

# 电子结构计算的数值方法与理论<sup>\*1)</sup>

戴小英<sup>2)</sup>

(LSEC, 中国科学院数学与系统科学研究院, 计算数学与科学与工程计算研究所, 北京 100190;  
中国科学院大学, 北京 100049)

## 摘要

第一原理电子结构计算已成为探索与研究物质机理、理解与预测材料性质的重要手段和工具。虽然第一原理电子结构计算取得了巨大的成功, 但是如何利用高性能计算机又快又好地计算大规模体系, 如何从数学角度理解电子结构模型的合理性与计算的可靠性和有效性, 依然充满各种挑战。基于密度泛函理论的第一原理电子结构计算的核心数学模型为 Kohn-Sham 方程或相应的 Kohn-Sham 能量泛函极小问题。近年来, 人们分别从非线性算子特征值问题的高效离散及 Kohn-Sham 能量泛函极小问题的最优化方法设计两个方面对电子结构计算的高效算法设计及分析展开了诸多研究。本文重点介绍我们小组在电子结构计算的方法与理论方面的一些进展, 同时简单介绍该领域存在的困难与挑战。

**关键词:** 电子结构; Kohn-Sham 方程; Kohn-Sham 总能极小问题; 特征值问题; 有限元方法; 最优化方法; 自适应; 并行计算

**MR (2010) 主题分类:** 35Q55, 65N15, 65N25, 65N30, 81Q05

## 1. 引言

第一原理电子结构计算的基本理论是奠基于二十世纪初的量子力学理论。根据量子力学理论, 在不考虑相对论的情况下, 描述多电子体系电子结构的基本数学模型为如下的定态 Schrödinger 方程<sup>[36]</sup>

$$\hat{H}\psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N) = E\psi(\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{r}_1, \dots, \mathbf{r}_N), \quad (1.1)$$

其中

$$\hat{H} = H_{ee} + V_{ne} + H_{nn}.$$

\* 2020年4月24日收到。

<sup>1)</sup> 基金项目: 国家重点研发计划项目 (2019YFA0709601)、国家自然科学基金 (91730302、11671389) 和中国科学院前沿科学重点研究项目 (QYZDJ-SSW-SYS010)。

<sup>2)</sup> 作者简介: 戴小英, 中国科学院数学与系统科学研究院研究员。2003年本科毕业于湘潭大学, 2008年在中国科学院数学与系统科学研究院获博士学位, 博士毕业后入职中国科学院数学与系统科学研究院工作至今, 期间于2008-2010年在法国巴黎六大 Jacques-Louis Lions 实验室从事博士后研究。主要研究领域包括数值分析、计算材料科学、并行计算等。一直围绕第一原理计算核心数学模型展开相关研究, 包括特征值问题及相应优化问题的高效算法设计及分析, 作为主要成员与课题组一起研制了一套第一原理电子结构实空间并行自适应计算程序 RealSPACES。曾获中国科学院数学与系统科学研究院“陈景润未来之星”称号、中国数学会计算数学分会“青年创新奖”。

这里

$$\begin{aligned} H_{ee} &= T_e + V_{ee} \\ T_e &= -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}_i}^2, \quad V_{ee} = \frac{1}{2} \sum_{i,j=1, i \neq j}^N \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}, \\ V_{ne} &= -\sum_{i=1}^N \sum_{I=1}^M \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|}, \\ H_{nn} &= T_n + V_{nn} = -\sum_{I=1}^M \frac{\hbar^2}{2M_I} \nabla_{\mathbf{R}_I}^2 + \frac{1}{2} \sum_{I,J=1, I \neq J}^M \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}, \end{aligned}$$

其中  $m_e$  是电子的质量,  $\{\mathbf{R}_I : I = 1, \dots, M\}$  是原子核的位置,  $M$  为原子核的个数,  $M_I$  是第  $I$  个原子核的质量,  $Z_I$  是第  $I$  个原子核的电荷数,  $\{\mathbf{r}_i : i = 1, \dots, N\}$  是电子的位置,  $N$  是电子的个数,  $e$  是电子的电荷数. 这里  $T_e$  是电子的动能算符,  $V_{ee}$  是电子间 Coulomb 排斥势,  $H_{nn}$  是核与核之间的作用算符,  $V_{ne}$  则是核与电子间的 Coulomb 吸引势.

多粒子体系的 Schrödinger 方程 (1.1) 是一个定义在  $\mathbb{R}^{3N+3M}$  空间上的算子含奇性的线性特征值问题. 只有极少数的问题能得到其精确的解析解, 其它问题都需要数值求解. 然而由于问题的过高维数带来极其庞大的计算量, 直接对 Schrödinger 方程进行数值求解是十分困难的, 目前只能对非常少量电子的体系进行求解. 因此, 人们对 Schrödinger 方程进行一些等价转换或近似, 在保证所需求精度的前提下尽量降低计算复杂度, 使其相对容易求解. 其中最基本的近似就是 Born 和 Oppenheimer 提出的绝热近似 (也称 Born-Oppenheimer 近似)<sup>[9]</sup>. 其基本出发点为: 由于原子核的质量  $M_I$  大约是电子质量  $m$  的 1836 倍, 速度比电子的速度小得多, 因此可考虑将电子的运动和原子核的运动分开. 目前第一原理电子结构计算的各种方法大都是在这个近似基础上进行的. 通过绝热近似, 可得到多电子系统的定态 Schrödinger 方程, 其具体形式如下:

$$(T_e + V_{ee} + V_{ne})\psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\psi(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (1.2)$$

如果采用原子单位, 即:  $e = 1, m_e = 1, \hbar = 1$ , 长度用 bohr 作单位 ( $1 \text{ bohr} = 0.529177249 \text{ \AA}$ ), 能量单位为 hartree ( $1 \text{ hartree} = 2 \text{ rydbergs} = 27.211 \text{ eV}$ ), 则

$$T_e = -\sum_{i=1}^N \frac{1}{2} \nabla_{\mathbf{r}_i}^2, \quad V_{ee} = \frac{1}{2} \sum_{i,j=1, i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad V_{ne} = -\sum_{i=1}^N \sum_{I=1}^M \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|}.$$

此后本文均采用原子单位.

方程 (1.2) 是一定义在  $\mathbb{R}^{3N}$  上的线性特征值问题. 虽然比起定义在  $\mathbb{R}^{3N+3M}$  上的线性特征值问题 (1.1), 方程 (1.2) 简单了不少, 但它仍然是一高维问题. 模型的高维度带来计算复杂度“灾难”. 因此, Schrödinger 方程 (1.2) 本质上是不可计算模型. 人们需要寻找其等价或简化的可计算模型. 目前已发展出诸如波函数类方法、密度泛函理论、量子 Monte Carlo 等多种方法. 这些等价或简化的可计算模型中最典型的包括波函数方法和密度泛函理论方法. 前者是量子化学领域的主流, 后者在凝聚态物理、材料计算领域广泛使用. 第一原理方法, 是指不使用任何经验参数和实验数据, 只利用电子和原子的质量、电荷、Plank 常数和 Bohr 半径计算 Schrödinger 方程或其等价模型的方法.

波函数方法的思想是通过特定方式构造单电子波函数的组合形式近似多体波函数, 把高维多体问题转化为低维单体问题求解. 常见的波函数方法包括 Hartree 方法<sup>[53]</sup>、Hartree-Fock 方法<sup>[44]</sup>、组态相互作用 (Configuration Interaction) 方法<sup>[81]</sup>、多组态自洽场 (Multi-Configuration Self-Consistent Field) 方法及耦合簇 (Coupled Cluster) 方法<sup>[22]</sup> 等. 这些波函数方法通过调整单电子波函数组合方式可以获得较高精度数值结果, 然而总体而言计算量仍然十分巨大, 故目前波函数方法仅适用于处理较小体系的计算. 围绕波函数方法的核心模型 Hartree-Fock 方程, 林霖在最近的综述文章<sup>[64]</sup> 中详细介绍了其数值求解的一些进展.

密度泛函理论的核心思想是用电子密度

$$\rho(\mathbf{r}) = N \int \dots \int \psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_2 \dots d\mathbf{r}_N$$

作为体系的唯一变量, 其他物理量都表示成电子密度的函数<sup>[56, 70, 72]</sup>. 与波函数不同, 电子密度只是位置的函数. 于是, 原本的高维问题降至 3 维问题. 密度泛函的思想最早可以追溯到 1927 年 Thomas 和 Fermi 提出的 Thomas-Fermi 理论<sup>[41, 42, 86]</sup>. 1964 年, Hohenberg 和 Kohn 提出 Hohenberg-Kohn 定理和 Hohenberg-Kohn 变分定理<sup>[56]</sup>, 这两个定理为密度泛函理论提供了理论基础.

Hohenberg-Kohn 定理指出除一可加常数外, 体系所处外势由体系基态电子密度唯一确定. 对包括 Coulomb 势在内的一类外势, 周爱辉<sup>[103]</sup> 给出了 Hohenberg-Kohn 定理严格的数学证明. 最近, 周爱辉<sup>[105]</sup> 又对更一般的一类外势, 利用唯一延拓定理给出了 Hohenberg-Kohn 定理严格的数学证明. Hohenberg 和 Kohn 在 [56] 还给出了确定系统基态的途径, 即将能量泛函对电子密度函数进行变分. 对 Hohenberg-Kohn 变分定理的数学证明可见 [63].

然而, 能量泛函的具体形式未能给出. 1965 年, Kohn 和 Sham<sup>[61]</sup> 引入一个假想的无相互作用系统, 其导出的密度与原相互作用系统的基态密度相同, 从而将原本的多体 Schrödinger 方程转化为一个有效单体问题—Kohn-Sham 方程.

密度泛函已取得巨大成果, 在计算凝聚态物理、计算材料科学、量子化学、量子生物学和许多工业技术部门等得到了成功的应用. Kohn 和 Pople 也因其分别在第一原理计算的密度泛函理论与波函数方法方面的贡献而共同获得 1998 年诺贝尔化学奖.

根据密度泛函理论, 电子结构计算的核心变为 Kohn-Sham 方程或相应的带正交约束的 Kohn-Sham 能量泛函极小问题的求解. 尽管相比于原来的 Schrödinger 方程, 这些模型简单了不少, 但是其数值求解依然困难. Kohn-Sham 方程是一非线性算子方程, Kohn-Sham 算子含奇性、所需特征对数量多、特征值间间隙小或简并、Kohn-Sham 能量泛函的非凸性、轨道的正交约束等, 这些给 Kohn-Sham 方程或 Kohn-Sham 能量泛函极小问题的算法设计及数值分析带来很大困难和挑战. 电子结构计算方法研究涉及离散方法、非线性迭代法、代数特征值求解器、带正交约束优化问题求解、相关算法有效性及可靠性分析等多方面.

根据离散空间选取的不同, 现有的 Kohn-Sham 的离散格式大体上可分为三类: 平面波方法、局部基集法以及实空间方法. 这些方法各有其优势和不足, 其中平面波方法和局部基集法研究相对成熟, 而实空间方法尚处于发展阶段. 实空间方法主要包括有限差分方法、小波方法、有限体方法与有限元方法等. 实空间方法具有很好的局部性, 适应并行计算, 在高性能计算机上实现时能极小全局通信时间, 越来越受到重视. 尤其是有限元方法, 由于其基函数构造简单、基底完备、基函数具有很好的局部性, 易实现并行计算及自适应计算等, 越来越多地被应用到电子结构计算中. 关于离散方法的更多的介绍见文献 [33].

无论选用哪种离散方法, 最终都归结为一些大规模代数特征值问题的求解. 对代数特征值问题的求解, 目前已发展出有很多求解方法 (详细参见 [3, 78]), 如 Lanczos 方法 [12]、LOBPCG 方法 [60]、Jacobi-Davidson 方法 [24, 80] 等以及由这些方法加以改进得到的方法, 如显式重新开始 Lanczos 方法、隐式重新开始 Lanczos 方法、块 Jacobi-Davidson 等.

围绕电子结构计算的核心模型—Kohn-Sham 方程或相应的带正交约束的 Kohn-Sham 能量泛函极小问题, 人们展开了各种研究. 近二十年来, 我们课题组也围绕这一核心模型的高效离散及数值分析展开了长期的研究. 本文着重介绍我们课题组在电子结构计算的方法设计和理论分析方面的研究进展. 关于电子结构计算数值方法的更多介绍, 可参考综述性文章 [33, 34, 65, 79]. 关于电子结构计算的数学基础及数学问题可参考 [102, 104].

