v_1\|_{H^{3/2}(\mathcal{D})}. \end{aligned} \tag{3.17}$$

Using the same argument as that in (3.16), we can derive

$$\|u - v_1\|_{H^{3/2}(\mathcal{D})} \lesssim K^{-(s-3/2)} \|u\|_{H^s(\mathcal{D})},$$

which together with (3.16) and (3.17) leads to

$$\|u_1 - v_1\|_{\delta} \leq C(L^{-(s-3/2)} + K^{-(s-3/2)}) \|u\|_{H^s(\mathcal{D})}. \tag{3.18}$$

For approximation of u_2 , we define the projection $P_N : H^1([0, R]) \rightarrow \Psi_N = \text{span}\{\chi_n, 0 \leq n \leq N - 1\}$ satisfying

$$(\nabla(v - P_N v), \nabla \psi) = 0 \quad \forall \psi \in H^1([0, R])$$

and the projection $P_L : L^2(S^2) \rightarrow M_L = \text{span}\{Y_{lm}, 0 \leq l \leq L, |m| \leq l\}$ by

$$P_L \varphi(\hat{\mathbf{r}}) = \sum_{lm}^L \check{\varphi}_{lm} Y_{lm}(\hat{\mathbf{r}}) \quad \text{with} \quad \check{\varphi}_{lm} = \int_{S^2} \varphi(\hat{\mathbf{r}}) Y_{lm}(\hat{\mathbf{r}}).$$

Define the projection $\mathbb{P}^{NL} : H^1([0, R] \times S^2) \rightarrow \Psi_N \times M_L$ by $\mathbb{P}^{NL} = P_N \circ P_L$. Using standard error estimates for polynomial and spherical harmonics approximations, we obtain that for any $u_2 \in H_0^1(\mathcal{C}) \cap H^s([0, R] \times S^2)$,

$$\|u_2 - \mathbb{P}^{NL} u_2\|_{H^1([0,R] \times S^2)} \leq C(L^{-(s-1)} + N^{-(s-1)}) \|u_2\|_{H^s([0,R] \times S^2)} \quad \forall s > 1. \tag{3.19}$$

Taking $v_2 = \mathbb{P}^{NL} u_2$, we have from Lemmas 2.4 and (3.19) that

$$\begin{aligned} \|u_2 - v_2\|_{H^1(\mathcal{C})} &\leq \|u_2 - v_2\|_{H^1([0,R] \times S^2)} \\ &\leq C(L^{-(s-3/2)} + N^{-(s-3/2)}) \|u_2\|_{H^{s-1/2}([0,R] \times S^2)} \quad \forall s > 3/2. \end{aligned} \tag{3.20}$$

Note that

$$\begin{aligned} \|u_2\|_{H^{s-1/2}([0,R] \times S^2)} &\leq \|u\|_{H^{s-1/2}([0,R] \times S^2)} + \|u(\hat{\mathbf{r}}R) \frac{r}{R}\|_{H^{s-1/2}([0,R] \times S^2)} \\ &\lesssim \|u\|_{H^{s-1/2}([0,R] \times S^2)} + \|u\|_{H^{s-1/2}(\Gamma)} \\ &\lesssim \|u\|_{H^s([0,R] \times S^2)} + \|u\|_{H^s(\mathcal{D})}. \end{aligned}$$

We can derive (3.12) by taking (3.18), (3.20) and the fact $v_1 + v_2 \in V_\varrho$ into account. This completes the proof. \square

Remark 3.2. We shall mention that the error estimate obtained in Lemma 3.1 may not be optimal, one would expect $\varrho^{-(s-1)}$ instead. The reason lies in that the approximation of u we construct in the proof is not a direct projection onto the finite dimensional space V_ϱ . Instead, we project $u|_{\mathcal{D}}$ onto \mathcal{P}_K , and the approximation of $u|_{\mathcal{E}}$ in the sphere is partially determined by the approximation outside the sphere through the surface Γ . This process may lead to a loss of accuracy.

Nevertheless, if the function u has sufficiently high regularity (for instance, if u is asymptotically well behaved, then s can be arbitrarily large by Lem. 2.2), then we may be satisfied with this convergence rate.

Define the operator $T : L^2_{\#}(\Omega) \rightarrow H^1_{\#}(\Omega)$ such that for $f \in L^2_{\#}(\Omega)$,

$$a(Tf, v) = (f, v) \quad \forall v \in H^1_{\#}(\Omega). \tag{3.21}$$

Then (2.2) is equivalent to the operator form $Tu = \lambda^{-1}u$. We also define the operator $T_\varrho : L^2_{\#}(\Omega) \rightarrow V_\varrho$ such that for $f \in L^2_{\#}(\Omega)$,

$$a_\delta(T_\varrho f, v) = (f, v) \quad \forall v \in V_\varrho. \tag{3.22}$$

We have that (3.5) is equivalent to $T_\varrho u_\varrho = \lambda_\varrho^{-1}u_\varrho$. Note that T and T_ϱ are well-defined due to (3.8). One can prove that T and T_ϱ are self-adjoint operators on $L^2_{\#}(\Omega)$ and satisfy the following lemma.

Lemma 3.3. *If $Tf \in \tilde{H}^s_L(\Omega)$ for $s > 3/2$ and $f \in L^2_{\#}(\Omega)$, then there exists a constant C such that*

$$\|Tf - T_\varrho f\|_\delta \leq C\varrho^{-(s-3/2)}\|Tf\|_{\tilde{H}^s_L(\Omega)}. \tag{3.23}$$

Proof. Let $u = Tf$ and $u_\varrho = T_\varrho f$. Processing in a standard way (see, e.g. [11, 12, 42]), we get

$$\|u - u_\varrho\|_\delta \leq C \left(\inf_{v_\varrho \in V_\varrho} \|u - v_\varrho\|_\delta + \sup_{v_\varrho \in V_\varrho} \frac{a_\delta(u - u_\varrho, v_\varrho)}{\|v_\varrho\|_\delta} \right) \tag{3.24}$$

and

$$a_\delta(u - u_\varrho, v_\varrho) = b(v_\varrho, \varphi - \psi_L) \quad \forall v_\varrho \in V_\varrho, \quad \forall \psi_L \in M_L$$

with $\varphi = \frac{\partial u}{\partial n}$, which together imply

$$\|u - u_\varrho\|_\delta \leq C \left(\inf_{v_\varrho \in V_\varrho} \|u - v_\varrho\|_\delta + \sup_{v_\varrho \in V_\varrho} \inf_{\psi_L \in M_L} \frac{b(v_\varrho, \varphi - \psi_L)}{\|v_\varrho\|_\delta} \right). \tag{3.25}$$

The first term in the right-hand side of (3.25) is the best approximation error, which has been obtained in Lemma 3.1. The second term is nothing else than the consistency error, which indicates the variational crime committed by the nonconforming discretization.

For the consistency error, we have

$$\begin{aligned} \inf_{\psi_L \in M_L} b(v_\varrho, \varphi - \psi_L) &= \inf_{\psi_L \in M_L} \int_\Gamma (v_\varrho^+ - v_\varrho^-)(\varphi - \psi_L) \\ &\leq \inf_{\psi_L \in M_L} \|v_\varrho^+ - v_\varrho^-\|_{H^{1/2}(\Gamma)} \|\varphi - \psi_L\|_{H^{-1/2}(\Gamma)} \\ &\leq CL^{-(s-1)} (\|v_\varrho^+\|_{H^1(\mathcal{E})} + \|v_\varrho^-\|_{H^1(\mathcal{D})}) \left\| \frac{\partial u}{\partial n} \right\|_{H^{s-3/2}(\Gamma)} \\ &\leq CL^{-(s-1)} \|v_\varrho\|_\delta \|u\|_{H^s(\mathcal{D})}, \end{aligned} \tag{3.26}$$

where the trace inequality is used.

Taking (3.12), (3.25) and (3.26) into account, we obtain (3.23), which completes the proof. \square

Denote by $\vartheta(T)$ the spectrum and $\rho(T)$ the resolvent set respectively of the solution operator T . For any $z \in \mathbb{C}$ in $\rho(T)$, we define the resolvent operator $R_z(T) = (z - T)^{-1}$. Let λ^{-1} be an eigenvalue of T and γ be a circle in the complex plane centered at λ^{-1} , which does not enclose any other point of $\vartheta(T)$. We define the operators \mathcal{E} and \mathcal{E}_ϱ by

$$\mathcal{E} = \mathcal{E}(\lambda^{-1}) = \frac{1}{2\pi i} \int_\gamma R_z(T) dz, \quad \mathcal{E}_\varrho = \mathcal{E}_\varrho(\lambda^{-1}) = \frac{1}{2\pi i} \int_\gamma R_z(T_\varrho) dz.$$

If K, N, L are sufficiently large, then \mathcal{E} and \mathcal{E}_ϱ are the spectral projectors (relative to λ^{-1}) of T and T_ϱ respectively.

To evaluate the distance between eigenspaces, we need a suitable notation. Let X and Y be two closed subspaces of $H_{\#L}(\Omega)$. We define (see, e.g. [41])

$$\hat{\delta}(X, Y) := \sup_{x \in X, \|x\|_{L^2(\Omega)}=1} \inf_{y \in Y, \|y\|_{L^2(\Omega)}=1} \|x - y\|_\delta$$

and denote the gap between X and Y by

$$\mathfrak{G}(X, Y) = \max\{\hat{\delta}(X, Y), \hat{\delta}(Y, X)\}.$$

The following theorem is the main result of this section, which states the convergence rate of the nonconforming eigenpair approximations.

Theorem 3.4. *Let λ be an eigenvalue of (2.2) with $\dim(\mathcal{R}(\mathcal{E})) = m$, where \mathcal{R} denotes the range. If (2.7) is satisfied and ϱ is sufficiently large, then there exist m eigenvalues $\lambda_{1,\varrho}, \dots, \lambda_{m,\varrho}$ of (3.5) such that for any $s > 3/2$,*

$$\sup_{1 \leq i \leq m} |\lambda - \lambda_{i,\varrho}| + \mathfrak{G}(\mathcal{R}(\mathcal{E}), \mathcal{R}(\mathcal{E}_\varrho)) \leq C_s \varrho^{-(s-3/2)}, \tag{3.27}$$

where C_s is a constant depending on s .

Proof. Similar to (2.10), we have

$$-\frac{1}{2}\Delta + V_{\text{eff}} = -\frac{1}{2r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} - 2r^2 V_{\text{eff}} \right]$$

in spherical coordinates. For operators of this form, one can apply the regularity results in ([19], p. 257, Thm. 9 and (8.137)). More precisely, for any $f \in L_{\#}^2(\Omega)$, we have $Tf \in \tilde{H}_L^2(\Omega)$ and

$$\|Tf\|_{\tilde{H}_L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}. \tag{3.28}$$

For any $f \in L_{\#}^2(\Omega)$, we have the following very rough L^2 -error estimate

$$\|(T - T_\varrho)f\|_{L^2(\Omega)} \leq \|Tf - T_\varrho f\|_\delta \lesssim \varrho^{-1/2} \|Tf\|_{\tilde{H}_L^2(\Omega)} \lesssim \varrho^{-1/2} \|f\|_{L^2(\Omega)}, \tag{3.29}$$

where Lemmas 3.3 and (3.28) are used. Therefore, we have

$$\lim_{\varrho \rightarrow \infty} \|T - T_\varrho\|_{\mathcal{L}(L_{\#}^2(\Omega), L_{\#}^2(\Omega))} \leq C \lim_{\varrho \rightarrow \infty} \varrho^{-1/2} = 0. \tag{3.30}$$

Using (3.30) and the result in Theorem 1 of [41], we have the convergence of the eigenvalues (*i.e.*, there exist m eigenvalues $\lambda_{1,\varrho}, \dots, \lambda_{m,\varrho}$ of (3.5) such that $\lim_{\varrho \rightarrow \infty} |\lambda - \lambda_{i,\varrho}| = 0, i = 1, \dots, m$) and

$$\delta(\mathcal{R}(\mathcal{E}), \mathcal{R}(\mathcal{E}_\varrho)) \lesssim \|T - T_\varrho\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))}. \tag{3.31}$$

It is now only necessary for us to estimate the right-hand side of (3.31). Using Lemma 3.3, the regularity results in Lemma 2.2, 2.3, and the fact $Tv = \lambda^{-1}v$ for $v \in \mathcal{R}(\mathcal{E})$, we have

$$\begin{aligned} \|T - T_\varrho\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))} &\leq \sup_{v \in \mathcal{R}(\mathcal{E}), \|v\|_{L^2(\Omega)}=1} \|Tv - T_\varrho v\|_\delta \\ &\leq C\varrho^{-(s-3/2)} \sup_{v \in \mathcal{R}(\mathcal{E}), \|v\|_{L^2(\Omega)}=1} \|v\|_{\tilde{H}_L^s(\Omega)} \\ &\leq C_s \varrho^{-(s-3/2)} \quad \forall s > \frac{3}{2}. \end{aligned} \tag{3.32}$$

This completes the estimate for $\delta(\mathcal{R}(\mathcal{E}), \mathcal{R}(\mathcal{E}_\varrho))$.

