Acta Numerica (2019), pp. 405–539
doi:10.1017/S0962492919000047

© Cambridge University Press, 2019
Printed in the United Kingdom

Numerical methods for Kohn–Sham density functional theory

Lin Lin[*]
Department of Mathematics,
University of California, Berkeley,
and Computational Research Division,
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
E-mail: linlin@math.berkeley.edu

Jianfeng Lu[†]
Department of Mathematics,
Department of Physics, and Department of Chemistry,
Duke University, Durham, NC 27708, USA
E-mail: jianfeng@math.duke.edu

Lexing Ying[‡]
Department of Mathematics and
Institute for Computational and Mathematical Engineering,
Stanford University, Stanford, CA 94305, USA
E-mail: lexing@stanford.edu

Kohn–Sham density functional theory (DFT) is the most widely used electronic structure theory. Despite significant progress in the past few decades, the numerical solution of Kohn–Sham DFT problems remains challenging, especially for large-scale systems. In this paper we review the basics as well as state-of-the-art numerical methods, and focus on the unique numerical challenges of DFT.

[*] Partially supported by the US National Science Foundation via grant DMS-1652330, and by the US Department of Energy via grants DE-SC0017867 and DE-AC02-05CH11231.

[†] Partially supported by the US National Science Foundation via grants DMS-1454939 and ACI-1450280 and by the US Department of Energy via grant DE-SC0019449.

[‡] Partially supported by the US Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, the 'Scientific Discovery through Advanced Computing (SciDAC)' programme and the US National Science Foundation via grant DMS-1818449.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

CONTENTS

1. Introduction

Kohn–Sham density functional theory (Hohenberg and Kohn 1964, Kohn and Sham 1965) (DFT) is today the most widely used electronic structure theory in chemistry, materials science and other related fields, and was recognized by the Nobel Prize in Chemistry awarded to Walter Kohn in 1998. Despite being developed in the mid-1960s, Kohn–Sham DFT was not a popular approach until the beginning of the 1980s, when practical exchange-correlation functionals obtained from quantum Monte Carlo calculations (Ceperley and Alder 1980, Perdew and Zunger 1981) became available. Together with the development of more advanced exchange-correlation functionals (*e.g.* Lee, Yang and Parr 1988, Becke 1993, Perdew, Burke and Ernzerhof 1996*a*, Heyd, Scuseria and Ernzerhof 2003), Kohn–Sham DFT has been demonstrated to be capable of predicting chemical and material properties for a wide range of systems from *first principles*. More specifically, under the Born–Oppenheimer approximation that the nuclei are treated as classical particles, in principle the only input needed for Kohn–Sham DFT calculations is the most basic information concerning the system: atomic species and atomic positions.

The success of Kohn–Sham DFT cannot be separated from the vast improvement in numerical algorithms and computer architecture, particularly during the 1980s and 1990s. Some relatively basic ideas, such as using the conjugate gradient method for solving eigenvalue problems (Payne *et al.* 1992), efficient use of the planewave basis set (*i.e.* pseudospectral methods) (Kresse and Furthmüller 1996), and density matrix based methods (*i.e.* computing matrix functions) (Goedecker 1999), have been cited thousands to tens of thousands of times. Unfortunately, to a large extent applied mathematicians have missed this ‘golden era’ of developing numerical

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

methods for Kohn–Sham DFT, so most numerical methods used in practical calculations were developed by physicists and chemists. By the middle of the 2000s, many numerical methods widely used today had been implemented in commercial or open source software packages, such as ABINIT (Gonze *et al.* 2016), CASTEP (Clark *et al.* 2005), FHI-aims (Blum *et al.* 2009), Molpro (Werner *et al.* 2012), NWChem (Valiev *et al.* 2010), Qbox (Gygi 2008), QChem (Shao *et al.* 2015), Quantum ESPRESSO (Giannozzi *et al.* 2017), SIESTA (Soler *et al.* 2002) and VASP (Kresse and Furthmüller 1996), to name just a few. In fact, the availability of these robust and mature software packages is *the* driving force behind the wide range of applicability of Kohn–Sham DFT for practical physical and chemical systems.

On the other hand, as Kohn–Sham DFT is being applied to ever larger and more complex quantum systems, new numerical challenges have emerged and thus provided new opportunities for further algorithmic development. For instance, the cost of orthogonalizing Kohn–Sham orbitals is typically not noticeable for small systems. However, the computational cost of this step scales cubically with respect to the system size, and quickly dominates the computational cost for large systems, often known as the 'cubic scaling wall'. Due to the cubic scaling wall, most practical electronic structure calculations are performed for systems with tens to hundreds of atoms, and can occasionally reach a few thousand atoms. Since the orthogonalization step is so inherently wired into all eigensolvers, any algorithm with reduced asymptotic complexity must necessarily not seek for an eigen-decomposition of the Kohn–Sham Hamiltonian. It should be noted that most electronic structure software packages used today are designed *around* the eigensolver. Hence avoiding the eigensolver entirely implies major changes for the algorithmic and software flow. With massively parallel high performance computers (HPC) becoming more widely available, it is also crucial to develop scalable numerical algorithms that can be efficiently scaled to tens of thousands of computational processors to tackle challenging electronic structure problems within a reasonable amount of wall clock time. With the development of low complexity and scalable new algorithms in the past decade, it is now possible to carry out electronic structure calculations for systems with tens of thousands of atoms, even for difficult metallic systems (*i.e.* systems without an energy band gap).

Another trend in recent developments of DFT is to move towards more complex exchange-correlation functionals. This is because simple exchange-correlation functionals, such as those based on local density approximation (LDA) and generalized gradient approximation (GGA), cannot reach 'chemical accuracy' (error of the energy below 1 kcal/mol) needed for quantitative predictability for many important systems, such as transition metal oxides. To this end, more involved functionals on the fourth and fifth rungs of the 'Jacob's ladder' of exchange-correlation functionals (see the discussion

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

in Section 2.5) become useful, and sometimes indispensable, for treating such systems. These functionals not only have a more complex formulation but may also fundamentally change the mathematical and hence algorithmic structure of the Kohn–Sham DFT problem. For instance, Kohn–Sham DFT with rung-1 to rung-3 functionals can be solved as a nonlinear eigenvalue problem with differential operators. The operator becomes an integro-differential operator when rung-4 functionals (also called hybrid functionals) are used, and the very notion of self-consistency becomes difficult to define properly when rung-5 functionals are used. We remark that practical electronic structure calculations with such functionals have only been pursued for the past decade in the physics community, at least when a large basis set such as the planewave basis set is concerned. The computational scaling of hybrid functional calculations is also cubic with respect to the system size, the same as that of the LDA and GGA functionals. However, hybrid functionals involve the Fock exchange operator, and the computational time of hybrid functional calculations can often be more than 20 times larger than that of LDA and GGA calculations. Therefore hybrid functional calculations are typically done for systems with tens to hundreds of atoms. Rung-5 functional calculations are even more expensive. This brings new opportunities for designing better algorithms to overcome the computational bottleneck with more complex functionals. With the development of new algorithms, the cost of hybrid functional calculations can now be reduced to within twice the cost of GGA calculations, thus enabling hybrid functional calculations for systems with thousands of atoms.

As a final example, the surge of ‘high-throughput computing’, such as that driven by the materials genome initiative (https://www.mgi.gov), requires hundreds of thousands of calculations to be performed automatically to form a vast database. This requires *all steps* of numerical algorithms to be performed in a robust fashion. Many numerical algorithms in electronic structure calculations were not designed to meet this criterion, as they often require many ‘knobs’, *i.e.* tuning parameters to reach convergence. Sometimes the tuning parameters require heavy human intervention and expert knowledge of the system. Therefore, the design of numerical methods to perform the same task, but in a more *robust* and *automatic* fashion, also raises new challenges.

There have been numerous advances in the past two decades towards addressing such new challenges, including our own attempts to reduce computational complexity, improve parallel scalability and design robust algorithms in a number of directions. This review aims to introduce both basic concepts and recent developments in numerical methods for solving Kohn–Sham DFT to a mathematical audience. Although we will try to cover the whole landscape of DFT algorithms, the discussion is *heavily biased* towards our own work and view of the field, and some omissions are inevitable.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

For simplicity, unless otherwise specified, we will consider isolated, charge-neutral, spinless systems throughout this review. Most methods we discuss can be easily generalized to periodic systems with Γ-point sampling of the Brillouin zone (*i.e.* systems with periodic boundary conditions), and often further to periodic systems with general **k**-point sampling strategies, charged systems, as well as spin-polarized systems. We will also focus on the so-called 'single-shot' DFT calculations, *i.e.*, calculations for a single given atomic configuration. The algorithms can then be used naturally when calculations for multiple – or even massively many – atomic configurations are needed, such as in the context of geometry optimization and *ab initio* molecular dynamics simulation (Car and Parrinello 1985, Marx and Hutter 2009). For a more general introduction to density functional theory, we refer readers to books by Parr and Yang (1989), Dreizler and Gross (1990), Eschrig (1996), Kaxiras (2003), Martin (2008) and Lin and Lu (2019).

The basic theory and formalism of Kohn–Sham DFT will be introduced in Section 2. The rest of the paper discusses components of the DFT calculations: numerical discretization in Section 3, evaluation of the Kohn–Sham maps in Sections 5 and 6, and self-consistent iteration in Section 7. In addition, Section 4 is devoted to several algorithmic tools that have proved to be helpful in a number of contexts in electronic structure calculations. Section 8 concludes with an outlook for trends and future developments. The list of notations is given at the end of the paper for readers' convenience.

2. Theory and formulation

2.1. Quantum many-body problem

In principle, all particles in a physical system, including both nuclei and electrons, are quantum particles and should be treated using quantum mechanics. Since the mass of the lightest element in the periodic table (hydrogen) is around 2000 times larger than that of the electron, the commonly used Born–Oppenheimer approximation assumes that the nuclei can be described by classical mechanics. This is often an accurate approximation, and will be assumed throughout this review.

In this review, we are mostly concerned with isolated systems surrounded by a vacuum in $\mathbb{R}^3$. This is also called the 'free space' boundary condition. Other settings such as the Dirichlet boundary condition in a finite-sized box can be discussed similarly. We use atomic units (a.u.), *i.e.* $m_e = e = \hbar = 1/(4\pi\epsilon_0) = k_B = 1$, where m_e is the mass of an electron, e is the unit charge, $\hbar$ is the reduced Planck constant, $1/(4\pi\epsilon_0)$ is the Coulomb constant, and k_B is the Boltzmann constant. Under the Born–Oppenheimer approximation,

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the many-body Hamiltonian with M nuclei and N electrons in $\mathbb{R}^3$ is

$$\begin{aligned} H &= \sum_{i=1}^{N} -\frac{1}{2}\Delta_{\mathbf{r}_i} + \sum_{i=1}^{N} V_{\text{ext}}(\mathbf{r}_i; \{\mathbf{R}_I\}) + \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I<J}^{M} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &\equiv \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} + E_{\text{II}}, \end{aligned} \tag{2.1}$$

where the first three terms on the second line are the kinetic energy, external potential and electron–electron interactions respectively:

$$\hat{T} = \sum_{i=1}^{N} -\frac{1}{2}\Delta_{\mathbf{r}_i}, \quad \hat{V}_{\text{ext}} = \sum_{i=1}^{N} V_{\text{ext}}(\mathbf{r}_i; \{\mathbf{R}_I\}), \quad \hat{V}_{\text{ee}} = \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}.$$

For a given atomic configuration $\{\mathbf{R}_I\}$, the ion–ion interaction simply adds a constant shift to the Hamiltonian with

$$E_{\text{II}} = \sum_{I<J}^{M} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}.$$

We remark that the treatment of periodic systems (more specifically, with a Γ-point only sampling strategy of the Brillouin zone) is very similar to that for isolated systems. Even for isolated charge-neutral systems, periodic boundary conditions are often used, since they give reasonably accurate approximations as long as certain care is taken with the electrostatic energy (Makov and Payne 1995). Periodic boundary conditions are also more natural for certain numerical methods, such as planewave methods, even for molecular systems.

The ground state is often the most important state. This is because the energy gap $E_1 - E_0$ for many electron systems is of the order of electron volts (eV), or 10^4 Kelvin when the unit of measurement is $k_B T$ (k_B is the Boltzmann constant). This is much greater than room temperature (300 Kelvin). According to the Boltzmann distribution, the probability of the quantum state E_i being occupied is $\mathrm{e}^{-\beta E_i}/Z$, where β is the inverse temperature, and Z is a normalization factor called the partition function. Hence the ground state is often the dominating state. Even for certain systems where $E_1 - E_0$ is small or zero, the ground state can still be very important. This paper focuses on the calculation of the ground-state energy. In the discussion below, we will simply refer to E_0 as E for notational simplicity.

Let $\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ denote both the spatial and spin degrees of freedom for the ith electron (though electrons are indistinguishable quantum particles). The many-body ground-state wavefunction $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is associated with the smallest eigenvalue E of the linear eigenvalue problem

$$H\Psi = E\Psi. \tag{2.2}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Due to the Pauli exclusion principle for the identical electrons, Ψ is a totally anti-symmetric function:

$$\Psi(\mathbf{x}_1,\dots,\mathbf{x}_i,\dots,\mathbf{x}_j,\dots,\mathbf{x}_N) = -\Psi(\mathbf{x}_1,\dots,\mathbf{x}_j,\dots,\mathbf{x}_i,\dots,\mathbf{x}_N). \tag{2.3}$$

The Courant–Fisher minimax theorem states that eigenvalue problem (2.2) can also be viewed as an optimization problem,

$$E = \inf_{|\Psi\rangle\in\mathcal{A}_N,\langle\Psi|\Psi\rangle=1} \langle\Psi|H|\Psi\rangle, \tag{2.4}$$

where $\mathcal{A}_N$ is the set of totally anti-symmetric functions. Equation (2.4) is the *variational principle* for the ground state in quantum mechanics. Here

$$\langle\Psi|\Psi\rangle := \int \Psi^*(\mathbf{x}_1,\dots,\mathbf{x}_N)\Psi(\mathbf{x}_1,\dots,\mathbf{x}_N)\,\mathrm{d}\mathbf{x}_1\cdots\mathbf{x}_N = 1 \tag{2.5}$$

defines the normalization condition for the many-body wavefunction, and the integration with respect to a multi-index $\mathbf{x}$ should be interpreted as

$$\int f(\mathbf{x})\,\mathrm{d}\mathbf{x} := \sum_{\sigma\in\{\uparrow,\downarrow\}} \int_{\mathbb{R}^3} f(\mathbf{r},\sigma)\,\mathrm{d}\mathbf{r}. \tag{2.6}$$

In order to solve the quantum many-body problem, it appears at first sight that one has to know the quantum many-body wavefunction. Surprisingly, this is at least formally not the case if we are only interested in the ground state. First established by Hohenberg and Kohn (1964), density functional theory (DFT) discovered that the many-body electron density, defined by

$$\rho(\mathbf{r}) = N \sum_{\sigma\in\{\uparrow,\downarrow\}} \int |\Psi((\mathbf{r},\sigma),\mathbf{x}_2,\dots,\mathbf{x}_n)|^2\,\mathrm{d}\mathbf{x}_2\cdots\,\mathrm{d}\mathbf{x}_N, \tag{2.7}$$

is all one needs to determine the quantum many-body ground state. More specifically, assuming the ground state is non-degenerate, there is a one-to-one correspondence between the electron density ρ and the external potential V_{ext}, so that knowledge of ρ is sufficient to reconstruct the many-body wavefunction Ψ. Shortly thereafter, Mermin (1965) proposed the DFT formulation in the finite temperature set-up to include thermal effects as well. The most widely used form of DFT, called Kohn–Sham DFT, was proposed by Kohn and Sham (1965). Below we introduce DFT from the perspective of constrained minimization, which was first proposed by Levy (1979) and then rigorously established by Lieb (1983). We also remark that there is as yet no analogous mathematically rigorous theory of DFT for excited states.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

2.2. Kohn–Sham density functional theory

According to the variational principle (2.4), the ground-state energy can be written as a constrained minimization problem:

$$E = \inf_{\Psi\in\mathcal{A}_N, \langle\Psi|\Psi\rangle=1} \langle\Psi|H|\Psi\rangle = \inf_{\rho\in\mathcal{J}_N}\left\{ \inf_{\substack{\Psi\in\mathcal{A}_N\\ \Psi\mapsto\rho}} \langle\Psi|H|\Psi\rangle \right\}. \tag{2.8}$$

Here, on the right-hand side, a constrained minimization given the density ρ is first carried out for the wavefunction Ψ. The resulting value is then optimized as a functional of ρ within the function space

$$\mathcal{J}_N = \left\{\rho \geq 0 : \int \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} = N, \quad \nabla\sqrt{\rho} \in L^2(\mathbb{R}^3)\right\}. \tag{2.9}$$

Here the condition $\nabla\sqrt{\rho}\in L^2(\mathbb{R}^3)$ is implied by the finiteness of the kinetic energy, since

$$\begin{aligned}\frac{1}{2}\int |\nabla\sqrt{\rho}(\mathbf{r})|^2\,\mathrm{d}\mathbf{r} &\leq \frac{N}{2}\sum_{\sigma\in\{\uparrow,\downarrow\}}\int |\nabla_{\mathbf{r}}\Psi((\mathbf{r},\sigma),\mathbf{x}_2,\ldots,\mathbf{x}_n)|^2\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_N\\ &= \langle\Psi|\hat{T}|\Psi\rangle. \end{aligned} \tag{2.10}$$

It can also be proved that $\rho\in\mathcal{J}_N$ implies that there exists at least one $\Psi\in\mathcal{A}_N$ with finite kinetic energy such that $\Psi\mapsto\rho$ (Lieb 1983).

Therefore the constrained minimization procedure is well-defined, and we have

$$E = \inf_{\rho\in\mathcal{J}_N}\left\{ \inf_{\substack{\Psi\in\mathcal{A}_N\\ \Psi\mapsto\rho}} \langle\Psi|(\hat{T}+\hat{V}_{\mathrm{ee}})|\Psi\rangle + \int \rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})\,\mathrm{d}\mathbf{r}\right\} + E_{\mathrm{II}} \tag{2.11}$$

$$= \inf_{\rho\in\mathcal{J}_N}\left\{\mathcal{F}_{\mathrm{LL}}[\rho] + \int \rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})\,\mathrm{d}\mathbf{r}\right\} + E_{\mathrm{II}}. \tag{2.12}$$

Here, together with equation (2.7), we have used

$$\langle\Psi|\hat{V}_{\mathrm{ext}}|\Psi\rangle = N\int |\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)|^2 V_{\mathrm{ext}}(\mathbf{r}_1)\,\mathrm{d}\mathbf{x}_1\cdots\mathrm{d}\mathbf{x}_N = \int\rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{2.13}$$

The functional $\mathcal{F}_{\mathrm{LL}}[\rho]$ is *universal*, since it depends only on the kinetic and electron–electron repulsion but not on the external potential V_{ext}, which is specified by the atomic configuration. Another important consequence of density functional theory, often known as the Hohenberg–Kohn theorem (Hohenberg and Kohn 1964), is that if $\rho^\star$ is the minimizer of (2.12) and if $\Psi^\star$ is the unique minimizer that results in $\mathcal{F}_{\mathrm{LL}}[\rho^\star]$, then the many-body ground-state wavefunction $\Psi^\star$ is determined by the electron density $\rho^\star$. In other words, if the ground state is non-degenerate, then there is a one-to-

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

one mapping between the ground-state electron density and the many-body ground-state wavefunction. All that remains to make DFT useful is thus a good explicit expression that approximates the functional $\mathcal{F}_{\mathrm{LL}}[\rho]$.

Since the very early days of quantum mechanics, physicists have been seeking the approximation to $\mathcal{F}_{\mathrm{LL}}[\rho]$ (before it was connected to density functional theory) pioneered by Thomas (1927) and Fermi (1927). Until the 1960s, efforts were mainly restricted to uniform electron gas, that is, the system is in a periodic box with a constant external potential field, where many calculations can be done analytically. Despite significant progress in the past few decades, modelling $\mathcal{F}_{\mathrm{LL}}[\rho]$ remains a very difficult task. To appreciate the difficulty, just recall the atomic shell structure from the eigenfunctions of the Hamiltonian operator of a hydrogen atom. It is already highly non-trivial to find such a mapping for a single atom. Furthermore, in chemistry and materials science, the absolute value of the energy is usually not the most important quantity. It is the *relative* energy difference that determines whether or not a chemical process should occur. This often requires the ground-state energy to be calculated with 99.9% accuracy or higher.

The breakthrough of density functional theory is generally attributed to Kohn and Sham (1965), who proposed combining DFT with single-particle orbital structure. Given N one-particle orthonormal spin-orbitals $\{\psi_i(\mathbf{x})\}_{i=1}^N$, that is,

$$\langle\psi_i|\psi_j\rangle := \int \psi_i^*(\mathbf{x})\psi_j(\mathbf{x})\,\mathrm{d}\mathbf{x} = \delta_{ij}, \quad i,j = 1,\ldots,N, \tag{2.14}$$

the associated Slater determinant is given by

$$\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\det\begin{pmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{pmatrix}, \tag{2.15}$$

which is anti-symmetric by the property of determinants. The computation of the electron density (2.7) with respect to a Slater determinant can be simplified as

$$\rho(\mathbf{r}) = \sum_{i=1}^{N}\sum_{\sigma\in\{\uparrow,\downarrow\}} |\psi_i(\mathbf{r},\sigma)|^2. \tag{2.16}$$

Using constrained minimization over Slater determinants, the Kohn–Sham proposal can be interpreted as

$$\mathcal{F}_{\mathrm{LL}}[\rho] = \inf_{\substack{\Psi\in\mathcal{A}_N^0\\ \Psi\mapsto\rho}} \langle\Psi|\hat{T}|\Psi\rangle + \frac{1}{2}\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\mathrm{xc}}[\rho], \tag{2.17}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Figure 2.1. (Credit: Burke 2012.) A ‘zoo’ of exchange-correlation functionals in DFT.

where $\mathcal{A}_N^0$ is the set of Slater determinants with N orbitals. It turns out that for any $\rho \in \mathcal{J}_N$ there exists at least one $\Psi \in \mathcal{A}_N^0$ that gives the density ρ, and the constrained minimization of the kinetic energy term is well-defined (Lieb 1983). In equation (2.17), the first term corresponds to the kinetic energy from N non-interacting single-particle orbitals. The second term is the Hartree energy, which characterizes the electron–electron repulsion energy at the mean-field level. The last term $E_{\mathrm{xc}}[\rho]$, called the exchange-correlation functional, at first glance simply defines whatever we do not know about $\mathcal{F}_{\mathrm{LL}}[\rho]$. The insight from Kohn and Sham (1965) is that the kinetic and Hartree terms often account for more than 95% of the total energy. Therefore the approximation to $E_{\mathrm{xc}}[\rho]$, while still very difficult, is a much easier task than approximating $\mathcal{F}_{\mathrm{LL}}[\rho]$ directly. As a result, the approximation to the ground-state energy in the Kohn–Sham DFT is given by minimizing over the Kohn–Sham energy functional:

$$\begin{aligned} E^{\mathrm{KS}} = \inf_{\Psi \in \mathcal{A}_N^0} \biggl\{ \langle \Psi | \hat{T} | \Psi \rangle + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}' \\ + E_{\mathrm{xc}}[\rho] + \int \rho(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \biggr\} + E_{\mathrm{II}}, \end{aligned} \tag{2.18}$$

with ρ given by the Slater determinant Ψ. Writing the kinetic energy of the Slater determinant more explicitly, we have

$$E^{\mathrm{KS}} = \inf_{\{\psi_i\}_{i=1}^N, \langle \psi_i | \psi_j \rangle = \delta_{ij}} \mathcal{F}^{\mathrm{KS}}(\{\psi_i\}) + E_{\mathrm{II}} \tag{2.19}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

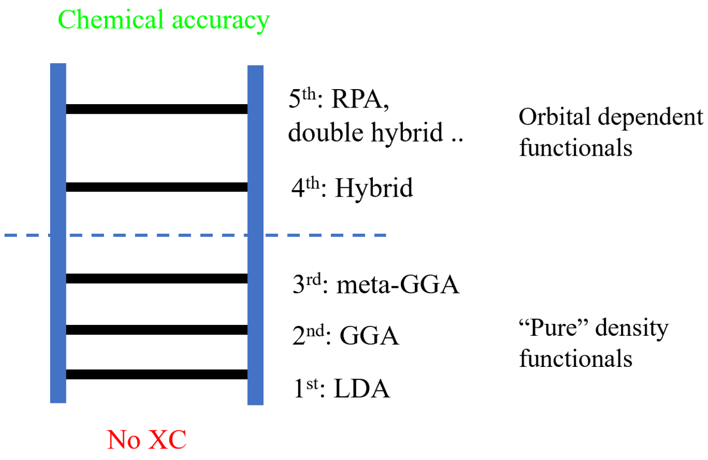


Figure 2.2. The 'Jacob's ladder' of exchange-correlation functionals.

with

$$\mathcal{F}^{\mathrm{KS}}(\{\psi_i\}) = \frac{1}{2}\sum_{i=1}^{N}\int |\nabla_{\mathbf{r}}\psi_i(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} + \int \rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\mathrm{xc}}[\rho]. \tag{2.20}$$

The Kohn–Sham DFT would in principle be exact if we had access to the exact exchange-correlation functional $E_{\mathrm{xc}}[\rho]$. The exchange-correlation functional is also universal, *i.e.* independent of the external potential V_{ext} and hence the atomic configuration. In order to use Kohn–Sham DFT in practice, the exchange-correlation functional E_{xc} must be approximated. Starting from the local density approximation proposed by Kohn and Sham, a 'zoo' of exchange-correlation functionals has been proposed: see an incomplete list in Figure 2.1 by Burke (2012).

According to Perdew (Perdew and Schmidt 2001, Perdew 2013), these exchange-correlation functionals can be organized using a 'Jacob's ladder' for exchange-correlation functionals (Figure 2.2). When no exchange-correlation functional is used, Kohn–Sham DFT is essentially a Hartree approximation (with Pauli's exclusion principle). In such a case, DFT is significantly less accurate than Hartree–Fock theory (Szabo and Ostlund 1989), and the latter is often regarded as the starting point of quantum chemistry. This is referred to as 'Hartree's hell'. Moving up the ladder, the accuracy of the DFT calculation *generally* improves towards the 'heaven of chemical accuracy' of 1 kcal/mol (or 1.6×10^{-3} Hartree per atom) when compared to experimental results. Correspondingly, the functional forms become increasingly more complex, which leads to higher computational costs.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

On the first rung of the ladder, we have the local density approximation (LDA), where E_{xc} is modelled locally by the electron density

$$E_{\mathrm{xc}}[\rho] = \int \epsilon_{\mathrm{xc}}(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r}. \tag{2.21}$$

Note that the integrand $\epsilon_{\mathrm{xc}}(\rho(\mathbf{r}))$ depends only on the electron density at $\mathbf{r}$. The most widely used LDA exchange-correlation functional is obtained by parametrizing the result from the quantum Monte Carlo simulation (QMC) obtained for the uniform electron gas system in the 1980s (Ceperley and Alder 1980, Perdew and Zunger 1981). Although most, if not all, real chemical and materials systems are very different from the uniform electron gas system, the Kohn–Sham DFT calculations with such LDA exchange-correlation functionals already perform surprisingly well for many systems.

In order to improve the accuracy of the exchange-correlation functional, on the next rung of the ladder we have the generalized gradient approximation (GGA) (Lee, Yang and Parr 1988, Becke 1988, Perdew, Burke and Ernzerhof 1996*a*), which in addition depends on information on the gradient of electron density,

$$E_{\mathrm{xc}}[\rho] = \int \epsilon_{\mathrm{xc}}(\rho(\mathbf{r}), \sigma(\mathbf{r}))\,\mathrm{d}\mathbf{r}, \tag{2.22}$$

where the functional depends only on the norm of the gradient

$$\sigma(\mathbf{r}) := |\nabla\rho(\mathbf{r})|^2$$

due to local rotational symmetry. The GGA functionals are currently the most widely used functionals since they achieve a balance between accuracy and computational cost.

The third rung of the ladder is the meta-GGA approximation (Staroverov, Scuseria, Tao and Perdew 2003, Sun, Ruzsinszky and Perdew 2015), where second-order derivative information is also added to the approximation, in particular the kinetic energy density

$$\tau(\mathbf{r}) := \frac{1}{2}\sum_{\sigma\in\{\uparrow,\downarrow\}}\sum_{i=1}^{N} |\nabla_{\mathbf{r}}\psi_i(\mathbf{r},\sigma)|^2.$$

The meta-GGA energy functional then takes the form

$$E_{\mathrm{xc}}[\{\psi_i\}] = \int \epsilon_{\mathrm{xc}}(\rho(\mathbf{r}), \sigma(\mathbf{r}), \tau(\mathbf{r}))\,\mathrm{d}\mathbf{r}. \tag{2.23}$$

Some other meta-GGA functionals also involve second-order information on the density, *i.e.* $\nabla^2\rho(\mathbf{r})$, and for simplicity we omit treatment of such terms. Since τ is the local kinetic energy of the orbitals, it may seem that the meta-GGA functional is no longer strictly a density functional. It turns out

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

that the Euler–Lagrange equation associated with the meta-GGA functional can still be written in the same form as that of the LDA and GGA energy functionals, as we will see in the next section. Hence, the LDA, GGA and meta-GGA functionals are known as *semi-local functionals*, and their numerical treatment is very similar, as will be discussed in Section 5.

Let us remark that the exchange-correlation functionals discussed so far do not explicitly depend on spin degrees of freedom (*i.e.* they are only functionals of the electron density rather than the spin-dependent electron density). Spin-dependent exchange-correlation functionals, such as local-spin-density approximation (LSDA) (von Barth and Hedin 1972), are developed for spin-polarized systems. From the perspective of numerical algorithms they are rather similar to the exchange-correlation functionals without spin dependence, so we will not go into the details.

The exchange-correlation functionals beyond the third rung are called non-local functionals, and will be discussed in Section 2.5. Before diving into these more complicated exchange-correlation functionals, we first introduce the Kohn–Sham equations, *i.e.* the Euler–Lagrange equation for Kohn–Sham density functional theory.

2.3. Kohn–Sham equations for semi-local functionals

In order to minimize the Kohn–Sham energy functional (2.20), we need to find the stationary point of the Lagrangian. Let us consider the case of the LDA exchange-correlation functional first. Note that

$$\frac{1}{2}\frac{\delta\mathcal{F}^{\mathrm{KS}}(\{\psi_i\})}{\delta\psi_i^*(\mathbf{x})} = \biggl(-\frac{1}{2}\Delta_{\mathbf{r}} + V_{\mathrm{ext}} + V_H[\rho] + V_{\mathrm{xc}}[\rho]\biggr)\psi_i(\mathbf{x}) =: H^{\mathrm{KS}}[\rho]\psi_i(\mathbf{x}), \tag{2.24}$$

where the effective Kohn–Sham Hamiltonian $H^{\mathrm{KS}}[\rho]$ is an operator acting on the orbitals which depends on ρ, given by the orbitals as in equation (2.16). The Hartree potential V_H is given by the Coulomb kernel via

$$V_H[\rho](\mathbf{r}) = (v_C\rho)(\mathbf{r}) := \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}', \tag{2.25}$$

and the exchange-correlation potential is defined to be the functional derivative of E_{xc}, that is,

$$V_{\mathrm{xc}}[\rho] = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}, \tag{2.26}$$

which for now can be thought of as a ρ-dependent potential acting on orbitals. V_{xc} of more general forms will be discussed further and made more explicit at the end of this section.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Taking into account the orthonormality condition (2.14), we obtain the Euler–Lagrange equations

$$H^{\mathrm{KS}}[\rho]\psi_i(\mathbf{x}) = \left(-\frac{1}{2}\Delta_{\mathbf{r}} + V_{\mathrm{ext}} + V_H[\rho] + V_{\mathrm{xc}}[\rho]\right)\psi_i(\mathbf{x}) = \sum_{j=1}^{N}\psi_j(\mathbf{x})\lambda_{ij}, \tag{2.27}$$

where the λ_{ij} are Lagrange multipliers. Let us further simplify the Euler–Lagrange equations to reveal the structure of the problem. First note that due to the orthonormality of $\{\psi_i\}$ and the self-adjointness of $H^{\mathrm{KS}}[\rho]$, we have

$$\lambda_{ij} = \langle\psi_j|H^{\mathrm{KS}}[\rho]|\psi_i\rangle = \langle\psi_i|H^{\mathrm{KS}}[\rho]|\psi_j\rangle^* = \lambda_{ji}^*. \tag{2.28}$$

Therefore $\Lambda = (\lambda_{ij})$ is a Hermitian matrix, and we may assume an eigendecomposition of Λ as

$$\Lambda = U\,\mathrm{diag}(\varepsilon_1,\ldots,\varepsilon_n)U^*, \tag{2.29}$$

where U is a unitary matrix.

Now consider a rotation of the orbitals, that is,

$$\varphi_i(\mathbf{x}) = \sum_{j=1}^{N}\psi_j(\mathbf{x})U_{ji}. \tag{2.30}$$

Since U is a unitary matrix, the transformation preserves the electron density and hence the effective Kohn–Sham Hamiltonian. The equations for the rotated orbitals $\{\varphi_i\}$ become

$$\begin{aligned} H^{\mathrm{KS}}[\rho]\varphi_i(\mathbf{x}) &= H^{\mathrm{KS}}[\rho]\sum_{j=1}^{N}\psi_j(\mathbf{x})U_{ji} = \sum_{k,j=1}^{N}\psi_k(\mathbf{x})\lambda_{kj}U_{ji} \\ &= \sum_{k=1}^{N}\psi_k(\mathbf{x})U_{ki}\varepsilon_i = \varphi_i(\mathbf{x})\varepsilon_i. \end{aligned} \tag{2.31}$$

Therefore, without loss of generality (up to a rotation of orbitals), it suffices to consider the Euler–Lagrange equations of the form

$$\left(-\frac{1}{2}\Delta_{\mathbf{r}} + V_{\mathrm{ext}} + V_H[\rho] + V_{\mathrm{xc}}[\rho]\right)\psi_i(\mathbf{x}) = \varepsilon_i\psi_i(\mathbf{x}), \quad i = 1,\ldots,N. \tag{2.32}$$

Since the operator H^{KS} depends on the orbitals $\{\psi_i\}$ via the electron density ρ, this is a set of nonlinear eigenvalue problems, known as the *Kohn–Sham equations*. Here the Hamiltonian is nonlinear with respect to the eigenvectors, rather than the eigenvalues.

The Kohn–Sham equations (2.32) must be solved self-consistently with respect to the electron density ρ. For a given electron density ρ, the Hamil-

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

tonian $H^{\mathrm{KS}}[\rho] = -\frac{1}{2}\Delta_{\mathbf{r}} + V_{\mathrm{eff}}(\mathbf{r})$ is a self-adjoint linear operator, where the effective potential induced by ρ is

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + V_H[\rho](\mathbf{r}) + V_{\mathrm{xc}}[\rho](\mathbf{r}). \tag{2.33}$$

The Kohn–Sham orbitals $\{\psi_i\}$ are thus eigenfunctions of $H^{\mathrm{KS}}[\rho]$. It should be noted that *a priori* there is no guarantee that $\{\psi_i\}_{i=1}^N$ should correspond to the lowest N eigenvalues (counting multiplicity) of $H^{\mathrm{KS}}[\rho]$ to achieve the global minimum of the Kohn–Sham energy functional (2.18). In practice, this is often assumed in solving the Kohn–Sham equations, known as the *aufbau principle*. It is known that the aufbau principle holds for non-interacting systems, as well as certain Hartree–Fock models. However, it can be violated for more complex models such as Kohn–Sham DFT. Assuming the aufbau principle, the first N eigenfunctions $\{\psi_i\}_{i=1}^N$ are called *occupied orbitals*, while the eigenfunctions of $H^{\mathrm{KS}}[\rho]$ with higher eigenvalues are called *virtual* or *unoccupied orbitals*.

For Kohn–Sham DFT with semi-local exchange-correlation functionals, the effective Kohn–Sham potential $V_{\mathrm{eff}}[\rho]$, and hence the Hamiltonian matrix, depends only on the density ρ through the Kohn–Sham potential. We refer to the mapping from V_{eff} to ρ as the *Kohn–Sham map*, denoted by

$$\rho = \mathcal{F}_{\mathrm{KS}}[V_{\mathrm{eff}}]. \tag{2.34}$$

The electron density ρ can be evaluated from the Kohn–Sham map by solving a linear eigenvalue problem, or using density matrix techniques, to be discussed in Section 5. Hence ρ and V_{eff} should be iteratively determined by each other until convergence. This is called the self-consistent field (SCF) iteration.

The numerical solution for solving the Kohn–Sham equations can thus be divided into three subproblems, as illustrated in Figure 2.3: discretization of the Kohn–Sham Hamiltonian using a finite basis set, evaluation of the Kohn–Sham map, and iteration until reaching self-consistency. The content of this review is thus largely organized around these three topics. Note that during the self-consistent iteration, one also needs to form the Kohn–Sham Hamiltonian by evaluating the Hartree and exchange-correlation potential. With some abuse of terminology, we consider this step as part of the discretization problem in Figure 2.3. We remark that the evaluation of the Hartree potential requires solving a Poisson equation. Depending on the choice of the basis set, the solution can be efficiently obtained using the fast Fourier transform or multigrid methods (see *e.g.* Brandt 1977, Brandt, McCormick and Ruge 1985, Briggs, Henson and McCormick 2000, Fattebert and Bernholc 2000).

Let us now come back to the exchange-correlation potential, given by the functional derivative of E_{xc} with respect to the density. For LDA, the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

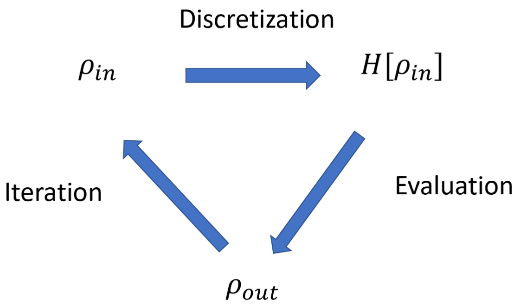


Figure 2.3. Three subproblems for the self-consistent solution of the Kohn–Sham equations. With some abuse of terminology, the step of forming the Kohn–Sham Hamiltonian is considered to be part of the discretization step in this figure.

variation of E_{xc} is given by

$$\delta E_{\mathrm{xc}} = \int \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}(\rho(\mathbf{r}))\delta\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}.$$

Hence

$$V_{\mathrm{xc}}[\rho](\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta \rho(\mathbf{r})} = \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}(\rho(\mathbf{r})).$$

For GGA, the variation with respect to the density gives

$$\delta E_{\mathrm{xc}} = \int \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}\delta\rho + \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \sigma}\delta\sigma\,\mathrm{d}\mathbf{r}.$$

Using the chain rule,

$$\delta\sigma = |\nabla(\rho+\delta\rho)|^2 - |\nabla\rho|^2 = 2\nabla\rho\cdot\nabla\delta\rho + O(|\delta\rho|^2),$$

we have

$$\delta E_{\mathrm{xc}} = \int \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}\delta\rho + 2\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \sigma}(\nabla\rho\cdot\nabla\delta\rho)\,\mathrm{d}\mathbf{r} = \int \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}\delta\rho - 2\nabla\cdot\left(\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \sigma}\nabla\rho\right)\delta\rho\,\mathrm{d}\mathbf{r}.$$

Thus

$$V_{\mathrm{xc}}[\rho] = \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho} - 2\nabla\cdot\left(\frac{\partial \epsilon_{\mathrm{xc}}}{\partial \sigma}\nabla\rho\right).$$

The derivation of the exchange-correlation potential is slightly different for meta-GGA, since the kinetic energy density $\tau(\mathbf{r})$ explicitly involves the Kohn–Sham orbitals $\{\psi_i\}$. We still start with

$$\delta E_{\mathrm{xc}} = \int \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \rho}\delta\rho + \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \sigma}\delta\sigma + \frac{\partial \epsilon_{\mathrm{xc}}}{\partial \tau}\delta\tau\,\mathrm{d}\mathbf{r}.$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Since

$$\delta\tau(\mathbf{r}) = \sum_{i=1}^{N} \frac{1}{2}\nabla\psi_i^*(\mathbf{r})\cdot\nabla\delta\psi_i(\mathbf{r}) + \frac{1}{2}\nabla\delta\psi_i^*(\mathbf{r})\cdot\nabla\psi_i(\mathbf{r}),$$

we can use integration by parts to obtain

$$\begin{aligned}\int \frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\delta\tau\,\mathrm{d}\mathbf{r} = &-\frac{1}{2}\sum_{i=1}^{N}\int \delta\psi_i^*\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\right)\psi_i\,\mathrm{d}\mathbf{r}\\ &-\frac{1}{2}\sum_{i=1}^{N}\int \psi_i^*\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\right)\delta\psi_i\,\mathrm{d}\mathbf{r}.\end{aligned} \tag{2.35}$$

Recall that the Kohn–Sham equation is obtained by variation with respect to $\delta\psi_i^*$, thus the exchange-correlation 'potential' $V_{\mathrm{xc}}[\rho]$ applied to the occupied orbital ψ_i can be viewed as

$$V_{\mathrm{xc}}[\rho]\psi_i = \left[\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\rho} - 2\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\sigma}\nabla\rho\right) - \frac{1}{2}\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\right)\right]\psi_i. \tag{2.36}$$

Therefore, the exchange-correlation functional $V_{\mathrm{xc}}[\rho]$ is still independent of the orbitals, and can be defined only from the electron density as

$$V_{\mathrm{xc}}[\rho] = \frac{\partial\epsilon_{\mathrm{xc}}}{\partial\rho} - 2\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\sigma}\nabla\rho\right) - \frac{1}{2}\nabla\cdot\left(\frac{\partial\epsilon_{\mathrm{xc}}}{\partial\tau}\nabla\right). \tag{2.37}$$

Strictly speaking this is no longer a potential, as it involves a differential operator $\nabla\cdot((\partial\epsilon_{\mathrm{xc}}/\partial\tau)\nabla)$ acting on the orbitals. On the other hand, this is still a local operator, and can be treated in a similar way to our previous discussions.