## 2. Kohn-Sham 模型

根据 Kohn-Sham 密度泛函理论, 多电子体系的总能可表示为 [61]

$$E(\Phi) = \int_{\mathbb{R}^3} \left( \frac{1}{2} \sum_{i=1}^N |\nabla \phi_i|^2 + V_{\text{ext}} \rho_{\Phi} + e_{\text{xc}}(\rho_{\Phi}) \right) + \frac{1}{2} \int_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_{\Phi}(\mathbf{r}) \rho_{\Phi}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (2.1)$$

其中  $\Phi = (\phi_1, \phi_2, \dots, \phi_N)$ ,  $\rho_{\Phi}(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2$  为电子密度 (为简单, 后文有时会省掉下标  $\Phi$ ), 这里我们假设每个占据态被 1 个电子占据, 从而占据 Kohn-Sham 轨道数为  $N$ ,  $e_{\text{xc}}(\rho_{\Phi})$  为交换关联能密度,  $V_{\text{ext}}$  为外势.

根据 Kohn-Sham 密度泛函理论, 体系基态可通过求解如下能量泛函极小问题得到

$$\begin{aligned} & \inf_{\Phi = (\phi_1, \phi_2, \dots, \phi_N) \in (H^1(\mathbb{R}^3))^N} E(\Phi) \\ & \text{s.t. } \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}, 1 \leq i, j \leq N. \end{aligned} \quad (2.2)$$

文献 [18, 84] 给出了能量泛函极小问题 (2.2) 解的存在性的一些讨论. 因此上述极小问题可写为

$$\begin{aligned} & \min_{\Phi = (\phi_1, \phi_2, \dots, \phi_N) \in (H^1(\mathbb{R}^3))^N} E(\Phi) \\ & \text{s.t. } \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}, 1 \leq i, j \leq N. \end{aligned} \quad (2.3)$$

如果能量泛函  $E(\Phi)$  可微, 上面能量泛函极小问题的求解可转化为 Euler-Lagrange 方程的求解: 求  $\phi_i(\mathbf{r}) \in H_0^1(\mathbb{R}^3)$ , 使得

$$\begin{cases} (-\frac{1}{2}\Delta + V_{\text{eff}}(\rho_{\Phi}))\phi_i(\mathbf{r}) = \sum_{j=1}^N \lambda_{ij} \phi_j(\mathbf{r}), \text{ in } \mathbb{R}^3, \\ \int_{\mathbb{R}^3} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N, \end{cases} \quad (2.4)$$

其中

$$V_{\text{eff}}(\rho_{\Phi}) = V_{\text{ext}} + V_{\text{H}}(\rho_{\Phi}) + V_{\text{xc}}(\rho_{\Phi}),$$

$$\Lambda = (\lambda_{ij})_{i,j=1}^N = \left( \int_{\mathbb{R}^3} \phi_j H_{\Phi} \phi_i \right)_{i,j=1}^N$$

为 Lagrange 乘子矩阵,  $H_{\Phi} : H^1(\mathbb{R}^3) \rightarrow H^{-1}(\mathbb{R}^3)$  为如下定义的 Kohn-Sham Hamiltonian 算子

$$\langle H_{\Phi} \phi, \varphi \rangle = \frac{1}{2} (\nabla \phi, \nabla \varphi) + (V_{\text{eff}}(\rho_{\Phi}) \phi, \varphi),$$

$\phi, \varphi \in H^1(\mathbb{R}^3)$ . 这里,

$$V_{\text{H}}(\rho_{\Phi})(\mathbf{r}) = \int_{\mathbb{R}^3} \frac{\rho_{\Phi}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$$

为 Hartree 势,  $V_{\text{xc}}(\rho_{\Phi}) = e'_{\text{xc}}(\rho_{\Phi})$  为交换关联势,  $V_{\text{ext}}$  取为核对电子的 Coulomb 吸引势, 即

$$V_{\text{ext}}(\mathbf{r}) = - \sum_{I=1}^M \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}. \quad (2.5)$$

由于电子密度  $\rho_{\Phi}$  在轨道  $\Phi$  做正交变换时保持不变, Lagrange 乘子矩阵  $\Lambda$  可被对角化. 具体地, 存在  $N$  阶正交矩阵  $U$  使得  $\Psi = \Phi U = (\psi_1, \psi_2, \dots, \psi_N)$  满足

$$\int_{\mathbb{R}^3} \psi_j H_{\Psi} \psi_i = \mu_i \delta_{ij}.$$

因此 Euler-Lagrange 方程 (2.4) 可变为

$$\begin{cases} \left( -\frac{1}{2} \Delta + V_{\text{eff}}(\rho_{\Psi}) \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}), \text{ in } \mathbb{R}^3, \\ \int_{\mathbb{R}^3} \psi_i(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N, \end{cases} \quad (2.6)$$

此即著名的 Kohn-Sham 方程. 它是一个定义在  $\mathbb{R}^3$  上的非线性特征值问题.

尽管相比于定义在  $\mathbb{R}^{3N}$  上的 Schrödinger 方程, 定义在  $\mathbb{R}^3$  上的 Kohn-Sham 方程求解的难度降低, 但它仍然充满各种困难与挑战. 比如, 它是一非线性算子特征值问题, 需要稳定的自洽场迭代格式; 核与电子间作用势  $V_{\text{ext}}$  含奇性, 从而这使得特征函数在离子实内部变化剧烈; 交换关联势  $V_{\text{xc}}$  并无精确解析表达式; 所需特征对数目会非常多, 特征值之间间隙小等.

## 2.1. 交换关联泛函选取

密度泛函理论导出的单电子 Kohn-Sham 方程是严格的, 没有作任何假设. 但是遗憾的是, 交换关联泛函  $e_{\text{xc}}(\rho)$  只是一个理论上存在的泛函, 其具体形式仍然是未知的. 而它包含了多粒子系统相互作用的全部复杂性, 在整个密度泛函理论中起着非常重要的作用. 在实际计算中, 我们常采用某种近似. 常见的近似有局域密度近似 (local density approximation, 简称 LDA) [61]、局域自旋密度近似 (local spin density approximation, 简称 LSDA)、广义梯度近似 (generalized gradient approximation, 简称 GGA) [38, 62, 74]、平均密度近似 (ADA) [51]、加权密度近似 (WDA) [51]、准粒子近似 [55]、杂化近似 [6, 7] 等近似方法. 交换关联能各种近似的比较详细的描述及比较可参见文献 [70]. 关于交换关联泛函的选取, 文献 [73, 75] 将其形容成爬 Jacob 天梯.

## 2.2. Coulomb 势处理

Kohn-Sham 方程中外势  $V_{ext}$  (见 (2.5)) 为一奇异算子. 算子的奇性导致波函数在离子实内部区域变化剧烈, 从而使得离散时需要在离子实内部需要用大量函数去逼近. 然而, 在很多情况下, 人们最关心的是价电子, 体系的物理和化学性质一般主要依赖于价电子而不是离子实内层的电子. 而恰恰是在价电子所在的离子实之间的区域, 波函数变换平滑, 与自由电子的平面波很相近. 由此人们提出赝势的概念, 即在离子实内部, 用光滑的势取代真实的势.

赝势不是唯一的, 它有多种具体形式. 目前, 在第一原理计算中最常用的是 Hamann 等<sup>[52]</sup> 提出后又经 Troullier 和 Martins<sup>[87]</sup> 进一步发展的模守恒赝势. 此外, 还有超软赝势 (ultrasoft pseudopotential)<sup>[8, 91]</sup>, 以及一些经验赝势和半经验赝势<sup>[95]</sup>. 关于赝势更详细的讨论参见文献 [70]

根据是否用赝势替换真实的 Coulomb 势 (2.5), 第一原理计算又分全势计算和赝势计算. 通常基于平面波离散的都是进行赝势计算, 如 Vasp<sup>[113]</sup>, Abinit<sup>[109]</sup> 及 Quantum Espresso<sup>[112]</sup>. 而基于原子轨道基函数方法离散的有全势计算, 如 Gaussian<sup>[110]</sup>, 亦有赝势计算, 如中国科技大学何力新课题组开发的 ABACUS<sup>[108]</sup>. 基于我们课题组长期以来的研究成果以及科学与工程计算国家重点实验室开发的并行自适应网格平台 PHG (Parallel Hierarchical Grid)<sup>[111]</sup>, 我们研制了一套第一原理实空间并行自适应计算程序 RealSPACES (Real Space Parallel Adaptive Calculation of Electronic Structure), 该程序能同时处理全势和赝势<sup>[34]</sup>.

## 2.3. 自洽场迭代

与原始的 Schrödinger 方程是一个线性特征值问题不同, Kohn-Sham 方程是一个定义在空间  $\mathbb{R}^3$  上的非线性特征值问题, 需要通过线性化迭代求解. 这在物理上称为自洽迭代算法. 自洽迭代通常通过电子密度或有效势迭代来实现. 这里, 我们只介绍基于电子密度的自洽迭代. 基于电子密度的自洽迭代求解步骤一般过程如下<sup>[70]</sup>.

---

### 算法 1 Kohn-Sham 方程的自洽场迭代

---

- 1: 给定  $\varepsilon > 0$ , 电子密度初值  $\rho_{in}$ . 记  $\rho_{out} \equiv 0$ .
- 2: **while**  $|||\rho_{out} - \rho_{in}||| > \varepsilon$  **do**
- 3: 通过某种方式由  $\rho_{in}$  和  $\rho_{out}$  得到新的  $\rho_{in}$ ;
- 4: 由  $\rho_{in}$  得到有效势  $V_{eff}(\rho_{in})$ , 然后求解线性特征值问题

$$\begin{cases} (-\frac{1}{2}\Delta + V_{eff}(\rho_{in}))\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \text{ in } \mathbb{R}^3, \\ \int_{\mathbb{R}^3} \psi_i(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N, \end{cases} \quad (2.7)$$

得到轨道  $\{\psi_i : i = 1, 2, \dots, N\}$ ;

- 5: 计算  $\rho_{out} = \sum_{i=1}^N |\psi_i|^2$ .
  - 6: **end while**
-

我们记

$$F(\rho_{in}) = \rho_{out} - \rho_{in}.$$

于是自洽场迭代的目的是寻找映射

$$F(\rho) = 0$$

的解. 我们记第  $m$  次迭代中的输入密度和输出密度分别为  $\rho_{in}^m$  和  $\rho_{out}^m$ . 在算法 1 中, 由  $\rho_{in}^m$  和  $\rho_{out}^m$  得到  $\rho_{in}^{m+1}$  的方法十分关键. 在实际计算中, 不同的方法决定着算法收敛与否以及算法收敛速度.

目前最常用的自洽场迭代方法是 Anderson 迭代<sup>[2]</sup>, 也被称为 Pulay 混合方法<sup>[76,77]</sup> 或者 DIIS (Direct Inversion Iterative Subspace) 方法. 此外, 包括一些拟牛顿法, 如 Broyden 迭代<sup>[10,59]</sup> 等, 也被经常使用. 近年来, 自洽场迭代的收敛性分析引起了人们的兴趣, 一些关于自洽场迭代算法的理论分析工作开始出现<sup>[11,68,69,96]</sup>. 这些工作中, 人们证明了当能量泛函的二阶导数一致有上界时, 如果占据态和非占据态之间的间隙大于某个常数, 则自洽场迭代收敛. 然而这些收敛性分析实际上并没有解决电子结构计算中的自洽场迭代的收敛性难题, 因为实际体系往往并不满足这个假设. 总的来说, 自洽场迭代的收敛性问题是一远未解决的难题, 人们在不断尝试一些加速自洽场迭代收敛的方法, 如<sup>[66,106]</sup>.

### 3. Kohn-Sham 模型的有限维离散及数值分析

基于密度泛函理论的第一原理电子结构计算取得了巨大成功, 但如何从数学角度理解电子结构计算的数值方法的可靠性和有效性却鲜有人涉足. 从理论角度理解数值逼近的收敛性和数值离散对计算结果的影响在实际计算中十分重要, 能为计算结果的可靠性和有效性提供理论依据. 然而, 由于 Kohn-Sham 模型本身的复杂性, 对 Kohn-Sham 模型有限维离散的数值分析工作非常少, 一直到最近十年才有了一些有限维逼近收敛性和收敛率方面的结果.

非线性特征值问题的数值分析最早的工作应是周爱辉在<sup>[100]</sup>中给出的. 该文给出了 Bose-Einstein 凝聚态的 Gross-Pitaevskii 方程的有限维逼近的收敛性. 接着, 周爱辉<sup>[101]</sup>研究了包括 Thomas-Fermi-von Weizsäcker 模型在内的刻画无轨道密度泛函理论模型的基态解有限维逼近. 除了收敛性分析外, 还得到了先验误差上界. Cancès 等<sup>[13]</sup>对凸能量泛函的 Thomas-Fermi-von Weizsäcker 模型得到了最优的先验误差估计. 2010 年 Suryanarayana 等<sup>[84]</sup>在一定条件下给出 Kohn-Sham 能量泛函有限元离散的收敛性. 2012 年 Cancès 等<sup>[14]</sup>在二阶最优性条件下对 Kohn-Sham 模型的平面波离散给出了先验误差分析. 陈华杰等人<sup>[18]</sup>于 2011 年对 Kohn-Sham 方程的任意有限维离散在局部同胚的假设下进行了系统的先验误差分析, 他们证明了在一定条件下任意有限维离散的极限都收敛到基态. 随后, 我们<sup>[17]</sup>又进一步对 Kohn-Sham 方程的有限元离散进行了后验误差分析, 这是关于 Kohn-Sham 方程有限维离散的首个后验误差分析结果. 下面我们简要介绍下 Kohn-Sham 方程有限维离散的先验误差分析及有限元离散的后验误差分析结果.