For eigenvalue estimate, we obtain by a simple calculation that

$$\lambda_\varrho - \lambda = a_\delta(u - u_\varrho, u - u_\varrho) - \lambda(u - u_\varrho, u - u_\varrho) + 2D_\varrho \tag{3.33}$$

with the consistency error

$$\begin{aligned} |D_\varrho| &= |a_\delta(u, u_\varrho) - \lambda(u, u_\varrho)| = \left| \int_\Gamma \frac{\partial u}{\partial n} (u_\varrho^+ - u_\varrho^-) \right| \\ &\leq \inf_{\psi_L \in M_L} \left\| \frac{\partial u}{\partial n} - \psi_L \right\|_{H^{-1/2}(\Gamma)} \|u_\varrho^+ - u_\varrho^-\|_{H^{1/2}(\Gamma)} \\ &\leq CL^{-(s-1)} \|u\|_{H^s(\mathcal{D})} \|u_\varrho\|_\delta, \end{aligned} \tag{3.34}$$

where similar argument as that in (3.26) is used.

For any eigenpair $(\lambda_{i,\varrho}, u_{i,\varrho})$ of (3.5) with $1 \leq i \leq m$, we can find an eigenfunction $u \in \mathcal{R}(\mathcal{E})$ and $\|u\|_{L^2(\Omega)} = 1$, such that $\|u_{i,\varrho} - u\|_\varrho \leq C\varrho^{-(s-3/2)}$. This together with (3.33) and (3.34) leads to $|\lambda - \lambda_{i,\varrho}| \leq C\varrho^{-(s-3/2)}$, which completes the proof of (3.27). \square

3.2. APW method

The APW method can be viewed as a modified method of the nonconforming scheme under the assumption that the effective potential is spherically symmetric inside the atomic sphere \mathcal{C} , say $V_{\text{eff}}(\mathbf{r}) = V(r)$ for $r \leq R$. It replaces the radial basis function $\chi_N(r)$ by $\chi_l(r, E)$, which is the regular solution of the following radial Schrödinger equation with parameter E :

$$-\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\chi_l}{dr} \right) + \left(\frac{l(l+1)}{r^2} + V(r) - E \right) \chi_l = 0 \quad \forall r \in [0, R]. \tag{3.35}$$

The approximation space is $V_{KL}^E = \text{span}\{\omega_{\mathbf{k}L}^E\}_{|\mathbf{k}| \leq K}$ with

$$\omega_{\mathbf{k}L}^E(\mathbf{r}) = \begin{cases} |\Omega|^{-\frac{1}{2}} e^{i\mathbf{k} \cdot \mathbf{r}} & \text{in } \mathcal{D}, \\ \sum_{lm}^L \alpha_{lm}^{\mathbf{k}} \chi_l(r, E) Y_{lm}(\hat{\mathbf{r}}) & \text{in } \mathcal{C}, \end{cases} \tag{3.36}$$

where the coefficients $\alpha_{lm}^{\mathbf{k}}$ are determined by the expansion (3.11) as

$$\alpha_{lm}^{\mathbf{k}} = 4\pi i^l j_l(kR) Y_{lm}^*(\hat{\mathbf{k}}) / \chi_l(R, E) \tag{3.37}$$

so that the constraint in (3.2) is satisfied. Since the effective potential is spherically symmetric, $\{\omega_{\mathbf{k}L}^E(\mathbf{r})\}_{|\mathbf{k}|<K}$ are solutions of the Schrödinger type equation inside the sphere, but only with the eigenvalue E . Consequently, E must be set equal to the eigenvalue to be calculated. This leads to a nonlinear eigenvalue problem: find $\lambda_{KL} \in \mathbb{R}$ and $0 \neq u_{KL} \in V_{KL}^{\lambda_{KL}}$, such that

$$a_\delta(u_{KL}, v) = \lambda_{KL}(u_{KL}, v) \quad \forall v \in V_{KL}^{\lambda_{KL}}, \tag{3.38}$$

where $a_\delta(\cdot, \cdot)$ is given by (3.6). Solving (3.38) is a very computationally demanding procedure. Using the root-tracing method (see, e.g. [39]), one has to choose a parameter E , solve the radial Schrödinger equation to obtain the APW basis functions and set up the matrix elements. The determinant has to be computed, which is expected to vanish according to (3.38) but does not with an incorrect parameter. One has to vary the trial parameters in an appropriate domain to numerically find the zeros of this determinant (see [44] or Example 2 in Sect. 5). This is the main drawback of the APW method, which works for simple systems with a few eigenvalues only.

Another, less serious, difficulty of the APW method is the so-called asymptote problem (see [45]): the eigenvalue dependent APW basis functions must be evaluated for a large number of parameters, and sometimes one might hit a parameter E for which $\chi_l(R, E)$ equals zero at the spherical surface. Inserting this in (3.37) will yield infinite coefficients $\alpha_{lm}^{\mathbf{k}}$ and can not match the basis functions at the spherical surface. Therefore, we shall always assume that $\chi_l(R, E) \geq \alpha$ with some constant $\alpha > 0$ for any $l \in \mathbb{N}$, which can be achieved by varying the radius of the atomic sphere.

In the following, we shall give a convergence analysis of the APW approximations. Note that we search the eigenvalue approximations in a finite domain that contains the eigenvalues to be calculated (in fact, we can derive that for each L , $\{\lambda_{KL}\}_{K \in \mathbb{N}}$ is a non-increasing sequence bounded below by $\lim_{K, N \rightarrow \infty} \lambda_{NL}^K$ with λ_{NL}^K being the eigenvalue of (3.5)). Consequently, we have limiting points in this domain when K and L go to infinity. Let λ_∞ be any accumulation point, i.e., there exists a subsequence of eigenpairs, which we still denote by $\{(\lambda_{KL}, u_{KL})\}_{K, L \in \mathbb{N}}$, such that

$$\lim_{K, L \rightarrow \infty} |\lambda_{KL} - \lambda_\infty| = 0. \tag{3.39}$$

We have from (3.8) that

$$\lambda_{KL} = a_\delta(u_{KL}, u_{KL}) \geq C \|u_{KL}\|_\delta^2,$$

which implies that $\{u_{KL}\}_{K, L \in \mathbb{N}}$ is bounded in $H_{\#\delta}$. Banach-Alaoglu Theorem yields that there exists a weakly convergent subsequence $\{u_{K_j L_j}\}_{j \in \mathbb{N}}$, such that

$$u_{K_j L_j} \rightharpoonup u_\infty \quad \text{in } H_{\#\delta}(\Omega) \quad \text{as } j \rightarrow \infty \tag{3.40}$$

for some $u_\infty \in H_{\#}^1(\Omega)$. For simplicity of notation, we shall denote the sequence $\{(\lambda_{K_j L_j}, u_{K_j L_j})\}_{j \in \mathbb{N}}$ by $\{(\lambda_j, u_j)\}_{j \in \mathbb{N}}$ afterwards. Using Sobolev's imbedding theorem, we have that u_j converges to u_∞ in $L^q(\Omega)$ for $1 \leq q < 6$, which leads to

$$\lim_{j \rightarrow \infty} \int_\Omega V_{\text{eff}}(\mathbf{r})(u_\infty^2 - u_j^2) \mathbf{d}\mathbf{r} = 0. \tag{3.41}$$

Therefore, we derive from (3.39) and (3.41) that $\|u_j\|_\delta \rightarrow \|u_\infty\|_\delta$, which together with (3.40) yields

$$u_j \rightarrow u_\infty \quad \text{in } H_{\#\delta}(\Omega).$$

Theorem 3.5. *If $V_{\text{eff}}(\mathbf{r})$ is a spherically symmetric potential in \mathcal{C} and satisfies (2.7), then the limiting pair $(\lambda_\infty, u_\infty)$ is an eigenpair of (2.2), that is*

$$a(u_\infty, v) = \lambda_\infty(u_\infty, v) \quad \forall v \in H_{\#}^1(\Omega). \tag{3.42}$$

Moreover, for the sequence of eigenpairs $\{(\lambda_j, u_j)\}_{j \in \mathbb{N}}$ that converges to $(\lambda_\infty, u_\infty)$, there exists a constant C_s such that

$$|\lambda_\infty - \lambda_j| + \|u_\infty - u_j\|_\delta \leq C_s(K_j^{-(s-1)} + L_j^{-(s-1)}), \tag{3.43}$$

where $s > 1$ and j is sufficiently large.

Proof. Since $V_{\text{eff}}(\mathbf{r}) = V(r)$ and $\chi_l(r, \lambda_j)$ satisfies the radial Schrödinger equation, we have that for $u_j \in V_{K_j L_j}^{\lambda_j}$, there holds

$$\left(-\frac{1}{2}\Delta + V_{\text{eff}}\right) u_j = \lambda_j u_j \quad \text{in } \mathcal{C}. \tag{3.44}$$

Denote by $a_{\mathcal{C}}(w, v) = \frac{1}{2} \int_{\mathcal{C}} \nabla w \cdot \nabla v + \int_{\mathcal{C}} V_{\text{eff}} w v$ and $a_{\mathcal{D}}(w, v) = \frac{1}{2} \int_{\mathcal{D}} \nabla w \cdot \nabla v + \int_{\mathcal{D}} V_{\text{eff}} w v$. For any $v \in C_{\#}^\infty(\Omega)$, we have

$$\begin{aligned} a_\delta(u_j, v) - \lambda_j(u_j, v) &= a_\delta(u_j, v_j) - \lambda_j(u_j, v_j) + a_\delta(u_j, v - v_j) - \lambda_j(u_j, v - v_j) \\ &= a_{\mathcal{C}}(u_j, v - v_j) + a_{\mathcal{D}}(u_j, v - v_j) - \lambda_j(u_j, v - v_j) \\ &= \frac{1}{2} \int_{\Gamma} \frac{\partial u_j|_{\mathcal{C}}}{\partial n} (v - v_j|_{\mathcal{C}}) + a_{\mathcal{D}}(u_j, v - v_j) - \lambda_j \int_{\mathcal{D}} u_j (v - v_j) \\ &\lesssim (\|v - v_j\|_{H^1(\mathcal{D})} + \|v_j|_{\mathcal{C}} - v_j|_{\mathcal{D}}\|_{H^{1/2}(\Gamma)}) \|u_j\|_\delta \quad \forall v_j \in V_{K_j L_j}^{\lambda_j}, \end{aligned}$$

where (3.44) is used for the third equality. Note that for any $v \in C_{\#}^\infty(\Omega)$, we can find $v_j \in V_{K_j L_j}^{\lambda_j}$ such that $\|v - v_j\|_{H^1(\mathcal{D})} \rightarrow 0$ as $j \rightarrow \infty$. This together with the fact $\lim_{j \rightarrow \infty} \|v_j|_{\mathcal{C}} - v_j|_{\mathcal{D}}\|_{H^{1/2}(\Gamma)} = 0$ leads to

$$\lim_{j \rightarrow \infty} a_\delta(u_j, v) - \lambda_j(u_j, v) = 0 \quad \forall v \in C_{\#}^\infty(\Omega). \tag{3.45}$$

Since $C_{\#}^\infty(\Omega)$ is dense in $H_{\#}^1(\Omega)$, we have that (3.45) holds for all $v \in H_{\#}^1(\Omega)$.