For simplicity, for the rest of the paper we will neglect the spin degrees of freedom and consider the 'spinless' electrons. Hence all single-particle orbitals can be written as $\psi_i(\mathbf{r})$ instead of $\psi_i(\mathbf{x})$. The numerical algorithms can be easily adapted to take into account the spin degrees of freedom.

2.4. Density matrix formulation

Consider a (one-body) Hamiltonian operator $H = -\frac{1}{2}\Delta + V$. Assume that H has a discrete spectrum and denote the eigenpairs of H as $\{(\varepsilon_i, \psi_i)\}$:

$$H\psi_i = \varepsilon_i\psi_i. \tag{2.38}$$

Assume that the system has N electrons that occupy the first N eigenstates according to Pauli's exclusion principle. Here, for definiteness, we assume that $\varepsilon_N < \varepsilon_{N+1}$ to avoid degeneracy.

The Kohn–Sham energy functionals are invariant with respect to unitary rotations of the orbitals. The unitary rotation matrix is often referred to as the *gauge* degrees of freedom. Hence, the physical quantity is the subspace

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

spanned by the occupied orbitals, instead of the individual eigenfunctions. This subspace $\operatorname{span}\{\psi_i\}_{i=1,\ldots,N}$ is known as the Kohn–Sham *occupied subspace* and can be represented by the *density matrix*

$$P=\sum_{i=1}^{N}|\psi_i\rangle\langle\psi_i|. \tag{2.39}$$

Since the $\{\psi_i\}$ are orthonormal, we have

$$P^2=\sum_{i,j=1}^{N}|\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j|=\sum_{i=1}^{N}|\psi_i\rangle\langle\psi_i|=P. \tag{2.40}$$

Hence P is self-adjoint and idempotent, and is the projection operator onto the occupied space. The kernel of P, viewed as an integral operator, is given by

$$P(\mathbf{r},\mathbf{r}')=\sum_{i=1}^{N}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}'). \tag{2.41}$$

In particular, we observe that the diagonal part of the kernel is just the electron density

$$P(\mathbf{r},\mathbf{r})=\sum_{i=1}^{N}\psi_i(\mathbf{r})\psi_i^*(\mathbf{r})=\sum_{i=1}^{N}|\psi_i(\mathbf{r})|^2=\rho(\mathbf{r}). \tag{2.42}$$

Moreover the trace of P is equal to the number of electrons:

$$\operatorname{Tr}P=\int P(\mathbf{r},\mathbf{r})\,\mathrm{d}\mathbf{r}=N. \tag{2.43}$$

The Kohn–Sham equations can be reformulated in terms of the density matrix. Compared with the orbital representation, the density matrix is more intrinsic, as it is invariant with respect to the unitary rotation of the orbitals. In terms of numerical algorithms, use of the density matrix can also be advantageous, especially for large-scale problems, as we will discuss in Section 5.

Note that P can be represented independent of the orbitals. Since H is a self-adjoint operator, for any Borel-measurable function f on the real line, a matrix function $f(H)$ can be defined using its spectral decomposition as

$$f(H)=\sum_{i}f(\varepsilon_i)|\psi_i\rangle\langle\psi_i|. \tag{2.44}$$

Recall that for simplicity we have assumed that the spectrum of H is discrete. Thus, assuming that $\varepsilon_N<\varepsilon_{N+1}$ and the parameter $\mu\in\mathbb{R}$ satisfies

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

$\varepsilon_N \le \mu < \varepsilon_{N+1}$, we have

$$\mathbb{1}_{(-\infty,\mu]}(H) = \sum_i \mathbb{1}_{(-\infty,\mu]}(\varepsilon_i)|\psi_i\rangle\langle\psi_i| = \sum_{i:\varepsilon_i\le\mu} |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i| = P. \tag{2.45}$$

We conclude with

$$P = \mathbb{1}_{(-\infty,\mu]}(H), \tag{2.46}$$

where the right-hand side is the spectral projection onto the interval $(-\infty,\mu]$. Here μ is known as the *Fermi level* or the *chemical potential.*

2.5. Non-local functionals

Let us now introduce the non-local exchange-correlation functionals, which are on the fourth and fifth rungs of Perdew's ladder above the semi-local ones introduced in Section 2.2. These more complex density functionals can improve the fidelity of DFT, but they may significantly increase the computational cost at the same time.

On the fourth rung are the so-called hybrid density functionals, such as B3LYP (Becke 1993), PBE0 (Perdew, Ernzerhof and Burke 1996*b*) and HSE (Heyd, Scuseria and Ernzerhof 2003). They have been shown to improve the accuracy of Kohn–Sham DFT calculations, such as the computation of adsorption energies for molecules on surfaces, and molecular frontier level alignment relative to metals and semiconductors. This is achieved by incorporating a fraction of the Hartree–Fock exact exchange or screened exchange operator into the Kohn–Sham Hamiltonian. The hybrid energy functionals depend not only on the density but also on the density matrix, and hence, strictly speaking, they are no longer 'pure' density functionals. A hybrid functional typically takes the following form of hybridization based on a density-based exchange energy and the exact exchange energy (Becke 1993, Heyd *et al.* 2003):

$$E_{\mathrm{xc}}[P] = E_{\mathrm{c}}[\rho] + (1-\alpha)E_{\mathrm{x}}[\rho] + \alpha E_{\mathrm{x}}^{\mathrm{EX}}[P]. \tag{2.47}$$

Here α is a parameter for the fraction of the non-local exchange contribution, and E_{x} and E_{c} are the exchange and correlation parts (see *e.g.* Martin 2008 for the separation of the exchange-correlation functional into the exchange and correlation components, respectively) from semi-local functionals, such as GGA functionals. $E_{\mathrm{x}}^{\mathrm{EX}}[P]$ is the (screened) Hartree–Fock exchange energy corresponding to the density matrix P:

$$E_{\mathrm{x}}^{\mathrm{EX}}[P] = -\frac{1}{2}\iint |P(\mathbf{r},\mathbf{r}')|^2 K(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}', \tag{2.48}$$

where $K(\mathbf{r},\mathbf{r}')$ is the kernel for the electron–electron interaction. For example, for Hartree–Fock exchange, $K(\mathbf{r},\mathbf{r}') = v_C(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ is

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the Coulomb kernel. In theories with screened Fock exchange interactions (Heyd *et al.* 2003), K can be a screened Coulomb kernel with $K(\mathbf{r},\mathbf{r}') = \operatorname{erfc}(\alpha_s|\mathbf{r}-\mathbf{r}'|)/|\mathbf{r}-\mathbf{r}'|$, where α_s is the inverse screening length parameter.

On the fifth rung of the ladder, we have functionals that depend not only on the density and occupied Kohn–Sham orbitals but also on other quantities such as the eigenvalues and virtual orbitals of the Kohn–Sham effective Hamiltonian. They are sometimes referred to as *orbital functionals*. Examples of such functionals include the double hybrid functionals (Grimme 2006, Zhang, Xu and Goddard III 2009, Goerigk and Grimme 2014), van der Waals functionals (Dion *et al.* 2004), random phase approximation (RPA) functionals (Ren, Rinke, Joas and Scheffler 2012*b*, Chen *et al.* 2017), and other functionals based on many-body perturbation theory (see *e.g.* Mori-Sánchez, Wu and Yang 2005, Ren, Rinke, Scuseria and Scheffler 2013, Zhang, Rinke and Scheffler 2016).

For an example of rung-5 functionals, let us consider the RPA functional, for which the exchange-correlation $E_{\mathrm{xc}} = E_{\mathrm{x}} + E_{\mathrm{c}}$ consists of the exact exchange and the random phase approximation (Bohm and Pines 1953, Gell-Mann and Brueckner 1957) of the correlation energy E_{c} given by

$$E_{\mathrm{c}}^{\mathrm{RPA}} = \frac{1}{2\pi}\int_0^\infty \operatorname{Tr}[\ln(1-\hat{\chi}^0(\mathrm{i}\omega)v_C) + \hat{\chi}^0(\mathrm{i}\omega)v_C]\,\mathrm{d}\omega, \tag{2.49}$$

where v_C is the Coulomb kernel in (2.25), and $\hat{\chi}^0$ is the dynamic polarizability operator

$$\hat{\chi}^0(\mathbf{r},\mathbf{r}',\mathrm{i}\omega) = \sum_i^{\mathrm{occ}}\sum_a^{\mathrm{vir}} \frac{\psi_i^*(\mathbf{r})\psi_a(\mathbf{r})\psi_a^*(\mathbf{r}')\psi_i(\mathbf{r}')}{\varepsilon_i-\varepsilon_a-\mathrm{i}\omega} + \text{c.c.} \tag{2.50}$$

Here ‘occ’ and ‘vir’ denote the set of occupied and virtual orbitals, respectively, and c.c. stands for the complex conjugate of the previous term. Approximation of the correlation is obtained by applying random phase approximation combined with the adiabatic connection (Langreth and Perdew 1975, Gunnarsson and Lundqvist 1976, Ren *et al.* 2012*b*). Note that equation (2.50) contains the virtual orbitals and also the Kohn–Sham orbital energies, which result from the dynamic linear response of the system.

The Kohn–Sham equations of non-local functionals involve non-local operators (hence the name non-local functionals). For rung-4 functionals, the Kohn–Sham equations take the form

$$\begin{gathered} H[P]\psi_i = \left(-\frac{1}{2}\Delta + V_{\mathrm{ext}} + V_{\mathrm{Hxc}}[\rho] + \alpha V_{\mathrm{x}}^{\mathrm{EX}}[P]\right)\psi_i = \varepsilon_i\psi_i, \\ \int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\,\mathrm{d}\mathbf{r} = \delta_{ij}, \quad P(\mathbf{r},\mathbf{r}') = \sum_{i=1}^N \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}'). \end{gathered} \tag{2.51}$$

Here V_{Hxc} is the Hartree and exchange-correlation contribution from the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

electron density ρ only, and $V_{\mathrm{x}}^{\mathrm{EX}}[P]$ is derived from the exact exchange functional $E_{\mathrm{x}}^{\mathrm{EX}}[P]$, with kernel

$$V_{\mathrm{x}}^{\mathrm{EX}}[P](\mathbf{r},\mathbf{r}') = -P(\mathbf{r},\mathbf{r}')K(\mathbf{r},\mathbf{r}'). \tag{2.52}$$

$V_{\mathrm{x}}^{\mathrm{EX}}[P]$ is often called the Fock exchange operator, and is negative semi-definite. Equation (2.51) is called the Hartree–Fock-like equation. In this case the Kohn–Sham map in equation (2.34) becomes the *generalized Kohn–Sham map* from V_{eff} to the density matrix P. The numerical treatment of the exact exchange term, as well as the Hartree–Fock-like equation, will be discussed in detail in Section 6.

The self-consistency for RPA functionals is much more complicated, since the functional involves virtual orbitals and orbital energies. Most current strategies are to first obtain the effective Hamiltonian and corresponding Kohn–Sham orbitals based on a semi-local or hybrid functional, and then calculate the RPA correlation in a post-processing step. Self-consistency calculations have been performed using the optimized effective potential (OEP) framework (Godby, Schlüter and Sham 1986, Godby, Schlüter and Sham 1988, Fukazawa and Akai 2015) or more recently the generalized optimized effective potential approach (Jin *et al.* 2017), but they come with considerable numerical effort. Finding an efficient approach for self-consistent treatment of RPA functionals is an active research area, and in this review we will not discuss self-consistency for rung-5 functionals in detail.

2.6. Finite temperature density functional theory

The Kohn–Sham density functional theory for ground-state quantum systems discussed so far can be extended to systems at finite temperature (Mermin 1965). Instead of a variational principle for the ground-state energy, the functional for free energy is minimized with respect to the density matrix, that is,

$$\begin{aligned}\mathcal{F}_\beta[\rho] = \inf_{\substack{P\in\mathcal{D}\\ P\mapsto\rho}} \biggl(\frac{1}{2}\operatorname{Tr}((-\Delta)P) + \beta^{-1}\operatorname{Tr}(P\ln P + (I-P)\ln(I-P)) \\ + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\mathrm{xc},\beta}[\rho]\biggr),\end{aligned} \tag{2.53}$$

where $\operatorname{Tr}(P\ln P + (I-P)\ln(I-P))$ is the Fermi–Dirac entropy (Parr and Yang 1989) and $\mathcal{D}$ is the set of all one-particle density matrices for an N-electron system:

$$\mathcal{D} = \{P \in \mathcal{B}(L^2(\mathbb{R}^3)) : P = P^*, 0 \preceq P \preceq I, \operatorname{Tr} P = N\}. \tag{2.54}$$

Here $\mathcal{B}(L^2(\mathbb{R}^3))$ stands for the bounded operators on $L^2(\mathbb{R}^3)$ and the constraint $0 \preceq P \preceq I$ comes from the Pauli exclusion principle. To see this, let

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the eigen-decomposition of P be

$$P = \sum_i f_i |\psi_i\rangle\langle\psi_i|. \tag{2.55}$$

The constraints $0 \preceq P \preceq I$ and $\operatorname{Tr} P = N$ imply that the occupation number $\{f_i\}$ satisfies

$$0 \leq f_i \leq 1 \quad \text{and} \quad \sum_i f_i = N. \tag{2.56}$$

Thus no state is occupied by more than one electron. Here the summation is performed over all single-particle orbitals, rather than only the occupied ones.

We can write the variational problem more explicitly in terms of the (f_i, ψ_i) as

$$F = \inf_{\substack{\{f_i\},\{\psi_i\} \\ 0\leq f_i\leq 1,\ \sum_i f_i = N \\ \langle\psi_i|\psi_j\rangle=\delta_{ij}}} F_\beta^{\mathrm{KS}}(\{f_i\},\{\psi_i\}), \tag{2.57}$$

with

$$\begin{aligned} &F_\beta^{\mathrm{KS}}(\{f_i\},\{\psi_i\}) \\ &\quad = \frac{1}{2}\sum_i \int f_i |\nabla\psi_i|^2 \,\mathrm{d}\mathbf{r} + \beta^{-1}\sum_i (f_i \ln f_i + (1-f_i)\ln(1-f_i)) \\ &\qquad + \int \rho(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) \,\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\mathrm{xc},\beta}[\rho], \end{aligned} \tag{2.58}$$

and the density is given by

$$\rho(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2. \tag{2.59}$$

In principle, the exchange-correlation functional for the finite temperature $E_{\mathrm{xc},\beta}[\rho]$ depends on β and also has a ladder of approximation schemes like the zero temperature case. However, most finite temperature DFT calculations in practice still use the temperature-independent exchange-correlation functional at the semi-local level. The zero and finite temperature functional approximations share the same mathematical structure. Hence, for the purpose of our discussions, we will not explicitly distinguish them below.

We now consider the Kohn–Sham equations for the finite temperature functional (2.53). As the functional involves minimization with respect to

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the density matrix P, it is more convenient to consider the combined minimization of density and the associated density matrix:

$$\begin{aligned}\inf_{P\in\mathcal{D}}\biggl(\frac{1}{2}\operatorname{Tr}((-\Delta)P)+\beta^{-1}\operatorname{Tr}(P\ln P+(I-P)\ln(I-P))\\ +\frac{1}{2}\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'+E_{\mathrm{xc},\beta}[\rho]\biggr),\end{aligned} \tag{2.60}$$

where ρ is given by the diagonal of the density matrix P. Taking the variation with respect to P, we obtain the Euler–Lagrange equation

$$H_\beta^{\mathrm{KS}}[\rho]+\beta^{-1}(\ln P-\ln(I-P))-\mu=0, \tag{2.61}$$

where μ is the Lagrange multiplier associated with the constraint $\operatorname{Tr}P=N$ and the effective Hamiltonian is given as in the zero temperature case (*cf.* (2.24)), that is,

$$H_\beta^{\mathrm{KS}}[\rho]=-\frac{1}{2}\Delta+V_{\mathrm{ext}}+V_H[\rho]+V_{\mathrm{xc},\beta}[\rho]. \tag{2.62}$$

Solving (2.61) for P, we get

$$P=[I+\exp(\beta(H_\beta^{\mathrm{KS}}[\rho]-\mu))]^{-1}. \tag{2.63}$$

Letting f_β be the *Fermi–Dirac distribution* function

$$f_\beta(\varepsilon)=\frac{1}{1+\exp(\beta\varepsilon)}, \tag{2.64}$$

we arrive at the self-consistent equation

$$P=f_\beta(H_\beta^{\mathrm{KS}}[\rho]-\mu)=\sum_i f_\beta(\varepsilon_i-\mu)|\psi_i\rangle\langle\psi_i|, \tag{2.65}$$

where ε_i and ψ_i are the eigenvalue and associated eigenfunction of $H_\beta^{\mathrm{KS}}[\rho]$ respectively. We see that the occupation number is given by

$$f_i=f_\beta(\varepsilon_i-\mu)=\frac{1}{1+\exp(\beta(\varepsilon_i-\mu))}. \tag{2.66}$$

Thus $f_i\in(0,1)$, and all eigenstates are occupied with a fractional number. Also, notice that as $\beta\to\infty$ (zero temperature limit), f_β converges to the function f_∞:

$$f_\infty(\varepsilon)=\begin{cases}1 & \varepsilon<0,\\ 1/2 & \varepsilon=0,\\ 0 & \varepsilon>0.\end{cases} \tag{2.67}$$

Let μ_β denote the Lagrange multiplier for the inverse temperature β, where we make the dependence on β explicit. In the limit, we can show that (as

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

before, we assume $\varepsilon_N < \varepsilon_{N+1}$)

$$\lim_{\beta\to\infty} \mu_\beta = \frac{1}{2}(\varepsilon_N + \varepsilon_{N+1}) =: \mu_\infty. \tag{2.68}$$

Therefore, the density matrix of finite temperature is consistent with the definition of the chemical potential at zero temperature. Compared to equation (2.46), we may also write

$$P = f_\infty(H - \mu_\infty).$$

2.7. Pseudopotential approximation

In typical electronic structure calculations, not all electrons play the same role. In a single-particle picture, it is well known that the core electrons (*i.e.* orbitals with relatively low energies) are barely affected by the chemical environment. It is therefore desirable to remove the core electrons from the actual computation, and solve the Kohn–Sham equations for valence electrons (*i.e.* orbitals with relatively high energies) only. Another difficulty in all-electron calculations lies in the treatment of the singular Coulomb interaction between nuclei and electrons, which introduces a cusp in the Kohn–Sham orbitals at each atomic position. The electron–nucleus cusp makes uniform basis functions such as planewaves (to be discussed in Section 3.1) very inefficient for all-electron calculations.

Hence the pseudopotential approximation is introduced to solve both issues simultaneously, which is widely used especially in solid-state physics and materials science. Although this is not a systematically controlled approximation, numerical results indicate that the error introduced by the use of the pseudopotential can be much smaller compared to other sources of error in Kohn–Sham DFT calculations. Simply speaking, the use of pseudopotentials has the following benefits.

(1) The number of Kohn–Sham orbitals depends only on the number of valence electrons. This is particularly important for heavy elements, where each atom involves several tens of electrons, while the number of valence electrons can be much smaller.

(2) The valence electron orbitals are typically smoother than the core electron orbitals, and hence require a smaller number of basis functions such as planewaves to resolve.

(3) After removing the core electrons, the resulting pseudo-valence electron orbitals have fewer nodes near the nuclei. This further enhances the smoothness of the orbitals and reduces computational cost.

Below we introduce the basic idea of the norm-conserving pseudopotential (Hamann, Schlüter and Chiang 1979, Troullier and Martins 1991) – the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

earliest and still the most widely used form of pseudopotential. With some abuse of notation, in this section we use Cartesian coordinates $(r_1, r_2, r_3)^\top$ and spherical coordinates $(r, \theta, \phi)^\top$ for the same electron position $\mathbf{r}$ interchangeably. When a function $f(\mathbf{r})$ depends only on the radial distance r, we will not distinguish between $f(\mathbf{r})$ and $f(r)$. Similarly, if $f(\mathbf{r})$ depends only on the angular variables θ and ϕ, we will not distinguish between $f(\mathbf{r})$ and $f(\theta, \phi)$.

For a single atom at the origin $\mathbf{R} = \mathbf{0}$, the electron–nucleus interaction becomes $V_{\text{ext}}(r) = -Z/r$, where Z is the charge of the nucleus. Consider the atomic Kohn–Sham equation

$$\begin{aligned}\left(-\frac{1}{2}\Delta + V_{\text{ext}}(r) + V_{\text{Hxc}}[\rho](r)\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \quad i = 1, \ldots, N,\\ \rho(\mathbf{r}) = \sum_{i=1}^{N}|\psi_i(\mathbf{r})|^2. \end{aligned} \tag{2.69}$$

Here V_{Hxc} includes the contribution from the Hartree and exchange-correlation interactions, which also depend only on the radial distance r if the electron density $\rho(\mathbf{r}) = \rho(r)$ satisfies spherical symmetry. However, some single-particle orbital $\psi_i(\mathbf{r})$ might not be spherically symmetric and thus still depend on all components of $\mathbf{r}$.

Let us separate the N Kohn–Sham orbitals into two groups: core electron orbitals $\{\psi_i\}_{i=1}^{N_c}$ and valence electron orbitals $\{\psi_i\}_{i=N_c+1}^{N}$. This also defines the number of valence orbitals $N_v := N - N_c$. The goal of the pseudopotential approximation is to find an operator V_{ps} which is defined to satisfy the following modified equation:

$$\begin{aligned}\left(-\frac{1}{2}\Delta + V_{\text{ps}} + V_{\text{Hxc}}[\widetilde{\rho}](r)\right)\widetilde{\psi}_i(\mathbf{r}) = \widetilde{\varepsilon}_i\widetilde{\psi}_i(\mathbf{r}), \quad i = 1, \ldots, N_v,\\ \widetilde{\rho}(\mathbf{r}) = \sum_{i=1}^{N_v}|\widetilde{\psi}_i(\mathbf{r})|^2. \end{aligned} \tag{2.70}$$

The orbitals $\{\widetilde{\psi}_i\}$ are called pseudo-valence orbitals, and the eigenvalues $\{\widetilde{\varepsilon}_i\}$ are called pseudo-valence eigenvalues, respectively. Note that we need only solve N_v instead of N orbitals. Furthermore, we require that the solution to equation (2.70) satisfies the following conditions, which are called the Hamann–Schlüter–Chiang (HSC) conditions (Hamann *et al.* 1979).

(1) The pseudo-valence eigenvalues agree with the real valence eigenvalues:

$$\widetilde{\varepsilon}_i = \varepsilon_{i+N_c}, \quad i = 1, \ldots, N_v. \tag{2.71}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

(2) The pseudo-valence orbitals agree with the real valence orbitals outside a given radius r_c:

$$\widetilde{\psi}_i(\mathbf{r}) = \psi_{i+N_c}(\mathbf{r}), \quad r \geq r_c, \quad i = 1, \dots, N_v. \tag{2.72}$$

(3) The pseudo-valence orbitals are normalized, that is,

$$\langle \widetilde{\psi}_i | \widetilde{\psi}_i \rangle = 1. \tag{2.73}$$

In particular, condition (3) is called the norm-conservation condition, which leads to the name of this class of pseudopotential approximation. Although the pseudo-valence orbitals should form an orthonormal set of functions, in the discussion below we will associate the orbitals with spherical harmonic functions that are orthogonal to each other. Hence the orthogonality condition will be automatically satisfied, and only the normalization condition is imposed here.

The choice of pseudopotential satisfying the HSC conditions is by no means unique. In order to construct a V_{ps} satisfying the HSC conditions, for simplicity let us first consider the case where $N_v = 1$. Note that the pseudo-valence orbital $\widetilde{\psi}_1$ now becomes the ground state of equation (2.70), and must therefore be a nodeless function (Lieb and Loss 2001). On the other hand, the real valence orbital ψ_1 should be orthogonal to all core orbitals, and therefore must have nodes within the core region $r \leq r_c$. Hence the real and pseudo-valence orbitals have qualitatively different shapes within the core region, but this does not affect the shape of the valence orbitals outside the core region due to HSC condition (2). When $N_v = 1$, the valence orbital is expected to be spherically symmetric (called an s-orbital):

$$\psi_{N_c+1}(\mathbf{r}) = \varphi(r), \tag{2.74}$$

where we use the notation $\varphi(r)$ to emphasize that the valence orbital depends only on the radial distance r and is real. Similarly the pseudo-valence orbital should also be spherically symmetric and we denote

$$\widetilde{\psi}_1(\mathbf{r}) = \widetilde{\varphi}(r). \tag{2.75}$$

In such a case, V_{ps} can be chosen to be a local potential that depends only on r, denoted by $V_{\mathrm{loc}}(r)$. To see this, we may simply choose a pseudo-valence orbital $\widetilde{\varphi}(r)$ such that HSC conditions (2) and (3) are satisfied. In order to satisfy HSC condition (1), we may invert the Schrödinger equation (2.70) as

$$V_{\mathrm{loc}}(r) = \varepsilon_1 + \frac{\Delta_{\mathbf{r}} \widetilde{\psi}_1(\mathbf{r})}{2\psi_1(\mathbf{r})} - V_{\mathrm{Hxc}}[\widetilde{\rho}](r) = \varepsilon_1 + \frac{1}{2r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \widetilde{\varphi}}{\partial r} \right) - V_{\mathrm{Hxc}}[\widetilde{\rho}](r). \tag{2.76}$$

Note that $V_{\mathrm{loc}}(r)$ depends only on the radial distance due to the assumptions of $\widetilde{\psi}_1$ and $\widetilde{\rho}$. Figure 2.4 shows an example comparing the valence 2s orbital of the fluorine atom (F) in an all-electron calculation, and the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

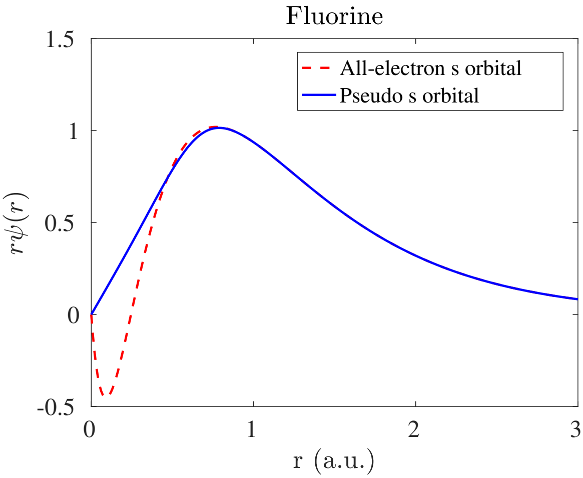


Figure 2.4. Comparison of the 2s orbital of the fluorine atom in an all-electron and a pseudopotential calculation. The cut-off radius r_c is set to 1.0 a.u.

pseudo-valence orbital using a Troullier–Martins pseudopotential (Troullier and Martins 1991). We find that the orbital becomes significantly smoother in the pseudopotential approximation. In particular, the pseudo-valence orbital is nodeless in the core region.

However, such a strategy becomes problematic when $N_v > 1$. To see this, following the notation of solutions to hydrogen-like atoms (Landau and Lifshitz 1991), we first relabel the real and pseudo-valence orbitals as (here $\hat{\mathbf{r}} = \mathbf{r}/r$)

$$\psi_{lm}(\mathbf{r}) = \varphi_l(r)Y_{lm}(\hat{\mathbf{r}}), \quad \widetilde{\psi}_{lm}(\mathbf{r}) = \widetilde{\varphi}_l(r)Y_{lm}(\hat{\mathbf{r}}). \tag{2.77}$$

The integers l, m are called the azimuthal and magnetic quantum numbers, respectively. For each azimuthal quantum number l, the admissible integer m should satisfy $-l \leq m \leq l$. Compared to the complete solution for hydrogen-like atoms, the choice of 'valence electron' fixes the principal quantum number implicitly. The radial parts $\varphi_l, \widetilde{\varphi}_l$ are real. $Y_{lm}(\hat{\mathbf{r}})$ is a spherical harmonic function defined on $\mathbb{S}^2$, which can also be viewed as a function $Y_{lm}(\theta, \varphi)$ that depends only on the angular degrees of freedom in spherical coordinates. The spherical harmonic functions satisfy the orthonormality condition

$$\begin{aligned}&\int_{\mathbb{S}^2} Y^*_{lm}(\hat{\mathbf{r}})Y_{l'm'}(\hat{\mathbf{r}})\,\mathrm{d}\hat{\mathbf{r}}\\ &\quad = \int_0^{\pi}\int_0^{2\pi} Y^*_{lm}(\theta,\varphi)Y_{l'm'}(\theta,\varphi)\sin^2\theta\,\mathrm{d}\varphi\,\mathrm{d}\theta = \delta_{ll'}\delta_{mm'}.\end{aligned} \tag{2.78}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Now if we choose a radial function $\widetilde{\varphi}_l$ for the pseudo-valence orbital $\widetilde{\psi}_{lm}$, we may invert the Schrödinger equation and obtain

$$\begin{aligned} V_{\mathrm{loc},l}(r) &= \varepsilon_l + \frac{\Delta_{\mathbf{r}}\widetilde{\psi}_{lm}(\mathbf{r})}{2\psi_{lm}(\mathbf{r})} - V_{\mathrm{Hxc}}[\widetilde{\rho}](r) \\ &= \varepsilon_l + \frac{1}{2r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\widetilde{\varphi}_l}{\partial r}\right) - \frac{l(l+1)}{2r^2}\widetilde{\varphi}_l(r) - V_{\mathrm{Hxc}}[\widetilde{\rho}](r). \end{aligned} \tag{2.79}$$

Due to the spherical symmetry, ε_l, $V_{\mathrm{loc},l}$ depend only on l but not on m. Again, $V_{\mathrm{loc},l}(r)$ is only a function of the radial distance r. Clearly V_{ps} cannot be equal to $V_{\mathrm{loc},l}(r)$ for all different choices of l.

Below we demonstrate that HSC condition (1) can be satisfied if we allow V_{ps} to be a *non-local potential*, *i.e.* an integral operator. We first choose the local potential to be V_{loc,l_0} for a particular angular momentum l_0. Then let the kernel of V_{ps} take the following form:

$$V_{\mathrm{ps}}(\mathbf{r},\mathbf{r}') = V_{\mathrm{loc},l_0}(r)\delta_{\mathbf{r},\mathbf{r}'} + \sum_{lm} Y_{lm}(\hat{\mathbf{r}})(V_{\mathrm{loc},l}(r) - V_{\mathrm{loc},l_0}(r))\delta_{r,r'}Y^*_{lm}(\hat{\mathbf{r}}')/r^2. \tag{2.80}$$

Applying equation (2.80) to a pseudo-valence orbital $\widetilde{\psi}_{lm}$, we have

$$\begin{aligned} (V_{\mathrm{ps}}\widetilde{\psi}_{lm})(\mathbf{r}) &= \int V_{\mathrm{ps}}(\mathbf{r},\mathbf{r}')\widetilde{\psi}_{lm}(\mathbf{r}')\,\mathrm{d}\mathbf{r}' \\ &= V_{\mathrm{loc},l_0}(r)\widetilde{\psi}_{lm}(\mathbf{r}) + Y_{lm}(\hat{\mathbf{r}})(V_{\mathrm{loc},l}(r) - V_{\mathrm{loc},l_0}(r))\widetilde{\varphi}_l(r) \\ &= V_{\mathrm{loc},l}(r)\widetilde{\psi}_{lm}(\mathbf{r}). \end{aligned} \tag{2.81}$$

Here we have used the form (2.77) and the orthonormality condition of spherical harmonics. Combined with equation (2.79), we find that HSC condition (1) is satisfied for any valence orbital $\widetilde{\psi}_{lm}$ under consideration.

Since the kernel of V_{ps} is local with respect to the radial variable, and non-local with respect to angular variables, equation (2.80) is called the semi-local form of the pseudopotential. However, for a general atomic configuration, the semi-local form cannot be treated efficiently other than being discretized as a dense matrix. Therefore the semi-local pseudopotential is rarely used in practice. In order to reduce the computational cost, we may note that it is only necessary for equation (2.81) to hold. Define

$$\delta V_{\mathrm{loc},l}(r) := V_{\mathrm{loc},l}(r) - V_{\mathrm{loc},l_0}(r). \tag{2.82}$$

We note that

$$\begin{aligned} \langle \delta V_{\mathrm{loc},l}\widetilde{\psi}_{lm}|\widetilde{\psi}_{l'm'}\rangle &= \int \delta V_{\mathrm{loc},l}(r)\widetilde{\varphi}^*_l(r)\widetilde{\varphi}_{l'}(r)Y^*_{lm}(\hat{\mathbf{r}})Y_{l'm'}(\hat{\mathbf{r}})\,\mathrm{d}\mathbf{r} \\ &= \delta_{ll'}\delta_{mm'}\int_0^\infty r^2\delta V_{\mathrm{loc},l}(r)|\widetilde{\varphi}_l(r)|^2\,\mathrm{d}r. \end{aligned} \tag{2.83}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Here we have used the orthonormality condition of the spherical harmonics. Now define the pseudopotential as

$$V_{\mathrm{ps}} = V_{\mathrm{loc},l_0} + \sum_{lm} \frac{1}{\langle \delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm} | \widetilde{\psi}_{lm} \rangle} |\delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm}\rangle \langle \delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm}|. \tag{2.84}$$

We may readily verify that

$$V_{\mathrm{ps}} \widetilde{\psi}_{lm} = V_{\mathrm{loc},l_0} \widetilde{\psi}_{lm} + \delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm} = V_{\mathrm{loc},l} \widetilde{\psi}_{lm}, \tag{2.85}$$

which means that V_{ps} also satisfies HSC condition (1). Compared to the semi-local form, the pseudopotential (2.84) consists of a local component, as well as a non-local component that can be stored as a low-rank matrix. The rank is equal to the number of pseudo-valence orbitals under consideration. This is called the Kleinman–Bylander form of non-local pseudopotential (Kleinman and Bylander 1982), and is used in almost all modern electronic structure software packages using norm-conserving pseudopotentials.

Let us define

$$b_{lm}(\mathbf{r}) = \frac{1}{\sqrt{|\langle \delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm} | \widetilde{\psi}_{lm} \rangle|}} \delta V_{\mathrm{loc}}(r) \widetilde{\varphi}_l(r) Y_{lm}(\hat{\mathbf{r}}), \tag{2.86}$$

which is called a projection vector. From HSC condition (2) and equation (2.79), we may readily find that $b_{lm}(\mathbf{r}) = 0$ if $r \geq r_c$, and hence is localized in real space. Defining $\gamma_l = \mathrm{sign}(\langle \delta V_{\mathrm{loc},l} \widetilde{\psi}_{lm} | \widetilde{\psi}_{lm} \rangle)$ and $V_{\mathrm{loc}}(\mathbf{r}) := V_{\mathrm{loc},l_0}(r)$, we may rewrite the kernel of the Kleinman–Bylander form as

$$V_{\mathrm{ps}}(\mathbf{r}, \mathbf{r}') = V_{\mathrm{loc}}(\mathbf{r}) \delta_{\mathbf{r},\mathbf{r}'} + \sum_{lm} \gamma_l b_{lm}(\mathbf{r}) b_{lm}^*(\mathbf{r}'). \tag{2.87}$$

Finally, the pseudopotential (2.87) only comes from one atom centred at the origin. For a general atomic configuration $\{\mathbf{R}_I\}_{I=1}^M$, the kernel of the non-local pseudopotential takes the form

$$\begin{aligned} V_{\mathrm{ps}}(\mathbf{r}, \mathbf{r}'; \{\mathbf{R}_I\}) &= \left\{ \sum_{I=1}^{M} V_{\mathrm{loc},I}(\mathbf{r} - \mathbf{R}_I) \right\} \delta_{\mathbf{r},\mathbf{r}'} \\ &\quad + \left\{ \sum_{I=1}^{M} \sum_{\ell=1}^{L_I} \gamma_{\ell,I} b_{\ell,I}(\mathbf{r} - \mathbf{R}_I) b_{\ell,I}^*(\mathbf{r}' - \mathbf{R}_I) \right\} \\ &= V_{\mathrm{loc}}(\mathbf{r}; \{\mathbf{R}_I\}) \delta_{\mathbf{r},\mathbf{r}'} + V_{\mathrm{nl}}(\mathbf{r}, \mathbf{r}'; \{\mathbf{R}_I\}). \end{aligned} \tag{2.88}$$

Here we have combined lm into a multi-index ℓ. Furthermore, $V_{\mathrm{loc},I}, b_{\ell,I}$ and the number of projection vectors L_I depend on the atom type and hence have the added index I. In this notation, V_{loc} and V_{nl} denote the collection of local and non-local parts of the pseudopotential, respectively. Equation (2.88) is the general form of norm-conserving pseudopotential and will be assumed throughout this review.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

More explicitly, the Kohn–Sham energy under the pseudopotential approximation becomes

$$E^{\mathrm{KS}} = \inf_{\{\psi_i\}_{i=1}^N, \langle\psi_i|\psi_j\rangle=\delta_{ij}} \tag{2.89}$$
$$\Biggl\{\frac{1}{2}\sum_{i=1}^{N}\int |\nabla\psi_i(\mathbf{r})|^2\,\mathrm{d}\mathbf{r} + \sum_{i=1}^{N}\int \psi_i^*(\mathbf{r})V_{\mathrm{nl}}(\mathbf{r},\mathbf{r}';\{\mathbf{R}_I\})\psi_i(\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'$$
$$+\int \rho(\mathbf{r})V_{\mathrm{loc}}(\mathbf{r};\{\mathbf{R}_I\})\,\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\mathrm{xc}}[\rho]\Biggr\} + E_{\mathrm{II}}.$$

The corresponding Kohn–Sham Hamiltonian becomes

$$H^{\mathrm{KS}}[\rho] = -\frac{1}{2}\Delta + V_{\mathrm{loc}} + V_{\mathrm{nl}} + V_H[\rho] + V_{\mathrm{xc}}[\rho]. \tag{2.90}$$

When the electron density is given, the effective potential (2.33) becomes

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{loc}}(\mathbf{r}) + V_H[\rho](\mathbf{r}) + V_{\mathrm{xc}}[\rho](\mathbf{r}), \tag{2.91}$$

where we have separated out the contribution from the non-local pseudopotential V_{nl}.

Remarks

Although not discussed here, another major reason for using pseudopotential is to take into account relativistic effects, which are non-negligible for heavy elements such as lead and gold. Since relativistic effects mainly affect the behaviour of core electrons, we may solve the relativistic Dirac–Kohn–Sham equation (Thaller 1992, Belpassi, Tarantelli, Sgamellotti and Quiney 2005) for a single atom to obtain the relativistically corrected valence orbitals ψ_{lm} and energies ε_l. The rest of the procedure for generating the pseudopotential is the same as above.