#### 3.1. 一些记号、定义及假设

物理上 Kohn-Sham 方程定义在  $\mathbb{R}^3$  上, 然而, 对于团簇等孤立体系, 由于基态波函数和电荷密度都具有指数衰减性<sup>[82]</sup>, 实际计算时常在一有限区域  $\Omega \subset \mathbb{R}^3$  上进行. 在后文叙述中, 我们令  $\Omega = \mathbb{R}^3$  或者  $\Omega \subset \mathbb{R}^3$  为一多边形区域,  $\partial\Omega$  是其边界. 我们使用 Sobolev 空间中的标准符

号  $W^{s,p}(\Omega)$  以及相应的范数  $\|\cdot\|_{s,p,\Omega}$  和半范数  $|\cdot|_{s,p,\Omega}$  ( $s=1,2,\dots, 1 \leq p \leq \infty$ ) [1]. 当  $p=2$  时, 我们记  $H^s(\Omega) = W^{s,2}(\Omega)$  及相应的范数为  $\|\cdot\|_{s,\Omega} = \|\cdot\|_{s,2,\Omega}$ , 且  $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$ , 其中  $v|_{\partial\Omega} = 0$  为迹意义下的.  $H^{-1}(\Omega)$  为  $H_0^1(\Omega)$  的对偶空间. 后文中  $\mathbb{N}$  表示所有正整数集合,  $\mathbb{N}_0$  表示所有非负整数集合.

字母  $C$  (或者带上标) 表示一个与网格尺寸无关的常数, 并且在不同的地方可能表示不同的大小. 为方便计, 我们将引入记号  $\lesssim$ ,  $A \lesssim B$  表示存在  $C$  使得  $A \leq CB$ .

为后文中描述理论结果方便, 我们还介绍几个函数空间. 令  $\mathcal{P}(p, (c_1, c_2))$  为满足如下条件的函数类:

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

其中  $c_1 \in \mathbb{R}$ ,  $c_2, p \in [0, \infty)$ . 令 Hilbert 空间  $\mathcal{H} = (H_0^1(\Omega))^N$ , 其上  $H^1$  内积定义如下

$$(\Phi, \Psi)_{\mathcal{H}} = \sum_{i=1}^N \int_{\Omega} \nabla \phi_i \cdot \nabla \psi_i + \phi_i \cdot \psi_i, \quad \Phi = (\phi_1, \phi_2, \dots, \phi_N), \Psi = (\psi_1, \psi_2, \dots, \psi_N) \in \mathcal{H}.$$

令  $\mathbb{Q}$  为一带正交约束的子空间, 也称 Stiefel 流形:

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

其中

$$\Phi^T \Psi = \left( \int_{\Omega} \phi_i \psi_j \right)_{i,j=1}^N \in \mathbb{R}^{N \times N} \quad (3.1)$$

为  $\Phi$  和  $\Psi$  的内积矩阵. 对  $\Phi \in \mathcal{H}$  及子区域  $\omega \subset \Omega$ , 我们记

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

后面的讨论中还会用到下面一些集合:

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

对任意  $\Phi \in \mathbb{Q}$ , 我们可将  $\mathcal{H}$  分裂为如下三个子空间的直和 (参见 [14]):

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

这里  $\mathcal{S}_{\Phi} = \Phi \mathcal{S}^{N \times N}$ ,  $\mathcal{A}_{\Phi} = \Phi \mathcal{A}^{N \times N}$ , 且  $\mathcal{T}_{\Phi} = \{\tilde{\Phi} \in \mathcal{H} : \tilde{\Phi}^T \Phi = 0 \in \mathbb{R}^{N \times N}\}$ .

定义算子  $\mathcal{F} : \mathbb{R}^{N \times N} \times \mathcal{H} \rightarrow \mathcal{H}^*$  如下:

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

$\mathcal{F}$  在  $(\Lambda, \Phi)$  处的关于  $\Phi$  的 Fréchet 导数  $\mathcal{F}'_{\Phi}(\Lambda, \Phi) : \mathcal{H} \rightarrow \mathcal{H}^*$  定义为

$$\begin{aligned} \langle \mathcal{F}'_{\Phi}(\Lambda, \Phi) \Psi, \Gamma \rangle &= \frac{1}{2} E''(\Phi)(\Psi, \Gamma) - \sum_{i,j=1}^N (\lambda_{ji} \psi_j, \gamma_i) \\ &= \sum_{i=1}^N (H_{\Phi} \psi_i - \sum_{j=1}^N \lambda_{ji} \psi_j, \gamma_i) + 2 \sum_{i,j=1}^N (e''_{xc}(\rho_{\Phi}) \phi_i \psi_i, \phi_j \gamma_j) + \sum_{i,j=1}^N 2D(\phi_i \psi_i, \phi_j \gamma_j), \end{aligned}$$



其中

$$D(f, g) = \int_{\Omega} \int_{\Omega} \frac{f(\mathbf{r})g(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'.$$

下面几个假设在后面的分析中要用到<sup>[18]</sup>:

**A1**  $|e'_{xc}(t)| + |te''_{xc}(t)| \in \mathcal{P}(p_1, (c_1, c_2))$  对  $p_1 \in [0, 2]$  成立.

**A2** 存在常数  $\alpha \in (0, 1]$  使得

$$|e''_{xc}(t)| + |te'''_{xc}(t)| \lesssim 1 + t^{\alpha-1} \quad \forall t > 0.$$

**A3**  $(\Lambda, \Phi)$  为 (2.4) 的解, 且存在依赖于  $(\Lambda, \Phi)$  的常数  $\beta > 0$  使得

$$\inf_{\Gamma \in \mathcal{T}_{\Phi}} \sup_{\Psi \in \mathcal{T}_{\Phi}} \frac{\langle \mathcal{F}'_{\Phi}(\Lambda, \Phi)\Psi, \Gamma \rangle}{\|\Psi\|_{1, \Omega} \|\Gamma\|_{1, \Omega}} \geq \beta. \quad (3.2)$$

### 3.2. Kohn-Sham 模型的有限维离散

根据 Kohn-Sham 密度泛函理论, 体系基态可通过求解如下带正交约束 Kohn-Sham 能量泛函极小问题得到

$$\min \{E(\Phi) : \Phi \in \mathbb{Q}\}, \quad (3.3)$$

亦可通过求解如下 Kohn-Sham 方程得到

$$\begin{cases} \left(-\frac{1}{2}\Delta + V_{\text{eff}}(\rho_{\Psi})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \text{ in } \Omega, \\ \int_{\Omega} \psi_i(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N. \end{cases} \quad (3.4)$$

定义双线性型  $a(\cdot; \cdot, \cdot)$  如下:

$$a(\rho_{\Phi}; \phi, \varphi) = \langle H_{\Phi}\phi, \varphi \rangle = \frac{1}{2}(\nabla\phi, \nabla\varphi) + (V_{\text{eff}}(\rho_{\Phi})\phi, \varphi) \quad \forall \phi, \varphi \in H^1(\Omega).$$

我们得到方程 (2.4) 的弱形式如下

$$\begin{cases} a(\rho_{\Phi}; \phi_i, \varphi) = \left(\sum_{j=1}^N \lambda_{ij}\phi_j, \varphi\right) \quad \forall \varphi \in H^1(\Omega), \\ \int_{\Omega} \phi_i(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N. \end{cases} \quad (3.5)$$

类似地, 方程 (3.4) 对应的弱形式如下: 求  $(\varepsilon_i, \psi_i) \in \mathbb{R} \times H_0^1(\Omega) (i = 1, \dots, N)$ , 使得

$$\begin{cases} a(\rho_{\Phi}; \psi_i, \varphi) = (\varepsilon_i\psi_i, \varphi) \quad \forall \varphi \in H^1(\Omega), \\ \int_{\Omega} \psi_i(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N. \end{cases} \quad (3.6)$$

注意到

$$E(\Phi) = E(\Phi U) = E\left(\left(\sum_{j=1}^N u_{ji}\phi_j\right)_{i=1}^N\right) \quad \forall U = (u_{ij})_{i,j=1}^N \in \mathcal{O}^{N \times N}, \quad (3.7)$$

这里  $\mathcal{O}^{N \times N}$  为  $N$  阶正交矩阵集合. 从 (3.7) 知如果  $\Phi$  为能量极小问题 (3.3) 的解, 则  $\Phi U$  亦为 (3.3) 的解, 其中  $U$  为任意  $N$  阶正交矩阵.

对任意  $\Phi \in \mathcal{H}$ , 我们定义等价类

$$[\Phi] = \{\Phi U : U \in \mathcal{O}^{N \times N}\}.$$

由于 Kohn-Sham 能量泛函非凸, 且在轨道正交变换下不变, 因此 Kohn-Sham 基态不唯一. 我们定义 Kohn-Sham 基态集合如下

$$\Theta = \left\{ (\Lambda, \Phi) \in \mathbb{R}^{N \times N} \times \mathbb{Q} : E(\Phi) = \min_{\tilde{\Phi} \in \mathbb{Q}} E(\tilde{\Phi}) \text{ 且 } (\Lambda, \Phi) \text{ 是 (3.5) 的解} \right\}, \quad (3.8)$$

以及基态解集

$$\mathcal{G} = \{\Phi \in \mathbb{Q} : E(\Phi) = \min_{\tilde{\Phi} \in \mathbb{Q}} E(\tilde{\Phi})\}.$$

定义满足 (3.5) 的态的集合为:

$$\mathcal{W} = \{(\Lambda, \Phi) \in \mathbb{R}^{N \times N} \times \mathcal{H} : (\Lambda, \Phi) \text{ 是 (3.5) 的解}\}.$$

易得  $\Theta \subsetneq \mathcal{W}$ . 在后面的讨论中, 我们有时会用假设基态能量与其他态对应的能量之间有间隙:

$$\min_{\tilde{\Phi} \in \mathbb{Q}} E(\tilde{\Phi}) < \inf_{(M, \tilde{\Phi}) \in \mathcal{W} \setminus \Theta} E(\tilde{\Phi}), \quad (3.9)$$

这个假设是合理的<sup>[17]</sup>.

问题 (3.3)、(3.5) 及 (3.6) 定义在无穷维空间  $H_0^1(\Omega)$  上, 通常我们需要对其作有限维逼近. 具体地, 构造  $H_0^1(\Omega)$  的有限维子空间  $V_n \subset H_0^1(\Omega)$ , 考虑问题 (3.3)、(3.5) 及 (3.6) 在  $V_n$  上的逼近, 分别得到如下离散问题:

$$\min \{E(\Phi_n) : \Phi_n \in (V_n)^N \cap \mathbb{Q}\}, \quad (3.10)$$

$$\begin{cases} a(\rho_{\Phi_n}; \phi_{i,n}, \varphi) = \left( \sum_{j=1}^N \lambda_{ij,n} \phi_{j,n} \right) \varphi \quad \forall \varphi \in V_n, \\ \int_{\Omega} \phi_{i,n}(\mathbf{r}) \phi_{j,n}(\mathbf{r}) d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N, \end{cases} \quad (3.11)$$

其中 Lagrange 乘子矩阵

$$\Lambda_n = (\lambda_{ij,n})_{i,j=1}^N = \left( \int_{\Omega} \phi_{j,n} H_{\Phi_n} \phi_{i,n} \right)_{i,j=1}^N,$$

及求  $(\varepsilon_{i,n}, \psi_{i,n}) \in \mathbb{R} \times V_n (i = 1, \dots, N)$ , 使得

$$\begin{cases} a(\rho_{\Psi_n}; \psi_{i,n}, \varphi) = (\varepsilon_{i,n} \psi_{i,n}, \varphi) \quad \forall \varphi \in V_n, \\ \int_{\Omega} \psi_{i,n}(\mathbf{r}) \psi_{j,n}(\mathbf{r}) d\mathbf{r} = \delta_{ij}, i, j = 1, 2, \dots, N. \end{cases} \quad (3.12)$$

这里,  $\Phi_n = (\phi_{1,n}, \dots, \phi_{N,n})$ ,  $\Psi_n = (\psi_{1,n}, \dots, \psi_{N,n})$ .