Note that

$$\begin{aligned} |a(u_\infty, v) - \lambda_\infty(u_\infty, v)| &= |a_\delta(u_\infty, v) - \lambda_\infty(u_\infty, v)| \\ &\leq |a_\delta(u_\infty - u_j, v)| + |\lambda_\infty(u_\infty - u_j, v)| + |(\lambda_j - \lambda_\infty)(u_j, v)| + |a_\delta(u_j, v) - \lambda_j(u_j, v)| \\ &\lesssim \|u_\infty - u_j\|_\delta + |\lambda_\infty - \lambda_j| + |a_\delta(u_j, v) - \lambda_j(u_j, v)| \quad \forall v \in H_{\#}^1(\Omega). \end{aligned}$$

Using (3.45), the convergence of $u_j \rightarrow u_\infty$ in $H_{\# \delta}(\Omega)$ and $\lambda_j \rightarrow \lambda_\infty$, we can conclude (3.42).

Define the operator $T_j : L_{\#}^2(\Omega) \rightarrow V_{K_j L_j}^{\lambda_j}$ such that for $f \in L_{\#}^2(\Omega)$,

$$a_\delta(T_j f, v) = (f, v) \quad \forall v \in V_{K_j L_j}^{\lambda_j}. \tag{3.46}$$

Let γ be a circle in the complex plane centered at λ_∞^{-1} , which does not enclose any other point of $\vartheta(T)$. Define the spectral projector of T_j by

$$\mathcal{E}_j = \mathcal{E}_j(\lambda_\infty^{-1}) = \frac{1}{2\pi i} \int_{\gamma} R_z(T_j) dz.$$

Note that the convergence of eigenvalues implies that any open set in the complex plane containing $\vartheta(T)$ contains $\vartheta(T_j)$ for sufficiently large j . Therefore, $R_z(T_j)$ is well defined and bounded on γ for sufficiently large j . Since

$$R_z(T) - R_z(T_j) = R_z(T_j)(T - T_j)R_z(T),$$

we have from the definition of spectral projector that

$$\|\mathcal{E} - \mathcal{E}_j\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))} \leq C\|T - T_j\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))}.$$

Therefore it is now only necessary for us to estimate $\|T - T_j\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))}$. For simplicity of notation, we will denote $(\lambda_\infty, u_\infty)$ by (λ, u) in the following.

Let \tilde{u} be the extension of $u|_{\mathcal{D}}$ given as in (3.13) and $\tilde{c}_{\mathbf{k}} = |\Omega|^{-1/2} \int_{\Omega} \tilde{u}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}$. Define

$$I_{j,\lambda}u = \sum_{|\mathbf{k}|\leq K_j} \tilde{c}_{\mathbf{k}} \omega_{\mathbf{k}L_j}^\lambda \quad \text{and} \quad I_{j,\lambda_j}u = \sum_{|\mathbf{k}|\leq K_j} \tilde{c}_{\mathbf{k}} \omega_{\mathbf{k}L_j}^{\lambda_j}. \quad (3.47)$$

Since $Tu = \lambda^{-1}u$ for $u \in \mathcal{R}(\mathcal{E})$, we obtain from (3.24) that

$$\|Tu - T_ju\|_\delta \lesssim \|u - I_{j,\lambda_j}u\|_\delta + \sup_{v_j \in V_{K_j L_j}^{\lambda_j}} \frac{a(Tu - T_ju, v_j)}{\|v_j\|_\delta}. \quad (3.48)$$

The first term on the right-hand side of (3.48) is estimated by

$$\|I_{j,\lambda_j}u - u\|_\delta \leq \|I_{j,\lambda_j}u - u\|_{H^1(\mathcal{D})} + \|I_{j,\lambda_j}u - I_{j,\lambda}u\|_{H^1(\mathcal{C})} + \|I_{j,\lambda}u - u\|_{H^1(\mathcal{C})}. \quad (3.49)$$

Using similar arguments as those in the proof of Lemma 3.1, we have

$$\|I_{j,\lambda_j}u - u\|_{H^1(\mathcal{D})} \leq CK_j^{-(s-1)} \|u\|_{H^s(\mathcal{D})}. \quad (3.50)$$

Note that similar to (3.44), $I_{j,\lambda}u$ and $I_{j,\lambda_j}u$ satisfy the following equations

$$\left(-\frac{1}{2}\Delta + V_{\text{eff}}\right) I_{j,\lambda}u = \lambda I_{j,\lambda}u \quad \text{and} \quad \left(-\frac{1}{2}\Delta + V_{\text{eff}}\right) I_{j,\lambda_j}u = \lambda_j I_{j,\lambda_j}u \quad \text{in } \mathcal{C}. \quad (3.51)$$

Since (3.47) implies that $I_{j,\lambda}u = I_{j,\lambda_j}u$ on Γ from both inside and outside the sphere, we have

$$\begin{cases} \left(-\frac{1}{2}\Delta + V_{\text{eff}} - \lambda_j\right) (I_{j,\lambda}u - I_{j,\lambda_j}u) = (\lambda - \lambda_j) I_{j,\lambda}u & \text{in } \mathcal{C}, \\ I_{j,\lambda}u - I_{j,\lambda_j}u = 0 & \text{on } \partial\mathcal{C}. \end{cases} \quad (3.52)$$

From the assumption on the asymptote problem, we have that λ_j is not an eigenvalue of the problem

$$\begin{cases} \left(-\frac{1}{2}\Delta + V_{\text{eff}}\right) v = \lambda v & \text{in } \mathcal{C}, \\ v|_{\partial\mathcal{C}} = 0. \end{cases}$$

Therefore, we have from (3.52) that

$$\|I_{j,\lambda}u - I_{j,\lambda_j}u\|_{H^1(\mathcal{C})} \lesssim |\lambda - \lambda_j| \|I_{j,\lambda}u\|_{L^2(\mathcal{C})} \lesssim |\lambda - \lambda_j| (\|u\|_{L^2(\mathcal{C})} + \|I_{j,\lambda}u - u\|_{L^2(\mathcal{C})}). \quad (3.53)$$

Similarly, we have from $(-\frac{1}{2}\Delta + V_{\text{eff}} - \lambda)(I_{j,\lambda}u - u) = 0$ in \mathcal{C} that

$$\begin{aligned} \|I_{j,\lambda}u - u\|_{H^1(\mathcal{C})} &\leq C\|I_{j,\lambda}u|_{\mathcal{C}} - u\|_{H^{1/2}(\Gamma)} \\ &\leq C(\|u - I_{j,\lambda}u|_{\mathcal{D}}\|_{H^{1/2}(\Gamma)} + \|I_{j,\lambda}u|_{\mathcal{C}} - I_{j,\lambda}u|_{\mathcal{D}}\|_{H^{1/2}(\Gamma)}) \\ &\leq C(K_j^{-(s-1)} + L_j^{-(s-1)}) \|u\|_{H^s(\mathcal{D})}. \end{aligned} \quad (3.54)$$

Therefore, we obtain from (3.49), (3.50), (3.53) and (3.54) that

$$\|I_{j,\lambda_j}u - u\|_\delta \lesssim (K_j^{-(s-1)} + L_j^{-(s-1)} + |\lambda - \lambda_j|)(\|u\|_{H^s(\mathcal{D})} + \|u\|_{L^2(\mathcal{C})}). \tag{3.55}$$

For the second term on the right-hand side of (3.48), we can obtain

$$\sup_{v_j \in V_{K_j L_j}^{\lambda_j}} \frac{a(Tu - T_j u, v_j)}{\|v_j\|_\delta} \lesssim L_j^{-(s-1)} \|u\|_{H^s(\mathcal{D})} \tag{3.56}$$

by using similar argument as that in (3.26). Taking the regularity results of Lemma 2.2, 2.3, (3.48), (3.55) and (3.56) into account gives rise to

$$\|T - T_j\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))} \leq C(K_j^{-(s-1)} + L_j^{-(s-1)} + |\lambda - \lambda_j|). \tag{3.57}$$

Using (3.33), we have $|\lambda - \lambda_j| \lesssim \|T - T_j\|_{\mathcal{L}(\mathcal{R}(\mathcal{E}), H_{\#\delta}(\Omega))}^2 + D_j$, where $|D_j| \lesssim L_j^{-(s-1)}$ can be obtained by using the same argument as that in (3.34). This together with (3.57) and the fact $\lambda_j \rightarrow \lambda$ leads to (3.43), which completes the proof. \square

Remark 3.6. In contrast to the basis functions we construct in Section 3.1, the APW method uses different radial basis functions for different angular components by solving (3.35). With this type of basis functions, we see from Theorem 3.5 that a half order is not lost in the estimate compared with Lemma 3.3 and Theorem 3.4.

Remark 3.7. When the radius R is well chosen, the assumption of spherically symmetric potential is reasonable for a lot of systems, *e.g.* close-packed crystals [39]. That is why the APW method has been successfully used in many computations.

3.3. LAPW(+lo) methods

We shall discuss LAPW(+lo) methods in this subsection. Since the accuracies of LAPW(+lo) methods depend heavily on the parameters chosen by experience, we can not prove the convergence of numerical approximations but only provide some conjectures on the error estimates.

The LAPW method uses the combination of $\chi_l(r, E)$ and its energy derivative $\dot{\chi}_l(r, E)$ with a fixed parameter E as the radial basis function. The energy derivative is defined by

$$\dot{\chi}_l(r, E) = \frac{\partial}{\partial \varepsilon} \chi_l(r, \varepsilon)|_{\varepsilon=E},$$

where χ_l is kept normalized to the same value in the sphere (the properties of $\dot{\chi}_l(r, E)$ can be read in [3, 39, 45]). The approximation space is $\tilde{V}_{KL}^E = \text{span}\{\tilde{\omega}_{\mathbf{k}L}^E\}_{|\mathbf{k}|\leq K}$ with

$$\tilde{\omega}_{\mathbf{k}L}^E(\mathbf{r}) = \begin{cases} |\Omega|^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{r}} & \text{in } \mathcal{D}, \\ \sum_{lm}^L [\alpha_{lm}^{\mathbf{k}} \chi_l(r, E) + \beta_{lm}^{\mathbf{k}} \dot{\chi}_l(r, E)] Y_{lm}(\hat{\mathbf{r}}) & \text{in } \mathcal{C}. \end{cases} \tag{3.58}$$

The coefficients $\alpha_{lm}^{\mathbf{k}}$ and $\beta_{lm}^{\mathbf{k}}$ are determined by (3.11) requiring that the basis functions match both the value and slope (in the weak sense of (3.2)). The LAPW method provides flexible basis functions to properly describe eigenfunctions that correspond to eigenvalues near E . In contrast to the APW method, this scheme allows us to obtain all eigenvalues by solving a linear eigenvalue problem.

Using the result of Theorem 3.5, it is only necessary for us to estimate the error introduced by linearization, say, the difference between the LAPW radial basis function with fixed parameter E

$$\tilde{\chi}_l(r, E) = \alpha_l \chi_l(r, E) + \beta_l \dot{\chi}_l(r, E) \tag{3.59}$$

and the APW radial basis function $\chi_l(r, \lambda_{KL})$. This is done by assuming that the APW radial basis functions $\chi_l(r, \lambda_{KL})$ and the LAPW combination (3.59) have the same logarithmic derivatives with respect to r at the spherical surface (see [3, 31, 39]), *i.e.*

$$\frac{\chi'_l(r, \lambda_{KL})}{\chi_l(r, \lambda_{KL})} \Big|_{r=R} = \frac{\tilde{\chi}'_l(r, E)}{\tilde{\chi}_l(r, E)} \Big|_{r=R}. \tag{3.60}$$

Let $\sigma = |\lambda_{KL} - E|$. It has been analyzed in [3, 31] that under the Assumption (3.60), the difference between $\tilde{\chi}_l(r, E)$ and $\chi_l(r, \lambda_{KL})$ is of order σ^2 , *i.e.*

$$\|\tilde{\chi}_l(r, E) - \chi_l(r, \lambda_{KL})\|_{H^1([0, R])} \lesssim \sigma^2.$$

Using similar arguments as those in the proof of Theorem 3.5, we may obtain the following *a priori* error estimate of LAPW approximations $(\tilde{\lambda}_{KL}, \tilde{u}_{KL})$

$$\|u - \tilde{u}_{KL}\|_\delta \leq C_s(K^{-(s-1)} + L^{-(s-1)} + \sigma^2), \tag{3.61}$$

$$|\lambda - \tilde{\lambda}_{KL}| \leq c_s(K^{-(s-1)} + L^{-(s-1)} + \sigma^4) \tag{3.62}$$

for some eigenpair (λ, u) of (2.2) and any $s > 1$.