There are a number of widely used norm-conserving pseudopotentials, such as the Troullier–Martins (TM) pseudopotential (Troullier and Martins 1991), the Hartwigsen–Goedecker–Hutter (HGH) pseudopotential (Hartwigsen, Goedecker and Hutter 1998) and the optimized norm-conserving Vanderbilt (ONCV) pseudopotential (Hamann 2013), to name just a few. All these pseudopotentials can be written in the form (2.88), though the *interpretation* of V_{loc}, b_ℓ and even the number of projectors L may be different. Recently the semi-local form of the pseudopotential has been analysed mathematically (Cancès and Mourad 2016). However, the Kleinman–Bylander form of pseudopotential may occasionally introduce ‘ghost states’, which are unphysical states with artificially low energies. In practice such ghost states can often be removed by choosing the local potential to have a different angular momentum l, and hence they are often not considered to be

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

a problem in most electronic structure calculations. However, the existence of ghost states introduces difficulty into the mathematical analysis.

It may appear that the norm-conservation condition should be naturally imposed since the pseudo-valence orbitals $\widetilde{\psi}_{lm}$ should be eigenfunctions of a modified Schrödinger operator. However, it has been found that by relaxing the norm-conservation condition, one can further improve the smoothness of the pseudopotential and hence reduce the computational cost. The most well-known example is the Vanderbilt ultrasoft pseudopotential (Vanderbilt 1990). Moreover, the widely used projected augmented wave (PAW) method (Blöchl 1994) can also be viewed as a pseudopotential, which also violates the norm-conservation condition. Compared to norm-conserving pseudopotentials, one drawback of such approaches is that the Kohn–Sham equations become inherently a generalized eigenvalue problem even when the basis set is orthonormal (see Section 3).

2.8. Physical quantities of interest

Once the Kohn–Sham equations have converged, the total energy can be evaluated readily from equation (2.89). It can also be evaluated from the Kohn–Sham eigenvalues as

$$E^{\mathrm{KS}} = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' - \int \rho(\mathbf{r})V_{\mathrm{xc}}[\rho](\mathbf{r})\,\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[\rho] + E_{\mathrm{II}}. \tag{2.92}$$

Here the summation of the eigenvalues $\sum_{i=1}^{N} \varepsilon_i$ is called the band energy. The difference between the total energy and the band energy (other than the nuclei interaction energy E_{II}) is called the *double-counting* term, which is due to the nonlinearity of the Hartree energy and the exchange-correlation energy functionals. When finite temperature effects are included, the entropy also needs to be evaluated in order to compute the free energy.

In Kohn–Sham DFT, besides the total energy and the electron density, we are often interested in computing the atomic force, which is necessary for performing geometry relaxation and *ab initio* molecular dynamics simulation. Once the SCF iteration reaches convergence, the force on the Ith atom can be computed as the negative derivative of the total energy with respect to the atomic position $\mathbf{R}_I$:

$$\mathbf{F}_I = -\frac{\partial E^{\mathrm{KS}}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I}. \tag{2.93}$$

The required derivative can be computed directly, for example via finite differences. However, even for first-order accuracy, the number of energy evaluations for a system containing M atoms is $3M+1$, that is, the Kohn–Sham equations must be solved $3M+1$ times independently. This approach

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

becomes prohibitively expensive as the system size increases. The cost of the force calculation is greatly reduced via the Hellmann–Feynman theorem (Martin 2008), which states that, at self-consistency, the partial derivative $\partial/\partial\mathbf{R}_I$ only needs to be applied to terms in equation (2.89) which depend *explicitly* on the atomic position $\mathbf{R}_I$. The Hellmann–Feynman (HF) force is then given by

$$\begin{aligned}\mathbf{F}_I = &-\int \frac{\partial V_{\mathrm{loc}}}{\partial \mathbf{R}_I}(\mathbf{r};\{\mathbf{R}_I\})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} - \sum_{i=1}^{N}\int \psi_i^*(\mathbf{r})\frac{\partial V_{\mathrm{nl}}}{\partial \mathbf{R}_I}(\mathbf{r},\mathbf{r}';\{\mathbf{R}_I\})\psi_i(\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\\ &+\sum_{J\neq I}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|^3}(\mathbf{R}_I-\mathbf{R}_J). \qquad (2.94)\end{aligned}$$

Note that equation (2.88) gives

$$\frac{\partial V_{\mathrm{loc}}}{\partial \mathbf{R}_I}(\mathbf{r};\{\mathbf{R}_I\}) = \frac{\partial V_{\mathrm{loc},I}}{\partial \mathbf{R}_I}(\mathbf{r}-\mathbf{R}_I) = -\nabla_{\mathbf{r}} V_{\mathrm{loc},I}(\mathbf{r}-\mathbf{R}_I),$$

and

$$\begin{aligned}&\frac{\partial V_{\mathrm{nl}}}{\partial \mathbf{R}_I}(\mathbf{r},\mathbf{r}';\{\mathbf{R}_I\})\\ &\quad= \sum_{\ell=1}^{L_I}\gamma_{I,\ell}\biggl(\frac{\partial b_{I,\ell}}{\partial \mathbf{R}_I}(\mathbf{r}-\mathbf{R}_I)b^*_{I,\ell}(\mathbf{r}'-\mathbf{R}_I) + b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)\frac{\partial b^*_{I,\ell}}{\partial \mathbf{R}_I}(\mathbf{r}'-\mathbf{R}_I)\biggr)\\ &\quad= -\sum_{\ell=1}^{L_I}\gamma_{I,\ell}(\nabla_{\mathbf{r}} b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)b^*_{I,\ell}(\mathbf{r}'-\mathbf{R}_I) + b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)\nabla_{\mathbf{r}'} b^*_{I,\ell}(\mathbf{r}'-\mathbf{R}_I)).\end{aligned}$$

Then the Hellmann–Feynman force in equation (2.94) can be written as

$$\begin{aligned}\mathbf{F}_I = &\int \nabla_{\mathbf{r}} V_{\mathrm{loc},I}(\mathbf{r}-\mathbf{R}_I)\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}\\ &+ 2\mathrm{Re}\sum_{i=1}^{N}\sum_{\ell=1}^{L_I}\gamma_{I,\ell}\biggl(\int \psi_i^*(\mathbf{r})\nabla_{\mathbf{r}} b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)\,\mathrm{d}\mathbf{r}\biggr)\biggl(\int b^*_{I,\ell}(\mathbf{r}'-\mathbf{R}_I)\psi_i(\mathbf{r}')\,\mathrm{d}\mathbf{r}'\biggr)\\ &+\sum_{J\neq I}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|^3}(\mathbf{R}_I-\mathbf{R}_J). \qquad (2.95)\end{aligned}$$

From the computational cost point of view, if we let N_g denote the number of grid points to discretize quantities such as $\rho(\mathbf{r})$ in the global domain to perform a quadrature, then the cost of computing the force due to the local potential $\int \nabla_{\mathbf{r}} V_{\mathrm{loc},I}(\mathbf{r}-\mathbf{R}_I)\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}$ is $O(N_g)$, since $V_{\mathrm{loc},I}(\mathbf{r}-\mathbf{R}_I)$ is a delocalized quantity in the global domain. On the other hand, each non-local projector $b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)$ is localized around $\mathbf{R}_I$, and the cost of evaluating the integral $(\int \psi_i^*(\mathbf{r})\nabla_{\mathbf{r}} b_{I,\ell}(\mathbf{r}-\mathbf{R}_I)\,\mathrm{d}\mathbf{r})$ or $(\int b^*_{I,\ell}(\mathbf{r}'-\mathbf{R}_I)\psi_i(\mathbf{r}')\,\mathrm{d}\mathbf{r}')$ is a constant N_l independent of the global number of grid points N_g. The computation

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

of the last term,

$$\sum_{J\neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|^3}(\mathbf{R}_I - \mathbf{R}_J),$$

involves only scalar operations, and its cost is usually negligibly small in electronic structure calculations. N_g and N_I are proportional to the number of electrons N. Hence, neglecting constant terms independent of N, we have that the computational cost of the Hellmann–Feynman force on each atom is $O(N_g + NL_I N_l) \sim O(N)$, and the cost of computing forces on all atoms is $O(N^2)$.

Remarks

The computational cost of evaluating equation (2.95) can scale as $O(N^3)$ if implemented straightforwardly. This is because, for each atom, the associated non-local pseudopotential needs to be applied to all orbitals via an integral. The cost is thus $MNN_g \sim O(N^3)$. Note that each projection vector $b_{\ell,I}$ has (at least approximately) compact support in real space and can thus be stored as a sparse vector, using the real-space representation (see Section 3.1.2). The number of non-zeros in the sparse vector is independent of the number of quadrature points in the global domain. The cost is then reduced to $O(N^2)$ as discussed above. We remark that use of the real-space representation may result in higher numerical error than the Fourier space representation (see Section 3.1.1), especially if the real-space grid is not dense enough. Hence several electronic structure software packages such as Quantum ESPRESSO still prefer the Fourier space representation, even though the cost scales as $O(N^3)$. In certain contexts, the computational cost can be further reduced to $O(N)$ using density matrix formalism (see Section 3.3).

The atomic force evaluated as in equation (2.95) uses the Hellmann–Feynman theorem, and hence is called the Hellmann–Feynman force. When numerical discretization as in Section 3 is under consideration, the derivative of the Kohn–Sham energy may involve the derivative with respect to a basis function. If the basis set depends on the atomic configuration, such as in the case of Gaussian-type orbitals to be discussed in Section 3.2, then the Hellmann–Feynman force does not agree with the negative derivative of the Kohn–Sham energy. The remaining difference is called the Pulay force (Pulay 1969). On the other hand, if the discretization is independent of atomic configuration, such as in the case of the planewave basis set in Section 3.1, the Pulay force vanishes and the Hellmann–Feynman force becomes exact. We also remark that even if the basis set depends on the atomic configuration, the magnitude of the Pulay force will systematically decrease as the basis set approaches the complete basis set limit.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

3. Numerical discretization

In order to solve Kohn–Sham DFT in practice, the Hamiltonian operator must first be discretized, for instance by a finite-sized basis set. The efficiency of a discretization scheme can be measured in terms of the number of degrees of freedom per atom, or the dimension of the discretized Hamiltonian matrix. Besides this standard metric, one special feature of electronic structure calculations is that the choice of the discretization scheme can directly affect the effectiveness of subsequent numerical methods to evaluate the Kohn–Sham map, which is to be discussed in detail in Sections 5 and 6. In the following discussion, we divide the numerical discretization schemes roughly into three categories: large basis sets (Section 3.1), small basis sets (Section 3.2) and adaptive basis sets (Section 3.3).

3.1. Large basis sets

In electronic structure calculations, a large basis set typically requires $100 \sim 10\,000$ basis functions per atom to achieve chemical accuracy, even when pseudopotentials are used. The size of the resulting Hamiltonian matrix is usually of the order of $10^3 \sim 10^6$. Hence it is not numerically efficient, or even feasible at all, to diagonalize the Hamiltonian matrix. In such a case, iterative algorithms should be used to compute the occupied orbitals, which will be discussed in Section 4.2. Most standard basis sets used to solve PDEs numerically fall into this category, for instance the planewave method (Payne *et al.* 1992, Kresse and Furthmüller 1996) (also known as the Fourier basis set, or more precisely, the pseudospectral method), the finite element method (Tsuchida and Tsukada 1995, Suryanarayana *et al.* 2010, Bao, Hu and Liu 2012, Chen *et al.* 2014) and the wavelet method (Genovese *et al.* 2008), to name just a few. Strictly speaking, the finite difference method (Chelikowsky, Troullier and Saad 1994) does not use a basis set. However, the number of degrees of freedom needed by the finite difference method is approximately in the same range, and hence finite difference can also be regarded as a large basis set method. The main advantage of using a large basis set is that physical quantities such as energies and forces can converge systematically with respect to refinement of the basis set, by tuning one or a few parameters.

3.1.1. Fourier basis set

For charge-neutral systems, periodic boundary conditions are often sufficiently accurate when the Coulomb energy is treated with care (Makov and Payne 1995). One of the key advantages of working with periodic boundary conditions is that they allow use of the Fourier basis set, which is arguably the most widely used large basis set for electronic structure calculations. In the Fourier basis set, the basis functions are of the form $\exp(\mathrm{i}\mathbf{g}\cdot\mathbf{r})$, where the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

parameter $\mathbf{g}$ is chosen from a specific set of frequencies. The Fourier basis set has several key advantages. First, the discretization of the Kohn–Sham Hamiltonian takes a rather simple form under this basis set. Second, as it is independent of the atom configurations, the Pulay force vanishes, thus making the force calculation rather straightforward. Third, by leveraging the fast Fourier transform (FFT), computation of the matrix–vector multiplication with the Kohn–Sham Hamiltonian can be carried out efficiently, and this allows for efficient iterative diagonalization while evaluating the Kohn–Sham map. However, the Fourier basis set also comes with a noticeable disadvantage: as a fixed basis set, it lacks the flexibility of local refinement, which can be a serious issue for all-electron calculations. Fortunately, introduction of the pseudopotentials addresses this problem by eliminating the singularity at the nuclei, as well as the highly localized core electrons. This allows one to focus solely on the smooth valence electrons. In the rest of this subsection, we shall work with the Kohn–Sham Hamiltonian under the pseudopotentials approximation as in equation (2.90).

Let us consider for simplicity a rectangular computational domain $\Omega = [0, L_1] \times [0, L_2] \times [0, L_3]$ with periodic boundary conditions. It is natural to associate with Ω a reciprocal lattice in frequency space given by

$$\mathbb{L}^* = \left\{ \mathbf{g} = \left(\frac{2\pi}{L_1} i_1, \frac{2\pi}{L_2} i_2, \frac{2\pi}{L_3} i_3 \right) : (i_1, i_2, i_3) \in \mathbb{Z}^3 \right\},$$

and the Fourier basis functions are the complex exponentials

$$\phi_{\mathbf{g}}(\mathbf{r}) = \frac{1}{\sqrt{|\Omega|}} \exp(\mathrm{i}\mathbf{g} \cdot \mathbf{r})$$

indexed by each $\mathbf{g} \in \mathbb{L}^*$. In order to obtain a finite set of basis functions, one typically introduces an energy cut-off E_{cut} and restricts to only the Fourier modes indexed by

$$\mathbb{G}_{\text{cut}} := \left\{ \mathbf{g} \in \mathbb{L}^* : \frac{1}{2} |\mathbf{g}|^2 \leq E_{\text{cut}} \right\} \subset \mathbb{L}^*.$$

We will use N_b to denote the cardinality of $\mathbb{G}_{\text{cut}}$, *i.e.* the number of basis functions used within the energy cut-off.

In order to present the Kohn–Sham Hamiltonian under this basis, it is convenient to introduce two Cartesian grids, one in frequency space and one in real space. The one in frequency space is

$$\begin{aligned}\mathbb{G} = \Bigg\{ \mathbf{g} = \left(\frac{2\pi}{L_1} i_1, \frac{2\pi}{L_2} i_2, \frac{2\pi}{L_3} i_3 \right) :& \\ -\frac{n_1}{2} \leq i_1 < \frac{n_1}{2}, -\frac{n_2}{2} \leq i_2 < \frac{n_2}{2}, -\frac{n_3}{2} \leq i_3 < \frac{n_3}{2} \Bigg\},\end{aligned}$$

where n_1, n_2, and n_3 are chosen in such a way that the width of $\mathbb{G}$ in each dimension is typically twice the width of $\mathbb{G}_{\text{cut}}$. The reason for such a choice

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

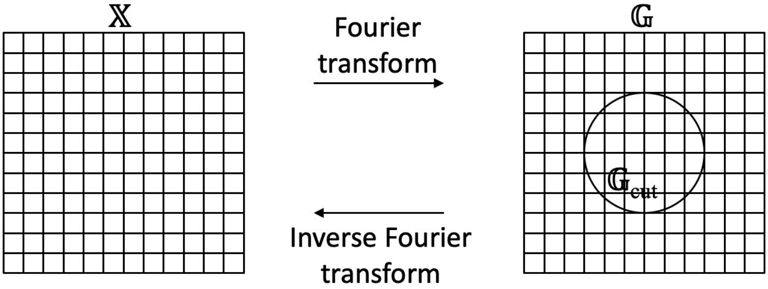


Figure 3.1. Illustration of the grids used for the Fourier basis set: $\mathbb{X}$, $\mathbb{G}$ and $\mathbb{G}_{\text{cut}}$.

will be discussed below. The Cartesian grid in real space is denoted by

$$\mathbb{X} = \left\{ \mathbf{r} = \left(\frac{i_1}{n_1} L_1, \frac{i_2}{n_2} L_2, \frac{i_3}{n_3} L_3 \right) : 0 \le i_1 < n_1, 0 \le i_2 < n_2, 0 \le i_3 < n_3 \right\}.$$

Note that the number of grid points in both grids is equal to $N_g = n_1 n_2 n_3$.

For a function $\psi(\mathbf{r})$ in the span of $\{\phi_{\mathbf{g}}(\mathbf{r}) : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$ with spanning coefficients $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$, its value at any point $\mathbf{r} \in \mathbb{X}$ is given by

$$\psi(\mathbf{r}) = \sum_{\mathbf{g} \in \mathbb{G}_{\text{cut}}} \hat{\psi}_{\mathbf{g}} \phi_{\mathbf{g}}(\mathbf{r}) = \sum_{\mathbf{g} \in \mathbb{G}_{\text{cut}}} \hat{\psi}_{\mathbf{g}} \frac{1}{\sqrt{|\Omega|}} \exp(\mathrm{i}\mathbf{g} \cdot \mathbf{r}). \tag{3.1}$$

Suppose now that $\psi(\mathbf{r})$ is a sufficiently smooth function. Then the basis coefficients $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}\}$ can be approximated using the samples of $\psi(\mathbf{r})$ at $\mathbb{X}$, computed via the inverse Fourier transform as

$$\begin{aligned} \hat{\psi}_{\mathbf{g}} &= \int_{\Omega} \psi(\mathbf{r}) \phi^*_{\mathbf{g}}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \\ &= \int_{\Omega} \psi(\mathbf{r}) \frac{1}{\sqrt{|\Omega|}} \exp(-\mathrm{i}\mathbf{g} \cdot \mathbf{r}) \, \mathrm{d}\mathbf{r} \approx \sum_{\mathbf{r} \in \mathbb{X}} \frac{\sqrt{|\Omega|}}{N_g} \psi(\mathbf{r}) \exp(-\mathrm{i}\mathbf{g} \cdot \mathbf{r}). \end{aligned} \tag{3.2}$$

Note that the last approximation is exact if $\psi(\mathbf{r})$ belongs to the span of $\{\phi_{\mathbf{g}}(\mathbf{r}) : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$.

Equations (3.1) and (3.2) show how to transform between the real-space and frequency-space representations for a function $\psi(\mathbf{r})$. Given the spanning coefficients $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$, the values $\{\psi(\mathbf{r}) : \mathbf{r} \in \mathbb{X}\}$ are computed by first extending $\hat{\psi}_{\mathbf{g}}$ from $\mathbb{G}_{\text{cut}}$ to $\mathbb{G}$ with zero-padding and then applying a discrete inverse Fourier transform. In the opposite direction, given the real-space values $\{\psi(\mathbf{r}) : \mathbf{r} \in \mathbb{X}\}$, the coefficients $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$ are computed by first applying a discrete Fourier transform and then restricting the result from $\mathbb{G}$ to $\mathbb{G}_{\text{cut}}$. In terms of computational efficiency, it is essential that both directions can be accelerated using the fast Fourier transform. Figure 3.1 illustrates the grids involved in these conversions.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

The Fourier basis set typically results in a dense matrix for the Kohn–Sham Hamiltonian in (2.89). Direct diagonalization of such dense matrices is often computationally too expensive. As a result, iterative diagonalization is often used instead (see Section 4.2). The key step of an iterative diagonalization algorithm is the application of the Kohn–Sham Hamiltonian to a given function represented in the Fourier basis set.

More precisely, for a function ψ given in terms of its Fourier coefficients $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$, the goal is to evaluate the Fourier coefficients of $H^{\text{KS}}\psi$ in $\mathbb{G}_{\text{cut}}$ effectively. Note from the definition (2.90) that the Kohn–Sham Hamiltonian H^{KS} can be decomposed into three parts:

- the kinetic energy part $-\frac{1}{2}\Delta$,
- the local potential part, that is,
$$V_{\text{loc}} + V_H[\rho] + V_{\text{xc}}[\rho],$$
where V_{loc} is the local part of the pseudopotential approximation,
- the non-local part of the pseudopotential approximation V_{nl}.

The overall strategy is to treat each part individually, either in real space or in frequency space depending on computational convenience.

Before describing in detail how each of these three parts is treated, we note that the Kohn–Sham Hamiltonian depends on the electron density $\rho(\cdot)$. In the actual numerical computation, one needs access to the values of $\rho(\cdot)$ at each $\mathbf{r} \in \mathbb{X}$. $\rho(\mathbf{r})$ is formed from the current copies of the Kohn–Sham orbitals $\{\psi_i(\mathbf{r})\}$ with $i = 1, \dots, N$. In the Fourier basis method, $\{\psi_i(\mathbf{r})\}$ are represented in the frequency domain $\hat{\psi}_{\mathbf{g};i}$ with $\mathbf{g} \in \mathbb{G}_{\text{cut}}$. The computation of $\rho(\mathbf{r})$ involves evaluating $\psi_i(\mathbf{r})$ for each i (*i.e.* one fast Fourier transform), squaring them, and summing the results over $i = 1, \dots, N$. In order to avoid the aliasing error coming from the squaring step, it is typically required that the sidelength of $\mathbb{G}$ is significantly larger than the diameter of the grid $\mathbb{G}_{\text{cut}}$ (for details see *e.g.* Fornberg 1998). In this case, the conventional choice in the electronic structure community is to set the sidelength of $\mathbb{G}$ to be at least twice the diameter of $\mathbb{G}_{\text{cut}}$, as mentioned above.

With $\rho(\mathbf{r})$ ready at each $\mathbf{r} \in \mathbb{X}$, let us now go through each of the three parts when applying the Kohn–Sham Hamiltonian H^{KS} to a function ψ. First, consider the kinetic energy part $-\frac{1}{2}\Delta\psi$. It is well known that a differential operator in real space is equivalent to a pointwise multiplication operator in frequency space. Therefore, applying the kinetic part $-\frac{1}{2}\Delta$ can be carried out by multiplying $\hat{\psi}_{\mathbf{g}}$ by $\frac{1}{2}|\mathbf{g}|^2$ for each $\mathbf{g} \in \mathbb{G}_{\text{cut}}$.

Second, for the computation of the local potential part, one first applies an inverse Fourier transform on $\{\hat{\psi}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\text{cut}}\}$ to obtain the real-space representation $\{\psi(\mathbf{r}) : \mathbf{r} \in \mathbb{X}\}$. As mentioned above, the local potential

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

consists of several parts,

$$(V_{\mathrm{loc}}(\mathbf{r}) + V_H[\rho](\mathbf{r}) + V_{\mathrm{xc}}[\rho](\mathbf{r}))\psi(\mathbf{r}), \tag{3.3}$$

The Hartree potential involves convolution in real space, and therefore it is convenient to treat it in the Fourier domain. More precisely, one first applies the Fourier transform to the electron density $\rho(\mathbf{r})$ with $\mathbf{r} \in \mathbb{X}$ to obtain $\{\widehat{\rho}_{\mathbf{g}} : \mathbf{g} \in \mathbb{G}_{\mathrm{cut}}\}$. Then each $\widehat{\rho}_{\mathbf{g}}$ is multiplied by $1/|\mathbf{g}|^2$, except that at $\mathbf{g} = 0$ it is kept at zero. This is due to the charge neutrality assumption, which allows the formally divergent contribution from the Fourier mode $\mathbf{g} = 0$ to be cancelled by the contribution from the electron–ion interaction. Finally, an inverse Fourier transform brings the result back to real space. When all the parts are ready, we compute the sum and multiply the result pointwise with $\psi(\mathbf{r})$. Note that the above procedure of computing derivative terms in frequency space and the multiplication terms in real space is the *pseudo-spectral method* for the numerical solution of partial differential equations (Fornberg 1998).

Third, the computation of the non-local potential part

$$\int \Biggl(\sum_{I=1}^{M} \sum_{\ell=1}^{L_I} \gamma_{\ell,I} b_{\ell,I}(\mathbf{r} - \mathbf{R}_I) b^*_{\ell,I}(\mathbf{r}' - \mathbf{R}_I)\Biggr) \psi(\mathbf{r}')\,\mathrm{d}\mathbf{r}' \tag{3.4}$$

can be treated either in frequency or in real space. When the basis functions $b_{\ell,I}(\mathbf{r}')$ are localized, the treatment in real space is often preferred due to its simplicity. More precisely, one evaluates the integral $\int b^*_{\ell,I}(\mathbf{r}')\psi(\mathbf{r}')\,\mathrm{d}\mathbf{r}'$ for each pair (ℓ, I) with a discrete sum over $\mathbb{X}$ and then scales $b_{\ell,I}(\mathbf{r})$ with this integral value before summing them up over (ℓ, I).

At this point, one holds the contribution from both local and non-local potentials in the spatial grid $\mathbb{X}$. In order to obtain its coefficients in frequency space on the grid $\mathbb{G}_{\mathrm{cut}}$, one simply applies a discrete fast Fourier transform.

3.1.2. Real-space representation

The above discussion is given in terms of the Fourier basis set. Most of the computation can be carried out almost equivalently in terms of a real-space basis function set. Therefore, for each $\mathbf{r}' \in \mathbb{X}$ we define the coefficients

$$\hat{\varphi}_{\mathbf{g};\mathbf{r}'} = \frac{1}{\sqrt{N_g}}\,\mathrm{e}^{-\,\mathrm{i}\mathbf{g}\cdot\mathbf{r}'}$$

with $\mathbf{g} \in \mathbb{G}$. Its inverse Fourier transform is a real-space function

$$\varphi_{\mathbf{r}'}(\mathbf{r}) = \sum_{\mathbf{g}\in\mathbb{G}} \hat{\varphi}_{\mathbf{g};\mathbf{r}'}\phi_{\mathbf{g}}(\mathbf{r}) = \frac{1}{\sqrt{N_g|\Omega|}} \sum_{\mathbf{g}\in\mathbb{G}} \exp(\mathrm{i}\mathbf{g}\cdot(\mathbf{r}-\mathbf{r}')). \tag{3.5}$$

This is called the periodic sinc function (Skylaris, Haynes, Mostofi and Payne 2005), or psinc function for short. In particular, the psinc functions

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

can be viewed as the numerical δ-function on the discrete set $\mathbb{X}$:

$$\varphi_{\mathbf{r}'}(\mathbf{r}) = \sqrt{\frac{N_g}{|\Omega|}}\delta_{\mathbf{r},\mathbf{r}'}, \quad \mathbf{r},\mathbf{r}' \in \mathbb{X}. \tag{3.6}$$

Hence a smooth function $\psi(\mathbf{r})$ can be expanded as

$$\psi(\mathbf{r}) \approx \sum_{\mathbf{r}' \in \mathbb{X}} \varphi_{\mathbf{r}'}(\mathbf{r})\psi(\mathbf{r}'). \tag{3.7}$$

The basis set $\{\varphi_{\mathbf{r}'}\}$ is often called the psinc basis set, or the planewave dual basis set. For a given function, the nodal representation (3.7) allows us to identify its function values evaluated at Cartesian grid points with the expansion coefficients under the psinc basis set. This is particularly convenient for describing many numerical algorithms below, such as the selected columns of the density matrix method (Section 4.3) and the interpolative separable density fitting method (Section 4.4). If a discretization scheme allows such a nodal representation, it is often referred to as a *real-space representation*. In such a case, the number of basis functions N_b can be identified with the number of grid points N_g. The finite difference discretization can also be viewed as a real-space representation. However, with some abuse of notation we shall use the words *psinc basis set* and *real-space representation* interchangeably in the following discussion. When the real-space representation is used, we may also slightly abuse the notation by using $\mathbf{r}$ to denote an element from a *discrete set* of points $\{\mathbf{r}_i\}_{i=1}^{N_g}$.

3.2. Small basis set

In electronic structure calculations, a small basis set typically only requires $10 \sim 100$ basis functions per atom to achieve chemical accuracy. This can be the case even for all-electron calculations in the absence of pseudopotentials. The reason why such a small basis set can be achieved without significant deterioration of the accuracy is that in quantum chemistry the electron orbitals usually do not vary arbitrarily in the presence of a chemical environment. Hence useful information can be extracted from the atomic limit, which gives rise to the atomic basis set. The size of the discretized Hamiltonian matrix is often of the order of $10^2 \sim 10^4$, which makes direct diagonalization of the Hamiltonian matrix a viable approach (and often the fastest option for small systems too) to solve Kohn–Sham DFT. For certain discretization schemes and exchange-correlation functionals, the discretized Hamiltonian is even a sparse matrix due to the spatial localization of the basis set. This permits the use of efficient numerical methods to reduce the asymptotic complexity.

In order to understand the atomic basis set, let us first recall the one-body Schrödinger equation for hydrogen-like atoms, given in spherical coordinates

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

(r, θ, φ) by

$$E\psi(\mathbf{r}) = \tag{3.8}$$
$$-\frac{1}{2}\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right)\psi(\mathbf{r})-\frac{Z}{r}\psi(\mathbf{r}).$$

The spherical part of the operator is given by

$$-\frac{1}{r^2}\boldsymbol{L}^2 := -\frac{1}{r^2}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right), \tag{3.9}$$

where $\boldsymbol{L}^2$ is the spherical Laplacian operator, whose eigenfunctions are given by the spherical harmonics $Y_{lm}(\theta, \varphi)$ parametrized by the azimuthal quantum number l and the magnetic quantum number m. For each non-negative integer l, the admissible m are given by $-l, -l+1, \ldots, l$. The eigenfunction of the one-body Schrödinger operator for hydrogen-like atoms is then given by

$$\psi_i(\mathbf{r}) = \frac{u_i(r)}{r}Y_{lm}(\hat{\mathbf{r}}), \tag{3.10}$$

where the term r in the denominator is used to capture the singularity of the solution near the origin $r = 0$. For notational simplicity we let l, m depend implicitly on the orbital index i.

The idea of the atomic basis set is then to use functions of the type (3.10) centred at each atom in the system to discretize the Kohn–Sham equations. Given the nuclei positions $\{\mathbf{R}_I\}$, the basis set is given by

$$\left\{\varphi_i(\mathbf{r}-\mathbf{R}_I) = \frac{u_i(|\mathbf{r}-\mathbf{R}_I|)}{|\mathbf{r}-\mathbf{R}_I|}Y_{lm}\left(\frac{\mathbf{r}-\mathbf{R}_I}{|\mathbf{r}-\mathbf{R}_I|}\right) : i = 1, \ldots, n_I,\ I = 1, \ldots, M\right\}, \tag{3.11}$$

where n_I atomic orbitals are used for the Ith nucleus. Such discretization has the clear advantage that the basis set is able to capture the singularity of the solutions to the Kohn–Sham problem near the nuclei. Hence only a small number of degrees of freedom are needed to discretize the Kohn–Sham orbitals. Moreover, the atomic basis set can be used even without the pseudopotential approximation.

Based on different choices of the radial part of the basis functions u_i, some widely used small basis sets are as follows.

(1) *Slater-type orbital (STO).* The radial part takes the form

$$u(r) = Cr^n\,\mathrm{e}^{-\zeta r}, \tag{3.12}$$

where the non-negative integer n plays the role of principal quantum number, ζ is a constant related to the effective charge of the nucleus, and C is the normalization factor. The physical motivation of STO is clear, as (3.12) gives the radial part of the eigenfunctions of hydrogen-like atoms.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

(2) *Gaussian-type orbitals (GTO).* Numerical integration involving STO can be difficult (we will further discuss the quadrature issues below). Gaussian-type orbitals – Gaussian functions or Gaussians multiplied by polynomials – are proposed as basis functions to avoid this difficulty, since many integrals can then be calculated explicitly. Many GTO basis sets have been proposed (Jensen 2013) in the quantum chemistry literature, starting from the idea of fitting STO using a few Gaussians (known as the STO-nG minimal basis), to the most widely used correlation-consistent basis sets (Dunning 1989).

(3) *Numerical atomic-orbitals (NAO).* Instead of using a predetermined analytical form of the basis functions, the idea of NAO is to obtain u_i by numerically solving a Schrödinger-like radial equation (after writing wavefunctions as products of the radial part and spherical harmonics, as in (2.77)). Thus we have

$$\left(-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{l(l+1)}{r^2}+v_i(r)+v_{\text{cut}}(r)\right)u_i(r)=\epsilon_i u_i(r), \tag{3.13}$$

where $(l(l+1))/r^2$ comes from the spherical Laplacian and the choice of spherical harmonics with azimuthal quantum number l, $v_i(r)$ is a radial potential chosen to control the main behaviour of u_i, and $v_{\text{cut}}(r)$ is a confining potential to ensure that u_i decays rapidly beyond a certain radius and can be treated as a compactly supported function. We refer readers to Blum *et al.* (2009) for details. The compact support ensures the locality of the resulting discrete Hamiltonian and facilitates numerical computations based on the NAO.

Let us denote the collection of atomic orbital basis functions by

$$\{\phi_p(\mathbf{r}) : p = 1, \dots, N_b\}.$$

In general, the basis set is not orthonormal. Therefore, by the usual Galerkin projection, the Kohn–Sham equation becomes a (nonlinear) generalized eigenvalue problem,

$$Hc_i = \varepsilon_i S c_i, \tag{3.14}$$

where H is the discrete Hamiltonian matrix

$$H_{pq} = \langle \phi_p | H | \phi_q \rangle, \tag{3.15}$$

and S is the overlap matrix (*i.e.* Gram matrix)

$$S_{pq} = \langle \phi_p | \phi_q \rangle. \tag{3.16}$$

Since the basis functions $\{\phi_p\}$ are by construction either compactly supported or decay rapidly, the resulting H and S matrices are sparse, which enables efficient algorithms for solving the Kohn–Sham eigenvalue problems, as will be discussed in Section 5.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

After solving (3.14), the Kohn–Sham orbitals are

$$\psi_i(\mathbf{r}) = \sum_p \phi_p(\mathbf{r}) c_{p,i}, \tag{3.17}$$

and thus the electron density is given by

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 = \sum_{i=1}^{N} \sum_{p,q=1}^{N_b} c^*_{p,i} c_{q,i} \phi^*_p(\mathbf{r}) \phi_q(\mathbf{r}). \tag{3.18}$$

Thus, to obtain the contribution to the matrix elements H_{ij} from the Hartree potential, the two-electron repulsion integral (we adopt the physicists' notation, as opposed to the chemists' notation) of the form

$$\langle pq|rs\rangle = \iint \frac{\phi^*_p(\mathbf{r})\phi^*_q(\mathbf{r}')\phi_r(\mathbf{r})\phi_s(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}' \tag{3.19}$$

is needed. In the case of Gaussian orbitals, the above integral can be obtained analytically by reducing the problem to integrals of Gaussians using the integral representation of the Coulomb kernel

$$\frac{1}{r} = \frac{1}{\pi^{1/2}} \int_{-\infty}^{+\infty} \exp(-r^2\xi^2) \, \mathrm{d}\xi. \tag{3.20}$$

For other basis functions, numerical quadratures are needed, for example by using the density fitting technique (see Section 4.4). We also refer readers to Reine, Helgaker and Lindh (2012) for a review of multi-electron integrals.

As discussed in Section 2.8, since small basis sets are typically constructed to be centred around each atom, the Hellmann–Feynman force may not be accurate enough for the atomic force, unless a large number of basis functions are used in the calculation. In such a case, the Pulay force involving derivative quantities $\{\partial\phi_p/\partial\mathbf{R}_I\}$ is needed. We refer readers to Soler *et al.* (2002), for example, for details of the implementation in the context of numerical atomic orbitals.

3.3. Adaptive basis set

As discussed in Section 3.1, a large basis set allows one to systematically improve the accuracy of the numerical discretization by tuning one or a handful of parameters (*e.g.* the kinetic energy cut-off E_{cut} in the planewave method and the grid spacing in the finite difference method). The disadvantage of a large basis set is the much higher number of degrees of freedom per atom. On the other hand, as shown in Section 3.2, a small basis set can significantly reduce the number of degrees of freedom per atom, but it can be more difficult to improve its quality in a systematic fashion, and the improvement may rely heavily on the practitioner's experience with the underlying chemical system. An adaptive basis set aims to combine the best of

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

both worlds, namely to achieve systematic improvability while significantly reducing the number of degrees of freedom per atom.

While adaptive basis sets are not yet as widely used as large or small basis sets, there have been a number of developments along this direction in the past few decades. Examples include the non-orthogonal generalized Wannier function (NGWF) approach in the ONETEP package (Skylaris *et al.* 2005), the adaptive minimal basis approach in the BigDFT package (Mohr *et al.* 2014), the filter diagonalization approach (Rayson and Briddon 2009), the localized spectrum slicing approach (Lin 2017) and the adaptive local basis set approach (Lin, Lu, Ying and E 2012*a*), to name just a few. Below we briefly introduce the adaptive local basis set (ALB) (Lin *et al.* 2012*a*) approach, which uses a discontinuous basis set for electronic structure calculations.

The basic idea of ALB is to partition the global domain into a number of subdomains (called elements), and to generate basis functions that are adapted to the Kohn–Sham solutions for each element. This allows the resulting basis functions to capture the structure of atomic orbitals but also effects from the chemical environment. Each basis function is continuous inside its associated supporting element but discontinuous across element boundaries. Then the discontinuous Galerkin (DG) method (see *e.g.* Babuška and Zlámal 1973, Arnold 1982, Cockburn, Karniadakis and Shu 2000) is used to construct a finite-dimensional projected Kohn–Sham Hamiltonian matrix with significantly reduced dimension. The projected Kohn–Sham problem can be solved efficiently in a similar way to the case when a small basis set is used.

Let Ω denote the global computational domain with periodic boundary conditions, and let $\mathcal{K}$ be a collection of quasi-uniform rectangular partitions of Ω into non-overlapping elements:

$$\mathcal{K} = \{\kappa_1, \kappa_2, \dots, \kappa_M\}. \tag{3.21}$$

For $\kappa \in \mathcal{K}$, we let $\overline{\kappa}$ denote the closure of κ. The periodic boundary condition on Ω implies that the partition is regular across the boundary $\partial\Omega$.

We let $H^1(\kappa)$ denote the standard Sobolev space of $L^2(\kappa)$-functions such that the first partial derivatives are also in $L^2(\kappa)$. We denote the set of piecewise H^1-functions by

$$H^1(\mathcal{K}) = \{v \in L^2(\Omega) : v|_\kappa \in H^1(\kappa),\ \text{for all } \kappa \in \mathcal{K}\},$$

which is also referred to as the broken Sobolev space. For $v, w \in H^1(\mathcal{K})$, we define the L^2-inner product as

$$(v, w)_{\mathcal{K}} = \sum_{\kappa\in\mathcal{K}} (v, w)_\kappa := \sum_{\kappa\in\mathcal{K}} \int_\kappa v^*(\mathbf{r}) w(\mathbf{r})\,\mathrm{d}\mathbf{r}, \tag{3.22}$$

which induces a norm $\|v\|_{\mathcal{K}} = (v, v)_{\mathcal{K}}^{1/2}$. For $v, w \in H^1(\mathcal{K})$ and $\kappa, \kappa' \in \mathcal{K}$,

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

define the average and jump operators on a face $\overline{\kappa} \cap \overline{\kappa}'$ by

$$\langle v \rangle = \frac{1}{2}(v|_\kappa + v|_{\kappa'}), \quad \langle \nabla v \rangle = \frac{1}{2}(\nabla v|_\kappa + \nabla v|_{\kappa'}), \tag{3.23}$$

and

$$[\![v]\!] = v|_\kappa \mathbf{n}_\kappa + v|_{\kappa'} \mathbf{n}_{\kappa'}, \quad [\![\nabla v]\!] = \nabla v|_\kappa \cdot \mathbf{n}_\kappa + \nabla v|_{\kappa'} \cdot \mathbf{n}_{\kappa'}, \tag{3.24}$$

where $\mathbf{n}_\kappa$ denotes the exterior unit normal of the element κ.