定义 Kohn-Sham 基态有限维逼近集合如下:

$$\Theta_n = \left\{ (\Lambda_n, \Phi_n) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap (V_n)^N) : E(\Phi_n) = \min_{\tilde{\Phi} \in \mathbb{Q} \cap (V_n)^N} E(\tilde{\Phi}), (\Lambda_n, \Phi_n) \text{ 是 (3.11) 的解} \right\}$$

及基态解的有限维逼近集合

$$\mathcal{G}_n = \{ \Phi_n \in \mathbb{Q} \cap (V_n)^N : E(\Phi_n) = \min_{\tilde{\Phi}_n \in \mathbb{Q} \cap (V_n)^N} E(\tilde{\Phi}_n) \}.$$

### 3.3. 先验误差分析

陈华杰等<sup>[18]</sup> 对任意满足如下逼近性条件 (见文献 [18] 的 (3.1)) 的有限维子空间  $V_n \in H_0^1(\Omega)$  给出了 Kohn-Sham 模型有限维逼近的先验误差分析

$$\lim_{n \rightarrow \infty} \inf_{\psi \in V_n} \|\psi - \phi\|_{1,\Omega} = 0 \quad \forall \phi \in H_0^1(\Omega). \quad (3.13)$$

逼近性 (3.13) 是一个很弱的假设, 常用的典型有限维空间如平面波函数张成的有限维空间、分片多项式有限元函数空间均满足这一条件<sup>[21]</sup>.

文献 [18] 证明了如下收敛性定理, 该定理表明所有有限维逼近解的极限点均是基态解.

**定理 1 (收敛性).**

$$\begin{aligned} \lim_{n \rightarrow \infty} d_{\mathcal{H}}(\Theta_n, \Theta) &= 0, \\ \lim_{n \rightarrow \infty} E_n &= \min_{\Phi \in \mathbb{Q}} E(\Phi), \end{aligned} \quad (3.14)$$

其中  $E_n = E(\Phi_n)$ ,  $\Phi_n \in \mathcal{G}_n$ , 集合  $X$  到  $Y (X, Y \subset \mathbb{R}^{N \times N} \times \mathcal{H})$  的距离定义如下

$$d_{\mathcal{H}}(X, Y) = \sup_{(\Lambda, \Phi) \in X} \inf_{(\mu, \Psi) \in Y} (|\Lambda - \mu| + \|\Phi - \Psi\|_{1,\Omega}). \quad (3.15)$$

此外, 文献 [18] 还证明了基态能量有限维逼近是二次收敛的.

**定理 2 (收敛率).** 令  $E_0$  为问题 (3.3) 的极小能,  $E_n$  为离散问题 (3.10) 的极小能, 对所有  $\Phi \in \mathcal{G}$  有  $E_0 = E(\Phi)$ , 对所有  $\Phi_n \in \mathcal{G}_n$ , 有  $E_n = E(\Phi_n)$ . 在假设 A1 成立的条件下, 有

$$|E_0 - E_n| \lesssim \mathcal{D}_{\mathcal{H}}^2(\mathcal{G}, \mathcal{G}_n), \quad (3.16)$$

其中集合  $X$  到  $Y (X, Y \subset \mathcal{H})$  的距离定义如下

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

对  $(\Lambda, \Phi) \in \mathbb{R}^{N \times N} \times \mathcal{H}$ , 定义  $B_{\delta}(\Lambda, \Phi)$  如下

$$B_{\delta}(\Lambda, \Phi) := \{(\mu, \Psi) \in \mathbb{R}^{N \times N} \times \mathcal{H} : |\Lambda - \mu| + \|\Phi - \Psi\|_{1,\Omega} \leq \delta\}.$$

令

$$X_{\Phi,n} = \mathcal{S}^{N \times N} \times (V_n \cap (\mathcal{S}_{\Phi} \oplus \mathcal{T}_{\Phi})),$$

文献 [18] 给出了如下先验误差估计.

**定理 3 (收敛率).** 令  $(\Lambda, \Phi) \in \Theta$ . 如果假设 A2 成立, 且  $(\Lambda, \Phi)$  满足假设 A3, 则存在  $\delta > 0$ , 使得对足够大的  $n$ , (3.11) 有唯一局部解  $(\Lambda_n, \Phi_n) \in X_{\Phi, n} \cap B_\delta((\Lambda, \Phi))$ , 且有如下估计

$$\|\Phi - \Phi_n\|_{1, \Omega} \lesssim \inf_{\tilde{\Phi} \in V_n} \|\Phi - \tilde{\Phi}\|_{1, \Omega}, \quad (3.17)$$

$$|\Lambda_n - \Lambda| \lesssim \|\Phi_n - \Phi\|_{1, \Omega}^2 + \|\Phi_n - \Phi\|_{0, \Omega}, \quad (3.18)$$

$$\|\Phi - \Phi_n\|_{0, \Omega} \lesssim r_n \|\Phi - \Phi_n\|_{1, \Omega}, \quad (3.19)$$

其中当  $n \rightarrow \infty$  时  $r \rightarrow 0$ .

### 3.4. 后验误差分析

有限元离散是一类典型的实空间离散方法, 其基函数构造简单、基底完备, 易实现自适应计算, 已被成功应用到各类科学工程计算相关问题离散中. 近年来越来越多的研究者将其引入电子结构计算中 [4, 16, 23, 39, 45, 54, 58, 83, 93].

记  $\mathcal{T}_h(\Omega)$  为  $\Omega$  的一个正则协调的有限元网格剖分: 存在常数  $\gamma^*$  使得

$$\frac{h_\tau}{\rho_\tau} \leq \gamma^* \quad \forall \tau \in \mathcal{T}_h, \quad (3.20)$$

其中  $h_\tau$  为单元  $\tau$  的直径,  $\tau \in \mathcal{T}_h$ ,  $\rho_\tau$  为  $\tau$  的内接圆的直径,  $h = \max\{h_\tau : \tau \in \mathcal{T}_h\}$ .  $\mathcal{E}_h$  为  $\mathcal{T}_h$  内部面 (或边) 的集合.

定义  $S^{h,r}(\Omega)$  为  $\Omega$  上相应于剖分  $\mathcal{T}_h(\Omega)$  的一连续函数空间, 它满足对任一函数  $v \in S^{h,r}(\Omega)$ ,  $v$  限制到每一个单元  $T$  上是一个次数不超过  $r$  的多项式, 即:

$$S^{h,r}(\Omega) = \{v \in C(\bar{\Omega}) : v|_T \in P_T^r \quad \forall T \in \mathcal{T}_h(\Omega)\},$$

其中  $P_T^r$  是单元  $T$  上的  $r$  次多项式空间. 为简单, 我们记  $S^{h,r}(\Omega)$  为  $S^h(\Omega)$ . 我们选  $S^h(\Omega)$  为标准的 Lagrange 有限元空间. 令  $S_0^h(\Omega) = S^h(\Omega) \cap H_0^1(\Omega)$ .

将 (3.10), (3.11) 及 (3.12) 中的  $V_n$  取为  $S_0^h(\Omega)$  即得到 Kohn-Sham 方程的有限元离散形式. 为与标准有限元离散符号一致, 在讨论 Kohn-Sham 方程有限元离散时前面所有有限维逼近中的下标  $n$  改为网格尺寸  $h$ , 其他不变. 即得到如下问题

$$\min \{E(\Phi_h) : \Phi_h \in (S_0^h(\Omega))^N \cap \mathbb{Q}\}, \quad (3.21)$$

$$\begin{cases} a(\rho_{\Phi_h}; \phi_{i,h}, \varphi) = \left( \sum_{j=1}^N \lambda_{ij,h} \phi_{j,h}, \varphi \right) \quad \forall \varphi \in S_0^h(\Omega), \\ \int_{\Omega} \phi_{i,h}(\mathbf{r}) \phi_{j,h}(\mathbf{r}) d\mathbf{r} = \delta_{ij}, \quad i, j = 1, 2, \dots, N, \end{cases} \quad (3.22)$$

及求  $(\varepsilon_{i,h}, \psi_{i,h}) \in \mathbb{R} \times V_h (i = 1, \dots, N)$ , 使得

$$\begin{cases} a(\rho_{\Psi_h}; \psi_{i,h}, \varphi) = (\varepsilon_{i,h} \psi_{i,h}, \varphi) \quad \forall \varphi \in S_0^h(\Omega), \\ \int_{\Omega} \psi_{i,h}(\mathbf{r}) \psi_{j,h}(\mathbf{r}) d\mathbf{r} = \delta_{ij}, \quad i, j = 1, 2, \dots, N. \end{cases} \quad (3.23)$$

关于线性特征值问题的有限元后验误差分析研究较多, 但是 Kohn-Sham 方程的后验误差分析是在文献 [17] 首次给出的. 我们构造了如下的残量型误差估计子. 首先, 构造单元残量  $\mathcal{R}_\tau(\Phi_h)$  及单元跳量  $J_e(\Phi_h)$  如下

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

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

这里  $e$  为单元  $\tau_1$  和  $\tau_2$  的公共面, 其外法向分别是  $\vec{n}_1$  和  $\vec{n}_2$ .

令  $\omega_h(e)$  为与面  $e$  有交集的单元的集合,  $\omega_h(\tau)$  为与单元  $\tau$  有公共面的所有单元的集合. 对  $\tau \in \mathcal{T}_h$ , 定义局部误差指示子  $\eta_h(\Phi_h, \tau)$  及振荡  $\text{osc}_h(\Phi_h, \tau)$  如下

$$\eta_h^2(\Phi_h, \tau) = h_\tau^2 \|\mathcal{R}_\tau(\Phi_h)\|_{0,\tau}^2 + \sum_{e \in \mathcal{E}_h, e \subset \partial\tau} h_e \|J_e(\Phi_h)\|_{0,e}^2, \quad (3.24)$$

$$\text{osc}_h(\Phi_h, \tau) = h_\tau \|\mathcal{R}_\tau(\Phi_h) - \overline{\mathcal{R}_\tau(\Phi_h)}\|_{0,\tau}, \quad (3.25)$$

这里  $\bar{w}$  为的  $w \in L^2(\Omega)$  在  $\tau$  或  $e$  上  $k$  次多项式空间上的  $L^2$ - 投影. 给定子区域  $\omega \subset \Omega$ , 定义其上的误差估计子  $\eta_h(\Phi_h, \omega)$  及振荡  $\text{osc}_h(\Phi_h, \omega)$  如下

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

文献 [17] 对 Kohn-Sham 模型有限元逼近误差的给出了如下误差估计.

**定理 4.** 设  $h_0 \ll 1$  且  $h \in (0, h_0]$ . 令  $(\Lambda, \Phi) \in \Theta$ . 如果假设 **A2** 和 **A3** 满足, 则存在只依赖网格正则性常数  $\gamma^*$  (见 (3.20)) 的正常数  $C_1, C_2$  及  $C_3$ , 使得

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

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

其中  $(\Lambda_h, \Phi_h) \in X_{\Phi,h}$  为满足 (3.17)-(3.19) 的有限元逼近解.

进一步, 我们 [17] 还给出了基态有限元逼近的集合到基态集合的距离给出了如下后验估计.

**定理 5.** 设  $h_0 \ll 1$  且  $h \in (0, h_0]$ . 令  $(\Psi_h, \Psi_h)$  为 (3.23) 的解. 如果假设 **A2** 和 **A3** 均满足, 则

$$d_{\mathcal{H}}^2(\Theta_{\Psi_h}, \Theta) \lesssim \eta_h^2(\Psi_h, \Omega),$$

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

这里

$$\Theta_{\Psi_h} = \{(\Lambda_h, \Phi_h) \in \mathbb{R}^{N \times N} \times (\mathbb{Q} \cap (S_0^h(\Omega))^N) : \Phi_h \in [\Psi_h], \quad \Lambda_h = \Phi_h^T H_{\Phi_h} \Phi_h\}$$

且  $\Theta_{\Psi_h} \subseteq \Theta_h$ .

## 4. Kohn-Sham 方程高效离散格式

### 4.1. 自适应有限元方法

由于 Coulomb 势含奇性, 不管是全势计算还是采用赝势计算, Kohn-Sham 轨道在离子实内部剧烈振荡, 而在远离离子实的区域则变化缓慢 [5, 25, 50]. Kohn-Sham 方程解的这一特点使得采用多分辨离散或自适应离散非常自然也非常必要. 尽管有不少关于使用自适应有限元方法来进行电子结构计算的工作 [4, 23, 25, 54, 83, 88, 89, 93, 98], 关于 Kohn-Sham 模型的自适应有限元离散的数值分析结果直到我们的文章 [17] 才首次给出. 在此之前, 文献 [27, 30] 对线性特征值问题的自适应有限元逼近进行了分析, 文献 [19, 20] 给出了无轨道密度泛函模型的自适应有限元数值分析.