Remark 3.8. The *a priori* error estimates (3.61) and (3.62) rely on the Assumption (3.60), which has not been proved (see a comment in [33]). Nevertheless, the numerical experiments in Section 5 support the exponential convergence rates.

Although the error of the eigenvalue approximation relies on the choice of E , the high order of σ in (3.62) results in that the LAPW method forms a good basis set over a relatively large eigenvalue region. In most materials, it is quite adequate to choose E near the center of the region of eigenvalues to be computed. However, in a few instances, there is no single choice of E that is adequate for all the eigenvalues that must be considered. Then the eigenvalue region may be divided into a few windows and separate computations with different parameters should be carried out for each.

In order to have enough variational flexibility in the radial basis functions, the LAPW+lo method adds new local basis functions to the LAPW method. The approximation space is spanned by $\{\tilde{\omega}_{\mathbf{k}L}^E\}_{|\mathbf{k}| \leq K}$ given in (3.58) and local orbitals $\{\varphi_i\}_{1 \leq i \leq N}$ as

$$\varphi_i(\mathbf{r}) = \begin{cases} 0 & \text{in } \mathcal{D}, \\ \sum_{lm}^L [\alpha_{i,lm} \chi_{\varepsilon_i}(r, \varepsilon_i) + \beta_{i,lm} \dot{\chi}_{\varepsilon_i}(r, \varepsilon_i)] Y_{lm}(\hat{\mathbf{r}}) & \text{in } \mathcal{C}, \end{cases} \tag{3.63}$$

where the coefficients $\alpha_{i,lm}$ and $\beta_{i,lm}$ do not depend on \mathbf{k} , but are determined by the requirement that φ_i is zero at the sphere boundary and normalized. Note that different parameters ε_i can be chosen so that different states can be described simultaneously.

As shown by the numerical experiments in [38], the LAPW+lo method converges practically to identical result as the LAPW method, but allows a significantly smaller basis set (up to 50%) and thus reduces the computational cost drastically.

4. ERROR ESTIMATES FOR KOHN–SHAM EQUATIONS

We have given an *a priori* error estimate for linear Schrödinger type eigenvalue problems in the previous section. In this section, we shall investigate the nonlinear KS equations (1.1), considering the convergence and *a priori* error estimate of the nonconforming approximations.

We shall introduce some notations first. For $\kappa \in \mathbb{R}^{N \times N}$, we denote its Frobenius norm by $|\kappa|$. Define

$$\mathcal{H} \equiv (H_{\#}^1(\Omega))^N = \{(\phi_1, \phi_2, \dots, \phi_N) : \phi_i \in H_{\#}^1(\Omega), i = 1, 2, \dots, N\},$$

which is a Hilbert space associated with the induced norm $\|\Phi\|_{1,\Omega} = \left(\sum_{i=1}^N \|\phi_i\|_{H^1(\Omega)}^2 \right)^{1/2}$. In our discussion, we shall use the following spaces:

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

and

$$\mathcal{Q} = \{\Phi \in \mathcal{H} : \langle \Phi^T \Phi \rangle = I^{N \times N}\} \quad \text{with} \quad \langle \Phi^T \Psi \rangle = \left(\int_{\Omega} \phi_i \psi_j \right)_{i,j=1}^N \in \mathbb{R}^{N \times N}.$$

For any $\Phi \in \mathcal{H}$, we may decompose \mathcal{H} into a direct sum of three subspaces (see, e.g. [18, 37]):

$$\mathcal{H} = \mathcal{S}^{\Phi} \oplus \mathcal{A}^{\Phi} \oplus \mathcal{T}^{\Phi}, \quad (4.1)$$

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

For the nonconforming framework, we define $\mathcal{H}_L = (H_{\#L}^1(\Omega))^N$ with the induced norm $\|\cdot\|_{\delta,\Omega}$, and $\mathcal{Q}_L = \{\Phi \in \mathcal{H}_L : \langle \Phi^T \Phi \rangle = I^{N \times N}\}$. For any $\Phi \in \mathcal{H}_L$, the space \mathcal{H}_L can be decomposed into three subspaces like (4.1)

$$\mathcal{H}_L = \mathcal{S}_L^{\Phi} \oplus \mathcal{A}_L^{\Phi} \oplus \mathcal{T}_L^{\Phi}. \quad (4.2)$$

We also need the discrete space $\mathcal{V}_{\varrho} = (V_{\varrho}(\Omega))^N \subset \mathcal{H}_L$, where ϱ and V_{ϱ} are defined in Section 3.1.

Remark 4.1. Following [18, 43], we can use Grassmann manifold to interpret the equivalence classes of orthonormal bases spanning the same N -dimensional subspace. Define the equivalence class of Φ by

$$[\Phi] = \{\Psi \in \mathcal{Q}_L : \Psi = \Phi U, U \text{ is an orthogonal matrix}\}.$$

\mathcal{T}_L^{Φ} is the tangent space of the Grassmann manifold at $[\Phi]$.

4.1. Abstract Kohn–Sham model and nonconforming discretization

The ground state solutions of the periodic KS-LDA model can be obtained by minimizing the KS energy functional

$$E(\Phi) = \frac{1}{2} \sum_{i=1}^N \int_{\Omega} |\nabla \phi_i(x)|^2 dx + \int_{\Omega} v_{\text{ext}}(x) \rho_{\Phi}(x) dx + \frac{1}{2} D(\rho_{\Phi}, \rho_{\Phi}) + \int_{\Omega} \mathcal{E}(\rho_{\Phi}(x)) dx \quad (4.3)$$

with respect to KS orbitals $\Phi = \{\phi_i\}_{i=1}^N \in \mathcal{H}$ under the orthogonality constraints $\langle \Phi^T \Phi \rangle = I^{N \times N}$. The function $\mathcal{E}(\rho)$ denotes the exchange-correlation energy per unit volume in an electron gas with density ρ and $\mathcal{E}'(t) = v_{\text{xc}}(t)$. $D(\rho_{\Phi}, \rho_{\Phi})$ models the periodic electron-electron Coulomb energy, where

$$D(f, g) = 4\pi \sum_{\mathbf{k} \in \mathcal{R}^* \setminus \{0\}} |\mathbf{k}|^{-2} \hat{f}_{\mathbf{k}}^* \hat{g}_{\mathbf{k}}.$$

The existence of a minimizer of the energy functional can be found in [1, 35]. The Euler–Lagrange equation associated with the minimization problem is

$$\begin{cases} (H_{\Phi} \phi_i, v) = \left(\sum_{j=1}^N \lambda_{ij} \phi_j, v \right) \quad \forall v \in H_{\#}^1(\Omega), \quad i = 1, 2, \dots, N, \\ \int_{\Omega} \phi_i \phi_j = \delta_{ij}, \end{cases} \quad (4.4)$$

where $\Lambda \equiv (\lambda_{ij})_{i,j=1}^N = \left(\int_{\Omega} \phi_j H_{\Phi} \phi_i \right)_{i,j=1}^N$ is the Lagrange multiplier and

$$(H_{\Phi} u, v) = \frac{1}{2} \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (v_{\text{ext}}(\mathbf{r}) + v_H(\rho_{\Phi}(\mathbf{r})) + v_{\text{xc}}(\rho_{\Phi}(\mathbf{r}))) uv$$

with $v_H(\rho_{\Phi}(\mathbf{r}))$ defined by (2.8). Note that the Lagrange multiplier Λ can be diagonalized since the energy functional and the Hamiltonian are invariant under any unitary transform of the KS orbitals, *i.e.*, for any $\Phi \in \mathbb{Q}$,

$$E(\Phi) = E(\Phi U) \quad \forall U = (u_{ij})_{i,j=1}^N \in \mathcal{O}^{N \times N}, \tag{4.5}$$

where $\mathcal{O}^{N \times N}$ is the set of orthogonal matrices.

To obtain the *a priori* error estimate of the nonconforming approximations, we shall represent the KS equation in the following setting. Define

$$Y = \mathbb{R}^{N \times N} \times \mathcal{H}$$

with the associated norm $\|(\Lambda, \Phi)\|_Y = |\Lambda| + \|\Phi\|_{1,\Omega}$ for $(\Lambda, \Phi) \in Y$. We may rewrite (4.4) as a nonlinear problem:

$$F((\Lambda, \Phi)) = 0 \in Y^*, \tag{4.6}$$

where $F : Y \rightarrow Y^*$ is given by

$$\langle F((\Lambda, \Phi)), (\varkappa, \Gamma) \rangle = \sum_{i=1}^N \left(H_{\Phi} \phi_i - \sum_{j=1}^N \lambda_{ij} \phi_j, \gamma_i \right) + \sum_{i,j=1}^N \chi_{ij} \left(\int_{\Omega} \phi_i \phi_j - \delta_{ij} \right) \tag{4.7}$$

with $\varkappa = (\chi_{ij})_{i,j=1}^N \in \mathbb{R}^{N \times N}$ and $\Gamma = (\gamma_1, \gamma_2, \dots, \gamma_N) \in \mathcal{H}$. The Fréchet derivative of F at $(\Lambda, \Phi) : Y \rightarrow Y^*$ is defined by

$$\langle F'_{(\Lambda, \Phi)}((\mu, \Psi)), (\varkappa, \Gamma) \rangle = a_{(\Lambda, \Phi)}(\Psi, \Gamma) - \sum_{i,j=1}^N (\mu_{ij} \phi_j, \gamma_i) + \sum_{i,j=1}^N \chi_{ij} \int_{\Omega} (\psi_i \phi_j + \phi_i \psi_j) \tag{4.8}$$

with $(\mu, \Psi) \in Y$ ($\mu = (\mu_{ij})_{i,j=1}^N \in \mathbb{R}^{N \times N}$, $\Psi = (\psi_1, \psi_2, \dots, \psi_N) \in \mathcal{H}$), $(\varkappa, \Gamma) \in Y$, and

$$\begin{aligned} a_{(\Lambda, \Phi)}(\Psi, \Gamma) = & \sum_{i=1}^N \left(\frac{1}{2} (\nabla \psi_i, \nabla \gamma_i) + (V_{\text{ext}} \psi_i, \gamma_i) + (\mathcal{E}'(\rho_{\Phi}) \psi_i, \gamma_i) + D(\rho_{\Phi}, \psi_i \gamma_i) \right. \\ & \left. - \left(\sum_{j=1}^N \lambda_{ij} \psi_j, \gamma_i \right) + \left(2\phi_i \mathcal{E}''(\rho_{\Phi}) \sum_{j=1}^N \phi_j \psi_j, \gamma_i \right) + \sum_{j=1}^N 2D(\phi_j \psi_j, \phi_i \gamma_i) \right). \end{aligned} \tag{4.9}$$

We shall then address the nonconforming form of (4.6). Let

$$Y_L = \mathbb{R}^{N \times N} \times \mathcal{H}_L$$

with the associated norm $\|(\mu, \Psi)\|_{Y_L} = |\mu| + \|\Psi\|_{\delta, \Omega}$ and $F_L : Y_L \rightarrow Y_L^*$ be the operator defined by

$$\begin{aligned} \langle F_L((\Lambda_L, \Phi_L)), (\varkappa_L, \Gamma_L) \rangle = & \frac{1}{2} \sum_{i=1}^N \left(\int_{\mathcal{C}} \nabla \phi_{L,i} \nabla \gamma_{L,i} + \int_{\mathcal{D}} \nabla \phi_{L,i} \nabla \gamma_{L,i} \right) \\ & + \sum_{i=1}^N ((v_{\text{ext}} + v_H(\rho_{\Phi_L}) + v_{\text{xc}}(\rho_{\Phi_L})) \phi_{L,i}, \gamma_{L,i}) - \sum_{i=1}^N \left(\sum_{j=1}^N \lambda_{L,ij} \phi_{L,j}, \gamma_{L,i} \right) \\ & + \sum_{i,j=1}^N \chi_{L,ij} \left(\int_{\Omega} \phi_{L,i} \phi_{L,j} - \delta_{ij} \right) \quad \forall (\Lambda_L, \Phi_L), (\varkappa_L, \Gamma_L) \in Y_L. \end{aligned} \tag{4.10}$$

We denote the derivative of F_L at $(\Lambda_L, \Phi_L) \in Y_L$ by $F'_{L,(\Lambda_L, \Phi_L)} : Y_L \rightarrow Y_L^*$ as

$$\begin{aligned} \langle F'_{L,(\Lambda_L, \Phi_L)}((\mu_L, \Psi_L)), (\varkappa_L, \Gamma_L) \rangle &= a_{(\Lambda_L, \Phi_L)}^\delta(\Psi_L, \Gamma_L) - \sum_{i,j=1}^N (\mu_{L,ij} \phi_{L,j}, \gamma_{L,i}) \\ &+ \sum_{i,j=1}^N \chi_{L,ij} \int_{\Omega} (\psi_{L,i} \phi_{L,j} + \phi_{L,i} \psi_{L,j}) \quad \forall (\mu_L, \Psi_L), (\varkappa_L, \Gamma_L) \in Y_L, \end{aligned} \tag{4.11}$$

where the bilinear form $a_{(\Lambda, \Phi)}^\delta(\Psi, \Gamma)$ is defined as (4.9) by replacing $(\nabla \psi_i, \nabla \gamma_i)$ with $\int_{\mathcal{C}} \nabla \psi_i \nabla \gamma_i + \int_{\mathcal{D}} \nabla \psi_i \nabla \gamma_i$.