In order to solve Kohn–Sham DFT in the broken Sobolev space $H^1(\mathcal{K})$, we also need to identify a basis set which spans a subspace of $H^1(\mathcal{K})$. Let N_κ be the number of DOFs on κ, and the total number of DOFs is $N_b = \sum_{\kappa \in \mathcal{K}} N_\kappa$. Let $\mathbb{V}(\kappa) = \text{span } \{\phi_{\kappa,j}\}_{j=1}^{N_\kappa}$, where each $\phi_{\kappa,j}$ is a function defined on Ω compactly supported in κ. Hence $\mathbb{V}(\kappa)$ is a subspace of $H^1(\mathcal{K})$ and is associated with a finite-dimensional approximation for $H^1(\kappa)$. Then $\mathbb{V} = \bigoplus_{\kappa \in \mathcal{K}} \mathbb{V}(\kappa)$ is a finite-dimensional approximation to $H^1(\mathcal{K})$. We also assume all functions $\{\phi_{\kappa,j}\}$ form an orthonormal set in the sense that

$$(\phi_{\kappa,j}, \phi_{\kappa',j'})_{\mathcal{K}} = \delta_{\kappa,\kappa'} \delta_{j,j'}, \quad \text{for all } \kappa, \kappa' \in \mathcal{K}, 1 \leq j \leq N_\kappa, 1 \leq j' \leq N_{\kappa'}. \tag{3.25}$$

For the moment, we assume that the basis set $\mathbb{V}$ has been given. For $w, v \in \mathbb{V}$, we introduce the following bilinear form:

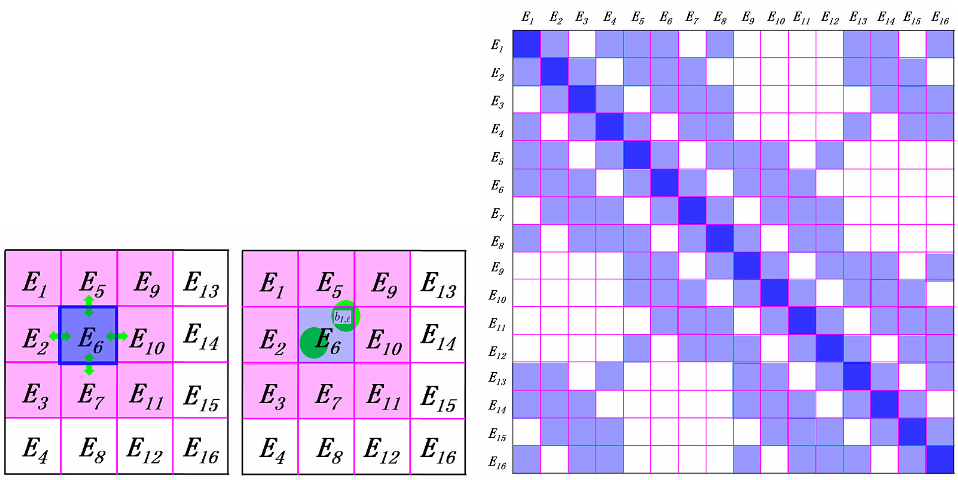
$$\begin{aligned} a(w, v) = \sum_{\kappa \in \mathcal{K}} \biggl[& \frac{1}{2}(\nabla w, \nabla v)_\kappa + ((V_{\text{eff}} + V_{\text{nl}})w, v)_\kappa \\ & - (\nabla w, [\![\psi_v]\!])_{\partial\kappa} - ([\![w]\!], \nabla v)_{\partial\kappa} + \alpha([\![w]\!], [\![v]\!])_{\partial\kappa} \biggr]. \end{aligned} \tag{3.26}$$

Here the first term on the right-hand side is the kinetic energy, and the second term is the effective local potential and the non-local potential as in equation (2.91). The third and fourth terms are obtained from integration by parts of the Laplacian operator, which cures the ill-defined operation of applying the Laplacian operator to discontinuous functions on the global domain. The last term is a penalty term to guarantee the stability condition (Arnold, Brezzi, Cockburn and Marini 2002). The penalty parameter needs to be sufficiently large, and the value of α depends on the choice of basis set $\mathbb{V}$.

In order to find the approximate Kohn–Sham orbitals $\{\psi_i\}_{i=1}^N \subset \mathbb{V}$, that is,

$$\psi_i(\mathbf{r}) = \sum_\kappa \sum_{j=1}^{N_\kappa} c_{\kappa,j;i} \phi_{\kappa,j}(\mathbf{r}), \tag{3.27}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047


E_1 E_5 E_9 E_{13}
E_2 E_6 E_{10} E_{14}
E_3 E_7 E_{11} E_{15}
E_4 E_8 E_{12} E_{16}
$b_{1,4}$
E_1 E_2 E_3 E_4 E_5 E_6 E_7 E_8 E_9 E_{10} E_{11} E_{12} E_{13} E_{14} E_{15} E_{16}

(a) (b) (c)

Figure 3.2. (Credit: Hu, Lin and Yang 2015.) A model system in 2D partitioned into 16 (4 × 4) equal-sized elements. The elements are denoted by E_i.

we minimize the following energy functional:

$$\mathcal{E}^{\mathrm{DG}}(\{\psi_i\}) = \sum_{i=1}^{N} a(\psi_i, \psi_i), \tag{3.28}$$

subject to the orthonormal constraint for $\{\psi_i\}$. The associated Euler–Lagrange equation corresponding to this minimization problem gives rise to the following eigenvalue problem:

$$\sum_{\kappa',j'} H^{\mathrm{DG}}_{\kappa,j;\kappa',j'} c_{\kappa',j';i} = \varepsilon_i^{\mathrm{DG}} c_{\kappa,j;i}, \quad i = 1, \ldots, N, \tag{3.29}$$

where the projected Hamiltonian matrix is given by

$$H^{\mathrm{DG}}_{\kappa,j;\kappa',j'} = a(\phi_{\kappa,j}, \phi_{\kappa',j'}). \tag{3.30}$$

The H^{DG} matrix can be naturally partitioned into matrix blocks as sketched in Figure 3.2. The kinetic energy and the local pseudopotential can only contribute to the diagonal matrix blocks. The terms from boundary integrals can contribute to both the diagonal and the off-diagonal blocks of H^{DG} corresponding to adjacent elements. Each boundary term involves only two neighbouring elements by definition, as plotted in Figure 3.2(a). For the non-local pseudopotential, since a projection vector of the non-local pseudopotential is spatially localized, we require the dimension of every element along each direction (usually around 6 ∼ 8 Bohr) to be larger than

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the size of the non-zero support of each projection vector (usually around $2 \sim 4$ Bohr). Thus, the non-zero support of each projection vector can overlap with at most 2^d elements as shown in Figure 3.2(b) ($d = 1, 2, 3$). As a result, each non-local pseudopotential term may also contribute to both the diagonal and the off-diagonal blocks corresponding to adjacent elements. In summary, H^{DG} is a sparse matrix and the non-zero matrix blocks correspond to interactions between neighbouring elements (Figure 3.2(c)).

From the solution of the eigenvalue problem (3.29), we may compute the electron density as

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \sum_{\kappa} \left| \sum_{j} c_{\kappa,j;i} \phi_{\kappa,j}(\mathbf{r}) \right|^2 . \tag{3.31}$$

Note that the computation of the electron density can be performed locally in each element due to the locality of the basis functions $\phi_{\kappa,j}$. Then the electron density can be fed back to the next step of the SCF iteration.

There are several methods to generate the basis set $\mathbb{V}$ in the broken Sobolev space. Here we introduce the adaptive local basis set in Lin, Lu, Ying and E (2012*a*), which is obtained by solving local Kohn–Sham problems. For each element κ, we form an *extended element* $\widetilde{\kappa}$ around κ, and we refer to $\widetilde{\kappa}\backslash\kappa$ as the buffer region for κ. On $\widetilde{\kappa}$ we solve the eigenvalue problem

$$\left(-\frac{1}{2}\Delta + V_{\mathrm{eff}}^{\widetilde{\kappa}} + V_{\mathrm{nl}}^{\widetilde{\kappa}}\right)\widetilde{\phi}_{\kappa,j} = \lambda_{\kappa,j}\widetilde{\phi}_{\kappa,j}. \tag{3.32}$$

with certain boundary conditions on $\partial\widetilde{\kappa}$. Here we define $V_{\mathrm{eff}}^{\widetilde{\kappa}} = V_{\mathrm{eff}}|_{\widetilde{\kappa}}$ to be the restriction of the effective potential at the current SCF step to $\widetilde{\kappa}$, and $V_{\mathrm{nl}}^{\widetilde{\kappa}} = V_{\mathrm{nl}}|_{\widetilde{\kappa}}$ to be the restriction of the non-local potential to $\widetilde{\kappa}$. This eigenvalue problem can be solved using a standard basis set such as finite difference, finite elements or planewaves. Note that the size of the extended element $\widetilde{\kappa}$ is independent of the size of the global domain, and so is the number of basis functions per element. We then restrict $\{\widetilde{\phi}_{\kappa,j}\}$ from $\widetilde{\kappa}$ to κ. The truncated functions are not necessarily orthonormal. Therefore, we apply a singular value decomposition (SVD) to obtain $\{\phi_{\kappa,j}\}$. The SVD procedure can ensure the orthonormal constraint of the basis functions inside each element, as well as eliminating the approximately linearly dependent functions in the basis set. We then extend each $\phi_{k,j}$ to the global domain by setting it to zero outside κ to generate the basis set $\mathbb{V}$. As a result, the overlap matrix corresponding to the adaptive local basis set is an identity matrix.

There are a number of possible ways to set the boundary conditions for the local problem (3.32). In practice, we use the periodic boundary condition for all eigenfunctions $\{\widetilde{\phi}_{\kappa,j}\}$ in $\widetilde{\kappa}$. In some sense, the details of the boundary condition do not affect the accuracy of the adaptive local basis set much as the buffer size increases. The periodic boundary condition permits the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

use of Fourier basis sets for solving the local problem. Note that while the eigenfunctions are required to satisfy periodic boundary conditions, the potential is obtained from the restriction of the global Kohn–Sham potential to the extended element. The size of each extended element should be chosen to balance between the effectiveness of the basis functions and the computational cost of obtaining them. For a typical choice used in practice, the elements are chosen to be of the same size, and each element contains on average a few atoms. The partition does not need to be updated when the atomic configuration is changed, as in the case of structure optimization and molecular dynamics.

Once the SCF iteration reaches convergence, we may evaluate the atomic forces. Recall the discussion in Section 2.8 that the atomic force generally consists of two components: the Hellmann–Feynman force and the Pulay force. Since each ALB depends on the atomic configuration, the Pulay force does not vanish as in the planewave basis set. Compared to atomic orbitals, the Pulay force for the adaptive basis set is much more difficult to evaluate. Nonetheless, the adaptive basis sets are constructed to be aware of the environment of each atom, and can be much closer to a complete basis set for describing occupied states than numerical atomic orbitals. Numerical results indicate that the Pulay force for the atomic basis set is indeed small, and use of the Hellman–Feynman formula can readily achieve chemical accuracy (Zhang *et al.* 2017). We also remark that using the spatial locality of the basis set $\mathbb{V}$, the computational cost of evaluating the force for all atoms can be reduced from $O(N^2)$ to $O(N)$. Compared to equation (2.95), the main reduction of the computational cost comes from (i) the localized pseudocharge associated with the local pseudopotential $V_{\text{loc},I}$ (the pseudocharge is defined as $-\Delta V_{\text{loc},I}/4\pi$), and (ii) the density matrix formulation for the contribution from the non-local pseudopotential. We refer readers to Zhang *et al.* (2017) for more details.

The adaptive local basis set and the DG formulation have been implemented in the DGDFT (discontinuous Galerkin density functional theory) software package. DGDFT is a massively parallel electronic structure software package designed for large-scale DFT calculations involving up to tens of thousands of atoms (Hu, Lin and Yang 2015). As an illustrating example, Figure 3.3 shows the ALBs of a 2D phosphorene monolayer with 140 phosphorus atoms (P_{140}). This is a two-dimensional system, and the global domain is partitioned into 64 equal-sized elements along the Y and Z directions, respectively. We show the isosurfaces of the first three ALB functions for the element denoted by E_{10} in Figure 3.3(a–c). Each ALB function shown is strictly localized inside E_{10} and is therefore discontinuous across the boundary of elements. On the other hand, each ALB function is delocalized across a few atoms inside the element since they are obtained from eigenfunctions of local Kohn–Sham Hamiltonian. Although the basis

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

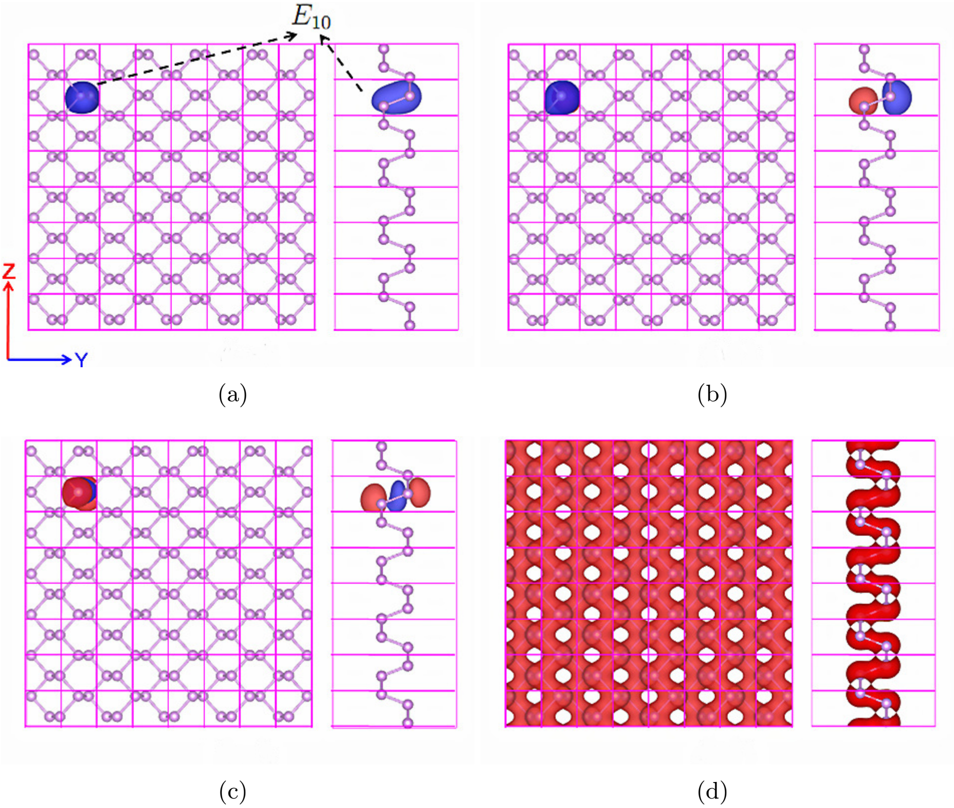
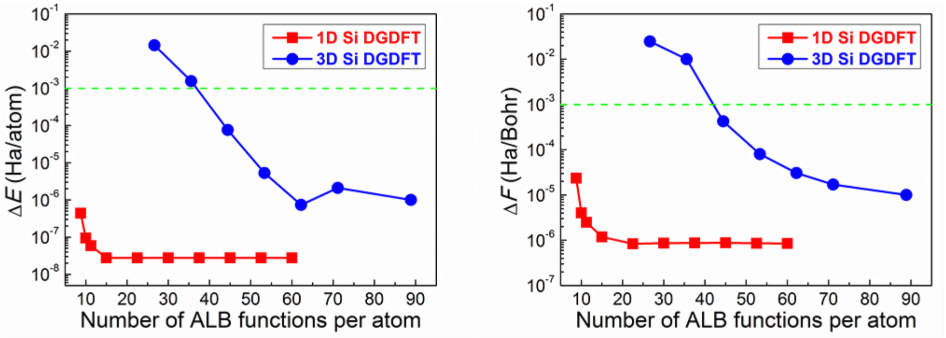


Figure 3.3. (Credit: Hu, Lin and Yang 2015.) The P_{140} system, with the isosurfaces (0.04 Hartree/Bohr^3) of the first three ALB functions belonging to the tenth element E_{10}: (a) ϕ_1, (b) ϕ_2, (c) ϕ_3 and (d) the electron density ρ in the top and side views in the global domain in the example of P_{140}. The red and blue regions indicate positive and negative isosurfaces, respectively. There are 64 elements and 80 ALB functions in each element, corresponding to 37 basis functions per atom.

functions are discontinuous, the electron density is well-defined and is very close to being a continuous function in the global domain (Figure 3.3(d)).

Figure 3.4 shows the convergence of the total energy and atomic forces for the quasi-1D and 3D Si systems with an increasing number of ALB functions per atom. First, we see that chemical accuracy is obtained with less than ten basis functions per atom for the quasi-1D system and a few tens of basis functions per atom for the 3D system. Furthermore, for the quasi-1D Si system, we find that the total energy error can be as small as 2.78×10^{-8} Ha/atom and the maximum error of the atomic forces can be as small as 8.47×10^{-7} Ha/Bohr when 22.5 ALB functions per atom are used.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047



(a) (b)

Figure 3.4. (Credit: Zhang *et al.* 2017.) Convergence of DGDFT total energy and forces for quasi-1D and 3D Si systems to reference planewave results with increasing number of ALB functions per atom. (a) Total energy error per atom ΔE (Ha/atom). (b) Maximum atomic force error ΔF (Ha/Bohr). The dashed green line corresponds to chemical accuracy. We set $E_{\mathrm{cut}} = 60$ Ha and penalty parameter $\alpha = 40$. When the number of basis functions per atom is sufficiently large, the error is well below the target accuracy (green dashed line).

For the bulk 3D Si system, the total energy error can be as small as 2.10×10^{-6} Ha/atom and the maximum error in atomic forces can be as small as 1.70×10^{-5} Ha/Bohr when 71.11 ALB functions per atom are used.

Remarks

The adaptive basis set is closely related to the ‘hybrid basis set’, which directly combines the large basis set and the small basis set. The idea of such a combination dates back to Slater in his augmented planewave method (APW) (Slater 1937), which was later improved to become the linearized APW method (Andersen 1975). The idea is to use an atomic basis set near the atom to resolve the singularity due to the electron–nucleus interaction in all-electron calculations, and use smooth functions such as planewaves in the interstitial region. Since planewaves are used without compression, the number of degrees of freedom per atom in the hybrid basis set is still relatively large. Recently there have been further developments along this direction using a partition of unity approach (Cai, Bai, Pask and Sukumar 2013) and the discontinuous Galerkin approach (Lu, Cai, Xin and Guo 2013).

The interior penalty formulation of DG requires setting a penalty parameter. For general non-polynomial basis functions, the value of the penalty parameter is not known *a priori*. We have demonstrated that the adjustable

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

penalty parameter is mainly used to ensure the stability of the numerical scheme, and has relatively little effect on the accuracy of the scheme when it takes a large range of values (Lin *et al.* 2012*a*, Hu *et al.* 2015). Another possible solution can be found in Lin and Stamm (2016), which provides a formula for evaluating α on the fly for general non-polynomial basis sets based on the solution of eigenvalue problems restricted to each element κ. This further leads to rigorous *a posteriori* error estimates for using non-polynomial basis functions in the DG setting to solve linear PDEs and eigenvalue problems (Kaye, Lin and Yang 2015, Lin and Stamm 2016, Lin and Stamm 2017). This allows adaptive refinement of the basis set to improve its efficiency for heterogeneous systems.

The adaptive local basis set has been improved in a number of directions. As in the NGWF approach, the adaptive local basis set can be obtained via numerical optimization (Lin, Lu, Ying and E 2012*b*), which rigorously eliminates the Pulay force when the minimizer is achieved. The basis functions can also be obtained from solutions to linear systems on the global domain, which removes the need to impose artificial boundary conditions on the extended element and simultaneously improves the accuracy of the basis set (Li and Lin 2019). The adaptive local basis set can also be constructed by partitioning the discretized basis indices instead of the continuous domain (Xu, Suryanarayana and Pask 2018). This is called the discrete discontinuous basis projection method. Formally, this is a Galerkin method instead of a discontinuous Galerkin approach, and the advantage of this method is that it removes the need to impose penalty conditions on the basis set.

4. Algorithmic tools

This section introduces a few algorithmic tools that are useful in electronic structure calculations. We introduce direct and iterative eigensolvers in Sections 4.1 and 4.2, respectively. These are the standard methods for solving the eigenvalue problem obtained in Kohn–Sham DFT after discretization. In Section 4.3 we discuss how to compress information on the Kohn–Sham orbitals through localization. Localized representation of the occupied space forms the basis for linear scaling algorithms for insulating systems to be discussed in Section 5.2. Then in Section 4.4 we introduce the interpolative separable density fitting (ISDF) technique to compress the information stored in the pair products of Kohn–Sham orbitals, which will become useful in reducing the computational cost of Kohn–Sham DFT calculations with rung-4 and rung-5 functionals. We remark that eigensolvers, localization and density fitting are generic techniques that are useful in a number of other contexts of scientific computing. Hence we summarize these techniques separately in this section.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

4.1. Direct diagonalization methods

When a small basis set is used and the system size is not too large, it is most efficient to solve the standard or generalized eigenvalue problem (3.14) using a generic eigensolver for dense matrices. It can be performed in serial using the LAPACK software package (Anderson *et al.* 1999) or in parallel using the ScaLAPACK software package (Blackford *et al.* 1997). Recently dense solvers specialized for solving large-scale Hermitian eigenvalue problems, such as ELPA (Marek *et al.* 2014), have been demonstrated to be more efficient and to scale massively in parallel to more than 100 000 computational cores. As a rule of thumb, dense eigensolvers are most efficient when the matrix dimension is 10 000 or less, but are not commonly used when the matrix dimension is beyond 100 000, due to the steep increase of the computational cost that scales as $O(N_b^3)$.

4.2. Iterative diagonalization methods

When a large basis set is used, iterative diagonalization methods that rely on the matrix–vector product of $H^{\mathrm{KS}}[\rho]$ with arbitrary vectors become particularly attractive, especially when the matrix–vector product can be performed efficiently. The two common cases are when $H^{\mathrm{KS}}[\rho]$ is sparse and when $H^{\mathrm{KS}}[\rho]$ is constructed using the Fourier basis set so that the fast Fourier transform can be used. Among the iterative diagonalization methods, we focus on Krylov subspace methods here. The discussion of another class of methods called filtering methods will be deferred to Section 5.1.

The key idea of Krylov subspace methods is to form a low-dimensional subspace and approximate the eigenpairs with those of the projected Hamiltonian into this subspace. Of these methods, the Davidson–Liu method (Davidson 1975, Liu 1978) and the LOBPCG method (Knyazev 2001) are widely used due to their simplicity and efficiency.

Consider for simplicity the standard eigenvalue problem

$$HX = X\Lambda,$$

where the columns of the $N_b \times N$ matrix X are the eigenvectors and the diagonal entries of the $N \times N$ diagonal matrix Λ give the eigenvalues.

In its simplest version, the Davidson–Liu method solves this problem by iteratively minimizing $\mathrm{Tr}[X^*HX]$ over a matrix X that satisfies the orthogonality constraint $X^*X = I$ and takes its columns from a subspace spanned by a collection of $2N$ vectors. These $2N$ vectors are set to be the columns of the matrix X from the previous step and another $N_b \times N$ matrix W, which is a preconditioned residual of X. More precisely, in each iteration one updates X via

$$X \leftarrow XC_X + WC_W, \quad W = TR := T(HX - X(X^*HX)),$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 1: Davidson–Liu method for solving the Kohn–Sham DFT eigenvalue problems $H\psi_i = \varepsilon_i \psi_i$

```
Input: Hamiltonian matrix $H$ and initial wavefunctions $\{\psi_i^0\}_{i=1}^N$.
Output: Eigenvalues $\{\varepsilon_i\}_{i=1}^N$ and wavefunctions $\{\psi_i\}_{i=1}^N$.
Initialize $X$ by $\{\psi_i^0\}_{i=1}^N$ and orthonormalize $X$.
while convergence not reached do
    Compute the preconditioned residual $W \leftarrow T(HX - X(X^*HX))$,
    where $T$ is a preconditioner.
    Update the trial subspace $S \leftarrow [X, W]$.
    Solve the projected eigenvalue problem $S^*HSC = S^*SC\Lambda$ and
    obtain the coefficients $C = \begin{bmatrix} C_X \\ C_W \end{bmatrix}$.
    Compute $X \leftarrow SC$.
end while
Update $\{\psi_i\}_{i=1}^N \leftarrow X$.
```

where $R = HX - X(X^*HX)$ is the residual and T is a preconditioner that is ideally close to $(H - \lambda I)^{-1}$ with a certain shift λ. The original version of the Davidson–Liu method allows one to use many copies of the history; here, in order to simplify the discussion, we keep only the most recent copy of X and W. The coefficients C_X and C_W are obtained by computing the lowest N eigenpairs of the projected $2N \times 2N$ generalized eigenvalue problem

$$S^*HSC = S^*SC\Lambda, \quad C = \begin{bmatrix} C_X \\ C_W \end{bmatrix} \in \mathbb{R}^{2N\times N},$$

where $S = [X, W] \in \mathbb{R}^{N_b \times 2N}$ represents the trial subspace and C contains the optimal coefficients. The Davidson–Liu method is summarized in Algorithm 1.

The LOBPCG algorithm improves on the Davidson–Liu algorithm by introducing a conjugate search direction. More precisely, it solves $HX = X\Lambda$ by iteratively minimizing $\mathrm{Tr}[X^*HX]$ over X that satisfies the orthogonality constraint $X^*X = I$ and takes its columns from a subspace spanned by $3N$ vectors. These $3N$ vectors are the columns of the matrix X of the previous step, its preconditioned residual W and an extra matrix P that consists of certain conjugate directions. The coefficients C_X, C_W and C_P are obtained by computing the lowest N eigenpairs of the projected $3N \times 3N$ generalized eigenvalue problem

$$S^*HSC = S^*SC\Lambda, \quad C = \begin{bmatrix} C_X \\ C_W \\ C_P \end{bmatrix} \in \mathbb{R}^{3N\times N},$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 2: LOBPCG method for solving the Kohn–Sham DFT eigenvalue problems $H\psi_i = \varepsilon_i\psi_i$

```
Input: Hamiltonian matrix $H$ and initial wavefunctions $\{\psi_i^0\}_{i=1}^N$.
Output: Eigenvalues $\{\varepsilon_i\}_{i=1}^N$ and wavefunctions $\{\psi_i\}_{i=1}^N$.
Initialize $X$ by $\{\psi_i^0\}_{i=1}^N$ and orthonormalize $X$.
while convergence not reached do
    Compute the preconditioned residual $W \leftarrow T(HX - X(X^*HX))$, where $T$ is a preconditioner.
    Update the trial subspace $S \leftarrow [X, W, P]$.
    Solve the projected eigenvalue problem $S^*HSC = S^*SC\Lambda$ and obtain the coefficients $C = \begin{bmatrix} C_X \\ C_W \\ C_P \end{bmatrix}$.
    Compute the conjugate gradient direction $P \leftarrow WC_W + PC_P$.
    Compute $X \leftarrow XC_X + P$.
end while
Update $\{\psi_i\}_{i=1}^N \leftarrow X$.
```

where $S = [X, W, P]$ is the trial subspace. The LOBPCG method is outlined in Algorithm 2.

Note that when N is relatively small ($N \sim$ 10–1000), the computational cost of the $2N \times 2N$ or $3N \times 3N$ projected eigenvalue problem in the Rayleigh–Ritz step is negligible. However, when N is relatively large ($N \sim$ 1000–10 000), the cost of solving such projected eigenvalue problems can become dominant. Furthermore, even with recent advances in dense eigensolvers such as ELPA (Marek *et al.* 2014), the Rayleigh–Ritz step can still limit the parallel scalability and dominate the wall clock time of the Davidson–Liu and LOBPCG algorithms when a large number of processors are used.

More recently, the projected preconditioned conjugate gradient (PPCG) (Vecharynski, Yang and Pask 2015) algorithm has been proposed to reduce the cost of the Rayleigh–Ritz step. The main idea of PPCG is to replace the $3N \times 3N$ projected eigenvalue problem of LOBPCG by N subproblems, each of size 3×3. Compared to the solution of the subspace problem from a full Rayleigh–Ritz step, PPCG relaxes the optimality of the search direction at each step, but the advantage is that all N subproblems can be solved independently. The computational cost of this step scales linearly with N, and is negligibly small even when N is large. The orbitals still need to be orthogonalized at each step to avoid degeneracy. We remark that the idea of the PPCG method is closely related to another variant of the LOBPCG method called LOBPCG II (Knyazev 2001).

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Remarks

Algorithms 1 and 2 update all columns of X simultaneously. This is often referred to as a block eigensolver, which can efficiently utilize BLAS 3 level matrix–matrix multiplication operations, and is suitable for parallelization. The disadvantage of a block eigensolver is that the cost of the Rayleigh–Ritz step can be expensive when N becomes large. Different eigenvectors often converge with different rates, and the subspace problem may thus become degenerate. In such a case, the converged eigenvectors should be deflated from the unconverged ones to avoid instability. Another strategy is to solve the eigenpairs one by one, starting from the smallest eigenvalue. Such a method is often referred to as a single-band method (Payne *et al.* 1992) in electronic structure calculation. Since each eigenvector is directly orthogonalized with all previously converged eigenvectors, the single-band method can be numerically more stable. Besides the eigensolvers mentioned in this section, the linear eigenvalue problem can also be solved via optimization methods under orthogonality constraints (Payne *et al.* 1992, Edelman, Arias and Smith 1998, Yang, Meza and Wang 2006, Wen and Yin 2013). We also remark that optimization techniques can be used to minimize the nonlinear Kohn–Sham energy functional directly with respect to the Kohn–Sham orbitals.

The choice of the preconditioner T can significantly affect the convergence rate of eigensolvers. In the context of the planewave basis set, the most commonly used preconditioner (Teter, Payne and Allan 1989) is diagonal in frequency space, $T = F^{-1}\hat{T}F$, where F is the discrete Fourier transform matrix, with the coefficients $\hat{T}$ given by

$$\hat{T}(\mathbf{g},\mathbf{g}') = \delta_{\mathbf{g},\mathbf{g}'}\left(\frac{27 + 18E_{\mathbf{g}} + 12E_{\mathbf{g}}^2 + 8E_{\mathbf{g}}^3}{27 + 18E_{\mathbf{g}} + 12E_{\mathbf{g}}^2 + 8E_{\mathbf{g}}^3 + 16E_{\mathbf{g}}^4}\right), \tag{4.1}$$

where $E_{\mathbf{g}} = |\mathbf{g}|^2/2$ is the kinetic energy associated with the Fourier mode $\mathbf{g}$. The purpose of the choice (4.1) is that the Fourier modes of long wavelengths (*i.e.* $|\mathbf{g}|$ is small) remain approximately unchanged, while Fourier modes of short wavelengths (*i.e.* $|\mathbf{g}|$ is large) are dampened by $1/E_{\mathbf{g}}$, which is approximately $2(-\Delta)^{-1}$. The rationale is that the Laplacian operator is an unbounded operator, while the pseudopotential, Hartree and exchange-correlation potentials are usually bounded operators. As the planewave basis set is being refined, the Laplacian operator dominates the spectral radius of the Hamiltonian, especially along the Fourier modes with short wavelengths. Hence, along such modes, the preconditioner T approximates H^{-1} well.

The preconditioner (4.1) can be efficiently implemented in the context of the planewave basis set using only two FFTs per orbital, but it becomes considerably more difficult in other large basis sets such as the finite

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

element method, and the finite difference method. In such a case, one often performs preconditioning by directly taking $T = -\Delta^{-1}$, and applying the preconditioner to the residual vector r amounts to solving a Poisson equation.

4.3. Localization methods

Localized representations of electronic wavefunctions have a wide range of applications in quantum physics, chemistry and materials science. As they require significantly less cost for both storage and computation, localized representations are the foundation of so-called 'linear scaling methods' (Kohn 1996, Goedecker 1999, Bowler and Miyazaki 2012) for solving quantum problems (see more discussion in Section 5.2). They can also be used to analyse the chemical bonding in complex materials, interpolate the band structure of crystals, accelerate ground and excited state electronic structure calculations, and form reduced-order models for strongly correlated many-body systems (Marzari *et al.* 2012).

The localization problem can be stated as follows. The Kohn–Sham orbitals $\{\psi_i\}_{i=1}^N$ are eigenfunctions of a Hamiltonian matrix, and are generally delocalized across the entire system, *i.e.* with significant magnitude in large portions of the computational domain. Nonetheless, if we apply a unitary rotation $U \in \mathbb{C}^{N\times N}$ to the Kohn–Sham orbitals

$$w_i(\mathbf{r}) = \sum_{j=1}^{N} \psi_j(\mathbf{r}) U_{ji}, \tag{4.2}$$

the density matrix and hence the electron density are invariant with respect to such rotation. However, for certain systems there exist unitary transformations for which the resulting orthonormal functions $\{w_i\}_{i=1}^N$ are approximately only supported on a small portion of the computational domain. For isolated systems, $\{w_i\}_{i=1}^N$ are often called Boys orbitals (Foster and Boys 1960), while for periodic systems they are often called Wannier functions (Wannier 1937). Below we will simply refer to the localized orbitals as Wannier functions.

For isolated systems surrounded by a vacuum, the occupied orbitals always decay exponentially to zero as $|\mathbf{r}| \to \infty$ (Lieb and Loss 2001). Hence qualitatively, both the Kohn–Sham orbitals ψ_i and the localized orbitals w_i decay exponentially, and the localization is defined only in a quantitative sense.

On the other hand, for periodic crystals, the Kohn–Sham orbitals satisfy the Bloch boundary condition on each unit cell, and hence are delocalized on a macroscopic scale. However, for insulating systems with a finite band gap, the kernel of the density matrix decays exponentially along the off-diagonal direction (Kohn 1959, Blount 1962, Nenciu 1983, E and Lu 2011, Benzi,

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Boito and Razouk 2013, Lin and Lu 2016), in the sense that

$$|P(\mathbf{r},\mathbf{r}')| \lesssim \mathrm{e}^{-C|\mathbf{r}-\mathbf{r}'|}, \quad |\mathbf{r}-\mathbf{r}'| \to \infty. \tag{4.3}$$

This is related to the *near-sightedness principle* in the physics literature (Kohn 1996, Prodan and Kohn 2005).

If the system is also topologically trivial, then one may further find a proper choice of the gauge to construct a set of exponentially localized Wannier functions (Kohn 1959, Blount 1962, Kivelson 1982, Nenciu 1983, Brouder *et al.* 2007, E, Li and Lu 2010, Panati and Pisante 2013). Therefore, localized representation can make a qualitative difference. In the past few decades, many numerical algorithms have been designed to compute such localized orbitals (Marzari and Vanderbilt 1997, Koch and Goedecker 2001, Gygi 2009, Damle, Lin and Ying 2015, Mustafa, Coh, Cohen and Louie 2015, Cancès, Levitt, Panati and Stoltz 2017, Damle, Lin and Ying 2017, Damle and Lin 2018).

For isolated systems, the localized representation can be identified by minimizing the following *spread functional* (Foster and Boys 1960, Marzari and Vanderbilt 1997):

$$\begin{aligned}\inf_{U}\ \Omega[\{w_i\}_{i=1}^{N}] &= \sum_{i=1}^{N}\bigl[\langle w_i|\mathbf{r}^2|w_i\rangle - (\langle w_i|\mathbf{r}|w_i\rangle)^2\bigr]\\ \text{subject to } w_i &= \sum_{j=1}^{N}\psi_j U_{ji},\ U^*U = I_N.\end{aligned} \tag{4.4}$$

In three-dimensional space, the formula for the functional Ω should be interpreted as

$$\sum_{\alpha=1}^{3}[\langle w_i|\mathbf{r}_\alpha^2|w_i\rangle - (\langle w_i|\mathbf{r}_\alpha|w_i\rangle)^2] := \langle w_i|(\mathbf{r}-\langle\mathbf{r}\rangle_i)\cdot(\mathbf{r}-\langle\mathbf{r}\rangle_i)|w_i\rangle,$$

where $\langle\mathbf{r}\rangle_i := \langle w_i|\mathbf{r}|w_i\rangle$. Hence the spread functional can be written compactly as

$$\Omega[\{w_i\}_{i=1}^{N}] = \sum_{i=1}^{N}\langle w_i|(\mathbf{r}-\langle\mathbf{r}\rangle_i)^2|w_i\rangle.$$

The functional Ω characterizes the total spatial spread of the rotated orbitals $\{w_i\}$, in terms of the second moment around each centre $\langle\mathbf{r}\rangle_i$. A smaller spread value indicates a more localized representation of the Kohn–Sham occupied subspace. Numerically, the localization problem (4.4) can be solved as a constrained minimization problem. One main drawback of the localization procedure above is the reliance on a nonlinear optimization subproblem, and in practice such nonlinear optimization algorithms can

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

frequently get stuck at local minima, which can lead to a qualitatively different physical interpretation.

The selected column of the density matrix (SCDM) method (Damle *et al.* 2015) is a recently developed procedure to overcome this problem. For simplicity we focus on the case of isolated, insulating systems below, and the formulation can be generalized to periodic crystals as well as metallic systems (Damle and Lin 2018). We also assume that the real-space representation is used, and we do not distinguish a continuous vector $\psi(\mathbf{r})$ and the corresponding discretized vector $\{u(\mathbf{r}_i)\}_{i=1}^{N_g}$, where the grid point $\mathbf{r}_i$ also serves as a basis index in the real-space representation.

For an isolated system, the density matrix P is a spectral projector of rank N to the occupied space. The kernel of the density matrix remains the same under the unitary transformation

$$P(\mathbf{r},\mathbf{r}') = \sum_{i=1}^{N} w_i(\mathbf{r}) w_i^*(\mathbf{r}').$$

Thus, if there *exists* a set of localized orbitals $\{w_i\}$, $P(\mathbf{r},\mathbf{r}')$ also decays rapidly as $|\mathbf{r}-\mathbf{r}'| \to \infty$. Intuitively, if we can select a set of N points $\mathcal{C} := \{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N}$ so that the corresponding column vectors of the kernel $\{P(\mathbf{r},\hat{\mathbf{r}}_\mu)\}_{i=1}^{N}$ are the 'most representative' and well-conditioned column vectors of P, these vectors almost form the desired Wannier functions up to the orthonormality condition.

In order to select the set $\mathcal{C}$, using a real-space representation, we let $\Psi = [\psi_1, \dots, \psi_N] \in \mathbb{C}^{N_g \times N}$ be a matrix with orthonormal columns. The corresponding discretized density matrix kernel, still denoted by P, is given by $P = \Psi\Psi^*$. From a numerical linear algebra point of view, the most representative column vectors can be identified via a QR factorization with column-pivoting (QRCP) (Golub and Van Loan 2013) applied to P, that is,

$$P\Pi = \widetilde{Q}\widetilde{R}. \tag{4.5}$$

Here $\Pi \in \mathbb{R}^{N_g \times N_g}$ is a permutation matrix, $\widetilde{Q} \in \mathbb{C}^{N_g \times N_g}$ is a unitary matrix, and $\widetilde{R} \in \mathbb{C}^{N_g \times N_g}$ is an upper triangular matrix. The QRCP decomposition is not unique, but the absolute value of the diagonal entries $|\widetilde{R}_{ii}|$ should always follow a non-increasing order, *i.e.* $|\widetilde{R}_{11}| \geq |\widetilde{R}_{22}| \geq \cdots$. Note that P is a rank-N matrix, and hence we must have $\widetilde{R}_{n,n} = 0, n > N$. The row indices of the non-zero entries in the first N columns of Π directly give the set $\mathcal{C}$.

One drawback of this method is that the cost of QRCP applied to P scales as $O(N_g^3)$, or at least $O(N_g^2 N)$. This becomes prohibitively expensive when a large basis set representation is used. The SCDM method (Damle *et al.* 2015) proposes that the set $\mathcal{C}$ can be equivalently computed via the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

QRCP of the matrix Ψ^* as

$$\Psi^*\Pi = QR \equiv Q[R_1, R_2]. \tag{4.6}$$

Here $Q \in \mathbb{C}^{N\times N}$ is a unitary matrix of reduced dimension, $R_1 \in \mathbb{C}^{N\times N}$ is an upper triangular matrix, and $R_2 \in \mathbb{C}^{N\times(N_g-N)}$ is a rectangular matrix. The cost of the factorization (4.6) only scales as $O(N_gN^2)$, which is similar to the cost of the matrix orthogonalization step in iterative eigensolvers as discussed in Section 4.2. Once the factorization (4.6) is obtained, we readily have

$$P\Pi = \Psi\Psi^*\Pi = (\Psi Q)[R_1, R_2].$$

Hence QRCP applied to Ψ^* *implicitly* provides a QRCP for the density matrix P.