基于网格加密的自适应有限元方法的基本过程如下 [15, 17, 20, 27, 30]:

求解  $\rightarrow$  估计  $\rightarrow$  标记  $\rightarrow$  加密.

- **求解:** 在给定网格  $\mathcal{T}_{h_k}$  (后面简记为  $\mathcal{T}_k$ ) 求出原问题的有限元逼近解.
- **估计:** 对给定网格剖分  $\mathcal{T}_k$ , 根据求解步求解出来的有限元逼近解, 构造有限元后验误差估计子, 对解逼近的好坏作一个评估.
- **标记:** 根据计算得到的有限元后验误差估计子, 根据某种策略挑选出需要加密的单元, 也就是逼近得不好的单元. 常见的策略包括最大加密策略 [48, 49]、Dörfler 加密策略 [37] 等.

#### 最大加密策略

1. 给定加密参数  $0 < \theta < 1$ .
2. 挑选出所有满足如下条件的单元并标记以待加密

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

#### Dörfler 标记策略

1. 给定加密参数  $0 < \theta < 1$ .
2. 挑选出子集  $\mathcal{M}_k \subset \mathcal{T}_k$  使得其满足如下条件

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

- **加密:** 给定剖分  $\mathcal{T}_k$  和标记待加密的集合  $\mathcal{M}_k$ , 将  $\mathcal{M}_k$  中的单元至少加密一次得到一个新的协调网格剖分  $\mathcal{T}_{k+1}$ .

也就是说, 在某一给定网格上求出有限元逼近解, 由这些有限元解计算后验误差估计子, 根据某种标记策略及所得后验误差估计子找出需要加密的单元, 再根据某种加密方法加密这些单元得到新的网格, 然后在新的网格上再求出有限元逼近解; 重复前面的过程, 直到达到某一事先给定的精度.

将上述过程应用到 Kohn-Sham 方程, 并选取前一节定义的残量型后验误差估计子 (3.24), 分别选取最大加密策略和 Dörfler 策略, 我们便得到如下的 Kohn-Sham 方程的自适应有限元算法<sup>[17]</sup>.

---

#### 算法 2 Kohn-Sham 方程的自适应有限元算法 I

---

- 1: 给定初始网格  $\mathcal{T}_0$ , 并置  $k = 0$ .
  - 2: **while** 未收敛 **do**
  - 3: 在  $\mathcal{T}_k$  上求解 Kohn-Sham 方程得到  $(\mu_{i,k}, \psi_{i,k})(i = 1, \dots, N)$ .
  - 4: 计算后验误差指示子  $\eta_k(\Phi_k, \tau), \forall \tau \in \mathcal{T}_k$ .
  - 5: 由最大加密策略构造子集  $\mathcal{M}_k \subset \mathcal{T}_k, k = k + 1$ .
  - 6: **end while**
- 

---

#### 算法 3 Kohn-Sham 方程的自适应有限元算法 II

---

- 1: 给定初始网格  $\mathcal{T}_0$ , 并置  $k = 0$ .
  - 2: **while** 未收敛 **do**
  - 3: 在  $\mathcal{T}_k$  上求解 Kohn-Sham 方程得到  $(\mu_{i,k}, \psi_{i,k})(i = 1, \dots, N)$ .
  - 4: 计算后验误差指示子  $\eta_k(\Phi_k, \tau), \forall \tau \in \mathcal{T}_k$ .
  - 5: 由 Dörfler 策略构造子集  $\mathcal{M}_k \subset \mathcal{T}_k, k = k + 1$ .
  - 6: **end while**
- 

## 4.2. 收敛性及收敛率

对算法 2, 我们给出了收敛性分析结果<sup>[17]</sup>.

**定理 6 (收敛性定理).** 设初始网格  $\mathcal{T}_0$  足够细,  $\{\Theta_k\}_{k \in \mathbb{N}_0}$  为由算法 2 得到的一序列. 如果条件 (3.9) 及假设 A1 满足, 则

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

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

定理 6 表明 Kohn-Sham 方程自适应有限元逼近均收敛到基态. 最近, 杨斌等人<sup>[97]</sup> 去掉了初始网格尺寸足够小这一条件证明了同样的结论.

在介绍后面的结果前, 我们先介绍一些概念. 对  $(\Lambda, \Phi) \in \Theta$  及  $\Phi_h \in V_h$ , 我们称等价类  $[\Phi_h]$  逼近另一等价类  $[\Phi]$ , 如果

$$\mathcal{D}_{\mathcal{H}}([\Phi_h], [\Phi]) < \mathcal{D}_{\mathcal{H}}([\Phi_h], [\tilde{\Phi}]), \quad \forall (\tilde{\Lambda}, \tilde{\Phi}) \in \Theta \text{ 且 } [\Phi] \neq [\tilde{\Phi}].$$

我们称一子列  $\{[\Psi_k]\}_{k \in \mathbb{N}_0}$  收敛到等价类  $[\Phi]$  指的是存在一列酉矩阵  $\{U_k\}_{k \in \mathbb{N}_0}$ , 使得

$$\lim_{k \rightarrow \infty} \|\Psi_k U_k - \Phi\|_{1, \Omega} = 0.$$

对使用 Dörfler 策略的算法 3, 我们进一步得到了算法的误差估计 [17].

**定理 7 (误差下降率).** 令  $\theta \in (0, 1)$  且  $h_0 \ll 1$ . 令  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  为算法 3 得到的一列有限元逼近解. 如果  $[\Psi_{k_i}]$  逼近  $[\Phi]$ , 其中  $(\Lambda, \Phi) \in \Theta$ , 设记  $k_{i+1} (> k_i)$  为序列  $\{[\Psi_k]\}$  中最近一个逼近  $[\Phi]$  的解的迭代指标. 如果假设 A2 成立且  $(\Lambda, \Phi)$  满足假设 A3, 则

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

这里  $\Phi_{k_{i+1}} \in X_{\Phi, k_{i+1}}$  和  $\Phi_{k_i} \in X_{\Phi, k_i}$  满足 (3.17)-(3.19) (其中  $h$  分别换成  $h_{k_{i+1}}$  及  $h_{k_i}$ ),  $\gamma > 0$  和  $\xi \in (0, 1)$  为只依赖常数  $c_a$ 、正则性参数  $\gamma^*$  以及标记参数  $\theta$  的常数.

由定理 6 及定理 7 有如下收敛率定理, 具体证明可见 [17].

**定理 8 (收敛率).** 令  $\theta \in (0, 1)$  且  $h_0 \ll 1$ . 令  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  为算法 3 得到的一列有限元逼近解,  $\{[\Psi_{k_i}]\}_{i \in \mathbb{N}_0}$  为收敛到同一等价类  $[\Phi]$  的子列,  $(\Lambda, \Phi) \in \Theta$ . 如果假设 A2 成立且  $(\Lambda, \Phi)$  满足假设 A3, 则

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

这里  $\Phi_{k_{i+1}} \in X_{\Phi, k_{i+1}}$  和  $\Phi_{k_i} \in X_{\Phi, k_i}$  满足 (3.17)-(3.19), 只是其中的  $h$  分别换成  $h_{k_{i+1}}$  及  $h_{k_i}$ ,  $\gamma > 0$  和  $\xi \in (0, 1)$  为只依赖常数  $c_a$ 、正则性参数  $\gamma^*$  以及标记参数  $\theta$  的常数. 因此, 算法 3 的第  $k_m$  次迭代解满足

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

及

$$|\Lambda - \Lambda_{k_m}| \lesssim \xi^{2m}.$$

进一步, 有

$$d_{\mathcal{H}}(\Theta_{\Psi_{k_m}}, \Theta) \lesssim \xi^{2m}. \quad (4.1)$$

定理 8 表明 Kohn-Sham 方程自适应有限元逼近以一定收敛率收敛到体系基态. 进一步, 文献 [17] 还证明了自适应有限元逼近的最优复杂度. 为此, 引入函数类

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

其中  $\gamma > 0$  为一常数, 且

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

$\mathcal{T} \subset \mathcal{T}_0$  指的是  $\mathcal{T}$  为  $\mathcal{T}_0$  加密得到. 注意到当  $\gamma > 0$  时,  $\mathcal{A}_{\gamma}^s = \mathcal{A}_1^s$ . 因此, 为简单计我们记  $\mathcal{A}_{\gamma}^s$  为  $\mathcal{A}^s$ , 记  $|\Phi|_{s, \gamma}$  为  $|\Phi|_s$ .

在适当假设下, 文献 [17] 给出了算法的最优复杂性估计.

**定理 9.** 令  $\theta \in (0, 1)$  且  $h_0 \ll 1$ . 假设 A2 满足且在不考虑正交变换下 (3.5) 有  $m$  个解, 记为  $[\Phi^{(l)}] (l = 1, \dots, m)$ , 其中  $m \in \mathbb{N}$  且可取为  $\infty$ . 令  $\{\Psi_k\}_{k \in \mathbb{N}_0}$  为由算法 3 产生的一列有限元解. 则成立

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



其中  $\Phi^l \in [\Phi^{(l)}] \cap \mathcal{A}^s$  满足假设 (3.2),  $\Phi_{k_{n_l}}^l \in X_{\Phi^l, k_{n_l}}$  满足 (3.17) 和 (3.19), 只是其中  $h$  替换为  $h_{k_{n_l}}$ .

### 4.3. 并行轨道更新方法

现有的第一原理电子结构计算大都从 Kohn-Sham 方程出发, 通过自洽场迭代及有限维离散, 最终归结为一些大规模代数特征值问题的求解. 然而, 现有代数特征值问题求解算法通常计算量巨大且并行可扩展性低, 从而制约了电子结构计算的规模. 事实上, 大规模特征值问题的求解一直是第一原理电子结构计算的瓶颈.

我们近年来提出了第一原理电子结构计算的并行轨道更新方法<sup>[26, 107]</sup>. 其基本思想为通过迭代技巧将大规模特征值问题的求解分解为一系列相互独立的大规模源问题的求解以及一些刚度矩阵几乎对角的小规模特征值问题的求解, 从而从源头上避免了大规模代数特征值问题的产生, 转而求解一系列相互独立的源问题以及一些小规模的特征值问题.

算法框架如下, 具体细节见 [26].

---

#### 算法 4 并行轨道更新算法基本框架

---

- 1: 给定初始轨道猜测及初始有限维空间.
  - 2: **while** 未收敛 **do**
  - 3: 基于某些策略 (如自适应策略) 扩充有限维空间得到新的空间;
  - 4: 在扩充后的有限维空间中求解一系列相互独立的源问题以更新每个轨道;
  - 5: 在这些更新后的轨道张成的子空间中求解 Kohn-Sham 方程.
  - 6: **end while**
- 

上述并行轨道更新方法不局限于某一特定的离散方法, 也不局限于某一特定空间扩充策略, 具有普适性. 事实上, 在 [26] 的基础上, 文献 [71] 进一步提出了修正的并行轨道更新算法, 并成功应用到基于平面波离散的电子结构计算中. 下面是修正的并行轨道更新算法的计算框架, 具体算法见 [71].

---

#### 算法 5 修正的并行轨道更新算法

---

- 1: 给定初始轨道猜测及初始有限维空间.
  - 2: **while** 未收敛 **do**
  - 3: 基于某些策略 (如自适应策略) 扩充有限维空间得到新的空间;
  - 4: 在扩充后的有限维空间中求解一系列相互独立的源问题以更新每个轨道对应的残量;
  - 5: 在现有轨道及更新后的残量张成的子空间中求解 Kohn-Sham 方程.
  - 6: **end while**
- 

上述算法已成功应用到我们小组研发的基于实空间离散的第一原理计算程序 RealSPACES 及平面波离散的第一原理计算开源软件 Quantum ESPRESSO<sup>[112]</sup> 中. 数值实验表明, 这类算法在计算时间和并行可扩展性方面比 Quantum ESPRESSO 原有的算法均有优势. 并行轨道

更新方法具有天然两层并行的特点, 使其在大规模体系电子结构并行计算中有巨大潜力, 是值得推荐的电子结构计算方法.

#### 4.4. 基于对称性的两层并行离散

当体系或问题具有某种特性, 还可利用这些性质设计高效算法. 方俊等人<sup>[39,40]</sup>利用特征值问题的对称性设计分解方法, 将原特征值问题分解为一组特征值子问题. 在电子结构计算中, Kohn-Sham 方程的对称性源于晶体或团簇的宏观对称性.