For numerical discretization of (4.6), let

$$Y_\varrho = \mathbb{R}^{N \times N} \times \mathcal{V}_\varrho$$

and $F_\varrho : Y_\varrho \rightarrow Y_\varrho^*$ be defined as the restriction of F_L to Y_ϱ :

$$\langle F_\varrho((\Lambda_\varrho, \Phi_\varrho)), (\varkappa_\varrho, \Gamma_\varrho) \rangle = \langle F_L((\Lambda_\varrho, \Phi_\varrho)), (\varkappa_\varrho, \Gamma_\varrho) \rangle \quad \forall (\Lambda_\varrho, \Phi_\varrho), (\varkappa_\varrho, \Gamma_\varrho) \in Y_\varrho.$$

The nonconforming approximation of (4.6) can be written as

$$F_\varrho((\Lambda_\varrho, \Phi_\varrho)) = 0 \in Y_\varrho^*. \tag{4.12}$$

We denote the derivative of F_ϱ at $(\Lambda_\varrho, \Phi_\varrho)$ by $F'_{\varrho,(\Lambda_\varrho, \Phi_\varrho)} : Y_\varrho \rightarrow Y_\varrho^*$, which is the restriction of $F'_{L,(\Lambda_\varrho, \Phi_\varrho)}$ to Y_ϱ .

Note that there are infinitely many solutions of (4.6) and (4.12) due to (4.5). We shall define

$$\mathbb{Q}_L^\Phi = \{\Psi \in \mathbb{Q}_L : \|\Psi - \Phi\|_{0,\Omega} = \min_{U \in \mathcal{O}^{N \times N}} \|\Psi U - \Phi\|_{0,\Omega}\} \quad \text{for } \Phi \in \mathbb{Q}$$

to get rid of the redundancy (see [13]). In our analysis, the following lemma will be used, whose proof is given in [13, 37].

Lemma 4.2. *If $\Phi \in \mathbb{Q}$, then $\Psi \in \mathbb{Q}_L^\Phi$ can be represented by*

$$\Psi = \Phi + \Phi S + W$$

with $W \in \mathcal{T}_L^\Phi$ and $S \in \mathcal{S}^{N \times N}$.

4.2. A priori error estimates of the nonconforming approximations

Given $(\Lambda, \Phi) \in \mathcal{S}^{N \times N} \times \mathbb{Q}$, we define

$$X_\Phi = \mathcal{S}^{N \times N} \times (\mathcal{S}_L^\Phi \oplus \mathcal{T}_L^\Phi) \subset Y_L$$

and

$$X_{\Phi,\varrho} = \mathcal{S}^{N \times N} \times (\mathcal{V}_\varrho \cap (\mathcal{S}_L^\Phi \oplus \mathcal{T}_L^\Phi)) \subset Y_\varrho.$$

We assume here and hereafter that $y_0 \equiv (\Lambda_0, \Phi_0)$ is a ground state solution of (4.4), where $\Lambda_0 = (\lambda_{0,ij})_{i,j=1}^N$ and $\Phi_0 = (\phi_{0,1}, \dots, \phi_{0,N})$. We shall derive the existence of a unique local discrete solution $y_\varrho \in X_{\Phi_0,\varrho}$ of (4.12) in the neighborhood of y_0 and further obtain the a priori error estimate.

Our analysis will be carried out under the following two assumptions:

A1. There exists a constant $\alpha \in (0, 1]$ such that $|\mathcal{E}''(t)| + |t\mathcal{E}'''(t)| \lesssim 1 + t^{\alpha-1} \quad \forall t > 0$.

A2. There exists a positive constant γ depending on (Λ_0, Φ_0) such that for any L ,

$$a_{(\Lambda_0, \Phi_0)}^\delta(\Psi, \Psi) \geq \gamma \|\Psi\|_{\delta,\Omega}^2 \quad \forall \Psi \in \mathcal{T}_L^{\Phi_0}. \tag{4.13}$$

Remark 4.3. The assumption **A2** is true for a symmetric bilinear form $\tilde{a}(u, v) = (Au, v)$ if the operator A has N lowest eigenvalues $\lambda_1, \dots, \lambda_N$ and there is a gap between the lowest N th eigenvalue and the $(N + 1)$ th eigenvalue (see, e.g. [13, 43]).

Under the framework of Grassmann manifold [18, 43], the assumption **A2** can be viewed as imposing the elliptic condition on $a_{(\Lambda_0, \Phi_0)}^\delta(\cdot, \cdot)$ on the tangent space.

The following lemma will be used in our analysis, which can be proved under the assumption **A1** by using the same arguments as those in ([13], Lem. 4.6 and [16], Lem. 2.1).

Lemma 4.4. *Let $y_1 = (\Lambda_1, \Phi_1)$ and $y_2 = (\Lambda_2, \Phi_2) \in Y_L$ satisfy $\|y_1\|_{Y_L} + \|y_2\|_{Y_L} \leq \bar{C}$. If the assumption **A1** is satisfied, then there exists a constant C_F depending on \bar{C} such that*

$$\|F_L(y_1) - F_L(y_2)\|_{Y_L^*} \leq C_F \|y_1 - y_2\|_{Y_L} \quad \forall y_1, y_2 \in Y_L. \tag{4.14}$$

Moreover, there is a constant C'_F such that

$$\|F'_{L,y_1} - F'_{L,y_2}\|_{\mathcal{L}(Y_L, Y_L^*)} \leq C'_F (\|y_1 - y_2\|_{Y_L}^\alpha + \|y_1 - y_2\|_{Y_L}^2) \quad \forall y_1, y_2 \in Y_L. \tag{4.15}$$

Lemma 4.5. *If the assumption **A2** is satisfied, then $F'_{L,y_0} : X_{\Phi_0} \rightarrow X_{\Phi_0}^*$ is an isomorphism.*

Proof. It is sufficient to prove that equation

$$F'_{L,y_0}((\mu, \Psi)) = (\eta, g) \tag{4.16}$$

is uniquely solvable in X_{Φ_0} for every $(\eta, g) \in X_{\Phi_0}^*$. To this end we define the following bilinear forms $a_{\Phi_0} : \mathcal{H}_L \times \mathcal{H}_L \rightarrow \mathbb{R}$ and $b_{\Phi_0}, c_{\Phi_0} : \mathcal{H}_L \times \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$ by

$$\begin{aligned} a_{\Phi_0}(\Psi, \Gamma) &= a_{(\Lambda_0, \Phi_0)}^\delta(\Psi, \Gamma), \\ b_{\Phi_0}(\Psi, \varkappa) &= \sum_{i,j=1}^N \chi_{ij}(\phi_{0,i}, \psi_j), \\ c_{\Phi_0}(\Psi, \varkappa) &= \sum_{i,j=1}^N \chi_{ij}((\phi_{0,i}, \psi_j) + (\phi_{0,j}, \psi_i)). \end{aligned}$$

Using (4.11), we may rewrite (4.16) as follows: find $\mu \in \mathcal{S}^{N \times N}$ and $\Psi \in \mathcal{S}_L^{\Phi_0} \oplus \mathcal{T}_L^{\Phi_0}$ such that

$$\begin{cases} a_{\Phi_0}(\Psi, \Gamma) - b_{\Phi_0}(\Gamma, \mu) = (g, \Gamma) & \forall \Gamma \in \mathcal{S}_L^{\Phi_0} \oplus \mathcal{T}_L^{\Phi_0}, \\ c_{\Phi_0}(\Psi, \varkappa) = \sum_{i,j=1}^N \chi_{ij} \eta_{ij} & \forall \varkappa \in \mathcal{S}^{N \times N}. \end{cases} \tag{4.17}$$

For any given $\varkappa \in \mathcal{S}^{N \times N}$, we can choose $\tilde{\Psi} = \Phi_0 \varkappa$, and thus

$$c_{\Phi_0}(\tilde{\Psi}, \varkappa) = 2 \sum_{i,j=1}^N |\chi_{ij}|^2, \tag{4.18}$$

where $\langle \Phi_0^T \Phi_0 \rangle = I^{N \times N}$ is used. Note that a simple calculation leads to

$$\|\tilde{\Psi}\|_{\delta, \Omega} = \|\Phi_0 \varkappa\|_{1, \Omega} \lesssim \left(\sum_{i,j=1}^N |\chi_{ij}|^2 \right)^{1/2} \|\Phi_0\|_{1, \Omega}. \tag{4.19}$$

By taking into account (4.18), (4.19) and the fact $\|\Phi_0\|_{1,\Omega} \leq C$, we obtain

$$\inf_{\mathbb{X} \in \mathcal{S}^{N \times N}} \sup_{\Psi \in \mathcal{S}_L^{\Phi_0}} \frac{c_{\Phi_0}(\Psi, \mathbb{X})}{\|\Psi\|_{\delta, \Omega} |\mathbb{X}|} \geq \kappa_c \tag{4.20}$$

with $\kappa_c > 0$. Hence, there exists a unique solution $\Psi_S \in \mathcal{S}_L^{\Phi_0}$ such that

$$c_{\Phi_0}(\Psi_S, \mathbb{X}) = \sum_{i,j=1}^N \chi_{ij} \eta_{ij} \quad \forall \mathbb{X} \in \mathcal{S}^{N \times N}.$$

Therefore (4.17) is equivalent to: find $\Psi_0 \in \mathcal{T}_L^{\Phi_0}$ such that

$$a_{\Phi_0}(\Psi_0, \Gamma) = (g, \Gamma) - a_{\Phi_0}(\Psi_S, \Gamma) \quad \forall \Gamma \in \mathcal{T}_L^{\Phi_0}. \tag{4.21}$$

The unique solvability of (4.21) is a direct consequence of (4.13).

Using similar arguments as those from (4.18) to (4.20), we get

$$\inf_{\mathbb{X} \in \mathcal{S}^{N \times N}} \sup_{\Psi \in \mathcal{S}_L^{\Phi_0}} \frac{b_{\Phi_0}(\Psi, \mathbb{X})}{\|\Psi\|_{\delta, \Omega} |\mathbb{X}|} \geq \kappa_b$$

with $\kappa_b > 0$. This implies that equation

$$b_{\Phi_0}(\Gamma, \mu) = a_{\Phi_0}(\Psi_0 + \Psi_S, \Gamma) - (g, \Gamma) \quad \forall \Gamma \in \mathcal{S}_L^{\Phi_0}$$

has a unique solution $\mu_S \in \mathcal{S}^{N \times N}$.