Having chosen $\mathcal{C}$, we must now orthonormalize the localized column vectors $\{P(\mathbf{r}, \hat{\mathbf{r}}_\mu)\}_{\mu=1}^N$ without destroying their locality. Note that

$$P(\mathbf{r}, \hat{\mathbf{r}}_\mu) = \sum_{i=1}^N \psi_i(\mathbf{r})\Xi_{i,\mu},$$

where $\Xi \in \mathbb{C}^{N\times N}$ has matrix elements

$$\Xi_{i,\mu} = \psi_i^*(\hat{\mathbf{r}}_\mu). \tag{4.7}$$

One way to enforce the orthogonality is

$$w_\mu(\mathbf{r}) = \sum_{i=1}^N \psi_i(\mathbf{r})U_{i,\mu}, \quad U = \Xi(\Xi^*\Xi)^{-1/2}. \tag{4.8}$$

It is clear that $U \in \mathbb{C}^{N\times N}$ is a unitary matrix, and U is in fact the minimizer of the orthogonal Procrustes problem:

$$U = \operatorname*{arg\,min}_{V^*V=I} \|V - \Xi\|_F. \tag{4.9}$$

In the quantum chemistry literature, the solution of the orthogonal Procrustes problem via the matrix square-root transformation in equation (4.8) is called the Löwdin transformation (Löwdin 1950).

Note that

$$(\Xi^*\Xi)_{\mu,\nu} = \sum_{i=1}^N \psi_i(\hat{\mathbf{r}}_\mu)\psi_i^*(\hat{\mathbf{r}}_\nu) = P(\hat{\mathbf{r}}_\mu, \hat{\mathbf{r}}_\nu). \tag{4.10}$$

The decay properties of the matrix P imply that $\{P(\hat{\mathbf{r}}_\mu, \hat{\mathbf{r}}_\nu)\}$ is a localized matrix, but of smaller size $N \times N$. If the eigenvalues $(\Xi^*\Xi)^{-1/2}$ are bounded away from 0, then $(\Xi^*\Xi)^{-1/2}$ will itself be localized (E and Lu 2011, Benzi *et al.* 2013), and consequently $\{w_i\}_{i=1}^N$ will be localized, orthonormal Wannier functions. Pseudocode for the SCDM algorithm is given in Algorithm 3.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 3: Selected columns of the density matrix (SCDM) algorithm

```
Input: Kohn–Sham orbitals $\Psi \in \mathbb{C}^{N_g \times N}$
Output: Localized orbitals $W \in \mathbb{C}^{N_g \times N}$
1: Compute the selected columns set $\mathcal{C} := \{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N}$ using QRCP
   via (4.6).
2: Evaluate the $\Xi$ matrix via (4.7).
3: Perform orthogonalization via (4.8).
```

Below we demonstrate the performance of the SCDM method by computing localized basis functions for two three-dimensional systems. Figure 4.1(a) shows one of the orthogonalized SCDM orbitals, obtained from a silicon crystal with 512 atoms consisting of 4×4×4 unit cells with a diamond structure, and the dimension of each unit cell is 10.26 × 10.26 × 10.26 a.u. Figure 4.1(b) shows one of the orthogonalized SCDM orbitals, obtained from a water system with 64 molecules in a cubic supercell with dimension 22.08 × 22.08 × 22.08 a.u. The orbitals are very localized in real space and resemble the shape of maximally localized Wannier functions (Marzari and Vanderbilt 1997). In fact, in this case the method automatically finds the centres of all localized orbitals, which for the silicon crystal are in the middle of the Si–Si bond, and for water is closer to the oxygen atoms than to the hydrogen atoms.

Remarks

Besides using Löwdin transformation, the gauge matrix U can also be directly identified from the QRCP (4.6) via a simple choice $U = Q$. This in fact corresponds to another orthogonalization method called the Cholesky-QR factorization, and in this context the Cholesky decomposition can be performed explicitly or implicitly (Golub and Van Loan 2013, Damle *et al.* 2017). Numerical results indicate that the two orthogonalization methods have comparable performance. However, the Löwdin transformation can be better when the system has certain spatial symmetries, and can be more easily generalized to periodic crystals.

The numerical methods for finding localized orbitals from the set of occupied orbitals of an insulating systems can be directly generalized to any set of eigenvalues in a given energy interval $\mathcal{I}$ satisfying the band isolation condition:

$$\inf_{\varepsilon_i \in \mathcal{I}, \varepsilon_{i'} \notin \mathcal{I}} |\varepsilon_i - \varepsilon_{i'}| := \varepsilon_g > 0. \tag{4.11}$$

When this condition is violated (*i.e.* $\varepsilon_g = 0$), the eigenvalues in $\mathcal{I}$ become *entangled*. Entangled eigenvalues appear ubiquitously in metallic systems,

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

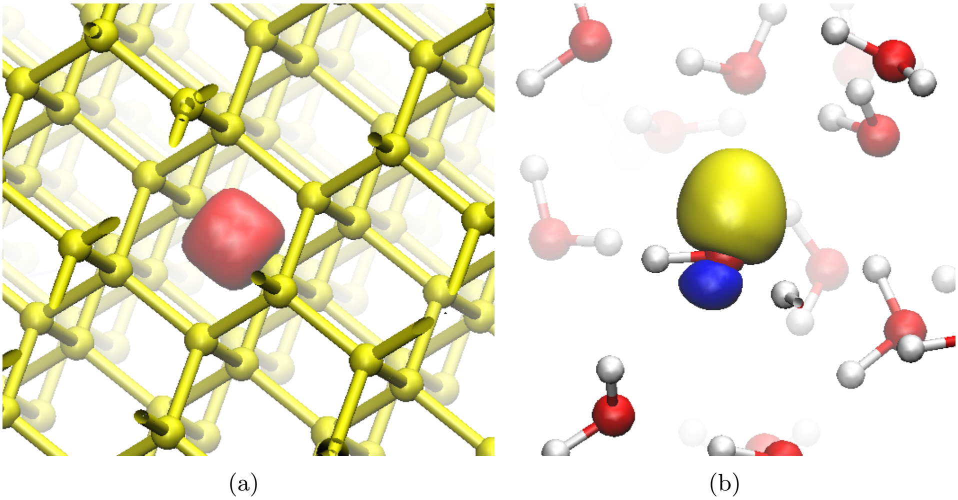


Figure 4.1. (Credit: Damle, Lin and Ying 2015.) Isosurface for an orthogonalized SCDM orbital. (a) A silicon crystal with 512 Si atoms (yellow balls). The red isosurface characterizes the localized orbitals located between two Si atoms. (b) A water system with 64 O atoms (red balls) and 128 H atoms (white balls). The yellow and blue isosurfaces characterize the positive and negative portions of localized orbitals.

and also in insulating systems when conduction bands or a selected range of valence bands are considered. The problem now becomes significantly more difficult: to identify a subspace $\mathcal{V}_w$ that admits a localized basis, and to construct such a basis.

The most widely used method to construct localized functions in this scenario is a *disentanglement* procedure (Souza, Marzari and Vanderbilt 2001). However, the result of the disentanglement procedure may depend sensitively on the initial guess. Often, detailed knowledge of the underlying physical system is required to obtain physically meaningful results. The SCDM method has been extended to handle systems with an entangled band structure in a more robust way (Damle and Lin 2018, Damle, Levitt and Lin 2019), using QRCP for a set of 'quasi' density matrices. Figure 4.2 demonstrates that the extended SCDM algorithm can accurately interpolate the band structure of graphene, even when zooming in on the region near the Dirac point.

4.4. Interpolative separable density fitting method

Kohn–Sham DFT calculations with rung-4 functionals and beyond often require computing quantities related to the *pair product of orbitals*, in the form $\{\varphi_i^*(\mathbf{r})\psi_j(\mathbf{r})\}_{i,j=1}^N$. Below we present a technique called interpolative

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

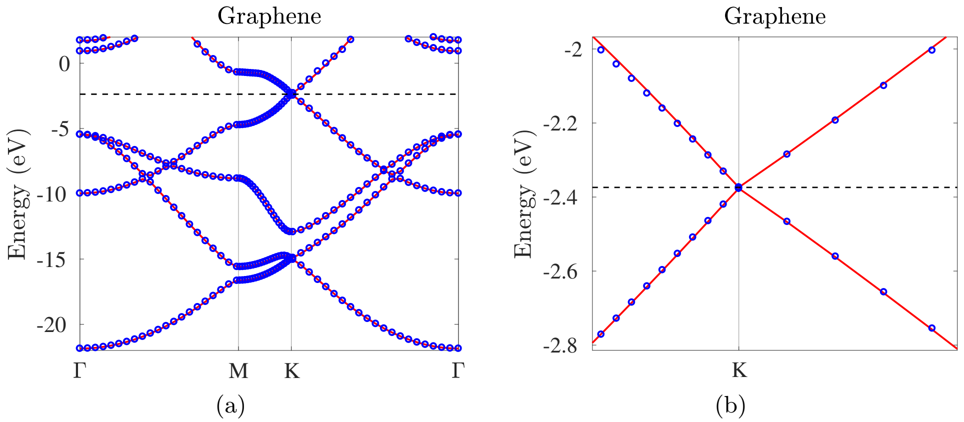


Figure 4.2. (Credit: Damle and Lin 2018.) Wannier interpolation with SCDM for the band structure for graphene (a) below the Fermi energy and (b) near the Dirac point. Direct calculation (red lines) and SCDM-based Wannier interpolation (blue circles). Here the chemical potential $\mu = -2.5$ eV, $\sigma = 4.0$ eV, and we use a $12 \times 12 \times 1$ **k**-grid to construct the Wannier functions.

separable density fitting (ISDF) (Lu and Ying 2015, Lu and Ying 2016), which allows us to reduce the number of pairs from $O(N^2)$ to $O(N)$ in general. This can be used to reduce the cost of rung-4 functional calculations with a large basis set (Hu, Lin and Yang 2017*a*, Dong, Hu and Lin 2018) (see Section 6.1), as well as RPA correlation energy calculations (Lu and Thicke 2017*a*) (see Section 6.3). It has also been shown to be useful in other contexts such as large-scale phonon calculations (Lin, Xu and Ying 2017).

To see why it is possible to find a compressed representation of the pair products, let us consider a real-space representation. Note that the number of grid points N_g to represent the orbital pairs scales linearly with respect to N. Hence, as N increases, the number of orbital pairs will eventually exceed N_g. This suggests that the numerical rank of $\{\varphi_i^*\psi_j\}$, viewed as a matrix of size $N_g \times N^2$, must scale asymptotically as $O(N)$. On the analytic level, consider the first N eigenfunctions $\{\psi_i\}$, $i = 1, \ldots, N$ of an effective Hamiltonian on a compact domain or closed manifold. Lu, Sogge and Steinerberger (2018) proved that for any ϵ there exists a subspace B_N of dimension $O_\delta(\epsilon^{-\delta}N^{1+\delta})$ such that

$$\|\psi_i\psi_j - \Pi_{B_N}(\psi_i\psi_j)\|_{L^2} \leq \epsilon, \tag{4.12}$$

where δ is an arbitrary positive constant. In other words, we can approximate the N^2 pair products of eigenfunctions with almost $O(N)$ auxiliary basis functions, regardless of the discretization level.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Motivated by the above observation, the ISDF method uses the following compression format:

$$\varphi_i^*(\mathbf{r})\psi_j(\mathbf{r}) \approx \sum_{\mu=1}^{N_\mu} \varphi_i^*(\hat{\mathbf{r}}_\mu)\psi_j(\hat{\mathbf{r}}_\mu)\zeta_\mu(\mathbf{r}). \tag{4.13}$$

Here $\{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N_\mu}$ is a subset of real-space grid points $\{\mathbf{r}_i\}_{i=1}^{N_g}$ on which the orbitals are evaluated. We will refer to $\{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N_\mu}$ as the interpolation points, and $\{\zeta_\mu(\mathbf{r})\}_{\mu=1}^{N_\mu}$ sampled on $\{\mathbf{r}_i\}_{i=1}^{N_g}$ the interpolation vectors.

The ISDF decomposition can also be understood from the perspective of interpolation. Let $\{\hat{\mathbf{r}}_\mu\}$ denote a set of grid points in real space, and let $\zeta_\mu(\mathbf{r})$ be the Lagrange interpolation function on these grid points satisfying

$$\zeta_\mu(\hat{\mathbf{r}}_{\mu'}) = \begin{cases} 1 & \mu = \mu', \\ 0 & \text{otherwise}. \end{cases} \tag{4.14}$$

Then the ISDF decomposition would become sufficiently accurate as one systematically refines the set $\{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N_\mu}$. In the worst case, all grid points need to be selected and $N_\mu = N_g$. However, numerical evidence and analytical results suggest that the decomposition often becomes sufficiently accurate when $N_\mu \ll N_g$, especially when both $\{\varphi_i(\mathbf{r})\}$ and $\{\psi_i(\mathbf{r})\}$ consist of functions that are sufficiently smooth.

We first assume that the interpolation points $\{\hat{\mathbf{r}}_\mu\}$ are given, and discuss how to find the interpolation vectors. Define

$$\{Z_{ij}(\mathbf{r}) := \varphi_i^*(\mathbf{r})\psi_j(\mathbf{r})\}_{1\le i\le N, 1\le j\le N}, \tag{4.15}$$

and equation (4.13) can be written as

$$Z = \Theta C, \tag{4.16}$$

where each column of Z is defined by equation (4.15) sampled on real-space grids $\{\mathbf{r}_i\}_{i=1}^{N_g}$. $\Theta = [\zeta_1, \zeta_2, \ldots, \zeta_{N_\mu}]$ contains the interpolating vectors, and each column of C with a multi-index (i, j) is given by

$$[\varphi_i^*(\hat{\mathbf{r}}_1)\psi_j(\hat{\mathbf{r}}_1), \ldots, \varphi_i^*(\hat{\mathbf{r}}_\mu)\psi_j(\hat{\mathbf{r}}_\mu), \ldots, \varphi_i^*(\hat{\mathbf{r}}_{N_\mu})\psi_j(\hat{\mathbf{r}}_{N_\mu})]^\top.$$

Equation (4.16) is an over-determined linear systems with respect to the interpolation vectors Θ. One possible way to solve the over-determined system is to impose the Galerkin condition

$$ZC^* = \Theta CC^*. \tag{4.17}$$

It follows that the interpolating vectors can be obtained from

$$\Theta = ZC^*(CC^*)^{-1}. \tag{4.18}$$

Note that the solution given by equation (4.18) is a least-squares approximation to the solution of equation (4.13).

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 4: Interpolative separable density fitting (ISDF) method

```
Input: Functions $\{\varphi_i(\mathbf{r})\}$ and $\{\psi_j(\mathbf{r})\}$
Output: Interpolation points $\{\hat{\mathbf{r}}_\mu\}$, $C$, and $\Theta$
1: Form a matrix $Z$ of size $N_g \times N^2$ with column $ij$ given by
   $\{Z_{ij}(\mathbf{r}) := \varphi_i^*(\mathbf{r})\psi_j(\mathbf{r})\}_{1\leq i\leq N, 1\leq j\leq jN}$,
2: Perform QRCP to $Z^*$ to obtain $Z^*\Pi = QR$. Here $Q$ is an $N^2 \times N_g$
   matrix that has orthonormal columns, $R$ is an upper triangular
   matrix, and $\Pi$ is a permutation matrix. Although a naive
   implementation takes $O(N^2 N_g^2)$ steps, randomized projection
   using the pair product structure of $Z$ can reduce the complexity to
   $O(N N_g^2)$.
3: The locations of the non-zero entries of the first $N_\mu$ columns of $\Pi$
   give $\{\hat{\mathbf{r}}_\mu\}$ and $C$.
4: $\Theta = ZC^*(CC^*)^{-1}$.
```

It may appear that the cost of matrix–matrix multiplications ZC^* and CC^* scales as $O(N^4)$, because the size of Z is $N_g \times N^2$ and the size of C is $N_\mu \times N^2$. However, both multiplications can be carried out with fewer operations due to the separable structure of Z and C, and the computational complexity of computing the interpolation vectors is therefore $O(N^3)$.

In order to optimize the set of interpolation points, a general strategy is to use the QR factorization with column pivoting (QRCP), as in the case of the SCDM method in Section 4.3. Pseudocode for the ISDF algorithm is given in Algorithm 4.

To demonstrate the interpolation points selected from the QRCP procedure, we consider a water molecule, with $N_\mu = 8$ interpolation points distributed in real space (see Figure 4.3). The choice of interpolation points agrees with chemical intuition: the locations of these points are consistent with the distribution of the electron density, and the points are relatively well separated, so that the set of interpolation vectors do not become linearly dependent.

The disadvantage of using equation (4.16) directly is that the storage requirement for the matrix Z is $O(N^3)$ and the computational cost associated with a standard QRCP procedure is $O(N^4)$. One possibility is to lower the cost of QRCP by using a random matrix to subsample columns of the matrix Z to form a smaller matrix $\widetilde{Z}$ of size $N_g \times \widetilde{N}_\mu$, where $\widetilde{N}_\mu$ is only slightly larger than N_μ (Lu and Ying 2015, Lu and Ying 2016). The reduced matrix size allows the computational cost of the QRCP procedure to be reduced to $O(N^3)$. Since the QRCP algorithm has been implemented in standard linear algebra software packages such as LAPACK and ScaLAPACK, the implementation and parallelization of ISDF is then relatively straightforward.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

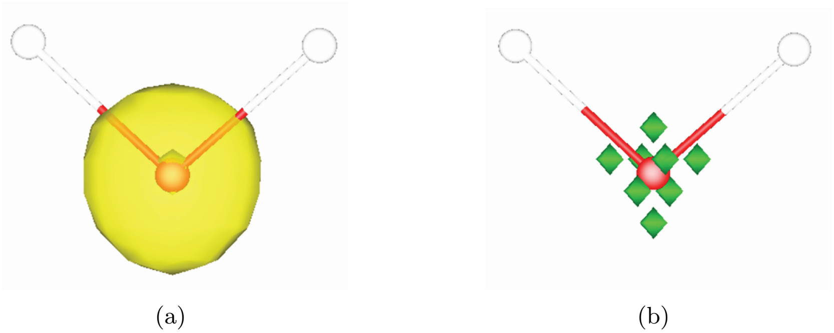


Figure 4.3. (Credit: Hu, Lin and Yang 2017*a*.) (a) The electron density (yellow isosurfaces) and (b) the interpolation points (green squares) $\{\hat{\mathbf{r}}_\mu\}_{\mu=1}^{N_\mu}$ $(N_\mu = 8)$ selected from the real-space grid points $\{\mathbf{r}_i\}_{i=1}^{N_g}$ $(N_g = 66^3)$ for a water molecule in a 10 Å × 10 Å × 10 Å box. The white and red balls denote hydrogen and oxygen atoms, respectively.

Another possibility for choosing interpolation points is to use a heuristic strategy. Note that an effective choice of the set of interpolation points should satisfy the following two conditions.

(i) The distribution of the interpolation points should roughly follow the distribution of the electron density. In particular, there should be more points when the electron density is high, and fewer or even zero points if the electron density is very low.

(ii) The interpolation points should not be very close to each other. Otherwise matrix rows represented by the interpolation points are nearly linearly dependent, and the matrix formed by the interpolation vectors become highly ill-conditioned.

The QRCP procedure satisfies both (i) and (ii) simultaneously. On the other hand, the conditions above can also be satisfied via a much simpler centroidal Voronoi tessellation (CVT) procedure applied to a weight vector, such as the electron density. More specifically, in the centroidal Voronoi tessellation (CVT) approach, we may partition the grid points in the global domain into N_μ Voronoi cells. The interpolation points can then be simply chosen to be the centroids corresponding to each cell. The CVT procedure can be effectively implemented through a K-Means algorithm (MacQueen 1967). Besides reduction of the computational cost, the use of a K-Means algorithm also produces a smoother potential energy surface, particularly in the context of *ab initio* molecular dynamics. We refer to Dong, Hu and Lin (2018) for more details of this approach.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Remarks
As in the SCDM method in Section 4.3, the ISDF method also relies on the real-space representation and smoothness of the orbitals. When small basis sets such as Gaussian-type orbitals are used, one possible way of using ISDF is to first interpolate the orbitals onto a real-space grid. However, in the context of all-electron calculations, the grid size needed to reach chemical accuracy can be exceedingly large. Hence in the quantum chemistry literature, the compression is often performed using the density fitting method (a.k.a. resolution of identity) (Weigend 2002, Ren *et al.* 2012*a*), applied to the basis functions $\{\phi_p\}_{p=1}^{N_b}$ directly. More specifically, density fitting seeks the following decomposition:

$$\phi_p^*(\mathbf{r})\phi_q(\mathbf{r}) \approx \sum_{\mu=1}^{N_\mu} C_\mu^{pq}\zeta_\mu(\mathbf{r}), \tag{4.19}$$

where the number of auxiliary functions N_μ is expected to be much smaller than the number of pairs N_b^2. The density fitting method can be viewed as an approximate solution to the singular value decomposition (SVD) of the matrix formed by the product of pairs of orbitals.

We remark that there is an important technical difference between ISDF and standard density fitting methods. In ISDF, the interpolation points $\{\hat{\mathbf{r}}_\mu\}$ are chosen first, and the auxiliary functions $\{\zeta_\mu\}$ are decided accordingly via a least-squares procedure. The interpolation coefficients in ISDF are obtained directly with a factorized form

$$C_\mu^{pq} = \phi_p^*(\hat{\mathbf{r}}_\mu)\phi_q(\hat{\mathbf{r}}_\mu). \tag{4.20}$$

In standard density fitting methods, the auxiliary functions are often chosen *a priori*, and part of the reason is to facilitate the computation of the two-electron integrals in subsequent calculations. The interpolation coefficients $\{C_\mu^{pq}\}$ are then determined via a least-squares procedure. Since $\{C_\mu^{pq}\}$ is generally a dense three-way tensor, the computational cost of density fitting scales generally as $O(N^4)$ unless additional locality constraints are enforced (Ren *et al.* 2012*a*).

Another technique to reduce the cost of the density fitting method is the tensor hypercontraction (THC) method (Parrish, Hohenstein, Martínez and Sherrill 2012, Parrish, Hohenstein, Martínez and Sherrill 2013), which is closely related to the ISDF method. Conceptually, THC aims at compressing the two-electron integrals directly, without going through intermediate decomposition of the form (4.19) with separable coefficients as in (4.20). However, since the two-electron integrals is a four-way tensor to start with, the computational cost associated with the THC decomposition can be $O(N^5)$. Once the decomposition is obtained, the THC format can then be used to reduce the cost of post-Hartree–Fock calculations, such as MP2 and coupled cluster theories.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

5. Evaluation of the Kohn–Sham map: semi-local functional

When solving Kohn–Sham DFT using the self-consistent field iteration, the most computationally expensive step is the evaluation of the Kohn–Sham map. For a given effective Hamiltonian at the current iteration, one needs to determine the electronic structure, in the form of occupied orbitals or the electron density. In this section we discuss the evaluation of the Kohn–Sham map for semi-local functionals. Given an effective Hamiltonian as a matrix, the straightforward way to obtain the corresponding density is to use conventional direct diagonalization (see Section 4.1) or iterative diagonalization (see Section 4.2) algorithms to obtain the low-lying eigenfunctions and then the electron density. These methods inherently display cubic scaling with respect to the system size, and become computationally expensive or even infeasible for large-scale systems.

To reduce the computational cost, it is important to realize that we require only the electron density in the Kohn–Sham map, but not the eigenfunctions or eigenvalues. More efficient algorithms can thus be developed for the Kohn–Sham map. In Section 5.1, we review filtering methods that are iterative methods *not* trying to converge to each individual eigenvector, but instead to the occupied subspace, which can often be much more efficient than traditional iterative solvers for eigenvalue problems.

Since we do not need each individual eigenvector, we can use an alternative representation for the occupied subspace, such as the density matrix or localized orbitals. By exploiting the sparsity, it is even possible to achieve linear scaling for insulating systems. We will review such methods in Section 5.2. While linear scaling methods are in principle desirable, it is often challenging to control their accuracy unless the system has a relatively large gap. The pole expansion and selected inversion method (PEXSI) has been developed as a reduced scaling method that yields accurate approximation for general systems, which will be discussed in Section 5.3.

5.1. *Filtering methods*

Instead of using Krylov-type methods to approximate the low-lying eigenmodes, one can also identify the low eigenmodes by evaluating an appropriate polynomial of the Hamiltonian H. The main criterion for choosing such a polynomial $p(z)$ is that it should maximize the magnitudes at the eigenvalues $\lambda_1, \ldots, \lambda_N$ and at the same time minimize the magnitudes at the eigenvalues $\lambda_{N+1}, \ldots, \lambda_{N_b}$. One of the most popular choices is the shifted-and-scaled Chebyshev polynomial $P_k(z)$ that maps the standard interval $[-1, 1]$ to $[\lambda_{N+1}, \lambda_{N_b}]$, where the subscript k denotes the degree of this Chebyshev polynomial (Zhou, Saad, Tiago and Chelikowsky 2006). The absolute values of Chebyshev polynomials are bounded by 1 within the interval $[-1, 1]$ and grow rapidly outside this interval. Hence the shifted-and-scaled

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 5: Chebyshev filtering method for solving the Kohn–Sham DFT eigenvalue problems $H\psi_i = \varepsilon_i \psi_i$

```
Input: Hamiltonian matrix $H$ and an orthonormal matrix $X \in \mathbb{C}^{N_b \times N_s}$.
Output: Eigenvalues $\{\varepsilon_i\}_{i=1}^N$ and wavefunctions $\{\psi_i\}_{i=1}^N$.
1: Estimate $\lambda_{N+1}$ and $\lambda_{N_b}$ using a few steps of the Lanczos algorithm.
2: while convergence not reached do
3:     Apply the Chebyshev polynomial $P_k(H)$ to $X$: $Y = P_k(H)X$.
3:     Orthonormalize columns of $Y$.
4:     Compute the projected Hamiltonian matrix $\tilde{H} = Y^* HY$ and solve
       the eigenproblem $\tilde{H}\tilde{\Psi} = \tilde{\Psi}\tilde{D}$.
5:     Subspace rotation $X = Y\tilde{\Psi}$.
6: end while
7: Update $\{\psi_i\}_{i=1}^N$ from the first $N$ columns of $X$.
```

Chebyshev polynomial $P_k(z)$ only amplifies the occupied states of the H. In practice, one does not know *a priori* the values of λ_{N+1} and λ_{N_b}. These values need to be estimated from the Kohn–Sham Hamiltonian (*e.g.* applying a few steps of the Lanczos iteration (Zhou, Chelikowsky and Saad 2014)), and also updated on the fly. When applying the Chebyshev filtering algorithm, we also often choose the number of states N_s to be slightly larger than the number of occupied orbitals N, to accelerate the convergence. Pseudocode for the Chebyshev filtering algorithm is shown in Algorithm 5.

The Chebyshev filtering algorithm is presented here as a diagonalization procedure, that is, one iterates until the low-lying eigenvalues converge. However, in many implementations, due to the existence of the outer SCF iteration, one often performs a handful of iterations (applying Chebyshev polynomials and orthonormalization) and accepts whatever improvement it produces. Sometimes even a single iteration per outer SCF iteration may be sufficient. An equivalent view is that it combines the outer SCF loop with the inner diagonalization loop, without performing accurate diagonalization. For many problems, such a combination does not increase the number of outer SCF iterations and hence may significantly reduce the overall running time.

Application of the Chebyshev polynomial to all occupied states scales as $O(N_b N)$. For large problems, the most expensive step is the solution of the Rayleigh–Ritz problem in step 4 (also called the subspace diagonalization step) of Algorithm 5, which scales as $O(N^3)$. Although the cubic scaling cannot be avoided in the filtering-based approaches, the prefactors can be significantly reduced by the following complementary subspace method (Banerjee *et al.* 2018).

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

In the Chebyshev filtering approach, the density matrix is computed as

$$P = Y\tilde{P}Y^*, \tag{5.1}$$

where $\tilde{P} \in \mathbb{R}^{N_s \times N_s}$ is the projected density matrix, with its eigenvalues given by the occupation numbers $\{f_i\}_{i=1}^{N_s}$, evaluated according to the Fermi–Dirac distribution in equation (2.66), for example (we switch to the finite temperature case here). Since N_s is only slightly larger than N, most of the occupation numbers are approximately equal to 1. We denote states 1 to N_1 as those with occupation numbers equal to 1 within numerical tolerance. The remaining states, from $N_1 + 1$ to N_s, have occupation numbers less than 1. Let N_t be the number of these fractionally occupied states, *i.e.* $N_t = N_s - N_1$. The eigenvectors of the projected density matrix are $\{\tilde{\psi}_i\}_{i=1}^{N_s}$, and we may rewrite the expression for the projected density matrix as

$$\begin{aligned}\tilde{P} &= \sum_{i=1}^{N_s} f_i\, \tilde{\psi}_i\, \tilde{\psi}_i^* \\ &= \sum_{i=1}^{N_s} \tilde{\psi}_i\, \tilde{\psi}_i^* - \sum_{i=N_1+1}^{N_s} \tilde{\psi}_i\, \tilde{\psi}_i^* + \sum_{i=N_1+1}^{N_s} f_i\, \tilde{\psi}_i\, \tilde{\psi}_i^* \\ &= \tilde{\mathcal{I}} - \sum_{i=N_1+1}^{N_s} (1 - f_i)\, \tilde{\psi}_i\, \tilde{\psi}_i^* \,. \end{aligned} \tag{5.2}$$

In equation (5.2) above, $\tilde{\mathcal{I}}$ is the identity matrix of dimension $N_s \times N_s$. That the first term of equation (5.2) is the identity matrix follows from the resolution of the identity. Hence, if the N_t top eigenvectors $\tilde{\psi}_i$ and corresponding occupation numbers f_i are known, the projected density matrix $\tilde{P}$ may be computed. Thus, instead of determining the full $N_s \times N_s$ set of vectors, we need to determine only an extremal block of vectors (of dimension $N_s \times N_t$), corresponding to the states $i = N_1 + 1$ to N_s.

Moreover, physical quantities such as the electron density, energy, entropy and atomic force can all be computed by knowing only the top eigenstates. These eigenstates can be efficiently evaluated via several Chebyshev polynomial filtering steps applied to the projected Hamiltonian matrix $\tilde{H}$. Using this complementary subspace strategy with two levels of Chebyshev filtering (CS2CF), for insulating systems, the Rayleigh–Ritz step may be avoided altogether. For metallic systems, the cost of the Rayleigh–Ritz step can also be significantly reduced. Table 5.1 shows that the CS2CF strategy can be applied to insulating and metallic systems with $O(10^4)$ atoms and efficiently parallelized over $O(10^4)$ computational cores. The wall clock time to solution can be more than an order of magnitude faster than parallel dense diagonalization methods such as ELPA for large systems.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Table 5.1. (Credit: Banerjee *et al.* 2018.) Wall clock times for one SCF iteration large systems using the CS2CF strategy in DGDFT.

System	No. of atoms	No. of electrons	Comput. cores	CS2CF time (s)	CS2CF subspace time (s)	ELPA (s)
Electrolyte $3D_{3\times3\times3}$	8 586	29 808	3 456	34	19	647
SiDiamond $3D_{10\times10\times10}$	8 000	32 000	3 456	40	24	648
Graphene $2D_{8\times8}$	11 520	23 040	4 608	35	27	262
CuFCC $3D_{10\times10\times10}$	4 000	44 000	3 000	75	46	199
LiBCC $3D_{12\times12\times12}$	27 648	82 944	12 960	180	165	5844

Remarks

The idea of filtering the Hamiltonian matrix also appears in the spectral slicing approach (Zhang, Smith, Sternberg and Zapol 2007, Polizzi 2009, Schofield, Chelikowsky and Saad 2012, Aktulga *et al.* 2014). The basic idea is to project matrix functions $f_i(H)$ to a set of orbitals, where $f_i(\cdot)$ can be polynomials or rational functions approximately supported only on a small segment $\mathcal{I}_i$. Then one may perform diagonalization methods to compute the eigenvalues restricted to the segment. Finally, the eigenvalues obtained from different segments of the spectrum are merged together. The spectral splicing approach is naturally suited to massive parallelization. Each core or multiple cores together may focus on calculations associated with one segment, and may significantly reduce the wall clock time. Note that spectral slicing methods may involve multiple parameters, such as the partitioning strategy of the spectrum, the degrees of filtering polynomials, and the merging strategy of different segments of the spectrum. Finding an efficient and robust strategy to determine these parameters is still an active research area.

5.2. Linear scaling methods

Direct diagonalization and iterative methods for solving the Kohn–Sham eigenvalue problem scale cubically with system size. The cubic scaling is the bottleneck for electronic structure calculations of large systems. When a

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

set of localized basis functions such as atomic orbitals and real-space representation are used, the resulting Hamiltonian matrix from the Kohn–Sham eigenvalue problem is localized. For insulating systems, the density matrix is localized and there exist localized representations – Wannier functions – for the subspace spanned by the occupied orbitals, as discussed in Section 4.3. One might hope to obtain the electronic structure, in the form of either a density matrix or localized orbitals, with a computation cost that scales linearly with system size. Such algorithms are referred to as 'linear scaling methods'. Many algorithms have been proposed and investigated over the past 30 years.

To achieve linear scaling, one needs to bypass the diagonalization of the Kohn–Sham Hamiltonian matrix. The crucial observation is that in the self-consistent iteration, we do not need full information on the eigenvectors. Thus, it is possible to reformulate the Kohn–Sham eigenvalue problem in a way that allows us to exploit the locality of the density matrix or localized orbitals to reduce computational complexity. In the following, we review a few representative linear scaling methods. While the list does not cover all algorithms developed in the literature, the discussion covers most of the existing ideas and strategies to achieve linear scaling. We refer readers to Goedecker (1999) and Bowler and Miyazaki (2012) for a more extensive review of linear scaling methods.

5.2.1. Divide-and-conquer method

The divide-and-conquer method proposed by Weitao Yang (Yang 1991*a*, Yang 1991*b*) is the first linear scaling method proposed in the literature for Kohn–Sham DFT. The idea of divide-and-conquer is to divide the physical system into several subsystems. The effective Hamiltonian for each subsystem is then solved separately, and the electron density of the whole system is obtained by merging the results of the subsystems together.

The divide-and-conquer method for each SCF step is described in Algorithm 6. Here a real-space representation is assumed. The divide-and-conquer method was originally developed for discretization using localized basis functions (Yang 1991*a*, Yang 1991*b*), in which case the algorithm is slightly different but shares the same spirit.

The divide-and-conquer method relies on the near-sightedness of the electron matter, as formulated by Kohn (1996). Roughly speaking, the dependence of electron density at $\mathbf{r}$ on the effective potential at $\mathbf{r}'$ is negligible when $|\mathbf{r}-\mathbf{r}'|$ is large. Hence, in the interior of the domain Ω_κ, the electron density corresponding to the Hamiltonian H is well approximated by that of the truncated Hamiltonian H_κ, since H agrees with H_κ inside Ω_κ. To make this work, besides requiring that the original system H is insulating, we also require that the truncated system H_κ does not destroy the energy gap (which could happen in practice, *e.g.* for heterogeneous systems). See

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 6: Divide-and-conquer method for evaluating the Kohn–Sham map

```
Input: Hamiltonian matrix $H$.
Output: Electron density $\rho$ corresponding to $H$.
Construct an overlapping partition $\{\Omega_\kappa\}$ of the whole computational
  domain $\Omega$ and the associated partition of unity $\{p_\kappa\}$ satisfying
      $\sum_\kappa p_\kappa(\mathbf{r}) = 1, \quad \text{for all } \mathbf{r} \in \Omega \text{ and } \operatorname{supp} p_\kappa \subset \Omega_\kappa.$
$H_\kappa = H|_{\Omega_\kappa}$: restrict the Hamiltonian on the subdomain $\Omega_\kappa$ with
  appropriate boundary conditions.
Solve the eigenvalue problem in each subsystem
      $H_\kappa \psi_j^\kappa(\mathbf{r}) = \varepsilon_j^\kappa \psi_j^\kappa(\mathbf{r}), \quad \mathbf{r} \in \Omega_\kappa.$
Determine the Fermi energy $\mu$ by fixing the total number of
  electrons:
      $N = \sum_\kappa \sum_j f_\beta(\varepsilon_j^\kappa - \mu) \int_{\Omega_\kappa} p_\kappa(\mathbf{r}) |\psi_j^\kappa(\mathbf{r})|^2 \, \mathrm{d}\mathbf{r},$
  where $f_\beta$ is the Fermi–Dirac function.
Construct the electron density as
      $\rho(\mathbf{r}) = \sum_\kappa p_\kappa(\mathbf{r}) \sum_j f_\beta(\varepsilon_j^\kappa - \mu) |\psi_j^\kappa(\mathbf{r})|^2.$
```

Chen and Lu (2016) for a mathematical analysis of the divide-and-conquer method, based on the analysis tools developed in E and Lu (2011).

Further developments of the divide-and-conquer method and related domain decomposition type method can be found in Yang and Lee (1995), Wang, Zhao and Meza (2008), Zhao, Meza and Wang (2008), Barrault, Cancès, Hager and Le Bris (2007) and Bencteux *et al.* (2008). A great advantage of the method lies in the intrinsic parallelism of the computation for each subsystem, which has been utilized for large-scale calculations with more than 10^6 atoms and 10^{12} electronic degrees of freedom (Kobayashi and Nakai 2009, Ohba *et al.* 2012, Shimojo, Kalia, Nakano and Vashishta 2008, Shimojo *et al.* 2011).

The divide-and-conquer method is based on separating the total computational domain into smaller ones and solving the eigenvalue problem for each subdomain. There are other linear scaling methods based on reformulations of the Kohn–Sham problems to bypass the eigenvalue problem, as we will discuss below.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

5.2.2. Orbital minimization methods

According to the Courant–Fisher variational principle, the first N eigenvectors of the discretized Hamiltonian matrix H can be found by solving

$$E = \min_{X,\, X^*X=I} \operatorname{Tr}(X^*HX), \tag{5.3}$$

where $X \in \mathbb{C}^{N_b\times N}$ is the matrix of discretized orbitals. The orthogonality constraint $X^*X = I$ is a bottleneck for linear scaling, as the cost of both the Gram–Schmidt orthogonalization and the Rayleigh–Ritz step scale cubically with respect to the system size.

The orbital minimization method (OMM) (Mauri, Galli and Car 1993, Ordejón, Drabold, Grumbach and Martin 1993, Mauri and Galli 1994, Ordejón, Drabold, Martin and Grumbach 1995) is based instead on an unconstrained variational formulation,

$$E = \min_{X} \operatorname{Tr}((2I - X^*X)(X^*(H - \varepsilon_{\max})X)), \tag{5.4}$$

where $\varepsilon_{\max}$ is an upper bound of the eigenvalue of H. The shift of the Hamiltonian by $-\varepsilon_{\max}$ makes all its eigenvalues negative, so the objective function is bounded from below, while it does not affect the eigenvectors of H.

While the OMM variational problem looks rather different from (5.3), it can be shown that the global minimizer X of (5.4) spans the eigenspace of the lowest N eigenvalues of H (assuming non-degeneracy) (Mauri *et al.* 1993, Pfrommer, Demmel and Simon 1999). In fact, somewhat surprisingly, while the objective function (5.4) is non-convex, it is proved in Lu and Thicke (2017*b*) that every local minimum of (5.4) is in fact also a global minimum. Thus it suffices for an optimization algorithm to converge locally.

The OMM minimization can be used as an alternative to direct diagonalization, as a way to obtain the Kohn–Sham map. This is implemented in the library libOMM (Corsetti 2014) with the preconditioned conjugate gradient algorithm.

More commonly, OMM is combined with truncation of the iterates of X to achieve linear scaling by only keeping $O(1)$ entries for each column of X during the iteration (Mauri *et al.* 1993, Ordejón *et al.* 1993). The hope is that the algorithm converges to the localized representation of the occupied space for insulating systems. However, with truncation, the optimization procedure might get stuck. Strategies to alleviate the problem have been proposed, for example in Kim, Mauri and Galli (1995), Tsuchida (2007), Gao and E (2009) and Lu and Thicke (2017*b*). The OMM is used in the SIESTA package (Soler *et al.* 2002) to achieve linear scaling.

As in OMM, localization and truncation steps can also be combined with other iterative algorithms that only rely on the subspace (rather than individual eigenvectors) to achieve linear scaling. One such algorithm is the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

localized subspace iteration algorithm (E *et al.* 2010, Garcia-Cervera, Lu, Xuan and E 2009), based on the Chebyshev filtering method (Algorithm 5). The Rayleigh–Ritz step (*i.e.* step 4 of Algorithm 5) is replaced by a localization and truncation step to maintain sparsity.