下面以模型问题来简要描述这一算法的核心思想. 考虑如下的特征值问题

$$\begin{cases} Lu = \lambda u \text{ 在 } \Omega \text{ 内,} \\ u = 0 \text{ 在 } \partial\Omega \text{ 上.} \end{cases} \quad (4.2)$$

假设阶为  $g$  的有限群  $G = \{R\}$  是上面特征值问题的一相应的对称群. 记  $G$  的所有不等价、不可约的酉表示为:  $\Gamma^\nu : \nu = 1, 2, \dots, n$ ,  $\Gamma^\nu$  的维数为  $d_\nu$ . 那么特征值问题 (4.2) 可被分解为  $\sum_{\nu=1}^n d_\nu$  个子特征值问题. 记  $n_{sub} = \sum_{\nu=1}^n d_\nu$ . 对于任一  $\nu$ , 相应的  $d_\nu$  个子问题可表示如下:

$$\begin{cases} Lu_l^\nu = \lambda_l^\nu u_l^\nu \text{ 在 } \Omega \text{ 内,} \\ u_l^\nu = 0 \text{ 在 } \partial\Omega \text{ 上,} \\ u_l^\nu(Rx) = \sum_{m=1}^{d_\nu} \Gamma^\nu(R)_{lm}^* u_m^\nu(x), \text{ 在 } \Omega \text{ 内, } \forall R \in G, \end{cases} \quad l = 1, 2, \dots, d_\nu. \quad (4.3)$$

假设对 (4.2) 需要求解其最小的  $N$  个特征值及相应特征函数, 离散 (4.2) 所需自由度数为  $N_g$ . 转化为子问题后, 对每个子问题 (4.3), 则只需计算其最小的  $N/n_{sub}$  个特征值及相应特征函数, 且每个子问题只需存储和计算约化区域中的自由度. 这样描述子问题所需自由度数为  $N_g/g$ , 其中  $g$  为对称操作数. 通过简单分析可知道, 通过将问题 (4.2) 化为  $n_{sub}$  个子问题 (4.3) 的求解后, 所需计算量大大降低. 更为重要的是, 该算法具有天然两层并行特性: 由于  $n_{sub}$  个子问题相互独立, 可以本质并行求解; 对每个子问题, 仍然可采用传统并行方法来并行计算. 这一特性使得算法在大规模并行计算时优势明显.

### 5. Kohn-Sham 总能极小模型的优化算法

由 Kohn-Sham 密度泛函理论, 体系的基态电子结构既可通过求解 Kohn-Sham 方程来获得, 也可通过求解 Kohn-Sham 能量泛函极小问题得到. 在过去几十年, 电子结构计算绝大多数研究工作都是围绕求解 Kohn-Sham 方程展开. 然而, 自洽场迭代过程的收敛性问题一直是困扰人们的一个难题, 尤其是对小带隙或无带隙问题, 如金属体系. 近年来, 越来越多的研究者开始转向直接求解 Kohn-Sham 能量泛函极小问题, 各种优化方法被不断引入电子结构计算中, 如信赖域方法<sup>[43,85,94]</sup>、梯度型方法<sup>[46,90,92,99]</sup>、共轭梯度法<sup>[29]</sup>、拟牛顿法<sup>[57]</sup>、牛顿法<sup>[32]</sup>等. 虽然这些方法已广泛应用于各类优化问题, 但是由于 Kohn-Sham 能量泛函的复杂性及正交性的限制, 这些传统优化方法在电子结构计算中的应用和分析并不容易. 此外, 针对带正交约束的极小化问题的一些具有本质并行特性的优化算法也被相继提出<sup>[28,47,67]</sup>. 最近文献<sup>[31]</sup>提出针对带正交约束的极小化问题线搜索方法的自适应步长选取策略, 该策略能回避通常步长策略中的回溯过程, 从而节省计算量, 该方法在电子结构计算中效果显著. 下面简要介绍下我们在文献<sup>[29]</sup>提出的电子结构计算的共轭梯度法及文献<sup>[28,67]</sup>提出的基于并行轨道更新思想的电子结构计算优化方法.

### 5.1. 共轭梯度法

对  $\Phi = (\phi_1, \dots, \phi_N) \in (L^2(\Omega))^N$ , 定义范数

$$\|\Phi\| = (\text{tr}(\Phi^T \Phi))^{\frac{1}{2}},$$

其中  $\Phi^T \Phi$  为  $\Phi$  与  $\Phi$  的内积矩阵, 其定义见 (3.1).

记  $V$  为 Hilbert 空间  $H_0^1(\Omega)$  或其有限维逼近空间  $V_n$ . 记 Stiefel 流形

$$\mathcal{M} = \{\Phi \in V^N : \Phi^T \Phi = I^{N \times N}\}.$$

由前文可知, 体系基态可通过求解如下的带正交约束的 Kohn-Sham 能量泛函极小问题得到

$$\min_{\Phi=(\phi_1, \dots, \phi_N) \in \mathcal{M}} E(\Phi), \quad (5.1)$$

其中

$$E(\Phi) = \int_{\Omega} \left( \frac{1}{2} \sum_{i=1}^N |\nabla \phi_i|^2 + V_{\text{ext}} \rho_{\Phi} + e_{\text{xc}}(\rho_{\Phi}) \right) + \frac{1}{2} \int_{\Omega \times \Omega} \frac{\rho_{\Phi}(\mathbf{r}) \rho_{\Phi}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'.$$

由于  $E(\Phi) = E(\Phi U), \forall U \in \mathcal{O}^{N \times N}$  (见 (3.7)), 我们在 Grassman 流形  $\mathbb{G}$

$$\mathbb{G} = \mathcal{M} / \sim$$

上考虑问题 (5.1).  $\Phi \sim \Psi \iff \exists U \in \mathcal{O}^{N \times N}$  s.t.  $\Phi = \Psi U$ .

能量泛函  $E(\Phi)$  的梯度及在 Grassman 流形  $\mathbb{G}$  上的梯度分别为

$$\nabla E(\Phi) = H_{\Phi} \Phi, \quad \nabla_{\mathbb{G}} E(\Phi) = H_{\Phi} \Phi - \Phi \Lambda.$$

后面还要用到  $E(\Phi)$  在 Grassman 流形  $\mathbb{G}$  上的 Hesse 算子

$$\begin{aligned} \text{Hess}_{\mathbb{G}} E(\Phi)[\Psi, \tilde{\Psi}] &= \text{tr}(\Psi^T H_{\Phi} \tilde{\Psi}) - \text{tr}(\Psi^T \tilde{\Psi} \Lambda) \\ &+ 2 \int_{\Omega} \int_{\Omega} \frac{(\sum_i \phi_i(r) \psi_i(r)) (\sum_j \phi_j(r') \tilde{\psi}_j(r'))}{|r - r'|} dr dr' \\ &+ 2 \int_{\Omega} \frac{\delta^2(e_{\text{xc}} \rho_{\Phi}) \rho_{\Phi}}{\delta \rho_{\Phi}^2}(r) (\sum_i \phi_i(r) \psi_i(r)) (\sum_j \phi_j(r) \tilde{\psi}_j(r)) dr. \end{aligned}$$

当使用线搜索方法来求解这类带正交约束的能量泛函极小化问题时, 在每个迭代步  $k$ , 首先要确定搜索方向  $D_k$ , 然后选取合适的搜索步长  $t_k$ , 最后考虑如何保持正交性约束.

对于共轭梯度法, 搜索方向为负梯度与前一步的搜索方向的线性组合. 当目标泛函为非二次型时, 组合参数  $\beta_k$  有多种选择, 如 Fletcher-Reeves (FR) 公式、Polak-Ribière-Polyak (PRP) 公式、Hestenes-Stiefel (HS) 公式、Dai-Yuan (DY) 公式<sup>[35]</sup>. 我们选取 PRP 公式, 即

$$\beta_k = \frac{\text{tr}((\nabla_{\mathbb{G}} E(\Phi_k) - \nabla_{\mathbb{G}} E(\Phi_{k-1}))^T \nabla_{\mathbb{G}} E(\Phi_k))}{\|\nabla_{\mathbb{G}} E(\Phi_{k-1})\|^2}. \quad (5.2)$$

对于步长  $t_k$ , 我们通过将能量泛函  $E(\Phi^{(k)} + \tau D^{(k)})$  在  $\Phi^{(k)}$  的邻域内的值在  $\Phi^{(k)}$  处进行二阶 Taylor 展开

$$E(\Phi^{(k)} + \tau D^{(k)}) \approx E(\Phi^{(k)}) + \tau \text{tr}(\nabla_G E(\Phi^{(k)})^T D^{(k)}) + \frac{\tau^2}{2} \text{Hess}_G E(\Phi^{(k)})[D^{(k)}, D^{(k)}], \tau \leq \theta / \|\|D^{(k)}\|\|$$

来得到步长的初值,  $0 < \theta < 1$  为一给定常数. 得到结果如下:

$$\tau_1^{(k)} = \begin{cases} \min\left(-\frac{\text{tr}(\nabla_G E(\Phi^{(k)})^T D^{(k)})}{\text{Hess}_G E(\Phi^{(k)})[D^{(k)}, D^{(k)}]}, \frac{\theta}{\|\|D^{(k)}\|\|}\right), & \text{如果 } \text{Hess}_G E(U^{(k)})[D^{(k)}, D^{(k)}] > 0 \\ \frac{\theta}{\|\|D^{(k)}\|\|}, & \text{其他} \end{cases} \quad (5.3)$$

再结合步长回溯, 我们得到如下基于 Hesse 算子的步长选取策略.

#### 基于 Hesse 算子的步长策略 $(\theta, t, \eta)$

1. 通过 (5.3) 给定初始步长  $\tilde{\tau}_k$ .

2. 计算步长

$$\tau_k = t^{m_k} \tilde{\tau}_k,$$

其中  $m_k \in \mathbb{N}$  为使得如下不等式成立的最小非负整数

$$E(\Phi^{(n+1)}(\tau^{(n)})) \leq E(\Phi^{(n)}) + \eta \tau^{(n)} \text{tr}(\nabla_G E(\Phi^{(n)})^T D^{(n)}).$$

为算法描述方便, 我们引入一个记号  $\text{ortho}(\Phi, D, \tau)$ , 表示从流形  $\mathcal{M}$  上的一点  $\Phi$  出发, 沿着搜索方向  $D$  以步长  $\tau$  迭代到下一步, 且下一步依然保持在流形  $\mathcal{M}$  上. 我们选取了三种不同的正交化算法来保持迭代解在流形上.

- 类 Crank-Nicolson 方法

$$\begin{aligned} \text{ortho}(\Phi, D, \tau) &= \Phi + \tau D \left( I_N + \frac{\tau^2}{4} D^T D \right)^{-1} \\ &\quad - \frac{\tau^2}{2} \Phi \left( I_N + \frac{\tau^2}{4} D^T D \right)^{-1} (D^T D); \end{aligned}$$

- QR 方法

$$\text{ortho}(\Phi, D, \tau) = (\Phi + \tau D)L^{-T},$$

其中  $L$  为通过如下分解得到的下三角阵

$$LL^T = I_N + \tau^2 D^T D;$$

- PD 方法 (polar decomposition, 极坐标分解)

$$\text{ortho}(\Phi, D, \tau) = (\Phi + \tau D)(I_N + \tau^2 D^T D)^{-\frac{1}{2}}.$$

综合上面过程, 我们得到如下电子结构计算的共轭梯度法.

---

**算法 6 共轭梯度法**


---

- 1: 给定  $\epsilon, \theta, t, \eta \in (0, 1)$ , 初值  $\Phi_0$ , *s.t.*  $\Phi_0^T \Phi_0 = I_N$ ,  $\Phi_{-1} = \Phi_0$ ,  $F_{-1} = 0$ , 计算  $\nabla_G E(\Phi_0)$ , 令  $k = 0$ .
  - 2: **while** 未收敛 **do**
  - 3: 根据公式 (5.2) 计算共轭梯度参数  $\beta_k$ , 计算  $F_k = -\nabla_G E(\Phi_k) + \beta_k F_{k-1}$ ;
  - 4: 将搜索方向投影到  $\Phi_k$  的切空间:  $D_k = F_k - \Phi_k(\Phi_k^T F_k)$ ;
  - 5: 令  $F_k = -F_k \text{sign}(\text{tr}(\nabla_G E(\Phi_k)^T D_k))$ ,  $D_k = -D_k \text{sign}(\text{tr}(\nabla_G E(\Phi_k)^T D_k))$ ;
  - 6: 用基于 Hesse 算子的步长策略  $(\theta, t, \eta)$  计算步长  $\tau_k$ ;
  - 7: 更新  $\Phi_{k+1} = \text{ortho}(\Phi_k, D_k, \tau_k)$ ;
  - 8: 令  $k = k + 1$ , 计算梯度  $\nabla_G E(\Phi_k)$ .
  - 9: **end while**
- 

在我们的理论证明中, 我们需要以下几个假设.