We have proved that for any $(\eta, g) \in X_{\Phi_0}^*$, (4.17) has a unique solution $(\mu_S, \Psi_0 + \Psi_S)$. This indicates that F'_{L,y_0} is an isomorphism from X_{Φ_0} to $X_{\Phi_0}^*$ and completes the proof. \square

Before giving a discrete counterpart of Lemma 4.5, we shall introduce two projections. First, we define the projection $\tilde{\Pi}_\varrho : \mathbb{Q} \rightarrow \mathcal{V}_\varrho \cap \mathbb{Q}$ such that

$$\|\tilde{\Pi}_\varrho \Phi - \Phi\|_{\delta, \Omega} = \min_{\Psi \in \mathcal{V}_\varrho \cap \mathbb{Q}} \|\Psi - \Phi\|_{\delta, \Omega} \quad \forall \Phi \in \mathbb{Q}.$$

To project further onto $X_{\Phi, \varrho}$, we then define $\Pi_\varrho : \mathcal{S}^{N \times N} \times \mathbb{Q} \rightarrow X_{\Phi, \varrho}$ by

$$\Pi_\varrho(\Lambda, \Phi) = (\Lambda, (\tilde{\Pi}_\varrho \Phi) \tilde{U}) \quad \forall (\Lambda, \Phi) \in \mathcal{S}^{N \times N} \times \mathbb{Q},$$

where

$$\tilde{U} = \arg \min_{U \in \mathcal{O}^{N \times N}} \|(\tilde{\Pi}_\varrho \Phi)U - \Phi\|_{0, \Omega}.$$

From Lemma 4.2, we see that $\Pi_\varrho : \mathcal{S}^{N \times N} \times \mathbb{Q} \rightarrow X_{\Phi, \varrho}$ is well-defined. We obtain by a direct estimate (see [16], proof of Lem. 3.3) that for any $y = (\Lambda, \Phi) \in \mathcal{S}^{N \times N} \times \mathbb{Q}$,

$$\|\Pi_\varrho y - y\|_{Y_L} \lesssim \inf_{\Psi \in \mathcal{V}_\varrho} \|\Psi - \Phi\|_{\delta, \Omega}. \tag{4.22}$$

Note that $F'_{L,y_0} : X_{\Phi_0} \rightarrow X_{\Phi_0}^*$ being an isomorphism is equivalent to the following inf-sup condition

$$\inf_{y_1 \in X_{\Phi_0}} \sup_{y_2 \in X_{\Phi_0}^*} \frac{\langle F'_{L,y_0} y_1, y_2 \rangle}{\|y_1\|_{Y_L} \|y_2\|_{Y_L}} = \beta > 0 \tag{4.23}$$

with the constant satisfying $\beta^{-1} = \|F'_{L,y_0}\|^{-1} \|\mathcal{L}(X_{\Phi_0}^*, X_{\Phi_0})\|$. We can derive from (4.22) (see [16]) that for sufficiently large ϱ ,

$$\inf_{y_1 \in X_{\Phi_0, \varrho}} \sup_{y_2 \in X_{\Phi_0, \varrho}} \frac{\langle F'_{\varrho, y_0} y_1, y_2 \rangle}{\|y_1\|_{Y_L} \|y_2\|_{Y_L}} \geq \frac{\beta}{2},$$

which together with the fact that F'_{ϱ, y_0} satisfies the following inequality analogous to (4.15)

$$\|F'_{\varrho, y_0} - F'_{\varrho, \Pi_\varrho y_0}\|_{\mathcal{L}(Y_\varrho, Y_\varrho^*)} \lesssim \|y_0 - \Pi_\varrho y_0\|_{Y_L}^\alpha + \|y_0 - \Pi_\varrho y_0\|_{Y_L}^2$$

leads to the following discrete counterpart of Lemma 4.5.

Lemma 4.6. *If the Assumption A1 and A2 are satisfied, then there exists $n_0 \in \mathbb{N}$ such that $F'_{\varrho, \Pi_\varrho y_0} : X_{\Phi_0, \varrho} \rightarrow X_{\Phi_0, \varrho}^*$ is an isomorphism for all $\varrho \geq n_0$. Moreover, there is a constant $M > 0$ such that*

$$\|F'_{\varrho, \Pi_\varrho y_0}\|_{\mathcal{L}(X_{\Phi_0, \varrho}^*, X_{\Phi_0, \varrho})} \leq M \quad \forall \varrho \geq n_0.$$

Now we have the main result of this section as follows.

Theorem 4.7. *Let $y_0 \equiv (\Lambda_0, \Phi_0)$ be a ground state solution of (4.4). If the Assumptions A1 and A2 are satisfied, then there exist $\sigma > 0$ and $n_1 \in \mathbb{N}$ such that (4.12) has a unique local solution*

$$y_\varrho = (\Lambda_\varrho, \Phi_\varrho) \in X_{\Phi_0, \varrho} \cap B_\sigma(y_0)$$

for all $\varrho \geq n_1$. Moreover, we have the following a priori error estimate

$$|\Lambda_0 - \Lambda_\varrho| + \|\Phi_0 - \Phi_\varrho\|_{\delta, \Omega} \leq C \left(\inf_{\Psi \in \mathcal{V}_\varrho} \|\Phi_0 - \Psi\|_{\delta, \Omega} + \varrho^{-(s-1)} \|\Phi_0\|_{H^s(\mathcal{D})} \right). \tag{4.24}$$

Proof. The idea of this proof is to construct a contractive mapping whose fixed point is y_ϱ . We rewrite (4.12) as

$$F_\varrho(y_\varrho) - F_\varrho(\Pi_\varrho y_0) = -F_\varrho(\Pi_\varrho y_0). \tag{4.25}$$

Using the fact that (Λ_0, Φ_0) is a ground state solution and similar argument to that in (3.26), we have that for any $\mathfrak{X} \in \mathbb{R}^{N \times N}$ and $\Gamma = (\gamma_1, \dots, \gamma_N) \in \mathcal{V}_\varrho$,

$$\langle F_\varrho((\Lambda_0, \Phi_0)), (\mathfrak{X}, \Gamma) \rangle = \sum_{i=1}^N \int_\Gamma \frac{\partial \phi_{0,i}}{\partial n} (\gamma_i^+ - \gamma_i^-) \leq CL^{-(s-1)} \|\Phi_0\|_{H^s(\mathcal{D})} \|\Gamma\|_{\delta, \Omega},$$

which together with (4.14) leads to

$$\begin{aligned} \|F_\varrho(\Pi_\varrho y_0)\|_{Y_\varrho^*} &\leq \|F_\varrho(\Pi_\varrho y_0) - F_\varrho(y_0)\|_{Y_\varrho^*} + \|F_\varrho(y_0)\|_{Y_\varrho^*} \\ &\leq \|F_L(\Pi_\varrho y_0) - F_L(y_0)\|_{Y_L^*} + \sup_{(\mathfrak{X}_\varrho, \Gamma_\varrho) \in X_{\Phi_0, \varrho}} \frac{\langle F_\varrho((\Lambda_0, \Phi_0)), (\mathfrak{X}_\varrho, \Gamma_\varrho) \rangle}{\|(\mathfrak{X}_\varrho, \Gamma_\varrho)\|_{Y_L}} \\ &\lesssim \|\Pi_\varrho y_0 - y_0\|_{Y_L} + L^{-(s-1)} \|\Phi_0\|_{H^s(\mathcal{D})}. \end{aligned}$$

From (4.25) and Lemma 4.6, we may define the map $\mathcal{N} : B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho} \rightarrow X_{\Phi_0, \varrho}$ by

$$F'_{\varrho, \Pi_\varrho y_0}(\mathcal{N}(x) - \Pi_\varrho y_0) = -F_\varrho(\Pi_\varrho y_0) - (x - \Pi_\varrho y_0) \int_0^1 \left(F'_{\varrho, \Pi_\varrho y_0 + t(x - \Pi_\varrho y_0)} - F'_{\varrho, \Pi_\varrho y_0} \right) dt \tag{4.26}$$

when $\varrho \geq n_0$. We will show that \mathcal{N} is a contraction from $B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$ into $B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$ if R is chosen sufficiently small and ϱ is large enough.

First, we prove that \mathcal{N} maps $B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$ to $B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$ for sufficiently small R . For each $x \in B_R(\Pi_\varrho y_0)$ and $\varrho > n_0$, we have $\mathcal{N}(x) - \Pi_\varrho y_0 \in X_{\Phi_0, \varrho}$ and

$$\begin{aligned} \|\mathcal{N}(x) - \Pi_\varrho y_0\|_{Y_L} &\leq M(\|F_\varrho(\Pi_\varrho y_0)\|_{Y_\varrho^*} + R \int_0^1 \|F'_{\varrho, \Pi_\varrho y_0 + t(x - \Pi_\varrho y_0)} - F'_{\varrho, \Pi_\varrho y_0}\| dt) \\ &\leq CM(\|\Pi_\varrho y_0 - y_0\|_{Y_L} + R(R^\alpha + R^2) + L^{-(s-1)}\|\Phi_0\|_{H^s(\mathcal{D})}). \end{aligned}$$

Since $CM(\|\Pi_\varrho y_0 - y_0\|_{Y_L} + R^{1+\alpha} + R^3 + L^{-(s-1)}\|\Phi_0\|_{H^s(\mathcal{D})})$ can be estimated by R when R is sufficiently small and ϱ is sufficiently large, we have that $\mathcal{N}(x) \in B_R(\Pi_\varrho y_0)$.

Next, we show that \mathcal{N} is a contraction on $B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$. If $x_1, x_2 \in B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$, then

$$F'_{\varrho, \Pi_\varrho y_0}(\mathcal{N}(x_1) - \mathcal{N}(x_2)) = (x_1 - x_2) \int_0^1 \left(F'_{\varrho, \Pi_\varrho y_0} - F'_{\varrho, x_1 + t(x_2 - x_1)} \right) dt.$$

Thus, $\|\mathcal{N}(x_1) - \mathcal{N}(x_2)\|_{Y_L}$ can be estimated as

$$\begin{aligned} \|\mathcal{N}(x_1) - \mathcal{N}(x_2)\|_{Y_L} &\leq M\|x_1 - x_2\|_{Y_L} \int_0^1 \|F'_{\varrho, \Pi_\varrho y_0} - F'_{\varrho, x_1 + t(x_2 - x_1)}\| dt \\ &\leq CM(R^\alpha + R^2)\|x_1 - x_2\|_{Y_L}. \end{aligned}$$

We obtain that $CM(R^\alpha + R^2) < 1$ for sufficiently small R and hence \mathcal{N} is a contraction on $B_R(\Pi_\varrho y_0)$.

We are now able to use Banach's Fixed Point Theorem to obtain the existence and uniqueness of a fixed point y_ϱ of the map $\mathcal{N} : B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho} \rightarrow B_R(\Pi_\varrho y_0) \cap X_{\Phi_0, \varrho}$, which is the solution of $F_\varrho(y_\varrho) = 0$.

Take $x = y_\varrho$ in (4.26), we have

$$\begin{aligned} \|y_\varrho - \Pi_\varrho y_0\|_{Y_L} &\lesssim \|y_0 - \Pi_\varrho y_0\|_{Y_L} + L^{-(s-1)}\|\Phi_0\|_{H^s(\mathcal{D})} \\ &\quad + \|y_\varrho - \Pi_\varrho y_0\|_{Y_L} (\|y_\varrho - \Pi_\varrho y_0\|_{Y_L}^\alpha + \|y_\varrho - \Pi_\varrho y_0\|_{Y_L}^2), \end{aligned}$$

which together with the fact that $\|y_\varrho - \Pi_\varrho y_0\|_{Y_L}$ can be arbitrary small implies

$$\|y_\varrho - \Pi_\varrho y_0\|_{Y_L} \lesssim \|y_0 - \Pi_\varrho y_0\|_{Y_L} + L^{-(s-1)}\|\Phi_0\|_{H^s(\mathcal{D})}. \tag{4.27}$$

Using (4.22), (4.27) and the triangle inequality

$$\|y_\varrho - y_0\|_{Y_L} \leq \|y_\varrho - \Pi_\varrho y_0\|_{Y_L} + \|y_0 - \Pi_\varrho y_0\|_{Y_L},$$

we can obtain (4.24). This completes the proof. □

Remark 4.8. Note that Lemma 2.2 can not be applied to KS equations directly due to the singularity of the exchange-correlation potential v_{xc} . If the KS orbitals $\{\phi_i\}_{i=1}^N$ are assumed to be asymptotically well behaved, then we get from Theorems 3.2 and 4.7 that

$$|\Lambda_0 - \Lambda_\varrho| + \|\Phi_0 - \Phi_\varrho\|_{\delta, \Omega} \leq C_s \varrho^{-(s-3/2)} \quad \forall s > \frac{3}{2}.$$

Remark 4.9. The arguments in this section are related to the techniques in [13, 16, 34, 51]. We shall point out that [13, 16, 34] are also devoted to the DFT models, nevertheless, the theory in [13, 16] can not be applied to nonconforming methods, and [34] analyzes the orbital-free DFT models which consider the lowest eigenvalue only.