5.2.3. Density matrix minimization methods

Besides using orbitals, one can also use the density matrix in the variational approach for linear scaling algorithms. The starting point is the following variational principle in terms of density matrices:

$$\begin{aligned} &E = \min_P \operatorname{Tr}(PH) \\ &\text{subject to } P^2 = P,\ P = P^*,\ \operatorname{Tr} P = N. \end{aligned} \tag{5.5}$$

In the above minimization, the idempotency constraint $P^2 = P$ is not easy to deal with, similar to the orthogonality constraint for orbitals. The idea is to circumvent the constraint by using some alternative variational principles. One such variational principle is

$$\begin{aligned} &E = \min_P E^{\mathrm{DMM}}[P] := \min_P \operatorname{Tr}((3P^2 - 2P^3)(H - \mu)) \\ &\text{subject to } P = P^*, \end{aligned} \tag{5.6}$$

where μ is the chemical potential assumed to be known, whose determination can be done by solving the constraint that $\operatorname{Tr} P = N$ (we will further discuss how to determine the chemical potential at the end of Section 5.2.4). Note that in (5.6) we have replaced P in the trace term by the polynomial $3P^2 - 2P^3$ (this is known as the McWeeny purification polynomial, which will be discussed in more detail in Section 5.2.4) and used a constant shift of the Hamiltonian by the chemical potential. To minimize (5.6), one can typically start with an initial guess (*e.g.* $P_{(0)} = \frac{1}{2}I$) and employ the conjugate gradient method to minimize the energy.

The energy (5.6) is in fact not bounded from below. Let ε_i be one of the eigenvalues of H that is larger than μ; we can take $P = \sum_i f_i |\psi_i\rangle\langle\psi_i|$ and send $f_i \to \infty$ to make the energy go to $-\infty$. On the other hand, convergence can be proved with particular choices of initial conditions and minimization algorithms. The variation of (5.6) can be computed as

$$\delta E^{\mathrm{DMM}} = \operatorname{Tr}\left[\frac{\delta E^{\mathrm{DMM}}}{\delta P}\delta P\right], \tag{5.7}$$

where

$$\frac{\delta E^{\mathrm{DMM}}}{\delta P} = 3(PH + HP - 2\mu P) - 2(P^2H + PHP + HP^2 - 3\mu P^2). \tag{5.8}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

When P commutes with H, equation (5.8) can be simplified as

$$\frac{\delta E^{\mathrm{DMM}}}{\delta P} = (6P - 6P^2)(H-\mu) = 6P(I-P)(H-\mu). \tag{5.9}$$

Equation (5.9) suggests that DMM may have many stationary points. In fact, any 'pure' density matrix satisfying $P^2 = P$ is a stationary point. Nonetheless, if the initial guess P_0 is not a stationary point, and can be represented as

$$P_0 = \sum_i f_i \psi_i \psi_i^*,$$

where $0 < f_i < 1$ for all i, that is, all the states are fractionally occupied. Then

$$\frac{\delta E^{\mathrm{DMM}}}{\delta P_0} = \sum_i 6(\varepsilon_i - \mu) f_i (1-f_i) \psi_i \psi_i^*. \tag{5.10}$$

Thus, if $\varepsilon_i < \mu$, the gradient direction will increase f_i, and *vice versa*. The steepest descent algorithm will then converge to the density matrix corresponding to H, if the chemical potential μ is correctly chosen. A common choice of initial condition is $P_0 = \frac{1}{2}I$, which satisfies the condition $0 < f_i < 1$ in general.

We may also understand the OMM functional (5.4) from the perspective of density matrices. Setting $P = XX^*$, we can rewrite (5.4) as

$$\begin{aligned} E &= \min_P \operatorname{Tr}((2P - P^2)(H - E_{\max})) \\ &\text{subject to } P = P^*,\ 0 \preceq P \preceq I,\ \operatorname{rank} P \le N. \end{aligned} \tag{5.11}$$

The semi-definite constraint $0 \preceq P \preceq I$ is not easy to deal with (see Lai, Lu and Osher 2015, Lai and Lu 2016 for some algorithms for zero and finite temperature extensions), and hence in practice the OMM formulation in terms of orbitals is preferred.

Density matrix based algorithms provide a possible route to reduce the computational complexity for the evaluation of the Kohn–Sham map. Note that the gradient of the DMM energy (5.8) only involves matrix–matrix multiplication operations. Hence, if each matrix can be approximated by a sparse matrix, as in the case of insulating systems (see Section 4.3), the computational cost may be significantly reduced. For such systems, the number of non-zero entries of the density matrix only increases linearly with respect to N. Therefore with a proper implementation, all matrix–matrix multiplication operations can be carried out with $O(N)$ cost, which leads to linear scaling. The DMM algorithm is used by the ONETEP package (Skylaris *et al.* 2005) together with optimization of the localized basis functions.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

5.2.4. Density matrix purification

Besides density matrix minimization, another class of density matrix based algorithms is density matrix purification, which works for zero temperature systems, as it directly uses the idempotency of the density matrix.

Let us first consider the problem of constructing an idempotent matrix, starting from a given Hermitian matrix with eigenvalues close to 0 and 1 (but not exactly 0 and 1). One strategy is McWeeny purification (McWeeny 1960), which recursively applies the function $f_{\mathrm{McW}}(x) = 3x^2 - 2x^3$ to the matrix, starting from the initial Hermitian matrix P_0:

$$P_{n+1} = f_{\mathrm{McW}}(P_n) = 3P_n^2 - 2P_n^3. \tag{5.12}$$

To see why McWeeny purification works, note that the iterate remains Hermitian throughout the iteration and the eigenfunctions remain the same; it hence suffices to keep track of the eigenvalues. By focusing on a specific eigenvalue λ_0 of P_0, we have

$$\lambda_{n+1} = f_{\mathrm{McW}}(\lambda_n) = 3\lambda_n^2 - 2\lambda_n^3. \tag{5.13}$$

The fixed point of this map is given by the solution to

$$x = f_{\mathrm{McW}}(x) = 3x^2 - 2x^3, \tag{5.14}$$

whose three roots are given by $0, 1/2$ and 1. Calculating the derivative of f_{McW} shows that

$$f'_{\mathrm{McW}}(x) = 6x - 6x^2 = \begin{cases} 0 & x = 0, \\ 3/2 & x = 1/2, \\ 0 & x = 1. \end{cases} \tag{5.15}$$

Therefore, 0 and 1 are the stable fixed points, while $1/2$ is unstable. We also observe that

$$\begin{aligned} x &< f_{\mathrm{McW}}(x) \quad \text{for } x \in (0, 1/2), \\ x &> f_{\mathrm{McW}}(x) \quad \text{for } x \in (1/2, 1). \end{aligned}$$

Hence, the iteration converges to 0 if the initial condition lies in $[0, 1/2)$, while it converges to 1 if the initial condition lies in $(1/2, 1]$. (See Figure 5.1. In fact, the iteration converges to 0 if starting within $[-1/2, 1/2)$ and to 1 if starting within $(1/2, 3/2]$.)

Given these observations, let us now consider how to use purification to get the density matrix

$$P = \mathbb{1}_{(-\infty,0]}(H - \mu). \tag{5.16}$$

Note that P shares the same eigenfunctions as H, so this fits into the purification framework. We want to make all the eigenvalues of H below the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 7: McWeeny purification algorithm for density matrix

```
Input: Hamiltonian matrix $H$ and chemical potential $\mu$.
Output: Density matrix $P$.
Estimate lower and upper bounds of eigenvalues $\varepsilon_{\min}$ and $\varepsilon_{\max}$, such
  that $\mathrm{spec}(H) \subset [\varepsilon_{\min}, \varepsilon_{\max}]$.
Set initial density matrix
  $P \leftarrow \frac{\alpha}{2}(\mu - H) + \frac{1}{2}I,$
  with $\alpha = \min\{(\varepsilon_{\max} - \mu)^{-1}, (\mu - \varepsilon_{\min})^{-1}\}$.
while convergence not reached do
    $P \leftarrow f_{\mathrm{McW}}(P)$
end while
```

chemical potential μ converge to 1 and all the eigenvalues of H above converge to 0. Algorithm 7 starts from a rescaled version of the matrix $\mu - H$ and applies purification.

The initial density matrix

$$P_0 = \frac{\alpha}{2}(\mu - H) + \frac{1}{2}I$$

with the proper choice of α guarantees convergence. Observe that at the mth iteration, the density matrix is approximated by

$$P \approx P_m = f_{\mathrm{McW}} \circ f_{\mathrm{McW}} \circ \cdots \circ f_{\mathrm{McW}}(P_0). \tag{5.17}$$

This is a polynomial of the Hamiltonian H and gives a recursive polynomial approximation to the density matrix. Due to the recursion, the degree of the polynomial becomes quite high: in the mth iteration, the polynomial has highest degree monomial x^{3^m} where $2m$ matrix–matrix multiplication is needed. This is useful since our goal is to obtain a polynomial approximation of a Heaviside function (Fermi–Dirac function at zero temperature), hence a high-order polynomial is required to reduce the effect of the Gibbs phenomenon. In fact, the McWeeny purification method can be understood as the Newton–Schultz algorithm for the matrix sign function (Higham 2008), and thus gives a fast iterative scheme for the density matrix.

One drawback of McWeeny purification and the density matrix minimization discussed above is that they require the chemical potential μ as input. Extensions that do not rely on input chemical potential have been developed, such as the canonical purification (Paler and Manolopoulos 1998), trace correcting purification (Niklasson 2002), trace resetting purification (Niklasson, Tymczak and Challacombe 2003) and generalized canonical purification (Truflandier, Dianzinga and Bowler 2016). These algorithms are used in large-scale parallel implementations (Chow, Liu, Smelyanskiy

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

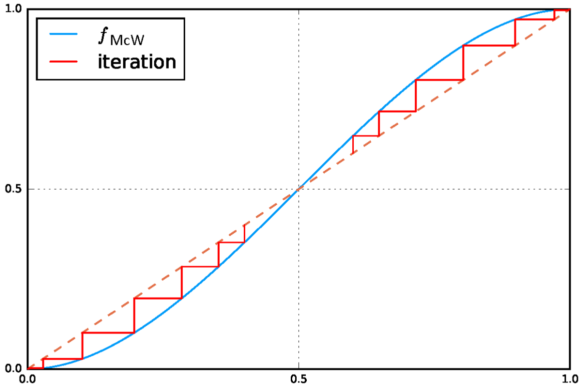


Figure 5.1. Shape of f_{McW} for McWeeny purification.

and Hammond 2015, Dawson and Nakajima 2018). We refer interested readers to Niklasson (2011) for a recent review on these methods.

5.2.5. Fermi operator expansion

Another approach to yielding linear scaling algorithms is the Fermi operator expansion (FOE) method. Consider the density matrix at the finite temperature

$$P = f_\beta(H - \mu). \tag{5.18}$$

The right-hand side is a matrix function with respect to the Hamiltonian matrix H. Instead of diagonalizing H and evaluating the matrix function using the eigen-decomposition, the basic idea of FOE is to expand the Fermi–Dirac function $f_\beta(\cdot)$ into an m-term expansion as

$$f_\beta(\varepsilon) \approx f_{\beta,m}(\varepsilon) = \sum_{n=1}^{m} g_n(\varepsilon). \tag{5.19}$$

The corresponding matrix function approximation is

$$f_\beta(H - \mu) \approx f_{\beta,m}(H - \mu) = \sum_{n=1}^{m} g_n(H - \mu). \tag{5.20}$$

The above formulation is general. We only require each term $g_n(H - \mu)$ to be a simple function, so that the corresponding matrix function can be evaluated directly without diagonalizing the matrix. For instance, g_n can be chosen to be a polynomial function or a rational function.

The approximation error of the matrix form (5.20) is directly related to that of the scalar form (5.19). Given the eigendecomposition

$$H\psi_i = \varepsilon_i \psi_i,$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

we have for any vector v

$$(f_\beta(H-\mu)-f_{\beta,m}(H-\mu))v=\sum_i(f_\beta(\varepsilon_i-\mu)-f_{\beta,m}(\varepsilon_i-\mu))\psi_i\langle\psi_i|v\rangle. \tag{5.21}$$

Thus,

$$\begin{aligned}
&\|(f_\beta(H-\mu)-f_{\beta,m}(H-\mu))v\|_2^2\\
&\quad=\sum_i|f_\beta(\varepsilon_i-\mu)-f_{\beta,m}(\varepsilon_i-\mu)|^2|\langle\psi_i|v\rangle|^2\\
&\quad\le\sup_i|f_\beta(\varepsilon_i-\mu)-f_{\beta,m}(\varepsilon_i-\mu)|^2\sum_i|\langle\psi_i|v\rangle|^2\\
&\quad=\sup_{t\in\operatorname{spec}(H-\mu)}|f_\beta(t)-f_{\beta,m}(t)|^2\|v\|_2^2.
\end{aligned} \tag{5.22}$$

The equation above can be rewritten as

$$\|f_\beta(H-\mu)-f_{\beta,m}(H-\mu)\|_2\le\|f_\beta(\cdot)-f_{\beta,m}(\cdot)\|_\infty, \tag{5.23}$$

where the left-hand side is the operator norm for matrices and the right-hand side is the L^∞-norm for scalar functions. Thus the error of the Fermi operator expansion (5.20) will be small as long as the corresponding approximation is small in the sense of expansions for scalar functions (5.19).

One example of FOE is to expand the Fermi–Dirac function into polynomials (Goedecker and Colombo 1994):

$$f_\beta(\varepsilon)\approx\sum_{n=1}^m c_n\varepsilon^{n-1}. \tag{5.24}$$

The corresponding matrix function version is

$$f_\beta(H-\mu)\approx\sum_{n=1}^m c_n(H-\mu)^{n-1}. \tag{5.25}$$

Note that each term of equation (5.25) is simply a matrix power $(H-\mu)^n$, which can be evaluated using only matrix–matrix multiplication recursively, without diagonalizing the matrix H. In order to implement the FOE (5.25) for a high-order polynomial, it is more efficient and stable to expand $f_\beta(t)$ using Chebyshev polynomials. For insulators, the number of terms needed in equation (5.25) scales as $\log\epsilon^{-1}$ to reach target accuracy ϵ (Trefethen 2008). In particular, the number of terms is independent of the system size and can be treated as a constant. When H is a sparse matrix, this means that the polynomial approximation to $f_\beta(H-\mu)$ is a sparse matrix as well, and the number of non-zeros only scales linearly with respect to the system size. Hence the FOE method (5.25) is a linear scaling algorithm.

Besides expansion using polynomials, another possibility is to approximate the Fermi–Dirac function using rational functions. A rational function

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

can be decomposed into a linear combination of terms of the form $(\varepsilon - z)^{-p}$, where $z \in \mathbb{C}$ and $p \geq 1$. In particular, if all terms use $p = 1$ the resulting expansion is called a simple pole expansion, or just the pole expansion. Compared to polynomial expansion, there are two main advantages in using the rational expansion. First, the number of terms needed for the rational expansion can be much smaller than that needed for polynomial expansion to achieve the same accuracy. This is particularly the case for systems with small gaps. Second, the use of the pole expansion can yield fast algorithms with reduced complexity even for metallic systems. This is called the pole expansion and selected inversion algorithm (PEXSI), which will be discussed in Section 5.3. To our knowledge, PEXSI is so far *the only* algorithm allowing such complexity reduction.

5.3. Pole expansion and selected inversion method

While linear scaling algorithms in principle yield fast algorithms for the evaluation of Kohn–Sham maps, their accuracy often crucially depends on the decay of orbitals or density matrices, and they are usually only suitable for insulating systems with a large gap. Another practical drawback is that they often require user input on support of truncation and other tuning parameters to achieve a balance between efficiency and accuracy. The pole expansion and selected inversion method (PEXSI) (Lin *et al.* 2009*b*, Lin, Chen, Yang and He 2013) is a reduced scaling algorithm with computational scaling at most $O(N^2)$ and smaller for lower-dimensional systems. While it has a worse computational scaling than linear, the PEXSI algorithm can be applied to general systems and gives accurate results.

The PEXSI algorithm is one type of FOE method, and uses the following pole expansion to approximate the Fermi–Dirac distribution:

$$P \approx \sum_{l=1}^{m} \omega_l (H - z_l)^{-1}. \tag{5.26}$$

If the band gap is small or zero, the number of terms needed in order for the polynomial expansion to reach a certain target accuracy ϵ scales as $O(\beta \Delta E)$, where ΔE is the spectral radius of the shifted operator $H - \mu$ (Goedecker 1999). Although the number of terms of a straightforward construction of the pole expansion also scales as $O(\beta \Delta E)$ (Baroni and Giannozzi 1992), it has subsequently been improved to $O((\beta \Delta E)^{1/2})$ (Ozaki 2007), $O((\beta \Delta E)^{1/3})$ (Ceriotti, Kühne and Parrinello 2008), and finally $O(\log(\beta \Delta E))$ (Lin, Lu, Ying and E 2009*a*).

In order to obtain such a pole expansion, one possibility is to use the Cauchy contour integral formulation. Note that the Fermi–Dirac function $f_\beta(\varepsilon) = 1/(1 + \mathrm{e}^{\beta\varepsilon})$ is a meromorphic function in $\mathbb{C}$, and the only poles are at $\varepsilon = (2n+1)\mathrm{i}\pi/\beta, n \in \mathbb{Z}$. Furthermore, $f_\beta(\varepsilon)$ can be expanded using the

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

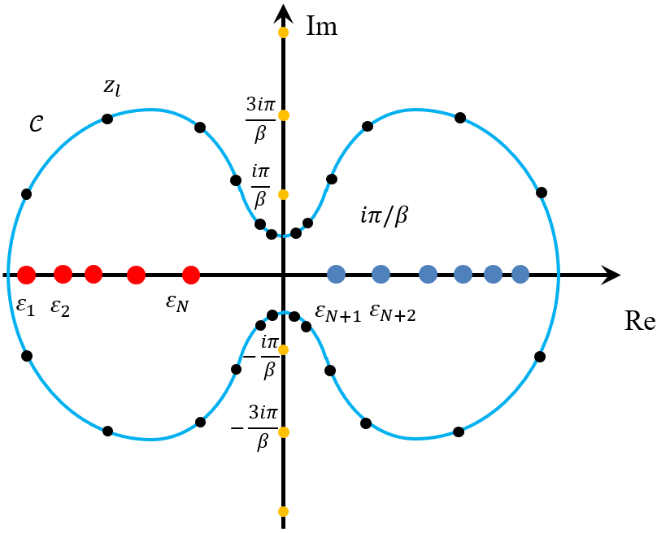


Figure 5.2. Contour integral representation and pole expansion for the density matrix at finite temperature.

following Matsubara expansion (Mahan 2000):

$$f_\beta(\varepsilon) = \frac{1}{2} - \frac{1}{\beta}\sum_{n\in\mathbb{Z}} \frac{1}{\varepsilon - (2n+1)\mathrm{i}\pi/\beta}. \tag{5.27}$$

Note that equation (5.27) only converges conditionally, and the infinite summation must be performed symmetrically with respect to the positive and negative choices of n. The number of terms needed in the direct truncation of the Matsubara series naturally scales as $O(\beta\Delta E)$.

The efficiency of the pole expansion can be improved by using a contour integral formulation:

$$f_\beta(H-\mu) = \frac{1}{2\pi\mathrm{i}}\oint_{\mathscr{C}} f_\beta(z)((z+\mu)I - H)^{-1}\,\mathrm{d}z. \tag{5.28}$$

Here the contour $\mathscr{C}$ should be chosen so that it includes all the (real) eigenvalues of $H-\mu$, without including any poles of $f_\beta(z)$, *i.e.* $(2n+1)\mathrm{i}\pi/\beta, n \in \mathbb{Z}$. This leads to the 'dumbbell-shaped' contour used in Lin, Lu, Ying and E (2009*a*) (Figure 5.2). The contour is symmetric with respect to the chemical potential μ. The discretization points are chosen to be denser around μ to resolve the sharp transition of the Fermi–Dirac function at μ. At finite temperature, the contour integral formulation remains well-defined for gapless systems, *i.e.* $\varepsilon_N = \varepsilon_{N+1}$.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Each term in the pole expansion corresponds to a matrix inverse, or Green's function $(z_l - H)^{-1}$, which can be evaluated directly without diagonalizing the matrix H. Equation (5.26) converts the problem of computing P to the problem of computing m Green's functions. In order to find the Kohn–Sham map, we do not need the entire density matrix P, but only the electron density which corresponds to the diagonal entries of P (again for simplicity we assume that the basis set is orthogonal and the overlap matrix S is an identity matrix). This amounts to the question of finding the diagonal entries of a Green's function. Note that even if H is a sparse matrix, the matrix inverse $G_l := (z_l - H)^{-1}$ can be a fully dense matrix. One naive method is to first evaluate each Green's function and extract its diagonal elements. However, when H is a sparse matrix, the computation of diagonal entries, and more generally the entries of G_l corresponding to the sparsity pattern of H, can be evaluated much more efficiently by means of the selected inversion method (Erisman and Tinney 1975, Lin *et al.* 2009*b*, Lin *et al.* 2011, Jacquelin, Lin and Yang 2016).

Although we assume H to be a general Hermitian matrix throughout the paper, for the simplicity of discussion in this section, we assume H to be a real symmetric matrix. This makes $A = z_l - H \in \mathbb{C}^{N_b \times N_b}$ a complex symmetric, non-singular matrix. For such a matrix, the standard approach to computing A^{-1} is to first decompose A as

$$A = LDL^\top, \tag{5.29}$$

where L is a unit lower triangular matrix and D is a diagonal or a block-diagonal matrix. Equation (5.29) is often known as the $LDL^\top$ factorization of A. Given such a factorization, one can obtain $A^{-1} = (x_1, x_2, \ldots, x_{N_b})$ by solving a number of triangular systems,

$$Ly = e_j, \quad Dw = y, \quad L^\top x_j = w, \tag{5.30}$$

for $j = 1, 2, \ldots, N_b$, where e_j is the jth column of the identity matrix I. The computational cost of such algorithms is generally $O(N_b^3)$. However, when A is sparse, we can exploit the sparsity structure of L and e_j to reduce the complexity of computing selected components of A^{-1}.

The selected inversion algorithm can be heuristically understood as follows (Lin *et al.* 2011). Let A be partitioned into the 2×2 block form

$$A = \begin{pmatrix} \alpha & b^\top \\ b & \widetilde{A} \end{pmatrix}. \tag{5.31}$$

The first step of an $LDL^\top$ factorization produces a decomposition of A that can be expressed by

$$A = \begin{pmatrix} 1 & \\ \ell & I \end{pmatrix} \begin{pmatrix} \alpha & \\ & \widetilde{A} - b\alpha^{-1}b^\top \end{pmatrix} \begin{pmatrix} 1 & \ell^\top \\ & I \end{pmatrix}, \tag{5.32}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

where α is often referred to as a pivot, $\ell = b\alpha^{-1}$ and $S = \widetilde{A} - b\alpha^{-1}b^\top$ is known as the *Schur complement*. The same type of decomposition can be applied recursively to the Schur complement S until its dimension becomes 1. The product of lower triangular matrices produced from the recursive procedure, which all have the form

$$\begin{pmatrix} I & & \\ & 1 & \\ & \ell^{(i)} & I \end{pmatrix},$$

where $\ell^{(1)} = \ell = b\alpha^{-1}$, yields the final L factor. At this last step the matrix in the middle becomes diagonal, which is the D matrix.

From equation (5.32), A^{-1} can be expressed by

$$A^{-1} = \begin{pmatrix} \alpha^{-1} + \ell^\top S^{-1}\ell & -\ell^\top S^{-1} \\ -S^{-1}\ell & S^{-1} \end{pmatrix}. \tag{5.33}$$

This expression suggests that once α and ℓ are known, the task of computing A^{-1} can be reduced to that of computing S^{-1}. Because a sequence of Schur complements is produced recursively in the $LDL^\top$ factorization of A, the computation of A^{-1} can be organized in a recursive fashion too. Clearly, the reciprocal of the last entry of D is the (N_b, N_b)th entry of A^{-1}. Starting from this entry, which is also the 1×1 Schur complement produced in the $(N_b - 1)$th step of the $LDL^\top$ factorization procedure, we can construct the inverse of the 2×2 Schur complement produced at the $(N_b - 2)$th step of the factorization procedure, using the recipe given by (5.33). This 2×2 matrix is the trailing 2×2 block of A^{-1}. As we proceed from the lower right corner of L and D towards their upper left corner, more and more elements of A^{-1} are recovered. Finally, we may recover all the diagonal entries of A^{-1} *exactly*. In fact, given the factorization $A = LDL^\top$, the selected inversion algorithm can be used to efficiently compute all entries

$$\{A^{-1}_{i,j} : (L + L^\top)_{i,j} \neq 0\}.$$

The validity of the selected inversion algorithm can be verified by the following statement. For any $1 \leq k < N_b$, define

$$\mathcal{C} = \{i : L_{i,k} \neq 0,\ i > k\}. \tag{5.34}$$

Then all entries $\{A^{-1}_{i,k} : i \in \mathcal{C}\}$, $\{A^{-1}_{k,j} : j \in \mathcal{C}\}$, and $A^{-1}_{k,k}$ can be computed using only the L, D factors and

$$\{A^{-1}_{i,j} : (L + L^\top)_{i,j} \neq 0,\ i, j > k\}.$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 8: Selected inversion algorithm based on $LDL^\top$ factorization

```
Input: $LDL^\top$ factorization of a symmetric matrix $A \in \mathbb{C}^{N_b \times N_b}$.
Output: Selected elements of $A^{-1}$, i.e. $\{A^{-1}_{i,j} : (L + L^\top)_{i,j} \neq 0\}$.
Calculate $A^{-1}_{N_b,N_b} \leftarrow (D_{N_b,N_b})^{-1}$.
for $k = N_b - 1, \dots, 1$ do
    Find the collection of indices $\mathcal{C} = \{i \mid i > k, L_{i,k} \neq 0\}$.
    Calculate $A^{-1}_{\mathcal{C},k} \leftarrow -A^{-1}_{\mathcal{C},\mathcal{C}} L_{\mathcal{C},k}$.
    Calculate $A^{-1}_{k,\mathcal{C}} \leftarrow (A^{-1}_{\mathcal{C},k})^\top$.
    Calculate $A^{-1}_{k,k} \leftarrow (D_{k,k})^{-1} - A^{-1}_{k,\mathcal{C}} L_{\mathcal{C},k}$.
end for
```

To see why this is the case, consider $\{A^{-1}_{i,k} : i \in \mathcal{C}\}$. As in equation (5.33) we can derive

$$A^{-1}_{i,k} = -\sum_{j=k+1}^{N_b} A^{-1}_{i,j} L_{j,k}, \quad i \in \mathcal{C}. \tag{5.35}$$

If $L_{j,k} \neq 0$, then $A^{-1}_{i,j}$ is needed in the sum. Since we are only interested in computing $A^{-1}_{i,k}$ for $i \in \mathcal{C}$, the i and j indices are constrained to satisfy the conditions $L_{j,k} \neq 0$ and $L_{i,k} \neq 0$. Due to the non-zero fill-in pattern in the $LDL^\top$ factorization, we have $(L + L^\top)_{i,j} \neq 0$ and the statement is proved for $\{A^{-1}_{i,k} : i \in \mathcal{C}\}$. The argument is the same for $\{A^{-1}_{k,j} : j \in \mathcal{C}\}$ due to symmetry. Finally for the diagonal entry, we have

$$A^{-1}_{k,k} = D^{-1}_{k,k} - \sum_{i=k+1}^{N_b} L_{i,k} A^{-1}_{i,k}, \tag{5.36}$$

which can be readily computed given $\{A^{-1}_{i,k} : i \in \mathcal{C}\}$ is available. This proves the statement above.

In order to understand the asymptotic complexity of the selected inversion algorithm, without loss of generality we assume the sparsity pattern of H is similar to that obtained by the second-order central difference discretization of a Laplace operator. The computational cost associated with the $LDL^\top$ factorization, as well as the selected inversion algorithm, scales as $O(N_b)$, $O(N_b^{1.5})$ and $O(N_b^2)$ for one-, two- and three-dimensional systems, respectively (Lin *et al.* 2009*a*). We remark that such a complexity count is robust to changes of the discretization scheme as long as local basis sets are used. Hence for quasi-1D systems (such as nanotubes) and quasi-2D systems (such as monolayer systems and surfaces), the computational cost also scales as $O(N_b)$ and $O(N_b^{1.5})$, respectively. Pseudocode for the selected inversion algorithm is given in Algorithm 8.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

In practice, a column-based sparse factorization and selected inversion algorithm as illustrated in Algorithm 8 may not be efficient due to the lack of level 3 BLAS operations. For a sparse matrix A, the columns of A and the L factor can be partitioned into supernodes. A supernode is a set of contiguous columns $\mathcal{J} = \{j, j+1, \dots, j+s\}$ of the L factor that have the same or similar non-zero sparsity structure below the $(j+s)$th row (Ashcraft and Grimes 1989). This allows use of matrix–matrix multiplications, which can significantly improve the efficiency.

The pole expansion and selected inversion (PEXSI) method (Lin *et al.* 2009*a*, Lin *et al.* 2011, Jacquelin *et al.* 2016) therefore combines the pole expansion and the selected inversion, and evaluates the Kohn–Sham map without solving any eigenvalues or eigenfunctions. The selected inversion method is an exact method if exact arithmetic is used, that is, the only error in the selected inversion method is due to round-off errors. Hence the accuracy of the PEXSI method is determined by the pole expansion, which can be systematically improved by increasing the number of poles m. The PEXSI method is ideally suited to massively parallel computers. The treatment of the poles can be parallelized in a straightforward fashion, with communication only needed at the end to construct the density matrix. The selected inversion method itself can also be massively parallelizable to thousands of processors (Jacquelin *et al.* 2016), and the total number of processors that can be efficiently used by PEXSI can be over 100 000.

The PEXSI software package (available at http://www.pexsi.org, distributed under the BSD license) has now been integrated into electronic structure packages such as BigDFT, CP2K, DGDFT, FHI-aims, QuantumWise ATK and SIESTA, and is part of the 'Electronic Structure Infrastructure' (ELSI) package (Yu *et al.* 2018).

In addition to computing the charge density at a reduced computational complexity in each SCF iteration, we can also use PEXSI to compute energy, free energy and the atomic forces efficiently without diagonalizing the Kohn–Sham Hamiltonian, using *the same set of poles* as those used for computing the charge density (Lin *et al.* 2013). Another numerical issue associated with the PEXSI technique, as well as the Fermi operator expansion techniques in general, is to determine the chemical potential, so that the condition

$$N = N_\beta(\mu) := \mathrm{Tr}[P] \tag{5.37}$$

is satisfied. Note that $N_\beta(\cdot)$ is a non-decreasing function of μ. Hence the chemical potential can be determined via a bisection strategy, or Newton's method. When Newton's method is used, the chemical potential converges rapidly near its correct value. However, the standard Newton's method may not be robust enough when the initial guess is far away from the correct chemical potential. It may give, for example, too large a correction when

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

$N'_\beta(\mu)$ is close to zero, such as when μ is near the edge or in the middle of a band gap.

One way to overcome the above difficulty is to efficiently approximate the function $N_\beta(\varepsilon)$ to narrow down the region in which the correct μ must lie. This function can be seen effectively as a (temperature smeared) cumulative density of states, counting the number of eigenvalues in the interval $(-\infty, \varepsilon)$. We can evaluate such a zero temperature limit, denoted by $N_\infty(\varepsilon)$, again without computing any eigenvalues of H. Instead, we perform a matrix decomposition of the shifted matrix $H - \varepsilon = LDL^\top$, where L is unit lower triangular and D is diagonal. It follows from Sylvester's law of inertia (Sylvester 1852) that the inertia (the number of negative, zero and positive eigenvalues) of a real symmetric matrix does not change under a congruent transform, that D has the same inertia as that of $H - \varepsilon$. Hence, we can obtain $N_\infty(\varepsilon)$ by simply counting the number of negative entries in D. Note that the matrix decomposition $H - \varepsilon = LDL^\top$ can be computed efficiently by using a sparse $LDL^\top$ or LU factorization in real arithmetic. It requires fewer floating point operations than the complex arithmetic direct sparse factorization used in PEXSI.

We apply the parallel PEXSI method to the DG_Graphene_2048 and DG_Graphene_8192 systems, which are disordered graphene systems with 2048 and 8192 atoms, respectively, and compare its performance with a standard approach that requires a partial diagonalization of (H, S). We use a ScaLAPACK subroutine pdsyevr (Vömel 2010), which is based on the multiple relatively robust representations (MRRR) algorithm, to perform the diagonalization.

Figure 5.3 shows that for graphene problems with 2048 and 8192 processors, the PEXSI technique is nearly two orders of magnitude faster than the ScaLAPACK routine pdsyevr, and can be scalable to a much larger number of processors. The advantage of PEXSI becomes even clearer for a disordered graphene system with 32 768 atoms (see Figure 5.4). For this case, the diagonalization routine is no longer feasible, while the time to solution for the PEXSI technique can be as small as 241 s.

Remarks

Unlike most linear scaling algorithms discussed in this section, the PEXSI method does not rely on the decay properties of the density matrix. Hence the efficiency of PEXSI is roughly the same for metallic, semiconducting and insulating systems. For insulating systems with a relatively large energy gap, PEXSI may become slightly more efficient due to the potentially smaller number of poles needed to approximate the density matrix. The computational cost of PEXSI scales as $O(N_b^\alpha)$, where $\alpha = 1, 3/2, 2$ for quasi-1D, 2D and 3D systems, respectively, where N_b is the number of basis functions.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

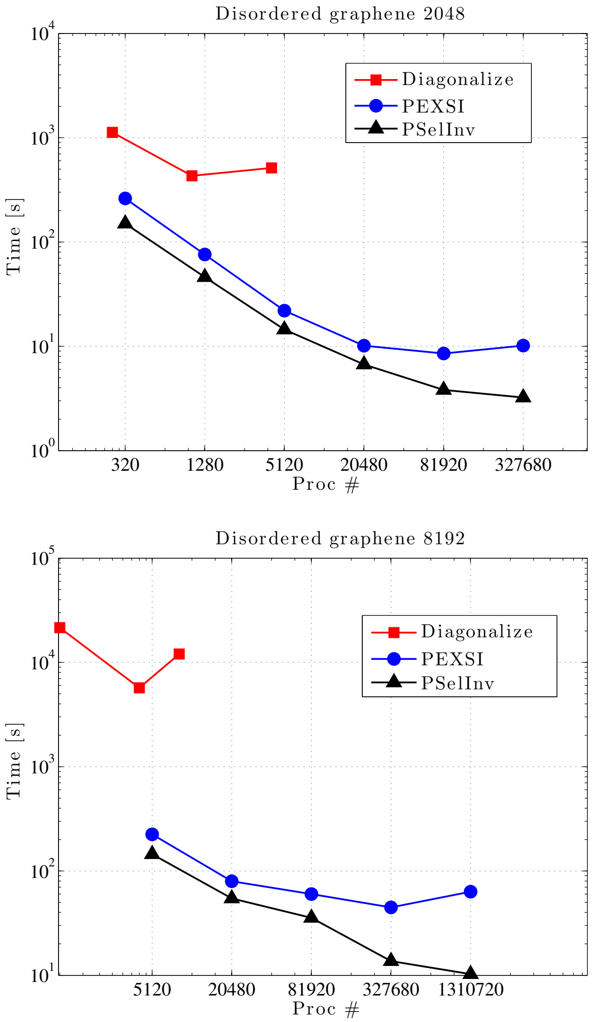


Figure 5.3. (Credit: Jacquelin, Lin and Yang 2016.) Wall clock time versus number of cores for a graphene system.

Compared to dense eigensolvers which scale as $O(N_b^3)$, the PEXSI method can be highly advantageous for large systems, as demonstrated in numerical results above. On the other hand, recall that iterative diagonalization methods discussed in Section 4.2 scale as $O(N_b N^2)$; we find that the cross-over point between PEXSI and iterative diagonalization methods is roughly at $N \sim N_b^{(\alpha-1)/2}$. This implies that PEXSI is most efficient when small basis sets such as Gaussian-type orbitals are used, and becomes less efficient for large basis sets such as finite elements. This statement is in fact true for most linear scaling methods discussed in Section 5.2.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

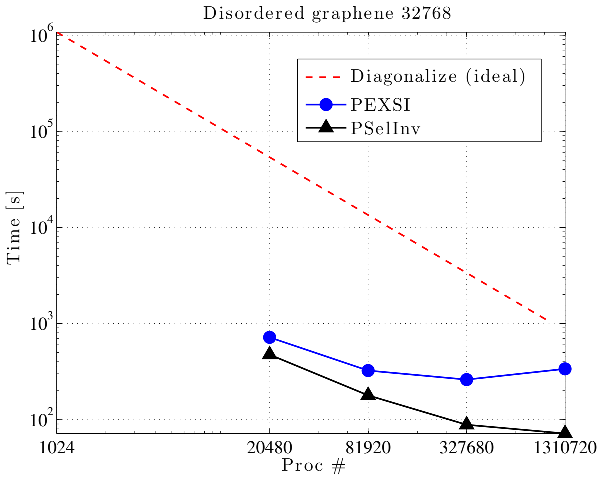


Figure 5.4. (Credit: Jacquelin, Lin and Yang 2016.) Wall clock time versus number of cores for a graphene system with 32 768 atoms.

In the discussion above, we have assumed that H is a real symmetric matrix, and hence $A = z_l - H$ is complex symmetric. This assumption is valid for molecular systems, as well as Γ point sampling of periodic systems. In a more general situation, such as in the context of **k**-point sampling for periodic systems, or when the system is magnetic, H is a Hermitian matrix, and $A = z_l - H$ is merely a *structurally symmetric* matrix. In such a case, the $LDL^\top$ factorization should be replaced by the LU factorization, and the non-symmetric version of the selected inversion algorithm should be used instead (Jacquelin, Lin and Yang 2018). Besides PEXSI, general-purpose selected inversion algorithms are also available in other software packages such as PARDISO (Schenk and Gartner 2006) and MUMPS (Amestoy, Duff, L'Excellent and Koster 2001). The selected inversion-type algorithms are also used in other contexts such as quantum transport calculations (Li, Ahmed, Klimeck and Darve 2008, Petersen *et al.* 2009).

Recently it has been shown that the pre-constant of the logarithmic scaling factor can be further reduced to be numerically near-optimal (Moussa 2016). This makes the pole expansion highly efficient for metallic systems, and the number of poles needed in practical calculations is typically only $10 \sim 40$.

During the self-consistency field iteration, the zero temperature limit $N_\infty(\varepsilon)$ evaluated from the inertia-counting procedure may be used to construct upper and lower bounds on the chemical potential. Then, coupled with a strategy to evaluate $N_\beta(\varepsilon)$ accurately using PEXSI over multiple energy points ε, we may accurately determine the chemical potential within

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

the window specified by the inertia-counting procedure. In particular, it is possible to perform PEXSI calculations over multiple energy points *only once per SCF iteration*, without sacrificing accuracy at convergence (Jia and Lin 2017).

6. Evaluation of the Kohn–Sham map: non-local functional

Kohn–Sham DFT calculations with non-local functionals, such as rung-4 functionals (hybrid functionals) and rung-5 functionals, can be considerably more costly than calculations with exchange-correlation functionals from the first three rungs of the ladder. More specifically, Kohn–Sham equations with local and semi-local functionals can be viewed as eigenvalue problems corresponding to differential operators. When a rung-4 functional is used, the Kohn–Sham Hamiltonian operator becomes an integro-differential operator due to the Fock exchange term. For rung-5 functionals, the self-consistency is computationally rather challenging, and most calculations are done as a post-processing step to obtain the correlation energy as a perturbation to the self-consistent solution for a semi-local functional.

Throughout the discussion in this section, we assume that a large basis set such as planewaves or real-space representation is used. We first discuss numerical methods for evaluating the Fock exchange operator in Section 6.1. We introduce an acceleration technique called the adaptive compression method in Section 6.2. We discuss a method for evaluating the RPA correlation energy as an example of the rung-5 functionals in Section 6.3.