**A4** 能量泛函的梯度  $\nabla E(\Phi)$  是 Lipschitz 连续的. 即, 存在常数  $L_0 > 0$  使得

$$\|\|\nabla E(\Phi) - \nabla E(\Psi)\|\| \leq L_0 \|\|\Phi - \Psi\|\|, \quad \forall \Phi, \Psi \in \mathcal{M}.$$

**A5** 存在  $\delta_1 > 0$ , 使得

$$\nu_1 \|\|D\|\|^2 \leq \text{Hess}_G E(\Phi)[D, D] \leq \nu_2 \|\|D\|\|^2, \quad \forall [\Phi] \in B([\Phi^*], \delta_1), \forall D \in \mathcal{T}_{[\Phi]}\mathbb{G},$$

其中  $[\Phi^*]$  为 (5.1) 的局部极小点,  $\nu_1, \nu_2 > 0$  为常数,

$$B([\Phi], \delta) := \{[\Psi] \in \mathbb{G} : \text{dist}([\Psi], [\Phi]) \leq \delta\}.$$

我们证明了对任意  $\delta_2 \in (0, \delta_1/(1 + L_1))$ , 其中  $L_1$  为一依赖  $L_0$  及  $\nu_1$  的常数, 存在  $E_0$  及相应水平集合

$$\mathcal{L} = \{[\Phi] \in \mathbb{G} : E(\Phi) \leq E_0\},$$

使得

$$\{[U] : [\Phi] \in \mathcal{L} \cap B([\Phi^*], \delta_1)\} \subset B([\Phi^*], \delta_2).$$

我们选取固定的  $\delta_2 \in (0, \delta_1/(1 + L_1))$ ,  $E_0$  及  $\mathcal{L}$ .

经过复杂的分析, 我们得到了算法 6 的收敛性<sup>[29]</sup>.

**定理 10 (收敛性).** 令假设 A4 和 A5 成立. 对算法 6 产生的序列  $\{\Phi_k\}_{k \in \mathbb{N}_0}$ , 其中更新步分别采用类 Crank-Nicolson 方法、QR 分解方法或 PD 方法. 如果  $[\Phi_0] \in B([\Phi^*], \delta_2) \cap \mathcal{L}$ , 则

$$\lim_{k \rightarrow \infty} \text{dist}([\Phi_k], [\Phi^*]) = 0,$$

$$\lim_{k \rightarrow \infty} \|\|\nabla_G E(\Phi_k)\|\| = 0.$$

## 5.2. 基于并行轨道更新的优化方法

受 [26] 中求解 Kohn-Sham 方程的并行轨道更新方法的启发, 文献 [28, 67] 将并行轨道更新的思想应用到电子结构计算的总能量极小模型的计算中, 得到了电子结构计算的基于并行轨道更新思想的本质并行的优化算法. 其基本思想如下: 利用迭代思想将  $N$  轨道变量的优化问题

$$\begin{aligned} & \min_{(\phi_1, \dots, \phi_N) \in (H^1(\Omega))^N} E(\phi_1, \phi_2, \dots, \phi_N), \\ & \text{s.t. } \int_{\Omega} \phi_i \phi_j = \delta_{ij}, 1 \leq i, j \leq N, \end{aligned} \quad (5.4)$$

解耦成  $N$  个单轨道变量的优化问题

$$\begin{aligned} & \min_{v \in H^1(\Omega)} E(\phi_1^l, \phi_2^l, \dots, \phi_{i-1}^l, v, \phi_{i+1}^l, \dots, \phi_N^l), \\ & \text{s.t. } \int_{\Omega} v^2 = 1, \end{aligned} \quad (5.5)$$

在第  $l+1$  层迭代时,  $\phi_1^l, \phi_2^l, \dots, \phi_{i-1}^l, \phi_{i+1}^l, \dots, \phi_N^l$  为已知.

由此可得到如下的并行轨道更新方法基本框架, 详细过程见 [28].

---

### 算法 7 基于并行轨道更新的优化方法

---

- 1: 给定初始有限维空间  $V_0$  及初始轨道.
  - 2: **while** 外迭代未收敛 **do**
  - 3:   基于某些策略 (如自适应策略) 扩充有限维空间得到  $V_{n+1}$ ;
  - 4:   **while** 内迭代未收敛 **do**
  - 5:     在  $V_{n+1}$  中并行地求解  $N$  个单轨道变量的优化问题;
  - 6:     更新轨道.
  - 7:   **end while**
  - 8: **end while**
- 

大量的实际算例测试表明基于并行轨道更新的优化方法能几乎保持算法原有的收敛性, 同时绝大多数计算量能本质并行, 在大规模体系电子结构并行计算方面很有潜力.

## 6. 总 结

本文介绍了基于密度泛函理论电子结构计算的核心数学模型, 以及围绕这些模型展开的数值方法与数值分析的一些进展. 重点介绍了本课题组在相关方面的一些进展, 包括 Kohn-Sham 方程有限维离散的先验误差分析、后验误差分析、高效算法及理论, 以及 Kohn-Sham 能量泛函极小问题的高效优化算法设计及理论分析等. 尽管经过多年的研究, 电子结构计算算法研究及数值分析方面取得了诸多进展, 但仍然存在一些问题亟待解决. 如: 小带隙或无带隙体系自洽场迭代的收敛性分析, 大规模代数特征值问题的并行求解, 周期性破缺体系的可计算建模及数值离散, 激发态计算的高效数值算法, 相关算法的有效性 & 可靠性分析等. 这些问题

的解决都亟需计算数学研究者的参与和投入. 由于篇幅和作者知识的限制, 很多电子结构计算方面的优秀工作没有被介绍和引用, 在此向同行表示歉意.

**致谢.** 本文的成稿参考了组内多篇论文, 在此对相关作者表示感谢. 同时, 也感谢两位匿名审稿人提出的宝贵意见. 此外, 还要特别感谢中国科学院数学与系统科学研究院周爱辉研究员对论文内容的帮助及其博士生张力维对论文的仔细检查.

## 参 考 文 献

- [1] Adams R A. Sobolev Spaces, Academic Press, New York, 1975.
- [2] Anderson D G. Iterative procedures for nonlinear integral equations[J]. J. Assoc. Comput. Mach., 1965, 12: 547–560.
- [3] Bai Z, Demmel J, Dongarra J, Ruhe A and van der Vorst H, editors. Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide. SIAM, Philadelphia, 2000.
- [4] Bao G, Hu G and Liu D. Numerical solution of the Kohn-Sham equation by finite element methods with an adaptive mesh redistribution technique[J]. J. Sci. Comput., 2013, 55: 372–391.
- [5] Beck T L. Real-space mesh techniques in density-function theory[J]. Rev. Mod. Phys., 2000, 72: 1041–1080.
- [6] Becke A D. A new mixing of Hartree-Fock and local density-functional theories[J]. J. Chem. Phys., 1993, 98: 1372–1377.
- [7] Becke A D. Density-functional thermochemistry. III. The role of exact exchange[J]. J. Chem. Phys., 1993, 98: 5648–5652.
- [8] Blöchl P E. Generalized separable potentials for electronic-structure calculations[J]. Phys. Rev. B, 1990, 41: 5414–5416.
- [9] Born M and Oppenheimer J R. Zur quantentheorie der Molekeln[J]. Ann. Physik, 1927, 84: 457–484.
- [10] Broyden C G. A class of methods for solving nonlinear simultaneous equations[J]. Math. Comput., 1965, 19: 577–593.
- [11] Cai Y, Zhang L, Bai Z, and Li R. On an eigenvector-dependent nonlinear eigenvalue problem[J]. SIAM J. Matrix Anal. Appl., 2018, 39: 1360–1382.
- [12] Calvetti D, Reichel L, Sorensen D C. An implicitly restarted Lanczos method for large symmetric eigenvalue problems[J]. Electron T. Numer. Ana., 1994, 2: 1–21.
- [13] Cancès E, Chakir R, and Maday Y. Numerical analysis of nonlinear eigenvalue problems[J]. J. Sci. Comput., 2010, 45: 90–117.
- [14] Cancès E, Chakir R, and Maday Y. Numerical analysis of the planewave discretization of some orbital-free and Kohn-Sham models[J]. M<sup>2</sup>AN, 2012, 46: 341–388.
- [15] Cascon J M, Kreuzer C, Nochetto R H, and Siebert K G. Quasi-optimal convergence rate for an adaptive finite element method[J]. SIAM J. Numer. Anal., 2008, 46: 2524–2550.
- [16] 陈华杰. 密度泛函理论的有限维逼近 [D]. 中国科学院研究生院博士论文, 2010.
- [17] Chen H, Dai X, Gong X, He L, and Zhou A. Adaptive finite element approximations for Kohn-Sham models[J]. Multiscale Model. Simul., 2014, 12: 1828–1869.
- [18] Chen H, Gong X, He L, Yang Z, and Zhou A. Numerical analysis of finite dimensional approximations of Kohn-Sham equations[J]. Adv. Comput. Math., 2013, 38: 225–256.

- [19] Chen H, Gong X, He L, and Zhou A. Adaptive finite element approximations for a class of nonlinear eigenvalue problems in quantum physics[J]. *Adv. Appl. Math. Mech.*, 2011, 3: 493–518.
- [20] Chen H, He L, and Zhou A. Finite element approximations of nonlinear eigenvalue problems in quantum physics[J]. *Comput. Methods Appl. Mech. Engrg.*, 2011, 200: 1846–1865.
- [21] Ciarlet P G. *The Finite Element Method for Elliptic Problems*. North-Holland, 1978.
- [22] Cramer C J. *Essentials of Computational Chemistry: Theories and Models*. Wiley, 2002.
- [23] 戴小英. 第一原理电子结构计算的有限元自适应及局部算法研究 [D]. 中国科学院研究生院博士论文, 2008.
- [24] 戴小英, 高兴誉, 周爱辉. 特征值问题的 Davidson 型方法及其实现技术 [J]. *数值计算与计算机应用*, 2006, 27: 218–240.
- [25] Dai X, Gong X, Yang Z, Zhang D, and Zhou A. Finite volume discretizations for eigenvalue problems with applications to electronic structure calculations[J]. *Multiscale Model. Simul.*, 2011, 9: 208–240.
- [26] Dai X, Gong X, Zhou A, and Zhu J. A parallel orbital-updating approach for electronic structure calculations[J]. *Multiscale Model. Simul.*, 2014, 12: 182–1869.
- [27] Dai X, He L, and Zhou A. Convergence rate and quasi-optimal complexity of adaptive finite element computations for multiple eigenvalues[J]. *IMA J Numer. Anal.*, 2015, 35: 1934–1977.
- [28] Dai X, Liu Z, Zhang X and Zhou A. A parallel orbital-updating based optimization method for electronic structure calculations. arXiv:1510.07230 (2015).
- [29] Dai X, Liu Z, Zhang L and Zhou A. A conjugate gradient method for electronic structure calculations[J]. *SIAM J. Sci. Comput.*, 2017, 39: A2702–A2740.
- [30] Dai X, Xu J, and Zhou A. Convergence and optimal complexity of adaptive finite element eigenvalue computations[J]. *Numer. Math.*, 2008, 110: 313–355.
- [31] Dai X, Zhang L, and Zhou A. An adaptive step size strategy for orthogonality constrained line search methods. arXiv:1906.02883 (2019).
- [32] Dai X, Zhang L and Zhou A. Pactical Newton methods for electronic structure calculations. arXiv:2001.09285 (2020).
- [33] 戴小英, 周爱辉. 电子结构计算的有限元方法 [J]. *中国科学: 化学*, 2015, 45: 800–811.
- [34] 戴小英, 周爱辉. 第一原理实空间并行自适应计算程序设计原理 [J]. *中国科学: 信息科学*, 2016, 46: 1421–1441.
- [35] Dai Y and Yuan Y. A nonlinear conjugate gradient method with a strong global convergence property[J]. *SIAM J. Optim.*, 1999, 10: 177–182.
- [36] Dirac P A M. *The Principles of Quantum Mechanics*. Oxford University Press, Oxford, 4th edition, 1988.
- [37] Dörfler W. A convergent adaptive algorithm for Poisson’s equation[J]. *SIAM J. Numer. Anal.*, 1996, 33: 1106–1124.
- [38] Ernzerhof M and Perdew J P. Generalized gradient approximation to the angle- and system-averaged exchange hole[J]. *J. Chem. Phys.*, 1998, 109: 3313–3320.
- [39] 方俊. 第一原理计算的若干实空间算法研究 [D]. 中国科学院大学博士论文, 2013.
- [40] Fang J, Gao X, Zhou A. A Symmetry-based decomposition approach to eigenvalue problems[J]. *J. Sci. Comput.*, 2013, 57: 638–669.
- [41] Fermi E. Un metodo statistics per la determinazione di alcune proprieta dell’atomo[J]. *Rend. Accad. Lincei*, 1927, 6: 602–607.
- [42] Fermi E. A statistical method for the determination of some atomic properties and the application