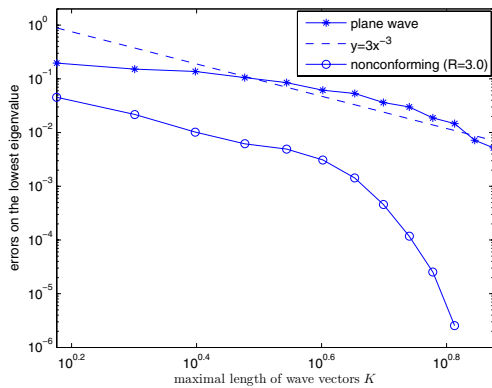


FIGURE 3. Eigenvalue errors of plane wave and nonconforming approximations.

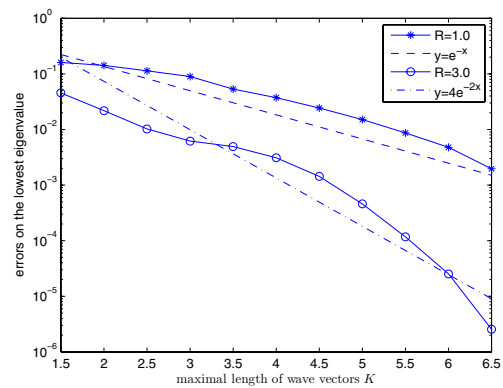


FIGURE 4. Eigenvalue errors of nonconforming approximations with different spherical radius.

5. NUMERICAL EXPERIMENTS

In this section, we shall test our theoretical results by some numerical experiments. Since the analytical solutions of the eigenvalue problems are not available, we use the numerical solutions on the finest discretizations as references to calculate the approximation errors.

Example 1. We solve the linear eigenvalue problem: find $\lambda \in \mathbb{R}$ and $u \in H_{\#}^1(\Omega)$ such that

$$-\frac{1}{2}\Delta u + V_{\text{ext}}u = \lambda u,$$

where $\Omega = [-5.0, 5.0]^3$ and the periodic external potential is

$$V_{\text{ext}}(\mathbf{r}) = \begin{cases} -\frac{1}{r} - \frac{1}{2r_0 - r} & \text{if } r < r_0, \\ -\frac{1}{r_0} & \text{if } r \geq r_0 \end{cases}$$

in Ω with $r_0 = 3.0$. We consider only the lowest eigenvalue of this problem.

We compare the eigenvalue errors of the plane wave approximations and our nonconforming approximations in Figure 3. It is observed that the convergence rate is significantly improved by our nonconforming method. The numerical errors obtained by different choices of atomic spheres are presented in Figure 4. We observe not only the exponential decay of the eigenvalue errors but also a faster decay with a larger atomic sphere. In fact, the eigenfunctions in the interstitial region with a larger sphere are less varying and hence more propitious to plane wave approximations. The eigenfunction on the x -axis is displayed in Figure 5, from which we observe that the nonconforming approximation can catch the cusp at the origin while the plane wave approximation can not.

We also present in Figure 6 the eigenvalue errors as a function of the order of polynomial radial basis functions. We observe that given a sufficiently large K , the errors of eigenvalue approximations converge exponentially. We mention that the condition of Lemma 2.2 is not strictly satisfied for this example since the potential V_{ext} is only continuous up to the first order derivative on the sphere with $r = r_0$. Nevertheless, we still have exponential decay of the numerical errors, which supports the theoretical results in this paper.

The numerical results of the following two examples are presented in atomic unit (a.u.).

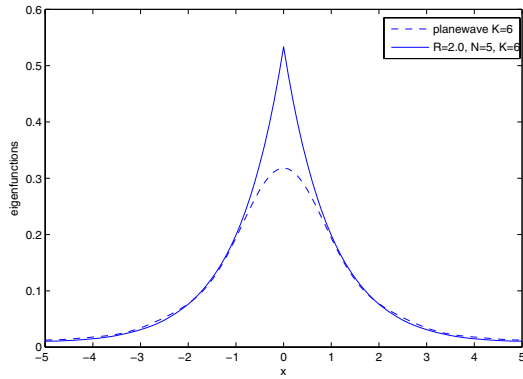


FIGURE 5. Eigenfunctions on x -axis obtained by plane wave and nonconforming methods.

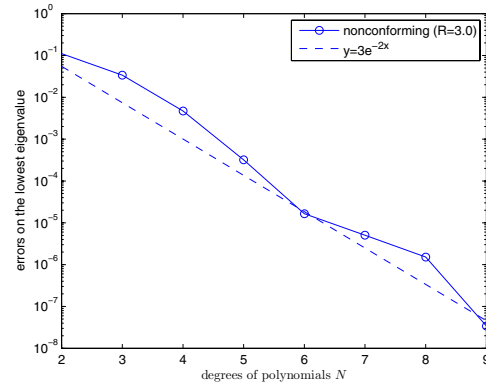


FIGURE 6. Eigenvalue errors of nonconforming approximations with respect to the order of polynomials with fixed $K = 6$.

Example 2. We consider the computation of the ground state of a hydrogen atom by the APW method in the supercell $\Omega = [-5.0, 5.0]^3$ with periodic boundary condition.

Since the APW basis functions depend on the eigenvalues and lead to nonlinear equations, the eigenvalues are normally solved by the “root tracing” technique that determines the eigenvalues by varying the parameter ε numerically to satisfy the condition

$$\det(H_\varepsilon - \varepsilon M_\varepsilon) = 0,$$

where H_ε and M_ε are Hamilton matrix and mass matrix generated by the basis set given in (3.36). Instead of calculating the determinant of $H_\varepsilon - \varepsilon M_\varepsilon$ (which varies strongly in this example and is difficult for root interpolation), we here compute the eigenvalues of $M_\varepsilon^{-1}H_\varepsilon$ for each parameter ε , and calculate the difference between the parameter and its closest eigenvalue λ_ε , say $\lambda_\varepsilon - \varepsilon$. The eigenvalues of the nonlinear problem can be found where the difference is 0 by interpolation (see Fig. 7). The errors on the lowest eigenvalue of plane wave and APW approximations are compared in Figure 8, from which we observe that the APW method is much more accurate.

We shall mention that the computational cost of this nonlinear problem is very large even for medium size systems. In contrast, LAPW(+lo) bases can result in straightforward linear eigenvalue problems and reduce the computational cost significantly.

Example 3. To examine the performance of LAPW approximations, we use the package *Exciting* [54] to calculate the aluminium (Al) and lithium-fluorine (LiF) crystals. *Exciting* is a full-potential DFT package based on the LAPW+lo methods and use a SCF iteration for solving the nonlinear KS equations. The numerical errors of the ground state energy approximations are presented in Figures 9 and 10, from which we observe exponential convergence rates with respect to both K and L .

Moreover, we compare the full-potential computations by using package *Exciting* and the pseudopotential computations by using package *Abinit* [25, 52] for the LiF crystal. We plot in Figure 11 the electron density on a plane (the structure is presented in the left of the figure) obtained by these two methods, from which we see that the true density including the core electrons is obtained by full-potential calculations, whereas the pseudopotential calculations can only have a pseudo electron density of valence electrons.

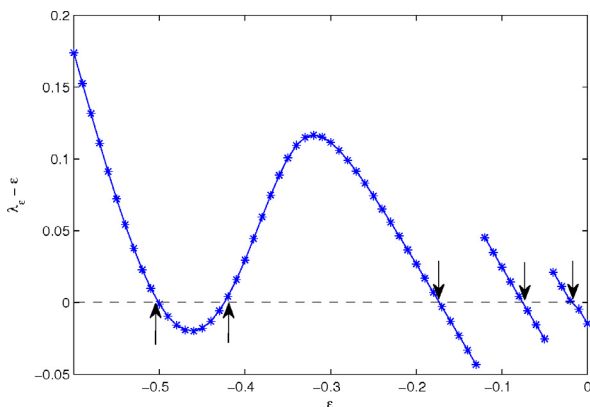


FIGURE 7. Search the eigenvalues of the nonlinear problem generated by APW basis set.

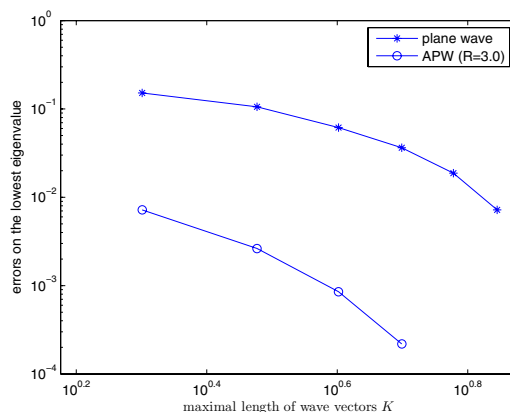


FIGURE 8. Eigenvalue errors of plane wave and APW approximations.

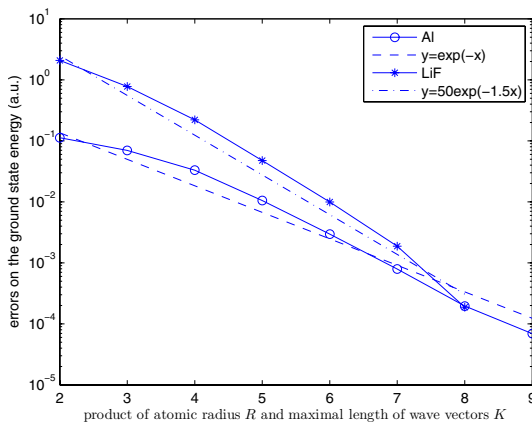


FIGURE 9. Errors on the ground state energy with respect to the maximal length of wave vectors K for Al and LiF using the package *Exciting*.

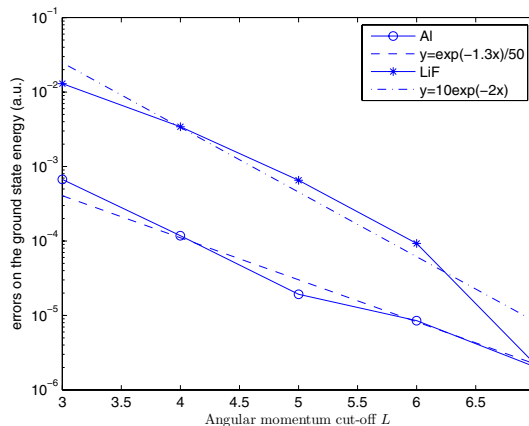


FIGURE 10. Errors on the ground state energy with respect to the angular momentum cut-off L for Al and LiF using the package *Exciting*.

6. CONCLUDING REMARKS

In this paper, we analyze the (L)APW methods which are widely used in full-potential electronic structure calculations. The (L)APW methods enable highly accurate and reliable electronic structure calculations, since they avoid the use of pseudopotentials that might be a source of modeling errors difficult to control. We introduce a nonconforming method based on the idea of the (L)APW methods, which use different types of basis functions inside and outside the atomic spheres. We obtain a *a priori* error estimate of the nonconforming approximations for both linear Schrödinger type equations and nonlinear KS equations, and present some numerical results to support our theory. Due to the fast convergence rate, which is the central result of this paper, the (L)APW methods provide high precision using only approximately the same number of degrees of freedom, or even less, than usually done in the plane wave codes.

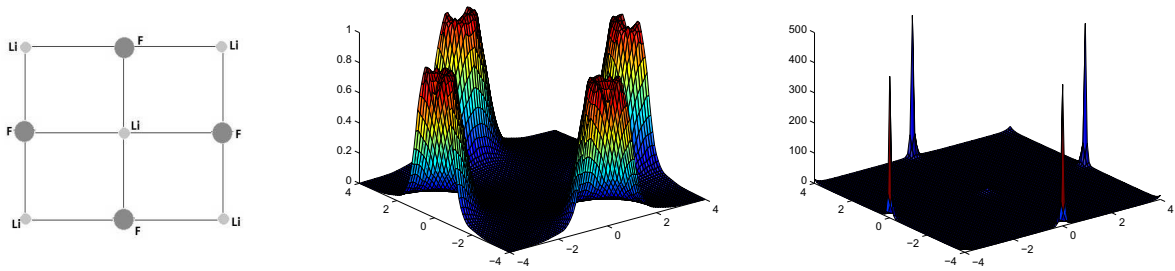


FIGURE 11. Electron density of LiF obtained by the packages *Exciting* and *Abinit*. Left: the structure of the plane. Middle: result of *Abinit*. Right: result of *Exciting*.