6.1. Fock exchange operator

The Fock exchange operator $V_{\mathrm{x}}^{\mathrm{EX}}[P]$ is an integral operator, introduced in equation (2.52), recalled here for readers' convenience:

$$V_{\mathrm{x}}^{\mathrm{EX}}[P](\mathbf{r},\mathbf{r}') = -P(\mathbf{r},\mathbf{r}')K(\mathbf{r},\mathbf{r}'),$$

where

$$P(\mathbf{r},\mathbf{r}') = \sum_{i=1}^{N} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}').$$

Here we assume the system is an insulating system for simplicity. The kernel of $V_{\mathrm{x}}^{\mathrm{EX}}[P]$ is not of low rank due to the Hadamard product (*i.e.* elementwise product) between the kernels of P and K. When a small basis set $\Phi = [\phi_1, \ldots, \phi_{N_b}]$ is used, each occupied orbital is expanded as

$$\psi_i(\mathbf{r}) = \sum_{p} \phi_p(\mathbf{r}) c_{p,i}, \quad i = 1, \ldots, N. \tag{6.1}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

The matrix elements of the Fock exchange operator become

$$
\begin{aligned}
(V_{\mathrm{x}}^{\mathrm{EX}}[P])_{pq} &= -\int \phi_p^*(\mathbf{r})P(\mathbf{r},\mathbf{r}')K(\mathbf{r},\mathbf{r}')\phi_q(\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \\
&= -\sum_{i=1}^{N}\sum_{r,s} c_{r,i}c_{s,i}^* \int \phi_p^*(\mathbf{r})\phi_r(\mathbf{r})\phi_s^*(\mathbf{r}')\phi_q(\mathbf{r}')K(\mathbf{r},\mathbf{r}')\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \\
&:= -\sum_{i=1}^{N}\sum_{r,s} c_{r,i}c_{s,i}^*\langle ps|rq\rangle. \qquad (6.2)
\end{aligned}
$$

Equation (6.2) defines the two-electron integral tensor $\langle ps|rq\rangle$, which is a fourth-order tensor. Hence the computational cost of constructing the Fock exchange matrix $(V_{\mathrm{x}}^{\mathrm{EX}}[P])_{pq}$ typically scales as $O(N^4)$.

It turns out that when a large basis set is used, the asymptotic complexity associated with the Fock exchange operator can be reduced to $O(N^3)$ *without approximation*. Note that in such a case, it is often prohibitively expensive to explicitly construct or to store the Fock exchange operator directly. It is only viable to apply it to an occupied orbital ψ_j as

$$
(V_{\mathrm{x}}^{\mathrm{EX}}[P]\psi_j)(\mathbf{r}) = -\sum_{i=1}^{N}\varphi_i(\mathbf{r})\int K(\mathbf{r},\mathbf{r}')\varphi_i^*(\mathbf{r}')\psi_j(\mathbf{r}')\,\mathrm{d}\mathbf{r}', \qquad (6.3)
$$

for

$$
P = \sum_{i=1}^{N} |\varphi_i\rangle\langle\varphi_i|.
$$

Here we deliberately distinguish the orbitals in the density matrix $(\{\varphi_i\})$ and the orbitals $V_{\mathrm{x}}^{\mathrm{EX}}$ acts on $(\{\psi_j\})$ to emphasize that they may correspond to different density matrices before self-consistency is achieved. The integral $\int K(\mathbf{r},\mathbf{r}')\varphi_i^*(\mathbf{r}')\psi(\mathbf{r}')\,\mathrm{d}\mathbf{r}'$ can be implemented by solving N Poisson-like equations. Let N_g denote the grid size, and the computational cost of solving each Poisson-like equation scales as $O(N_g \log N_g)$ when fast Fourier transform can be used. When applied to all pairs of occupied orbitals, the computational cost scales as $O(N_g \log N_g N^2)$ to $O(N^3)$. Note that the preconstant of this cubic scaling component can be very large. In practical Hartree–Fock calculations, the application of the Fock exchange operator can often take more than 95% of the overall computational time.

The Hartree–Fock-like equations require the density matrix P to be computed self-consistently. A common strategy is to solve the linearized Hartree–Fock equation by fixing the density matrix P so that $H[P]$ becomes a fixed operator. Then one solves a nonlinear fixed-point problem to obtain the self-consistent P. The most time-consuming step is to solve the linearized

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Hartree–Fock equation. We will describe the strategy for achieving self-consistency in Section 7.4.

In order to reduce the computational cost of evaluating the Fock exchange operator, one strategy is to compute in parallel with a large number of cores. Note that the N^2 Poisson-like equations associated with the pair products $\varphi_i^*(\mathbf{r})\psi_j(\mathbf{r})$ are independent of each other. Hence one may employ in principle $O(N^2)$ processors to distribute the work (Duchemin and Gygi 2010, Valiev *et al.* 2010). The disadvantage of this approach is that the rest of the components of a Kohn–Sham solver often cannot scale to such a large number of processors, and hence the excessive number of computational cores may not be used efficiently.

In Section 4.3, we have introduced localization techniques to find localized representations of the Kohn–Sham subspace via the density matrix and Wannier functions. These techniques can be used to reduce the computational cost of evaluating the Kohn–Sham map as in Section 5. They can also be used to reduce the cost of applying exchange operators to $O(N)$. More specifically, let $\{\varphi_i\}, \{\psi_j\}$ be exponentially localized Wannier functions after a proper rotation operation, and the pair product $\varphi_i^*(\mathbf{r})\psi_j(\mathbf{r})$ would vanish if the support sets of φ_i and ψ_j do not overlap with each other. This reduces the number of pairs from $O(N^2)$ to $O(N)$. Furthermore, each Poisson-like equation only needs to be solved on a computational domain that is independent of the system size. Taking both factors into account, we find that the cost of equation (6.3) is reduced to $O(N)$. We refer readers to Wu, Selloni and Car (2009) and Dawson and Gygi (2015), for example, for more details of linear scaling hybrid functional calculations. As in the discussion in Section 5.2, linear scaling algorithms typically have a large prefactor, and hence they become advantageous to the cubic scaling methods only for systems of relatively large sizes.

On the other hand, for metallic and semi-conducting systems, the decay rate of the density matrix along the off-diagonal direction can be very slow. This may also introduce a significantly larger number of non-zero entries in the Hamiltonian matrix. Hence Green's function-based methods, such as the PEXSI method, become less efficient in the context of hybrid functional calculations.

However, the interpolative separable density fitting (ISDF) (Lu and Ying 2015) technique introduced in Section 4.4 can be applicable to insulating, semiconducting and metallic systems. Following the ISDF decomposition equation (4.13), the Poisson-like equation only needs to be solved with $\{\zeta_\mu(\mathbf{r})\}$ on the right-hand side. This reduces the number of Poisson-like equations from $O(N^2)$ to $N_\mu \sim O(N)$. Of course, the solutions of these equations need to be reassembled through linear algebra operations to obtain $\{V_{\mathrm{x}}^{\mathrm{EX}}[P]\psi_j\}_{j=1}^N$. The cost of these linear algebra steps is still $O(N^3)$ but the preconstant can be effectively reduced (Hu *et al.* 2017*a*, Dong

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

et al. 2018). Hence ISDF is effective when it is relatively costly to solve Poisson-like equations, such as in the context of planewave methods, and even more so for finite difference and finite element methods.

6.2. Adaptive compression method

When a large basis set is used, the standard method for solving Hartree–Fock-like equations self-consistently is the two-level nested SCF method in Section 7.4. The motivation for this method is that the Fock exchange energy is only a small fraction (usually 5% or less) of the total energy, although such a contribution is sufficient to result in qualitatively different results for the study of many chemical systems. The main idea is then to separate the self-consistent field (SCF) iteration into two sets of SCF iterations. In the *inner SCF iteration*, the exchange operation $V_{\mathrm{x}}^{\mathrm{EX}}$ defined by a set of orbitals $\{\varphi_i\}_{i=1}^N$ is fixed, and the Hamiltonian operator depends only on the density $\rho(\mathbf{r})$. This allows the use of efficient charge mixing schemes to reduce the number of SCF iterations. In other words, the SCF iteration then proceeds as in Kohn–Sham DFT calculations with a fixed exchange operator. In the *outer SCF iteration*, the orbitals defining the exchange operator are updated via a fixed-point iteration. Furthermore, each inner SCF iteration requires the solution of a linear eigenvalue problem, which is to be performed with iterative diagonalization methods discussed in Section 4.2 and requires repeated application of the Hamiltonian operator to occupied orbitals. The application of the Fock exchange operator to occupied orbitals according to equation (6.3) appears in this innermost loop, which makes electronic structure calculations with rung-4 functionals much more costly than those with rung-1 to rung-3 functionals.

The adaptively compressed exchange operator (ACE) method (Lin 2016) accelerates hybrid functional calculations by reducing the frequency of applications of the Fock exchange operator without compromising the accuracy of the self-consistent solution. This is achieved by constructing a low-rank surrogate operator, denoted by $\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}$, to approximate the Fock exchange operator $V_{\mathrm{x}}^{\mathrm{EX}}$. Note that $V_{\mathrm{x}}^{\mathrm{EX}}$ is generally a dense, full-rank operator. Hence the low-rank surrogate cannot be expected to be accurate when applied to an arbitrary vector. Instead we only require $\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}$ to be accurate when applied to all occupied orbitals.

More specifically, for a given set of orbitals $\{\varphi_i\}_{i=1}^N$ defining the density matrix implicitly and thus the exchange operator, we first compute the application of the exchange operator to $\{\varphi_i\}_{i=1}^N$ using

$$W_i(\mathbf{r}) = (V_{\mathrm{x}}^{\mathrm{EX}}[\{\varphi\}]\varphi_i)(\mathbf{r}), \quad i = 1, \ldots, N. \tag{6.4}$$

The adaptively compressed exchange operator, denoted by $\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}$, should

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

satisfy the conditions

$$(\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}\varphi_i)(\mathbf{r}) = W_i(\mathbf{r}) \quad \text{and} \quad \widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}(\mathbf{r},\mathbf{r}') = (\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}(\mathbf{r}',\mathbf{r}))^*. \tag{6.5}$$

The choice of this surrogate operator is not unique. One possible choice satisfying the conditions (6.5) is given by

$$\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}(\mathbf{r},\mathbf{r}') = \sum_{i,j=1}^{N} W_i(\mathbf{r}) B_{ij} W_j^*(\mathbf{r}'), \tag{6.6}$$

where $B = M^{-1}$ is a negative definite matrix, and

$$M_{kl} = \int \varphi_k^*(\mathbf{r}) W_l(\mathbf{r}) \,\mathrm{d}\mathbf{r}.$$

Perform Cholesky factorization for $-M$, *i.e.* $M = -LL^*$, where L is a lower triangular matrix. Then we get $B = -L^{-*}L^{-1}$. Define the projection vector in the ACE formulation as

$$\xi_k(\mathbf{r}) = \sum_{i=1}^{N} W_i(\mathbf{r}) (L^{-*})_{ik}. \tag{6.7}$$

Then the adaptively compressed exchange operator is given by

$$\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}(\mathbf{r},\mathbf{r}') = -\sum_{k=1}^{N} \xi_k(\mathbf{r}) \xi_k(\mathbf{r}'). \tag{6.8}$$

The ACE can be readily used to reduce the computational cost of the exchange energy, without the need to solve any extra Poisson equations, using

$$\begin{aligned} E_{\mathrm{x}}^{\mathrm{EX}} &= \frac{1}{2} \sum_{i=1}^{N} \iint \psi_i^*(\mathbf{r}) \widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}(\mathbf{r},\mathbf{r}') \psi_i(\mathbf{r}') \,\mathrm{d}\mathbf{r} \,\mathrm{d}\mathbf{r}' \\ &= -\frac{1}{2} \sum_{i,k=1}^{N} \left| \int \psi_i^*(\mathbf{r}) \xi_k(\mathbf{r}) \,\mathrm{d}\mathbf{r} \right|^2. \end{aligned} \tag{6.9}$$

Note that once ACE is constructed, the cost of applying $\widetilde{V}_{\mathrm{x}}^{\mathrm{EX}}$ to any orbital ψ is similar to the application of a non-local pseudopotential operator, thanks to its low-rank structure. ACE only needs to be constructed once per outer iteration, and can be repeatedly used for all the subsequent inner SCF iterations for the electron density, and each iterative step for solving the linear eigenvalue problem. The ACE formulation has been integrated in electronic structure software packages such as Quantum ESPRESSO.

Table 6.1 demonstrates the accuracy of the ACE formulation for Kohn–Sham DFT calculations with the HSE06 (Heyd *et al.* 2003) hybrid functional, for silicon systems ranging from 64 to 1000 atoms. The ACE formu-

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Table 6.1. Comparison between the conventional hybrid DFT calculations and ACE-enabled hybrid DFT calculations in terms of the HF energy $E_{\mathrm{x}}^{\mathrm{EX}}$ (Hartree) and the energy gap E_{gap} (Hartree) for the Si_{64}, Si_{216}, Si_{512} and Si_{1000} systems. The corresponding relative errors of the HF energy are shown in parentheses.

Methods	ACE HSE06		Conventional HSE06	
Systems	$E_{\mathrm{x}}^{\mathrm{EX}}$	E_{gap}	$E_{\mathrm{x}}^{\mathrm{EX}}$	E_{gap}
Si_{64}	-13.541616 (10^{-6})	1.488335	-13.541629	1.488352
Si_{216}	-45.471192 (10^{-7})	1.449790	-45.471190	1.449790
Si_{512}	-107.698011 (10^{-7})	1.324901	-107.698016	1.324902
Si_{1000}	-210.300628 (10^{-6})	1.289162	-210.300524	1.289128

Table 6.2. Comparison between the conventional hybrid DFT calculation and an ACE-enabled hybrid DFT calculation in terms of the number of inner SCF iterations and wall clock time spent in each outer SCF iteration for Si_{1000} on 2000 cores.

Methods	ACE HSE06		Conventional HSE06	
#Outer SCF	#Inner SCF	Time (s)	#Inner SCF	Time (s)
1	6	356	6	2518
2	5	320	5	2044
3	5	308	4	1665

lation can evaluate the HF energy and energy gap accurately (total energy difference is under 10^{-4} Hartree) even for large systems, and the remaining difference is in fact due to the tolerance of the SCF iteration. Table 6.2 demonstrates that ACE can perform hybrid functional calculations at a fraction of the cost of conventional methods (Hu *et al.* 2017*c*).

The efficiency of the ACE formulation also rests on the assumption that the magnitude of the Fock exchange operator is relatively small compared to other components of the Hamiltonian. To be more specific, we demonstrate below that in the context of linearized Hartree–Fock-like equations, the ACE formulation leads to desirable convergence properties compared to standard iterative solvers.

To make the discussion more general, we consider the following linear eigenvalue problem (linearized Hartree–Fock-like equations take this form):

$$(A+B)v_i = \lambda_i v_i, \quad i = 1, \ldots, N. \tag{6.10}$$

Here $A, B \in \mathbb{C}^{N_b \times N_b}$ are Hermitian matrices. The eigenvalues $\{\lambda_i\}$ are real

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

and ordered non-decreasingly. Due to the Pauli exclusion principle we need to compute the eigenpairs (λ_i, v_i) corresponding to the lowest N eigenvalues, which are assumed to be separated from the rest of the eigenvalues by a positive spectral gap $\lambda_g := \lambda_{N+1} - \lambda_N$. The matrix A in (6.10) is obtained by discretizing the Hamiltonian operator excluding the exchange operator, while B as a discretization of the exchange operator is negative definite.

In order to reduce the number of matrix–vector multiplication operations Bv, the simplest idea is to fix $w_i := Bv_i$ at some stage, and to replace Bv_i with w_i for a number of iterations. This leads to the following subproblem:

$$Av_i + w_i = \lambda_i v_i, \quad i = 1, \dots, N. \tag{6.11}$$

Note that equation (6.11) is not an eigenvalue problem: if v_i is a solution to (6.11), then v_i multiplied by a constant c is typically not a solution. Equation (6.11) could be solved using optimization-based methods, but such a problem is typically more difficult than a Hermitian eigenvalue problem. In practice, software packages for solving Hartree–Fock-like equations are typically built around eigensolvers, which is another important factor that makes subproblem (6.11) undesirable.

The adaptive compression method re-uses the information in $\{w_i\}$ in a different way, retaining the structure of the eigenvalue problem (6.10). Define $V = [v_1, \dots, v_N]$, $W = [w_1, \dots, w_N]$, so $V, W \in \mathbb{C}^{N_b \times N}$, and construct

$$\widetilde{B}[V] = W(W^*V)^{-1}W^*. \tag{6.12}$$

Since $B \prec 0$, $W^*V \equiv V^*BV$ has only negative eigenvalues and is invertible. $\widetilde{B}[V]$ is Hermitian of rank N, and agrees with B when applied to V as

$$\widetilde{B}[V]V = W(W^*V)^{-1}W^*V = W = BV. \tag{6.13}$$

We shall refer to the operation from B to $\widetilde{B}[V]$ as an *adaptive compression*. It turns out that the adaptive compression $\widetilde{B}[V]$ is the unique rank-N Hermitian matrix that agrees with B on $\operatorname{span} V$. Furthermore, $B \preceq \widetilde{B}[V] \preceq 0$ (Lin and Lindsey 2019).

Note that the subspace $\operatorname{span} V$ is precisely the solution for (6.10) and is not known *a priori*. Therefore $\widetilde{B}$ needs to be constructed in an adaptive manner. Starting from some initial guess $V^{(0)}$, we will obtain a sequence $V^{(k)}$ and corresponding compressed operators $\widetilde{B}[V^{(k)}]$. More specifically, ACE uses a fixed-point iteration given by

$$(A + \widetilde{B}[V^{(k)}])v_i^{(k+1)} = \lambda_i^{(k+1)} v_i^{(k+1)}, \quad i = 1, \dots, N. \tag{6.14}$$

In each iteration, after $\widetilde{B}[V^{(k)}]$ is constructed, (6.14) can be solved via *any* iterative eigensolver to obtain $V^{(k+1)}$. The iterative eigensolver only requires application of A and the low-rank matrix $\widetilde{B}$ to vectors, and does not require any additional application of B until $V^{(k+1)}$ is obtained. If $\operatorname{span} V^{(k)}$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 9: Adaptive compression method for solving linear eigenvalue problem

```
Initialize $V^{(0)}$ by solving $Av_i^{(0)} = \lambda_i^{(0)} v_i^{(0)}, \quad i = 1, \ldots, N$.
while convergence not reached do
    Compute $W^{(k)} = BV^{(k)}$.
    Evaluate $[(W^{(k)})^* V^{(k)}]^{-1}$ to construct $\widetilde{B}[V^{(k)}]$ implicitly.
    Solve (6.14) to obtain $V^{(k+1)}$.
    Set $k \leftarrow k + 1$.
end while
```

converges to span V, then the consistency condition $\widetilde{B}[V]V = BV$ is satisfied, and the adaptive compression method is numerically exact. The adaptive compression method for solving the linear eigenvalue problem (6.10) is given in Algorithm 9, where we initialize $V^{(0)}$ by solving the eigenvalue problem in the absence of B.

At first glance, the advantage of converting a linear eigenvalue problem (6.10) into a nonlinear eigenvalue problem (6.14) is unclear. The advantage of the adaptive compression method comes from the decoupling of the matrix–vector multiplication operations Av and Bv, and asymptotically the number of Bv operations is independent of the spectral radius $\|A\|_2$. More specifically, starting from an initial density matrix $P^{(0)}$, the asymptotic convergence rate measured by the convergence of the density matrix at the kth iteration is given by

$$\|P - P^{(k)}\|_2 \lesssim \gamma^k \|P - P^{(0)}\|_2, \quad \text{where} \quad \gamma \leq \frac{\|B\|_2}{\|B\|_2 + \lambda_g}. \tag{6.15}$$

Furthermore, one may prove that the adaptive compression method converges globally starting from *almost every* initial guess. We refer readers to Lin and Lindsey (2019) for more details.

Remarks

An alternative way to adaptively reduce the rank of the exchange operator is via a projector-based compression of the exchange operator (Duchemin and Gygi 2010, Boffi, Jain and Natan 2016). Compared to the discussion in Section 2.7, we can also find that the Kleinman–Bylander form of the pseudopotential follows the same spirit as that in the ACE formulation to construct a low-rank approximation to the semi-local pseudopotential. However, in the case of the pseudopotential, the orbitals are fixed and are not computed adaptively. The adaptive compression method is also related to the Nyström sketching method in numerical optimization, and can be used to accelerate the convergence rate of optimization-based electronic structure solvers as well (Hu *et al.* 2018).

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

6.3. RPA correlation energy

After hybrid functionals that involve the exact exchange operator, the next family of approximations to the exchange-correlation functional involves virtual orbitals, such as the random phase approximation to the correlation functional discussed in Section 2.5; see (2.49). The approximate functional form on the fifth rung is still under active development and thus the efficient numerical algorithm for such functionals is also an active research field with many recent ideas and on-going work. It is beyond the scope of this work to review all these developments, and we will restrict our focus to recent work on cubic scaling algorithms for the computation of RPA correlation energy (Lu and Thicke 2017*a*) based on the interpolative separable density fitting, as discussed in Section 4.4.

Several cubic scaling methods to calculate the RPA correlation energy have been developed. Recall that the dynamic polarizability operator is defined as

$$\hat{\chi}^0(\mathbf{r},\mathbf{r}',\mathrm{i}\omega)=\sum_i^{\mathrm{occ}}\sum_a^{\mathrm{vir}}\frac{\psi_i^*(\mathbf{r})\psi_a(\mathbf{r})\psi_a^*(\mathbf{r}')\psi_i(\mathbf{r}')}{\varepsilon_i-\varepsilon_a-\mathrm{i}\omega}+\text{c.c.} \tag{6.16}$$

The general idea is to split up the i and a dependence in the computation of $\hat{\chi}^0$ by introducing a new integral. Two different integrals have been utilized for this purpose. The first is

$$-\int_0^\infty \mathrm{e}^{\varepsilon_i t}\,\mathrm{e}^{-\varepsilon_a t}\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t=\frac{1}{\varepsilon_i-\varepsilon_a-\mathrm{i}\omega}, \tag{6.17}$$

where $\varepsilon_a>\varepsilon_i$ since the former corresponds to a virtual orbital and the latter to an occupied orbital. Using this integral, one separates the dependence of i and a in (6.16) into a product of exponentials inside the integral. This leads to the Laplace transform cubic scaling methods. This idea was first applied to RPA calculations by Kaltak, Klimeš and Kresse (2014*a*, 2014*b*), using a projector augmented wave (PAW) basis. It was later extended to atomic orbitals (Schurkus and Ochsenfeld 2016, Luenser, Schurkus and Ochsenfeld 2017, Wilhelm, Seewald, Del Ben and Hutter 2016). Another integral decomposition that can be used to break up the i and a dependence is

$$\frac{1}{2\pi\mathrm{i}}\oint_{\mathscr{C}}\frac{1}{(\lambda-\varepsilon_i+\mathrm{i}\omega)(\lambda-\varepsilon_a)}\,\mathrm{d}\lambda=\frac{1}{\varepsilon_i-\varepsilon_a-\mathrm{i}\omega}, \tag{6.18}$$

where $\mathscr{C}$ is a positively oriented closed contour that encloses $\varepsilon_i-\mathrm{i}\omega$, but not ε_a: an example is shown in Figure 6.1. This idea was first used in the context of cubic scaling RPA in Moussa (2014), and combined with the ISDF approach in Lu and Thicke (2017*a*) to further reduce the computational cost by taking advantage of the 'low-rank' nature of $\hat{\chi}^0$. If the usual density fitting approximation is used, the $\hat{\chi}^0$ cannot be constructed in $O(N^3)$ since

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

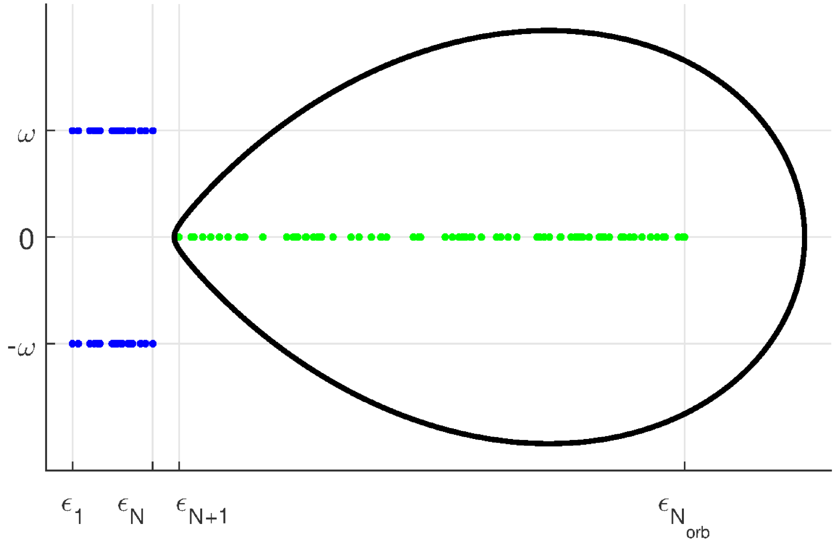


Figure 6.1. (Credit: Lu and Thicke 2017*a*.) An example of contour $\mathscr{C}$. The blue points represent $\{\varepsilon_i \pm \mathrm{i}\omega\}_{i=1}^{N}$ (for a particular choice of ω) and the green points represent $\{\varepsilon_a\}_{a=N+1}^{N_{\mathrm{orb}}}$.

i and a are coupled in the coefficients of the density fitting method. The solution to this is provided by the interpolative separable density fitting (ISDF) method (Lu and Ying 2015) which keeps the dependence on i and a separate. We remark that density fitting has been widely used to improve the computational efficiency in electronic structure theory, including rung-5 functionals (Feyereisen, Fitzgerald and Komornicki 1993, Weigend, Häser, Patzelt and Ahlrichs 1998, Ren *et al.* 2012*a*); the ISDF has the additional benefit of keeping indices separated to make it more flexible.

After we choose the contour $\mathscr{C}$ and apply Cauchy's integral formula to split up the dependence of i and a in the denominator of (6.16), we obtain a reformulation for $\hat{\chi}^0$ which can be computed in cubic time complexity,

$$\begin{aligned}\langle \mathbf{r}|\hat{\chi}^0(\mathrm{i}\omega)|\mathbf{r}'\rangle &= \sum_{i}^{\mathrm{occ}}\sum_{a}^{\mathrm{vir}} \frac{\psi_i^*(\mathbf{r})\psi_a(\mathbf{r})\psi_a^*(\mathbf{r}')\psi_i(\mathbf{r}')}{\varepsilon_j - \varepsilon_k - \mathrm{i}\omega} + \text{c.c.} \qquad (6.19)\\ &= \frac{1}{2\pi\,\mathrm{i}}\oint_{\mathscr{C}} \left(\sum_{i}^{\mathrm{occ}} \frac{\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}')}{\lambda - \varepsilon_i + \mathrm{i}\omega}\right)\left(\sum_{a}^{\mathrm{vir}} \frac{\psi_a(\mathbf{r})\psi_a^*(\mathbf{r}')}{\lambda - \varepsilon_a}\right) \mathrm{d}\lambda + \text{c.c.}\end{aligned}$$

Similar to that discussed in the case of PEXSI (see Section 5.3), the contour integral can be discretized and the number of quadrature points is logarithmic in $(\varepsilon_{N_{\mathrm{orb}}} - \varepsilon_N)/(\varepsilon_{N+1} - \varepsilon_N)$; we refer readers to Lu and Thicke (2017*a*) for details.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Note that the formula (6.19) already provides a cubic scaling method for calculating $\chi^0(\mathrm{i}\omega)$. In particular, ignoring logarithmic factors, $\chi^0(\mathrm{i}\omega)$ can be calculated with cost $O(N_{\mathrm{orb}}N_g^2)$, assuming we keep N_{orb} total orbitals in the summation (known as the virtual orbital cut-off). However, the number of grid points N_g could be much larger than the number of orbitals N_{orb}. This motivates the use of ISDF to reduce the prefactor in the computational cost. We approximate the $\hat{\chi}^0$ operator by an operator $\tilde{\chi}^0$ by using the ISDF approximation:

$$\begin{aligned}
&\langle \mathbf{r}|\hat{\chi}^0(\mathrm{i}\omega)|\mathbf{r}'\rangle \\
&= \frac{1}{2\pi\mathrm{i}}\oint_{\mathscr{C}} \Bigg(\sum_i^{\mathrm{occ}} \frac{\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}')}{\lambda-\varepsilon_i+\mathrm{i}\omega}\Bigg)\Bigg(\sum_a^{\mathrm{vir}} \frac{\psi_a(\mathbf{r})\psi_a^*(\mathbf{r}')}{\lambda-\varepsilon_a}\Bigg)\,\mathrm{d}\lambda + \text{c.c.} \\
&\approx \sum_{\mu\nu} \frac{1}{2\pi\mathrm{i}}\oint_{\mathscr{C}} \Bigg(\sum_i^{\mathrm{occ}} \frac{\psi_i^*(\hat{\mathbf{r}}_\mu)\psi_i(\hat{\mathbf{r}}_\nu)}{\lambda-\varepsilon_i+\mathrm{i}\omega}\Bigg)\Bigg(\sum_a^{\mathrm{vir}} \frac{\psi_a(\hat{\mathbf{r}}_\mu)\psi_a^*(\hat{\mathbf{r}}_\nu)}{\lambda-\varepsilon_a}\Bigg)\,\mathrm{d}\lambda\,\zeta_\mu(\mathbf{r})\zeta_\nu(\mathbf{r}') + \text{c.c.} \\
&=: \sum_{\mu\nu} \chi^0_{\mu\nu}(\mathrm{i}\omega)\zeta_\mu(\mathbf{r})\zeta_\nu(\mathbf{r}') =: \langle \mathbf{r}|\tilde{\chi}^0(\mathrm{i}\omega)|\mathbf{r}'\rangle, \qquad (6.20)
\end{aligned}$$

where the last line defines $\chi^0_{\mu\nu}$ and $\tilde{\chi}^0$. We note that in the above the separability of the ISDF coefficients into the i and a components is crucial. Without this separability (*e.g.* if a conventional density fitting were used) we would not be able to calculate $\tilde{\chi}^0$ in cubic time since the sums over i and a would not decouple.

We now return our attention to the RPA correlation energy under random phase approximation (2.49). Given $\tilde{\chi}^0$ sufficiently close to $\hat{\chi}^0$, we can approximate $\mathrm{Tr}[\ln(I-\hat{\chi}^0 v_C)]$ using (with a suitable choice of constant c for convergence of the series)

$$\begin{aligned}
\ln(I-\tilde{\chi}^0 v_C) &= \ln[cI-((c-1)I+\tilde{\chi}^0 v_C)] \\
&= \ln(c)I + \ln\bigg[I-\frac{1}{c}((c-1)I+\tilde{\chi}^0 v_C)\bigg] \\
&= \ln(c)I - \sum_{\ell=1}^{\infty} \frac{[(c-1)I+\tilde{\chi}^0 v_C]^\ell}{\ell c^\ell} \\
&= \ln(c)I - \sum_{\ell=1}^{\infty} \frac{1}{\ell c^\ell}\sum_{p=0}^{\ell}\binom{\ell}{p}(c-1)^{\ell-p}(\tilde{\chi}^0 v_C)^p, \qquad (6.21)
\end{aligned}$$

and represent each the polynomial in terms of $\tilde{\chi}^0 v_C$ on the right-hand side in the auxiliary basis. After some manipulation, this results in our final desired approximation,

$$\mathrm{Tr}[\ln(1-\hat{\chi}^0(\mathrm{i}\omega)v_C)+\hat{\chi}^0(\mathrm{i}\omega)v_C] \approx \mathrm{Tr}[\ln(1-\chi^0(\mathrm{i}\omega)v)+\chi^0(\mathrm{i}\omega)v], \quad (6.22)$$

where χ^0 and v are the matrix elements in the auxiliary basis.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 10: Cubic scaling calculation of the RPA correlation energy

```
Input: Kohn–Sham orbitals $\{\psi_k\}$ and corresponding energies $\{\varepsilon_k\}$.
Output: $E_{\mathrm{c}}^{\mathrm{RPA}}$
1: Use $\{\psi_k\}_{k=1}^{N_{\mathrm{orb}}}$ as the input to ISDF to obtain $\{\hat{\mathbf{r}}_\mu\}_{k=1}^{N_\mu}$ and $\{\zeta_\mu\}_{k=1}^{N_\mu}$.
2: Compute the matrix $v_{\mu,\nu} = \langle \zeta_\mu | v_C | \zeta_\nu \rangle$.
3: For each quadrature point $\omega_m$ for the contour integral on $\mathscr{C}$:
     (a) Compute $\chi^0_{\mu,\nu}(\mathrm{i}\omega_m)$ as defined in (6.20).
     (b) Compute $\frac{1}{2\pi}\operatorname{Tr}[\ln(1-\chi^0(\mathrm{i}\omega_m)v)+\chi^0(\mathrm{i}\omega_m)v]$.
4: Calculate
     $$E_{\mathrm{c}}^{\mathrm{RPA}} = \frac{1}{2\pi}\int_0^\infty \operatorname{Tr}[\ln(1-\chi^0(\mathrm{i}\omega)v)+\chi^0(\mathrm{i}\omega)v]\,\mathrm{d}\omega$$
   via numerical quadrature.
```

We present the cubic scaling algorithm for the calculation of the RPA correlation energy as in Algorithm 10. Details of the algorithm can be found in Lu and Thicke (2017*a*).

Our numerical results use the following test problem. Our two-dimensional spatial grid is $10 \times 10N$ equally spaced points. First, we solve for the Kohn–Sham orbitals of the periodic system with an external potential consisting of N Gaussian potential wells, the centres of which are randomly perturbed from the centres of their respective 10×10 box of grid points. Then the eigenvectors of H are used as the orbitals in the calculation of the RPA correlation energy.

We investigate the cubic scaling behaviour of the algorithm in Figure 6.2. The quartic scaling method using traditional density fitting is also plotted for comparison. In this example, we choose the virtual orbital cut-off $N_{\mathrm{orb}} = 0.2N_g$ and the system size is scaled up to $N = 160$. We observe that the cubic scaling algorithm greatly outperforms the quartic scaling algorithm for large system sizes.

7. Self-consistent field iteration

In this section we discuss numerical methods for performing self-consistent field (SCF) iterations. In Sections 7.1 and 7.2 we introduce basic SCF iteration techniques such as the fixed-point iteration and the simple mixing method, as well as acceleration techniques based on Newton and quasi-Newton methods. We introduce these techniques in the context of semi-local functionals. In Section 7.3 we discuss preconditioning techniques for SCF

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

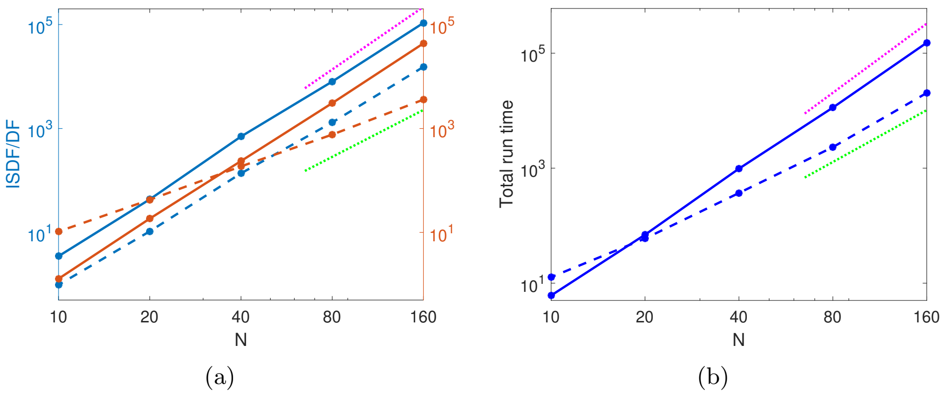


Figure 6.2. (Credit: Lu and Thicke 2017*a*.) Timing results for the quartic scaling method are plotted with solid lines, and results for the cubic scaling method are plotted with dashed lines. For reference, the purple and green dotted lines represent the slopes of N^4 and N^3 respectively. (a) Comparison of the time required to calculate χ^0 and the time to perform the respective density fitting schemes for each method. (b) The total running time to calculate $E_{\mathrm{c}}^{\mathrm{RPA}}$ for each method.

iterations specific to semi-local functionals using a large basis set. Similarly Section 7.4 introduces SCF techniques specifically for non-local functionals. Throughout the discussion, we will assume a large basis set is used.

7.1. Fixed-point iteration and simple mixing

For a fixed atomic configuration, the self-consistent iteration starts from certain initial electron density denoted by ρ_0. We let ρ_k, V_k denote the electron density and the effective potential V_{eff} at the kth SCF iteration, respectively. When pseudopotentials are used, since the non-local pseudopotential is independent of ρ, V_{eff} is still a local potential as in equation (2.91). Then the flow of the SCF iteration becomes

$$\cdots \to \rho_k \to V_k = V_{\mathrm{eff}}[\rho_k] \to \rho_{k+1} \to V_{k+1} = V_{\mathrm{eff}}[\rho_{k+1}] \to \cdots. \tag{7.1}$$

Depending on the starting point, the relation (7.1) can be viewed as a mapping from ρ_k to ρ_{k+1}, or from V_k to V_{k+1}. The former viewpoint is called *density mixing*, while the latter is called *potential mixing*. There is no qualitative difference between the two types of mixing schemes. However, the density mixing has the extra constraint that the density must be nonnegative everywhere, and must be normalized to have the correct number of electrons. In practice, this constraint can be easily satisfied by setting all the negative entries of ρ (usually these entries have very small magnitude) to 0 or a small positive number, followed by a normalization step. On the other hand, potential mixing is formally free of such constraints, and hence we will consider potential mixing below. We remark that both density mixing and

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

potential mixing strategies are widely used in electronic structure software packages, and the algorithm below for potential mixing can be used for density mixing as well.

When self-consistency is reached, the converged effective potential is denoted by $V_\star$ and satisfies the nonlinear equation

$$V_\star = V_{\text{eff}}[\mathcal{F}_{\text{KS}}[V_\star]]. \tag{7.2}$$

The simplest version of the SCF iteration is the fixed-point iteration, where the potential at the $(k+1)$th step is directly given by the output potential at the kth step:

$$V_{k+1} = V_{\text{eff}}[\mathcal{F}_{\text{KS}}[V_k]]. \tag{7.3}$$

Since the exchange-correlation functional is neither convex nor concave with respect to the electron density, rigorous study of the global convergence properties of SCF schemes in Kohn–Sham DFT calculations is difficult. Hence we consider the linear response regime, where we assume the initial effective potential V_0 is already close to $V_\star$. Let

$$e_k := V_k - V_\star$$

be the error of the potential at the kth iteration. In order to study the propagation of the error in the fixed-point iteration (7.3), we apply the chain rule

$$e_{k+1} = f_{\text{Hxc}}\chi_0 e_k + O(\|e_k\|^2). \tag{7.4}$$

Here $f_{\text{Hxc}} = \delta V_{\text{eff}}/\delta\rho$ is the kernel for the Hartree and exchange-correlation contributions, and $\chi_0 = \delta\mathcal{F}_{\text{KS}}/\delta V$ is called the independent particle polarizability operator. In the linear response regime, we assume the $O(\|e_k\|^2)$ term is small and it is omitted in the following discussion. Then after k steps

$$e_k \approx (f_{\text{Hxc}}\chi_0)^k e_0. \tag{7.5}$$

Hence in the linear response regime, the convergence of the fixed-point iteration requires that the spectral radius of the operator, denoted by $r_\sigma(f_{\text{Hxc}}\chi_0)$, is smaller than 1. Unfortunately, such a spectral radius is generally much larger than 1, and the error in the fixed-point iteration will therefore diverge, even if the initial potential is already very close to the self-consistent potential.

In order to achieve the self-consistent solution, the simplest practically usable scheme is the simple mixing method, which introduces a slight modification of the fixed-point iteration, shown in Algorithm 11. The iteration can also be written equivalently as

$$V_{k+1} = \alpha V_{\text{eff}}[\mathcal{F}_{\text{KS}}[V_k]] + (1-\alpha)V_k. \tag{7.6}$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 11: Simple mixing method

```
Input: Initial guess $V_0$ of the potential and relaxation parameter $\alpha$.
for $k = 0, \dots,$ until convergence do
    Apply the Kohn–Sham map to compute the density $\rho_k = \mathcal{F}_{\mathrm{KS}}[V_k]$.
    Form the residual error $r_k = V_k - V_{\mathrm{eff}}[\rho_k]$.
    $V_{k+1} = V_k - \alpha r_k$.
end for
```

In the simple mixing method, the error propagation follows

$$e_{k+1} = (I - \alpha \epsilon_d) e_k + O(\|e_k\|^2). \tag{7.7}$$

Here $\epsilon_d = I - f_{\mathrm{Hxc}} \chi_0$ is called the dielectric operator. Equation (7.7) can be iterated, yielding

$$e_k \approx (I - \alpha \epsilon_d)^k e_0. \tag{7.8}$$

When the exchange-correlation functional contribution is dropped from f_{Hxc}, the resulting approximation is called the random phase approximation (RPA).[1] Then ϵ_d can be transformed using a similarity transformation to a positive definite matrix. In particular, ϵ_d is diagonalizable with real positive eigenvalues. In order to achieve convergence in the linear response regime, we need each eigenvalue of ϵ_d, denoted by λ, to satisfy

$$|1 - \alpha\lambda| < 1,$$

or equivalently

$$0 < \alpha < \frac{2}{\lambda}.$$

Let $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalues of ϵ_d, respectively, and the spectral radius $r_\sigma(\epsilon_d)$ is then given by $\lambda_{\max}$. If

$$0 < \alpha < \frac{2}{r_\sigma(\epsilon_d)} \tag{7.9}$$

is satisfied, the simple mixing would converge.