- of this method to the theory of the periodic system of elements[J]. *Zeit. Fur Physik*, 1928, 48: 73–79.
- [43] Francisco J B, Martinez J M, and Martinez L. Globally convergent trust-region methods for self-consistent field electronic structure calculations[J]. *J. Chem. Phys.*, 2004, 121: 10863–10878.
- [44] Fock V. Näherungsmethode zur lösung des quantenmechanischen mehrkörper problems[J]. *Z. Phys.*, 1930, 61: 126–148.
- [45] 高兴誉. 第一原理电子结构计算的六面体有限元方法 [D]. 中国科学院研究生院博士论文, 2009.
- [46] Gao B, Liu X, Chen X, and Yuan Y. A new first-order framework for orthogonal constrained optimization problems[J]. *SIAM J. Optim.*, 2018, 28: 302–332.
- [47] Gao B, Liu X and Yuan Y. Parallelizable algorithms for optimization problems with orthogonality constraints[J]. *SIAM J. Sci. Comput.*, 2019, 41: A1949–A1983.
- [48] Garau E M. and Morin P. Convergence and quasi-optimality of adaptive FEM for Steklov eigenvalue problems[J]. *IMA J. Numer. Anal.*, 2011, 31: 914–946.
- [49] Garau E M, Morin P, and Zuppa C. Convergence of adaptive finite element methods for eigenvalue problems[J]. *M<sup>3</sup>AS*, 2009, 19: 721–747.
- [50] Gong X, Shen L, Zhang D, and Zhou A. Finite element approximations for Schrödinger equations with applications to electronic structure computations[J]. *J. Comput. Math.*, 2008, 23: 310–327.
- [51] Gunnarsson O, Jonson M, and Lundqvist B I. Descriptions of exchange and correlation effects in inhomogeneous electron systems[J]. *Phys. Rev. B*, 1979, 20: 3136–3164.
- [52] Hamann D R, Schlüter M and Chiang C, Norm-conserving pseudopotentials[J]. *Phys. Rev. Lett.*, 1979, 43: 1494–1497.
- [53] Hartree D R. The wave mechanics of an atom with non-coulombic central field: parts I, II, III[J]. in *Proc. Cambridge Phil. Soc*, 1928, 24: 89–110, 111–132, 426–437.
- [54] 何莲花. 第一原理电子结构计算研究: 数值分析与数值模拟 [D]. 中国科学院研究生院博士论文, 2012.
- [55] Hedin L. New method for calculating the one-particle Green’s function with application to the electron-gas problem[J]. *Phys. Rev.*, 1965, 139: A796–A823.
- [56] Hohenberg P and Kohn W. Inhomogeneous electron gas[J]. *Phys. Rev.*, 1964, 136: B864–B871.
- [57] Hu J, Jiang B, Lin L, Wen Z and Yuan Y. Structured quasi-Newton methods for optimization with orthogonality constraints[J]. *SIAM J. Sci. Comput.*, 2019, 41: A2239–A2269.
- [58] Hu G, Xie H, and Xu F. A multilevel correction adaptive finite element method for Kohn-Sham equation[J]. *J. Comput. Phys.*, 2018, 355: 436–449.
- [59] Johnson D D, Modified Broyden’ s method for accelerating convergence in self-consistent calculations[J]. *Phys. Rev. B*, 1988, 38: 12807–12813
- [60] Knyazev A V. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method[J]. *SIAM J. Sci. Comput.*, 2001, 23: 517–541.
- [61] Kohn W and Sham L J. Self-consistent equations including exchange and correlation effects[J]. *Phys. Rev. A*, 1965, 140: 1133–1138.
- [62] Lee C, Yang W, and Parr R G, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density[J]. *Phys. Rev. B*, 1988, 37: 785–789.
- [63] Lieb E H. Density functionals for Coulomb systems[J]. *Inter. J. Quantum Chem.*, 1983, 24: 243–277.
- [64] 林霖. 类 Hartree-Fock 方程的数值方法 [J]. *计算数学*, 2019, 41(2): 113–125.
- [65] Lin L, Lu J, Ying L. Numerical methods for Kohn-Sham density functional theory[J]. *Acta Numerica*, 2019, 28: 405–539.

- [66] Lin L and Yang C. Elliptic preconditioner for accelerating the self-consistent field iteration in Kohn-Sham density functional theory[J]. *SIAM J. Sci. Comput.*, 2013, 35: S277–S298.
- [67] 刘壮. 第一原理电子结构计算的优化算法若干研究 [D]. 中国科学院大学博士学位论文, 2016.
- [68] Liu X, Wang X, Wen Z, Yuan Y. On the convergence of the self-consistent field iteration in Kohn-Sham density functional theory[J]. *SIAM J. Matrix Anal. A.*, 2014, 35: 546–558.
- [69] Liu X, Wang X, Wen Z, Ulbrich M, and Yuan Y. On the analysis of the discretized Kohn-Sham density functional theory[J]. *SIAM J. Numer. Anal.*, 2015, 53: 1758–1785
- [70] Martin R M. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.
- [71] Pan Y, Dai X, S de Gironcoli, Gong X, Rignanese G M, Zhou A. A parallel orbital-updating based plane-wave basis method for electronic structure calculations[J]. *J. Comput. Phys.*, 2017, 348: 482–492.
- [72] Parr R G and Yang W T. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, Oxford, 1994.
- [73] Perdew J P. Climbing the ladder of density functional approximations[J]. *MRS Bull.*, 2013, 28: 743–750.
- [74] Perdew J P, Burke K and Ernzerhof M. Generalized gradient approximation made simple[J]. *Phys. Rev. Lett.*, 1996, 77: 3865–3868.
- [75] Perdew J P and Schmidt K. Jacob’s ladder of density functional approximations for the exchange-correlation energy[J]. *In AIP Conference Proceedings*, 2001, 577: 1–20.
- [76] Pulay P. Convergence acceleration of iterative sequences. the case of scf iteration[J]. *Chem. Phys. Lett.*, 1980, 73: 393–398.
- [77] Pulay P. Improved SCF convergence acceleration[J]. *J. Comput. Chem.*, 1982, 3: 556–560.
- [78] Saad Y. *Numerical Methods for Large Eigenvalue Problems*[M]. New York: Halstead Press, 1992.
- [79] Saad Y. Numerical methods for electronic structure calculations of materials[J]. *SIAM Rev.*, 2010, 52: 3–54.
- [80] Sleijpen G L G and van der Vorst H A. A Jacobi-Davidson iteration method for linear eigenvalue problems[J]. *SIAM J. Matrix Anal. Appl.*, 1996, 17: 401–425.
- [81] Sherrill C D and Schaefer III H F. The configuration interaction method: Advances in highly correlated approaches[J]. *Adv. Quantum Chem.*, 1999, 34: 143–269.
- [82] Simon B. Schrödinger operators in the twentieth century[J]. *J. Math. Phys.*, 2000, 41: 3523–3555.
- [83] 沈丽华. 基于密度泛函理论电子结构有限元并行自适应算法 [D]. 中国科学院研究生院博士论文, 2005.
- [84] Suryanarayana P, Gavini V, Blesgen T, Bhattacharya K, and Ortiz M. Non-periodic finite-element formulation of Kohn-Sham density functional theory[J]. *J. Mech. Phys. Solids*, 2010, 58: 256–280.
- [85] Thøgersen L, Olsen J, Yeager D, Jørgensen P, Sałek P, and Helgaker T. The trustregion self-consistent field method: Towards a black-box optimization in Hartree-Fock and Kohn-Sham theories[J]. *J. Chem. Phys.*, 2004, 121: 16–27.
- [86] Thomas L H. The calculation of atomic fields[J]. *Proc. Cambridge Phil. Soc.*, 1927, 23: 542–548. Reprinted in March(1975).
- [87] Troullier N and Martins J L. Efficient pseudopotentials for plane-wave calculations[J]. *Phys. Rev. B*, 1991, 43: 1993–2006.
- [88] Tsuchida E and Tsukada M. Adaptive finite-element method for electronic-structure calculations[J]. *Phys. Rev. B*, 1996, 54: 7602–7605.
- [89] Tsuchida E and Tesukada M. Large-scale electronic-structure calculations based on the adaptive

- finite element method[J]. *J. Phy. Soc. Jpn.*, 1998, 67: 3844–3858.
- [90] Ulbrich M, Wen Z, Yang C, Klöckner D, and Lu Z. A proximal gradient method for ensemble density functional theory[J]. *SIAM J. Sci. Comput.*, 2015, 37: A1975–A2002.
- [91] Vanderbilt D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism[J]. *Phys. Rev. B*, 1990, 41: 7892–7895.
- [92] Wen Z and Yin W. A feasible method for optimization with orthogonality constraints[J]. *Math. Program. Ser. A.*, 2013, 142: 397–434.
- [93] 杨章. 基于有限体积离散的第一原理电子结构计算 [D]. 中国科学院研究生院博士论文, 2011.
- [94] Yang C, Meza J C, and Wang L. A trust region direct constrained minimization algorithm for the Kohn-Sham equation[J]. *SIAM J. Sci. Comput.*, 2007, 29: 1854–1875.
- [95] 谢希德, 陆栋. 固体能带理论 [M]. 复旦大学出版社, 上海, 1998.
- [96] Yang C, Gao W, Meza J C. On the convergence of the self-consistent field iteration for a class of nonlinear eigenvalue problems[J]. *SIAM J. Matrix Anal. Appl.*, 2009, 30: 1773–1788.
- [97] Yang B and Zhou A. Eigenfunction behaviors and adaptive finite element approximations of nonlinear eigenvalue problems in quantum physics. arXiv:1907.03968(2019).
- [98] 张笛儿. 有限元方法在电子结构计算中的应用 [D]. 复旦大学博士论文, 2007.
- [99] Zhang X, Zhu J, Wen Z and Zhou A. Gradient type optimization methods for electronic structure calculations[J]. *SIAM J. Sci. Comput.*, 2014, 36: 265–289.
- [100] Zhou A. An analysis of finite-dimensional approximations for the ground state solution of Bose-Einstein condensates[J]. *Nonlinearity*, 2004, 17: 541–550.
- [101] Zhou A. Finite dimensional approximations for the electronic ground state solution of a molecular system[J]. *Math. Meth. Appl. Sci.*, 2007, 30: 429–447.
- [102] 周爱辉. 电子结构模型的数学基础. 中国科学院研究生院讲义, 2010.
- [103] Zhou A. Hohenberg-Kohn theorem for Coulomb type systems and its generalization[J]. *J. Math. Chem.*, 2012, 50: 2746–2754.
- [104] 周爱辉. 电子结构模型与计算的数学问题 [J]. *中国科学: 数学*, 2015, 45(6): 929–938.
- [105] Zhou A. A mathematical aspect of Hohenberg-Kohn theorem[J]. *Science China Mathematics*, 2019, 62: 63–68
- [106] Zhou Y, Wang H, Liu Y, Gao X, Song H. Applicability of Kerker preconditioning scheme to the self-consistent density functional theory calculations of inhomogeneous systems[J]. *Phys. Rev. E*, 2018, 97: 033305.
- [107] 朱金伟. 基于平均场模型与强关联理论的第一原理计算 [D]. 中国科学院大学博士论文, 2014.
- [108] ABACUS, <http://abacus.ustc.edu.cn/>.
- [109] ABINIT, <http://www.abinit.org>.
- [110] Gaussian, <http://www.gaussian.com/>.
- [111] PHG, <http://lsec.cc.ac.cn/phg/>.
- [112] Quantum Espresso, <http://www.quantum-espresso.org>.
- [113] VASP, <http://www.vasp.at>.

## NUMERICAL METHODS AND THEORIES FOR ELECTRONIC STRUCTURE CALCULATIONS

Dai Xiaoying

*(LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;  
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China)*

### Abstract

The first principles electronic structure calculations have become important tools for studying the material mechanism, understanding and predicting the material properties, and have achieved great success. However, it is still full of challenge for how to design highly efficient and highly accurate computational methods to deal with larger system, how to understand the reliability and efficiency of calculation from a mathematical point of view. Based on the Kohn-Sham DFT, the key mathematical modes for electronic structure calculations are the Kohn-Sham equation or the Kohn-Sham energy functional minimization problem. In the past decades, the highly efficient algorithms design and numerical analysis have attracted the attention of many distinguished mathematicians. Our group have also focused on this field and have done several works. In this paper, we introduce recent progresses in this field, mainly about those done by our group.

**Keywords:** electronic structure; Kohn-Sham equation; Kohn-Sham energy functional minimization problem; eigenvalue problem; finite element method; optimization method; adaptive; parallel computing

**2010 Mathematics Subject Classification:** 35Q55, 65N15, 65N25, 65N30, 81Q05