In order to set up the Galerkin system matrices and the overlap matrices, one has to integrate plane waves over a complicated interstitial region \mathcal{D} where the atomic spheres are excluded from the unit cell Ω . These integrals are computed by integrating over the whole unit cell and subtracting the integrals over the spheres. The explicit formulas for these integrals can be found in, *e.g.* [16, 39]. However, there arise some discretization errors when calculating these integrals numerically. In this paper, we have not considered this issue but assume that the integrals can be computed exactly. The numerical errors arising from the integrations may be investigated in our future works.

Instead of polynomials, more physically meaningful basis functions like Gaussian or Slater type functions may be incorporated in this nonconforming framework. However, the difficulties lie in that all χ_i do not vanish at the spherical surface, which may generate too many basis functions in computations. The discontinuous Galerkin (DG) methods may be a proper way to handle this problem, in which different areas can be approximated by completely separate basis functions and matched together by DG schemes [15].

Similar to the APW method, the Muffin-tin orbital (MTO) approach [3, 4, 39] is another type of atomic sphere methods. It exploits the same idea that divides the simulation domain and provides efficient representations for atomic-like features that are rapidly varying near each nucleus and smoothly varying between the atoms. The MTO method reformulates the multiple-scattering (MST, also called KKR since it was invented independently by Korringa *et al.* [39]) method, and leads to physically meaningful descriptions of the electronic states in terms of a small basis set of localized, augmented functions.

The MTO basis functions depend on the parameters κ and ε ,

$$\omega_{lm}^{\text{MTO}}(\mathbf{r}) = i^l Y_{lm}(\hat{\mathbf{r}}) \begin{cases} \chi_l(r, \varepsilon) + \kappa \cot(\eta_l(\varepsilon)) j_l(\kappa r) & \text{in } \mathcal{C}, \\ \kappa n_l(\kappa r) & \text{in } \mathcal{D}, \end{cases} \quad (6.1)$$

where $\chi_l(r, \varepsilon)$ is the solution of (3.35), j_l and n_l are spherical Bessel and Neumann functions respectively (for negative energies, the Neumann functions are replaced by Hankel functions), and the coefficients η_l are determined by requiring that the basis functions match the value and slope at the spherical surface. The equations for many-atom systems can be derived using an expansion theorem and the tail cancelation condition, which expresses the tail of a MTO basis extending into another sphere in terms of functions centered on that sphere (see, *e.g.* [4, 39]). This amounts to a transformation of the KKR method and (the eigenvalue-dependent basis set) would lead to a nonlinear eigenvalue problem. Since the solutions obtained by the MTO method satisfy the equation both inside and outside the spheres, the convergence of the MTO approximations can be proven using similar arguments as those in this paper (the detail proof will be given elsewhere), and the error is only determined by the truncation of angular momentum L .

For the linearized Muffin-tin orbital (LMTO) method (see, *e.g.* [39]), the error estimate is far too difficult compared with the LAPW method. The errors of the LAPW approximations are mainly induced by linearization, while the numerical integrations of basis functions can be done quite accurately both inside and outside the

atomic spheres. In contrast, the numerical integrations for the LMTO method are much more complicated, which apply the atomic sphere approximation (ASA) and use the space-filling cells (Wigner-Seitz cells, see, e.g. [3, 39]) neglecting the interstitial region. These error analyses shall be investigated in our future works.

Acknowledgements. We are indebted to two anonymous referees for carefully reading the original manuscript and for suggesting a number of corrections and improvements. The research for this paper has been enabled by the Alexander von Humboldt Foundation, whose support for the long term visit of H. Chen at Technische Universität Berlin is gratefully acknowledged.

REFERENCES

- [1] A. Anantharaman and E. Cancès, Existence of minimizers for Kohn–Sham models in quantum chemistry, *Ann. Inst. Henri Poincaré Anal. Non Lin.* **26** (2009) 2425–2455.
- [2] O.K. Andersen, Simple approach to the band structure problem. *Solid State Commun.* **13** (1973) 133–136.
- [3] O.K. Andersen, Linear methods in band theory. *Phys. Rev. B* **12** (1975) 3060–3083.
- [4] O.K. Andersen and R.V. Kasowski, Electronic states as linear combinations of muffin-tin orbitals. *Phys. Rev. B* **4** (1971) 1064–1069.
- [5] I. Babuška and J. Osborn, Eigenvalue Problems, in Finite Element Methods (Part 1). *Handb. Numer. Anal.* Edited by P.G. Ciarlet and J.L. Lions. In vol. 2. Elsevier Science Publishers, North-Holland (1991) 640–787.
- [6] I. Babuška and M. Rosenzweig, A finite element scheme for domains with corners. *Numer. Math.* **20** (1972) 1–21.
- [7] M. Bachmayr, H. Chen and R. Schneider, Error estimates for Hermite and even-tempered Gaussian approximations in quantum chemistry. *Numer. Math.* **128** (2014) 137–165.
- [8] F.B. Belgacem, The Mortar finite element method with Lagrange multipliers. *Numer. Math.* **84** (1999) 173–197.
- [9] C. Bernardi, N. Débit and Y. Maday, Coupling finite element and spectral methods: first results. *Math. Comput.* **54** (1990) 21–29.
- [10] C. Bernardi, Y. Maday and A.T. Patera, A new non conforming approach to domain decomposition: The Mortar element method, In *Collège de France Seminar*, edited by H. Brezis and J.L. Lions, Pitman (1990).
- [11] S.C. Brenner and L.R. Scott, *Mathematical Theory of Finite Element Methods*. Springer (2002).
- [12] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*. Springer-Verlag, New York (1991).
- [13] E. Cancès, R. Chakir and Y. Maday, Numerical analysis of the plane-wave discretization of some orbital-free and Kohn–Sham models. *ESAIM: M2AN* **46** (2012) 341–388.
- [14] C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, *Spectral Methods for Fluid Dynamics*. Springer-Verlag (1988).
- [15] H. Chen and R. Schneider, DG methods using radial bases functions and plane waves for full-potential electronic structure calculations, preprint.
- [16] H. Chen, X. Gong, L. He, Z. Yang and A. Zhou, Numerical analysis of finite dimensional approximations of Kohn–Sham models. *Adv. Comput. Math.* **38** (2013) 225–256.
- [17] M. Crouzeix and P.A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *Rev. Francaise Automat. Informat. Recherch Operationelle Sér. Anal. Numér.* **7** (1973) 33–75.
- [18] A. Edelman, T.A. Arias and S.T. Smith, The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* **20** (1998) 303–353.
- [19] Y.V. Egorov and B.W. Schulze, *Pseudo-differential Operators, Singularities, Applications*. Birkhäuser, Basel (1997).
- [20] H. Ehrenreich, F. Seitz and D. Turnbull, *Solid State Physics*. Edited by H. Ehrenreich, F. Seitz and D. Turnbull, New York, London (1971).
- [21] H.J. Flad, R. Schneider and B.W. Schulze, Asymptotic regularity of solutions to Hartree-Fock equations with Coulomb potential. *Math. Meth. Appl. Sci.* **31** (2008) 2172–2201.
- [22] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, The electron density is smooth away from the nuclei. *Commun. Math. Phys.* **228** (2002) 401–415.
- [23] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, Analyticity of the density of electronic wavefunctions. *Arkiv för Matematik* **42** (2004) 87–106.
- [24] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, Non-isotropic cusp conditions and regularity of the electron density of molecules at the nuclei. *Ann. Henri Poincaré* **8** (2007) 731–748.
- [25] X. Gonze, J.M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.Y. Raty and D.C. Allan, First-principles computation of material properties: the ABINIT software project. *Comput. Mater. Sci.* **25** (2002) 478–492.
- [26] P. Grisvard, Singularities in boundary value problems. In vol. 22 of *Research Appl. Math.* Masson, Paris (1992).
- [27] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, Electron wavefunctions and densities for atoms. *Ann. Henri Poincaré* **2** (2001) 77–100.
- [28] P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas. *Phys. Rev. B* **136** (1964) 864–871.
- [29] E. Hunsicker, V. Nistor and J.O. Sofo, Analysis of periodic Schrödinger operators: Regularity and approximation of eigenfunctions. *J. Math. Phys.* **49** (2008) 08350101–08350121.

- [30] B.M. Irons and A. Razzaque, Experience with the patch test for convergence of finite elements, in *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Part II*. Edited by A.K. Aziz. Academic Press, New York (1972) 557–587.
- [31] D.D. Koelling and G.O. Arbman, Use of energy derivative of the radial solution in an augmented plane wave method: application to copper. *J. Phys. F: Metal Phys.* **5** (1975) 2041–2054.
- [32] W. Kohn and L.J. Sham, Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* **140** (1965) 1133–1138.
- [33] E.E. Krasovskii, V.V. Nemoshkalenko and V.N. Antonov, On the accuracy of the wavefunctions calculated by LAPW method. *I. Phys. B* **91** (1993) 463–366.
- [34] B. Langwallner, C. Ortner and E. Süli, Existence and convergence results for the Galerkin approximation of an electronic density functional. *Math. Models Methods Appl. Sci.* **20** (2010) 2237–2265.
- [35] C. Le Bris, *Quelques problèmes mathématiques en chimie quantique moléculaire*. Ph.D. thesis, École Polytechnique (1993).
- [36] P. Lesaint, On the convergence of Wilson’s nonconforming element for solving the elastic problems. *Comput. Methods Appl. Mech. Engrg.* **7** (1976) 1–16.
- [37] Y. Maday and G. Turinici, Error bars and quadratically convergent methods for the numerical simulation of the Hartree-Fock equations. *Numer. Math.* **94** (2000) 739–770.
- [38] G.K.H. Madsen, P. Blaha, K. Schwarz, E. Sjöstedt and L. Nordström, Efficient linearization of the augmented plane-wave method. *Phys. Rev. B* **64** (2001) 1951341–1951349.
- [39] R.M. Martin, *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press (2005).
- [40] A. Messiah, *Quantum Mechanics*, vol. I. Wiley, New York (1964).
- [41] J.E. Osborn, Spectral approximation for compact operators. *Math. Comput.* **29** (1975) 712–725.
- [42] P.A. Raviart and J.M. Thomas, Primal hybrid finite element methods for 2nd order elliptic equations. *Math. Comput.* **31** (1977) 391–413.
- [43] 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) 360–387.
- [44] K. Schwarz, P. Blaha and G.K.H. Madsen, Electronic structure calculations of solids using the WIEN2k package for material sciences *Comput. Phys. Commun.* **147** (2002) 71–76.
- [45] D.J. Singh and L. Nordstrom, *Planewaves, Pseudopotentials, and the LAPW Method*. Springer-Berlin (2006).
- [46] E. Sjöstedt, L. Nordström and D.J. Singh, An alternative way of linearizing the APW method. *Solid State Commun.* **114** (2000) 15–20.
- [47] J.C. Slater, Wave functions in a periodic potential. *Phys. Rev.* **51** (1937) 846–851.
- [48] G. Strang. Variational crimes in the finite element methods, in *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Part II*. Edited by A.K. Aziz. Academic Press, New York (1972) 689–710.
- [49] 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. Solid.* **58** (2010) 256–280.
- [50] E.P. Wigner and F. Seitz, On the constitution of metallic sodium. *Phys. Rev.* **43** (1933) 804–810.
- [51] E. Zeidler, *Nonlinear Functional Analysis and Its Applications. I: Fixed-Point Theorems*, translated from the German by P.R. Wadsack. Springer-Verlag (1986).
- [52] ABINIT, <http://www.abinit.org/>
- [53] FLEUR: The Jülich FLAPW code family, <http://www.flapw.de/pm/>.
- [54] The exciting Code, <http://exciting-code.org/>.
- [55] WIEN2k, <http://www.wien2k.at/>.