It remains to determine the *optimal* choice of α satisfying the constraint (7.9). This requires the solution of the following minimax problem:

$$\min_{\alpha} \max_{\lambda} |1 - \alpha\lambda|. \tag{7.10}$$

[1] Here the meaning of ‘RPA’ is related to the RPA functional in Section 6.3, in the sense that the exchange-correlation kernel is dropped in both cases, but the usage has different origins.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

The optimal choice of α satisfies

$$1 - \alpha\lambda_{\min} = \alpha\lambda_{\max} - 1, \tag{7.11}$$

or

$$\alpha = \frac{2}{\lambda_{\min} + \lambda_{\max}}. \tag{7.12}$$

Substituting this choice of α into (7.7), we find that the optimal convergence rate of simple mixing is

$$\max_{\lambda}|1 - \alpha\lambda| = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} = \frac{\kappa(\epsilon_d) - 1}{\kappa(\epsilon_d) + 1}. \tag{7.13}$$

Here $\kappa(\varepsilon_d) = \lambda_{\max}/\lambda_{\min}$ is the condition number of the dielectric operator.

In practical calculations, the condition number might be very large and the simple mixing method converges with a very slow rate. Thus the simple mixing method is rarely used directly in practical electronic structure calculations.

7.2. Newton and quasi-Newton type methods

The convergence of the simple mixing method often requires a rather small mixing constant α. Hence the SCF procedure may take many iterations to converge. One possible acceleration can be achieved by using Newton's method, which can be written as

$$V_{k+1} = V_k - \mathcal{J}_k^{-1} r_k. \tag{7.14}$$

Here $\mathcal{J}_k$ is the Jacobian matrix for the residual map

$$\delta V \mapsto \delta V - V_{\text{eff}}\big[\mathcal{F}_{\text{KS}}[V_k + \delta V]\big].$$

Note that at converged potential $V_\star$, the Jacobian matrix $\mathcal{J}_\star = \epsilon_d$. Hence the simple mixing method can also be interpreted as approximating the inverse of the Jacobian matrix $\mathcal{J}_k^{-1}$ simply by αI in the evaluation of the Jacobian matrix for the composition map $V_{\text{eff}} \circ \mathcal{F}_{\text{KS}}$. For a system with N electrons, the evaluation of the Jacobian matrix requires in principle $O(N)$ evaluations of the Kohn–Sham map, which is prohibitively expensive.

The Jacobian-free Newton–Krylov method replaces the need for the explicit evaluation of the Jacobian matrix by solving a linear equation

$$\mathcal{J}_k \delta V_k = -r_k \tag{7.15}$$

to obtain the Newton update δV_k. This can be done using iterative methods for solving linear equations, such as the generalized minimal residual method (GMRES) (Saad and Schultz 1986). In order to compute the matrix–vector

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

multiplication related to the Jacobian matrix, one can use the finite difference approximation

$$\mathcal{J}_k \delta V_k \approx \delta V_k - \big(V_{\text{eff}}\left[\mathcal{F}_{\text{KS}}[V_k + \delta V_k]\right] - V_{\text{eff}}\left[\mathcal{F}_{\text{KS}}[V_k]\right]\big). \tag{7.16}$$

The finite difference calculation requires at least one additional function evaluation of $\mathcal{F}_{\text{KS}}(V_k + \delta V)$ per iteration step. Therefore, even though Newton's method may exhibit local quadratic convergence, each Newton iteration may take many inner iterations in order to solve the linear equation (7.15).

A widely used alternative to Newton's method is the class of quasi-Newton methods, which replace $\mathcal{J}_k^{-1}$ with an approximate matrix C_k that is easy to compute and apply. Then the updating strategy becomes

$$V_{k+1} = V_k - C_k r_k. \tag{7.17}$$

Using Broyden's techniques (Nocedal and Wright 1999), one can systematically approximate $\mathcal{J}_k$ or $\mathcal{J}_k^{-1}$. In Broyden's second method, C_k is obtained by performing a sequence of low-rank modifications to some initial approximation C_0 of the Jacobian inverse using a recursive formula (Fang and Saad 2009, Marks and Luke 2008). At each step, C_k is obtained by solving the following constrained optimization problem:

$$\begin{aligned} &\min_C \frac{1}{2}\|C - C_{k-1}\|_F^2 \\ &\text{subject to } S_k = C Y_k, \end{aligned} \tag{7.18}$$

where C_{k-1} is the approximation to the Jacobian constructed in the $(k-1)$th Broyden iteration. The matrices S_k and Y_k above are defined as

$$S_k = (s_k, s_{k-1}, \ldots, s_{k-\ell}), \quad Y_k = (y_k, y_{k-1}, \ldots, y_{k-\ell}), \tag{7.19}$$

where s_j and y_j are defined by $s_j = V_j - V_{j-1}$ and $y_j = r_j - r_{j-1}$ respectively. The number ℓ is the length of history used in Broyden's method, which is typically set to $5 \sim 20$ in practical calculations.

Equation (7.18) is a constrained optimization problem and can be solved using the method of Lagrange multipliers, given by

$$C_k = C_{k-1} + (S_k - C_{k-1} Y_k) Y_k^{\dagger}. \tag{7.20}$$

Here $Y_k^{\dagger}$ denotes the Moore–Penrose pseudo-inverse of Y_k, that is, $Y_k^{\dagger} = (Y_k^* Y_k)^{-1} Y_k^*$. We remark that in practice $Y_k^{\dagger}$ is not constructed explicitly since we need only apply $Y_k^{\dagger}$ to a residual vector r_k. This operation can be carried out by solving a linear least-squares problem with appropriate regularization (*e.g.* through a truncated singular value decomposition).

A variant of Broyden's method is Anderson's method (Anderson 1965), which is widely used in electronic structure software packages. Anderson's

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 12: Anderson's method

```
Input: Initial guess $V_0$ of the potential, relaxation parameter $\alpha$, history length $\ell$
for $k = 0, \ldots,$ until convergence do
    Apply the Kohn–Sham map to compute the density $\rho_k = \mathcal{F}_{\mathrm{KS}}[V_k]$.
    Form the residual error $r_k = V_k - V_{\mathrm{eff}}[\rho_k]$.
    Define $S_k = (s_k, s_{k-1}, \ldots, s_{k-\ell})$ with $s_j = V_j - V_{j-1}$ and
    $Y_k = (y_k, y_{k-1}, \ldots, y_{k-\ell})$ with $y_j = r_j - r_{j-1}$. For $k < \ell$, keep only
    the vectors starting from index 0 for both $S_k$ and $Y_k$ matrices.
    Form $d_k = Y_k^{\dagger} r_k$ by solving a linear least-squares problem with
    appropriate regularization.
    $V_{k+1} = V_k - \alpha(r_k - Y_k d_k) - S_k d_k$.
end for
```

method fixes C_{k-1} to the initial approximation C_0 when solving (7.18) during each iteration. It follows from equation (7.17) that Anderson's method updates the potential as

$$V_{k+1} = V_k - C_0(I - Y_k Y_k^{\dagger}) r_k - S_k Y_k^{\dagger} r_k. \tag{7.21}$$

In particular, if C_0 is set to αI, we obtain Anderson's method,

$$V_{k+1} = V_k - \alpha(I - Y_k Y_k^{\dagger}) r_k - S_k Y_k^{\dagger} r_k,$$

commonly used in Kohn–Sham DFT solvers. Pseudocode for Anderson's method is shown in Algorithm 12.

An alternative way to derive Broyden's method is through a technique called direct inversion of iterative subspace (DIIS). The technique was originally developed by Pulay to accelerate Hartree–Fock calculations (Pulay 1980). Hence it is often referred to as Pulay mixing. The motivation of Pulay's method is to minimize the residual $V - V_{\mathrm{eff}}[\mathcal{F}_{\mathrm{KS}}(V)]$ within the subspace spanned by $\{V_{k-\ell-1}, \ldots, V_k\}$. In Pulay's original work (Pulay 1980), the optimal approximation to V is expressed as

$$V_{\mathrm{opt}} = \sum_{j=k-\ell-1}^{k} \alpha_j V_j,$$

where V_j $(j = k - \ell - 1, \ldots, k)$ are previous approximations to V, and the coefficients α_j are chosen to satisfy the constraint $\sum_{j=k-\ell-1}^{k} \alpha_j = 1$.

When the V_j are all sufficiently close to the fixed-point solution,

$$V_{\mathrm{eff}}\left[\mathcal{F}_{\mathrm{KS}}\left(\sum_j \alpha_j V_j\right)\right] \approx \sum_j \alpha_j V_{\mathrm{eff}}[\mathcal{F}_{\mathrm{KS}}(V_j)]$$

holds approximately. Hence we may obtain α_j (and consequently V_{opt}) by

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

solving the following quadratic programming problem:

$$\min_{\{\alpha_j\}} \Bigg\| \sum_{j=k-\ell-1}^{k} \alpha_j r_j \Bigg\|_2^2 \quad \text{subject to} \quad \sum_{j=k-\ell-1}^{k} \alpha_j = 1, \tag{7.22}$$

where $r_j = V_j - V_{\text{eff}}[\mathcal{F}_{\text{KS}}(V_j)]$.

Note that (7.22) can be reformulated as an unconstrained minimization problem if V_{opt} is required to take the form

$$V_{\text{opt}} = V_k + \sum_{j=k-\ell}^{k} \beta_j (V_j - V_{j-1}),$$

where β_j can be any unconstrained real number. Again, if we assume $V_{\text{eff}}[\mathcal{F}_{\text{KS}}(V)]$ is approximately linear around V_j and let $b = (\beta_{k-\ell}, \dots, \beta_k)^\top$, minimizing $\|V_{\text{opt}} - V_{\text{eff}}[\mathcal{F}_{\text{KS}}(V_{\text{opt}})]\|$ with respect to $\{\beta_j\}$ yields $b = -Y_k^\dagger r_k$, where Y_k is the same as that defined in (7.19). Then Pulay's method for updating V is thus $V_{k+1} = V_{\text{opt}}$. More generally, we may introduce a C_0 matrix as

$$V_{k+1} = V_{\text{opt}} - C_0(V_{\text{opt}} - V_{\text{eff}}[\mathcal{F}_{\text{KS}}(V_{\text{opt}})]). \tag{7.23}$$

Substituting $V_{\text{opt}} = V_k - S_k Y_k^\dagger r_k$ into (7.23), together with the linearity assumption of $V_{\text{eff}}[\mathcal{F}_{\text{KS}}(V)]$, yields exactly Anderson's updating formula (7.21).

The matrix C_0 plays the role of a preconditioner, and hence a better C_0 can be chosen to accelerate the convergence of Anderson's method in practical electronic structure calculations, as will be discussed in Section 7.3.

It is worthwhile remarking that a variant of the Pulay method above can be derived by taking other definitions of the residual vector r. This leads to the commutator-DIIS (C-DIIS) method (Pulay 1982), which defines the residual as the commutator of the Hamiltonian operator and the density matrix. This is the most widely used method in quantum chemistry software packages for achieving self-consistency. On the other hand, the C-DIIS method requires explicit storage of the density matrix and the Hamiltonian matrix for a few iterations, and hence is only feasible for calculations with small basis sets.

In the C-DIIS method, the residual, now written as R, is the commutator between $H[P]$ and P, that is,

$$R[P] = H[P]P - PH[P]. \tag{7.24}$$

Note that the Hamiltonian matrix $H[P]$ and the effective potential $V_{\text{eff}}[P]$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

contain the same information. Let H_k denote the approximate Hamiltonian produced at step k. We define a new Hamiltonian $\tilde{H}_{k+1}$ at step $(k+1)$ as a linear combination of previous approximations to the Hamiltonian, that is,

$$\tilde{H}_{k+1} = \sum_{j=k-\ell-1}^{k} \alpha_j H_j, \tag{7.25}$$

where α_j satisfies the constraint $\sum_{j=k-\ell-1}^{k} \alpha_j = 1$. Each Hamiltonian matrix H_j defines a density matrix P_j via the generalized Kohn–Sham map. Before self-consistency is reached, the residual $R_j = R[P_j]$ defined by equation (7.24) is non-zero. However, when all the Hamiltonian matrices $\{H_j\}$ are close to the self-consistent Hamiltonian operator $H_\star$, it is reasonable to expect the residual associated with $\tilde{H}_{k+1}$ to be well approximated by $\tilde{R}_{k+1} \equiv \sum_{j=k-\ell-1}^{k} \alpha_j R_j$. The C-DIIS method determines $\{\alpha_j\}$ by minimizing $\tilde{R}_{k+1}$, that is, we solve the following constrained minimization problem in each C-DIIS iteration:

$$\begin{aligned} &\min_{\{\alpha_j\}} \left\| \sum_{j=k-\ell-1}^{k} \alpha_j R_j \right\|_F^2 \\ &\text{subject to} \sum_{j=k-\ell-1}^{k} \alpha_j = 1. \end{aligned} \tag{7.26}$$

Here the Frobenius norm is defined as $\|A\|_F^2 = \mathrm{Tr}[A^* A]$.

As in the comparison between Anderson's method and Pulay's method above, the constraint in (7.26) can be eliminated by rewriting equation (7.25) as

$$\tilde{H}_{k+1} = H_k + \sum_{j=k-\ell}^{k} \beta_j (H_{j-1} - H_j). \tag{7.27}$$

Define $Y_j = R_{j-1} - R_j$ for $k-\ell+1 \le j \le k$. Then the constraint minimization problem (7.26) becomes an unconstrained minimization problem:

$$\min_{\{\beta_j\}} \left\| R_k + \sum_{j=k-\ell+1}^{k} \beta_j Y_j \right\|_F^2. \tag{7.28}$$

As a result, equation (7.28) has an analytic solution,

$$\beta = -M^{-1} b, \tag{7.29}$$

where the $\ell \times \ell$ matrix M and the vector b are defined as

$$M_{ij} = \mathrm{Tr}[Y_i^* Y_j], \quad b_j = \mathrm{Tr}[Y_j^* R_k], \tag{7.30}$$

respectively.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

7.3. Preconditioning techniques

From the analysis of the simple mixing method we observe that the convergence rate is determined by the condition number of $\epsilon_d = \mathcal{J}_\star$. Hence we should examine the dependence of $\kappa(\mathcal{J}_\star)$ with respect to the system size. Here we mainly examine periodic systems, of which the system size is often characterized by the number of unit cells in the computational domain. To simplify our discussion, we assume the unit cell to be a simple cubic cell with a lattice constant L. For non-periodic systems such as molecules, we can construct a fictitious (cubic) supercell that encloses the molecule and periodically extend the supercell so that properties of the system can be analysed via Fourier analysis. In both cases, we assume the number of atoms in each supercell is proportional to L^3.

When $\mathcal{J}_\star$ is similar to a positive definite matrix, the smallest eigenvalue $\lambda_{\min}$ is often independent of the system size, but the largest eigenvalue $\lambda_{\max}$ may depend sensitively on the system size. For instance, for the jellium system (or uniform electron gas), $\mathcal{J}_\star$ becomes a diagonal matrix in the Fourier space, and the eigenvalues can be evaluated explicitly as (Ziman 1979)

$$\lambda_{\mathbf{q}} = 1 + \left(\frac{4\pi}{q^2} + \kappa^2\right)\gamma F_L(q), \quad \mathbf{q} \in \mathbb{R}^3, \tag{7.31}$$

where $q = |\mathbf{q}|$, γ, κ are constants, and $F_L(q)$ is known as the *Lindhard* response function, which satisfies

$$\lim_{q\to 0} F_L(q) = 1, \quad \lim_{q\to\infty} F_L(q) = 0. \tag{7.32}$$

The above formulation is formally valid when the system size is infinite. Note that the smallest possible value q is $2\pi/L$, and then $\lambda_{\max}$ is bounded by $1 + \gamma(L^2/\pi + \kappa^2)$. As a result, the convergence of SCF iteration schemes becomes slower as the system size increases. If the mixing parameters are not adjusted properly, the long wavelength modes (corresponding to small q) tend to be amplified as the SCF iteration proceeds. This phenomenon is called 'charge sloshing'.

In order to accelerate the convergence, one can replace the scalar α in the simple mixing method with a matrix C_0, and the convergence rate is then determined by the condition number of the matrix $C_0\mathcal{J}_\star$. As discussed in Section 7.2, the choice of the matrix C_0 in Anderson's method plays the role of a preconditioner. The ideal preconditioner is $C_0 = \mathcal{J}_\star^{-1}$ as $\kappa(C_0\mathcal{J}_\star) = 1$. This corresponds to Newton's method. To overcome charge sloshing, a widely used preconditioner is the Kerker preconditioner (Kerker 1981), which is much easier to apply. It assigns a smaller weight to the long wavelength Fourier modes in order to attenuate charge sloshing. More specifically, the Kerker preconditioner is a diagonal matrix in the Fourier

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

space, and its eigenvalues are given by

$$\lambda_{\mathbf{q}}^{\text{Kerker}} = 1 + \frac{4\pi\widetilde{\gamma}}{q^2}, \tag{7.33}$$

where $\widetilde{\gamma} > 0$ is an adjustable parameter. For the uniform electron gas, the matrix $C_0\mathcal{J}_\star$ is a diagonal matrix in the Fourier space, and its eigenvalues are

$$\widetilde{\lambda}_{\mathbf{q}} = \frac{q^2 + \gamma F_L(q)(4\pi + \kappa^2 q^2)}{q^2 + 4\pi\widetilde{\gamma}}. \tag{7.34}$$

The eigenvalue $\widetilde{\lambda}_{\mathbf{q}}$ becomes approximately 1 for large q and $F_L(q)\gamma/\widetilde{\gamma}$ for small q. Hence, after preconditioning, the eigenvalues can be bounded by constants independent of the system size. In practice it is found that the Kerker preconditioner is an effective preconditioner for simple metals such as sodium (Na) or aluminium (Al).

However, the Kerker preconditioner is not an appropriate preconditioner for insulating systems. Although in general the Jacobian associated with the insulating system cannot be diagonalized by the Fourier basis, it can be shown that $\mathrm{e}^{\mathbf{i}\mathbf{q}\cdot\mathbf{r}}$ is an approximate eigenfunction of the independent particle polarizability operator χ_0 with the corresponding eigenvalue $-\xi q^2$ (Pick, Cohen and Martin 1970, Ghosez, Gonze and Godby 1997) for small q, where $\xi > 0$ is a constant factor. If we neglect the contribution from the exchange-correlation kernel, $\mathrm{e}^{\mathbf{i}\mathbf{q}\cdot\mathbf{r}}$ is also an approximate eigenfunction of $\mathcal{J}_\star$ with the corresponding eigenvalue $1 + (4\pi/q^2)q^2\xi = 1 + 4\pi\xi$ for small q. If C_0 is chosen to be the Kerker preconditioner, then the corresponding eigenvalue of $C_0\mathcal{J}_\star$ for small q is approximately

$$\widetilde{\lambda}_{\mathbf{q}} = \frac{q^2}{q^2 + 4\pi\widetilde{\gamma}}(1 + 4\pi\xi).$$

As the system size L increases, the smallest eigenvalue $\widetilde{\lambda}_{\mathbf{q}} \sim L^{-2}$, and thus the condition number $\kappa(C_0\mathcal{J}_\star)$ increases as the system size increases. For simple insulating systems, one good preconditioner is simply a constant, as the convergence of the simple mixing method is already independent of the system size.

As in the above discussion, simple insulating and metallic systems call for different types of preconditioners to accelerate the convergence of a fixed-point iteration for solving the Kohn–Sham problem. A natural question one may ask is how we should construct a preconditioner for a complex material that may contain both insulating and metallic components or metal surfaces. This leads to the elliptic preconditioner (Lin and Yang 2013) as one promising strategy.

Note that under RPA (neglecting the exchange-correlation kernel), we may approximate $\mathcal{J}_\star^{-1}$ as

$$\widetilde{\mathcal{J}}_\star^{-1} = (v_C^{-1} - \chi_0)^{-1} v_C^{-1}.$$

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Since the Coulomb kernel $v_C^{-1} = -\Delta/(4\pi)$, applying $\widetilde{\mathcal{J}}_\star^{-1}$ to a residual vector r_k simply amounts to solving the equation

$$(-\Delta - 4\pi\chi_0)\tilde{r}_k = -\Delta r_k. \tag{7.35}$$

To construct a preconditioner, we will replace χ_0 with a simpler operator. In many cases, we can choose the approximation to be a local (diagonal) operator defined by a function $b(\mathbf{r})$, although other types of more sophisticated operators are possible. To compensate for the simplification of χ_0, we replace the Laplacian operator on the left of (7.35) by $-\nabla \cdot (a(\mathbf{r})\nabla)$ for some appropriately chosen function $a(\mathbf{r})$. This additional change yields the following elliptic PDE:

$$(-\nabla \cdot (a(\mathbf{r})\nabla) + 4\pi b(\mathbf{r}))\tilde{r}_k = -\Delta r_k. \tag{7.36}$$

Because our construction of the preconditioner involves solving an elliptic equation, equation (7.36) is called an elliptic preconditioner.

The elliptic preconditioner is naturally compatible with previous preconditioners for simple metals or insulators. For example, for uniform electron gas, setting $a(\mathbf{r}) = 1$ and $b(\mathbf{r}) = -\widetilde{\gamma}$ for some constant $\widetilde{\gamma} > 0$ yields

$$(-\Delta + 4\pi\widetilde{\gamma})\tilde{r}_k = -\Delta r_k. \tag{7.37}$$

The solution of the above equation is exactly the same as that produced by the Kerker preconditioner. For simple insulating system, setting $a(\mathbf{r}) = \alpha^{-1}$ and $b(\mathbf{r}) = 0$ yields

$$-\alpha^{-1}\Delta\tilde{r}_k = -\Delta r_k.$$

The solution to the above equation is simply

$$\tilde{r}_k = \alpha r_k. \tag{7.38}$$

Such a solution corresponds to simple mixing with a mixing parameter α.

For a complex system that consists of both insulating and metallic components, it is desirable to choose approximations to $a(\mathbf{r})$ and $b(\mathbf{r})$ that are spatially dependent. The asymptotic behaviour of χ with respect to the sizes of both insulating and metallic systems suggests that $a(\mathbf{r})$ and $b(\mathbf{r})$ should be chosen to satisfy $a(\mathbf{r}) \geq 1$ and $b(\mathbf{r}) \geq 0$.

The implementation of the elliptic preconditioner only requires solving an elliptic equation, for instance through an iterative linear solver, such as the conjugate gradient method. In particular, fast algorithms such as multigrid (Brandt 1977), the fast multipole method (FMM) (Greengard and Rokhlin 1987), hierarchical matrix solvers (Hackbusch 1999) and hierarchical semi-separable (HSS) matrix solvers (Chandrasekaran, Gu and Pals 2006) can be used to solve (7.36) in $O(N)$ arithmetic operations.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 13: Two-level nested SCF method for solving Hartree–Fock-like equations

```
while Exchange energy is not converged do
    while Electron density $\rho$ is not converged do
        Solve the linear eigenvalue problem $H\psi_i = \varepsilon_i\psi_i$ with any
          iterative eigensolver.
        Update $\rho^{\text{out}}(\mathbf{r}) \leftarrow \sum_{i=1}^{N}|\psi_i(\mathbf{r})|^2$.
        Update $\rho$ using $\rho^{\text{out}}$ and charge densities computed and saved
          from previous iterations using a charge mixing scheme.
    end while
    Compute the exchange energy.
    Update $\{\varphi_i\}_{i=1}^{N} \leftarrow \{\psi_i\}_{i=1}^{N}$.
end while
```

7.4. Rung-4 functionals

As discussed in Section 7.2, the C-DIIS method is the most widely used method for Kohn–Sham DFT calculations with non-local functionals using a small basis set, such as Gaussian-type orbitals and atomic orbitals. However, it requires the explicit storage of the density matrix and the Hamiltonian matrix, and thus cannot be used in the setting of a large basis set such as planewaves. On the other hand, one cannot simply take the output Kohn–Sham orbitals from one SCF iteration and use them as the input Kohn–Sham orbitals for the next SCF iteration. As analysed in Section 7.1, such a fixed-point iteration is vulnerable to large eigenvalues of the Jacobian matrix and can suffer from the 'charge sloshing' problem. In practice, the most commonly used method for converging Kohn–Sham DFT calculations with a large basis set is a two-level nested SCF procedure. This is implemented, for instance, in the Quantum ESPRESSO software package. The motivation for using a two-level SCF procedure is to apply advanced charge mixing schemes to the electron density in the inner iteration to mitigate charge sloshing, and to use a fixed-point iteration to update the Kohn–Sham orbitals and consequently the exchange potential in the outer iteration. The update of the exchange potential is more costly, even though its contribution to total energy is typically much smaller.

The two-level nested SCF method is summarized in Algorithm 13. In each outer iteration, the exchange operator $V_X[P]$ is updated. This is implicitly done by updating a set of orbitals $\{\varphi_i\}_{i=1}^{N}$ defining the density matrix as $P = \sum_{i=1}^{N} \varphi_i\varphi_i^*$. We remark that this set of orbitals may be different from the Kohn–Sham orbitals in the inner SCF iteration. The update is done through a fixed-point iteration, that is, $\{\varphi_i\}_{i=1}^{N}$ are given by the output Kohn–Sham orbitals in the previous outer iteration. In the inner SCF

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

iteration, with the exchange operator fixed, the Hamiltonian H depends only on the electron density ρ. Charge mixing schemes for ρ can be performed in a similar fashion to what is done in a standard Kohn–Sham DFT calculation without the exchange operator in the inner SCF iteration. Finally, within each inner iteration, with both P and ρ fixed, the Hartree–Fock-like equation becomes a linear eigenvalue problem and can be solved by an iterative eigensolver discussed in Section 4.2. The outer SCF iteration continues until convergence is reached, which can be monitored using the change in the exchange energy, for example.

Recently, a new method, called the projected commutator DIIS (PC-DIIS) method, has been introduced so that the C-DIIS method can be used for Kohn–Sham DFT calculations with non-local functionals performed with a large basis set (Hu, Lin and Yang 2017*b*). Since it is not possible to explicitly store or mix the density matrices in such calculations, it is tempting to perform the DIIS procedure on the $N_b \times N$ orbital matrix $\Psi = [\psi_1, \ldots, \psi_N]$. However, one key difference between the density matrix and the orbital matrix is that the former is gauge-invariant. That is, if we replace Ψ with ΨU, where U is an $N \times N$ unitary gauge matrix, the density matrix $P = \Psi\Psi^* = \Psi U U^* \Psi^*$ does not change. Therefore, it is completely safe to combine two density matrices constructed from Ψ that differ by a gauge transformation, as the total energy is gauge-invariant.

However, since the orbital matrix Ψ is not gauge-invariant, combining successive approximations to Ψ that differ by a gauge transformation may hinder the stability of the SCF iteration. To overcome this difficulty, we introduce an auxiliary orbital matrix Φ that spans the same subspace spanned by Ψ. This orbital matrix is obtained by applying the orthogonal projection operator associated with Ψ to a reference orbital matrix Φ_{ref} to be specified later. That is, Φ is chosen to be

$$\Phi = P\Phi_{\text{ref}} = \Psi(\Psi^*\Phi_{\text{ref}}). \tag{7.39}$$

We require Φ_{ref} to be fixed throughout the entire SCF procedure. Note that Φ is invariant to any gauge transformation applied to Ψ. Therefore, the auxiliary orbital matrices obtained in successive SCF iterations can be safely combined to produce a better approximation to the desired invariant subspace. The columns of Φ are generally not orthogonal to each other. However, as long as the columns of Φ are not linearly dependent, both Ψ and Φ span the range of the density matrix P, which can also be written as

$$P = \Phi(\Phi^*\Phi)^{-1}\Phi^*. \tag{7.40}$$

The PC-DIIS method constructs a new approximation to Φ in the kth SCF iteration by taking a linear combination of the auxiliary orbital matrices

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

$\Phi_{k-\ell}, \dots, \Phi_k$ obtained in the most recent $\ell+1$ iterations, that is,

$$\tilde{\Phi}_{k+1} = \sum_{j=k-\ell}^{k} \alpha_j \Phi_j. \tag{7.41}$$

The coefficients $\{\alpha_j\}$ in (7.41) are determined by minimizing the residual associated with $\tilde{\Phi}_{k+1}$, which, under the assumption that Φ_j are sufficiently close to the solution of the Kohn–Sham equations, is well approximated by $R \equiv \sum_{j=k-\ell}^{k} \alpha_j R_{\Phi_j}$, where the residual associated with an auxiliary orbital matrix Φ is defined by

$$R_\Phi = H[P]P\Phi_{\mathrm{ref}} - PH[P]\Phi_{\mathrm{ref}} = (H[P]\Psi)(\Psi^*\Phi_{\mathrm{ref}}) - \Psi((H[P]\Psi)^*\Phi_{\mathrm{ref}}). \tag{7.42}$$

Note that evaluation of the residual in equation (7.42) only requires multiplying $H[P]$ by Ψ and the multiplications of matrices of sizes $N_b \times N$ and $N \times N$ only. These operations are already used in iterative methods for computing the desired eigenvectors Ψ of H. The PC-DIIS algorithm does not require $P, H[P]$ or R to be constructed or stored explicitly.

An interesting observation is that if $\Phi_{\mathrm{ref}} = \Psi$ then $\Psi^* H[P]\Psi$ is a diagonal matrix denoted by Λ. Consequently, the projected commutator takes the form

$$R_\Phi = H[P]\Psi - \Psi\Lambda. \tag{7.43}$$

This expression coincides with the standard definition of the residual associated with an approximate eigenpair (Λ, Ψ). Hence the PC-DIIS method can also be viewed as an extension of an iterative eigensolver for nonlinear problems.

As in the reformulation of the constrained minimization problem into an unconstrained minimization problem in the C-DIIS method, the constrained minimization problem

$$\begin{gathered} \min_{\{\alpha_j\}} \left\| \sum_{j=k-\ell}^{k} \alpha_j R_{\Phi_j} \right\|_F^2 \\ \text{subject to } \sum_{j=k-\ell}^{k} \alpha_j = 1, \end{gathered} \tag{7.44}$$

to be solved in the PC-DIIS method can also be reformulated as an unconstrained minimization problem. Using the same change of variable as that presented in Section 7.2, we can write

$$\tilde{\Phi}_{k+1} = \Phi_k + \sum_{j=k-\ell+1}^{k} \beta_j (\Phi_{j-1} - \Phi_j). \tag{7.45}$$

If we let $Y_{\Phi_j} = R_{\Phi_{j-1}} - R_{\Phi_j}$, the coefficients β_j in (7.45) can be retrieved

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Algorithm 14: The PC-DIIS method for solving Hartree–Fock-like equations

```
Input: Reference orbitals $\Phi_{\mathrm{ref}}$.
Output: Approximate solution $\Psi = \{\psi_i\}$, $i = 1, 2 \ldots, N$.
Construct the initial Hamiltonian $H$ and evaluate the exchange energy using $\Phi_{\mathrm{ref}}$.
while Exchange energy is not converged do
    Solve the linear eigenvalue problem $H[P]\psi_i = \varepsilon_i \psi_i$ using an iterative eigensolver.
    Evaluate $\Phi, R_\Phi$ according to (7.39), (7.42).
    Perform the DIIS procedure according to (7.45) to obtain the new $\tilde{\Phi}$ which implicitly defines a density matrix $P$ via (7.40).
    Update the Hamiltonian $H[P]$.
    Compute the exchange energy.
end while
```

from the vector $\beta = -(M^\Phi)^{-1} b^\Phi$, where

$$M_{ij}^\Phi = \mathrm{Tr}[Y_{\Phi_i}^* Y_{\Phi_j}], \quad b_j^\Phi = \mathrm{Tr}[Y_{\Phi_j}^* R_{\Phi_k}]. \tag{7.46}$$

Once $\tilde{\Phi}_{k+1}$ is obtained, a density matrix associated with this orbital matrix is implicitly defined through equation (7.40). This implicitly defined density matrix allows us to construct a new Hamiltonian from which a new set of Kohn–Sham orbitals Ψ_{k+1} and auxiliary orbitals Φ_{k+1} can be computed.

We now discuss how to choose the gauge-fixing matrix Φ_{ref}. Note that in hybrid functional calculations, the contribution from the exchange operator is relatively small. Hence the density matrix associated with Kohn–Sham orbitals obtained from a DFT calculation that uses a local or semi-local exchange-correlation functional is already a good initial guess for the density matrix required in a hybrid functional calculation. Therefore, we may use these orbitals as Φ_{ref}. Compared to the two-level nested loop structure, the PC-DIIS method only requires one level of SCF iteration. The PC-DIIS method is summarized in Algorithm 14.

The discussion above is applicable when Ψ only contains the occupied orbitals. When Ψ also involves the virtual orbitals, we use the fact that the density matrix defining the Fock exchange operator only involves the occupied orbitals, and we need only apply the PC-DIIS method to the occupied orbitals. We also remark that the PC-DIIS method is not yet applicable for finite temperature calculations with fractionally occupied orbitals.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

8. Conclusion and future directions

In this paper, we have reviewed some basic aspects and recent developments of numerical strategies for solving Kohn–Sham DFT. Most of the numerical methods focus on large systems, either due to the presence of a large number of electrons (thousands to tens of thousands), or due to the use of a large basis set (such as the planewave basis set). In this sense, the numerical algorithms in this paper are more suitable for applications in quantum physics and materials science, which often favour a large basis set and relatively large system sizes. On the other hand, the quantum chemistry literature often favours a small basis set (such as Gaussian-type orbitals) and relatively small system sizes, partly due to their focus on post-Hartree–Fock and post-DFT methods, which can be much more accurate but also more costly than DFT calculations. Even so, we remark that there has been growing interest in large system sizes and large basis sets in quantum chemistry (Sun, Berkelbach, McClain and Chan 2017, Stoudenmire and White 2017, Mardirossian, McClain and Chan 2018), and the methods or ideas reviewed in this paper may then become applicable as well.

Although many advanced numerical methods have been implemented in mature DFT software packages, there are still many outstanding challenges for solving Kohn–Sham DFT. Below we provide our own perspectives organized according to the contents of this review.

(1) *Numerical discretization.* The pseudopotential approximation greatly facilitated the efficient discretization of the Kohn–Sham Hamiltonian. Nonetheless, it is not a systematic approximation, and is regarded by many as one of the 'artistic' components for solving Kohn–Sham DFT. Efficient discretizations that allow efficient all-electron calculations to be performed accurately, such as the numerical atomic basis sets, wavelets, and recent development of the Gausslet basis set (White 2017), are still very much of interest. A key challenge would be to effectively control the number of degrees of freedom, so that calculations can still be performed efficiently for large systems.

(2) *Evaluation of the Kohn–Sham map with semi-local functionals.* The cross-over point of linear scaling methods and reduced scaling methods over conventional cubic scaling methods is still large, especially for three-dimensional bulk systems. Note that some key factors affecting the cross-over point, such as the decay rate of the density matrix, are determined by physical systems under consideration rather than numerical algorithms. Hence one cannot realistically expect the cross-over point to be reduced to below several hundreds of atoms in general. Nonetheless, many physical processes, such as the battery degradation process, require large simulation of systems with thousands to tens

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

of thousands of atoms. Therefore further improvement of numerical algorithms and their parallel scalability, particularly for metallic systems, may enable a wide range of applications beyond reach today.

(3) *Evaluation of the Kohn–Sham map with non-local functionals.* The applicability range of Kohn–Sham DFT is ultimately determined by the choice of exchange-correlation functionals, and an increasing number of numerical calculations are now performed using rung-4 and rung-5 functionals. In some sense, the precise form of these functionals is still under development and somewhat debated in the literature. For instance, the self-consistently screened hybrid functional (Brawand, Vörös, Govoni and Galli 2016) was only developed in recent years, and there is as yet no consensus on the formulation of self-consistent rung-5 functionals. In particular, rung-5 functionals are closely related to Green's function methods such as those in the many-body perturbation theory (MBPT) (Hedin 1965, Onida, Reining and Rubio 2002), which are often treated as post-DFT methods. For strongly correlated systems, novel exchange-correlation functionals such as those based on the strictly correlated electron (SCE) limit (Seidl, Gori-Giorgi and Savin 2007, Malet and Gori-Giorgi 2012) offer new perspectives on the design of exchange-correlation functionals, though they have yet to be shown to be effective for practically relevant chemical systems.

(4) *Self-consistent field iterations.* For large-scale heterogeneous systems with a small or zero energy gap, the SCF iteration often converges slowly, or fails to converge at all. This is particularly problematic for geometry optimization and molecular dynamics simulation, where a large number of iterations need to be performed at each step. It remains challenging to design efficient and robust approaches for self-consistency for large-scale systems. From this perspective, optimization based algorithms with guaranteed reduction of energy at each step could be more effective and robust.

(5) *Numerical analysis.* Due to the nonlinearity of the exchange-correlation functional, numerical analysis of Kohn–Sham DFT can be very challenging. Much progress has been made for simplified models such as the Thomas–Fermi-type models and reduced Hartree–Fock models (see *e.g.* Cancès, Deleurence and Lewin 2008, Cancès and Lewin 2010, Cancès and Mourad 2014). Even so, there are still many basic questions from numerical analysis that remain to be answered, such as the convergence analysis of self-consistent field iterations beyond the linear response regime.

Finally, we would like to emphasize again the fact that although 'Kohn–Sham DFT' has become synonymous with 'electronic structure theory' in

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

many contexts, it is only *one* of the many branches of electronic structure theories. We have not touched upon most of the wavefunction methods in the quantum chemistry literature, nor the methods for excited states or time-dependent problems. The fact that Kohn–Sham DFT has become so widely used today is because it strikes the right balance between efficiency, accuracy and general applicability *so far*. This balance has been constantly reshaped together with the development of many other theories, and will continue to be further reshaped in the future. Let us mention two possibilities.

Recently, there has been rapid surge of interest in applying machine learning tools to electronic structure theories. In the past few years there has been considerable progress in building inter-atomic potentials, for which the quality can be comparable to a first-principles simulation (see *e.g.* Bartók, Payne and Csányi 2010, Zhang *et al.* 2018). Although machine learning based methods will still require electronic structure calculations to generate the training data, it is possible that in future we will not need to perform a monolithic, long-time first-principles molecular dynamics simulation for a large system. Instead it may be efficient enough to perform long molecular dynamics calculations for many small- to medium-sized systems in parallel, or for large systems but at many discontinuous snapshots. This may enable DFT calculations on a massively parallel scale, and many large systems beyond reach today might then become feasible. It may also be possible to use machine learning tools to build better density functionals or to more efficiently solve DFT. The interaction between physical modelling, data and numerical algorithms promises to further improve the predictive power of electronic structure theories for materials and chemical systems.

Another possibility is the fusion of DFT methods with other physical theories. In the quantum physics and chemistry literature, such multiscale-like theories are called ‘quantum embedding theories’. Kohn–Sham DFT can be combined with more coarse-grained theories such as molecular mechanics to simulate large-scale systems such as biological systems. These ‘quantum mechanics / molecular mechanics’ (QM/MM) methods began in the 1970s (Warshel and Levitt 1976) and were recognized by the Nobel Prize in Chemistry in 2013. Kohn–Sham DFT can also be combined with more accurate post-DFT methods, such as coupled cluster methods, density matrix renormalization group methods, quantum Monte Carlo methods and exact diagonalization methods (Sun and Chan 2016, Georges, Kotliar, Krauth and Rozenberg 1996, Kotliar *et al.* 2006, Knizia and Chan 2012, Manby, Stella, Goodpaster and Miller III 2012) to solve strongly correlated systems. As of today, these methods certainly still lack the ease of use and general applicability compared to Kohn–Sham DFT. There are challenges as well as opportunities for applied mathematicians to make contributions.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Notation

General conventions	
i	imaginary unit
z^*	complex conjugate of the complex number z
N	number of electrons
M	number of nuclei
β	inverse temperature
$\oint_{\mathscr{C}} \mathrm{d}\lambda$	contour integral
$\langle\psi\|$, $\|\psi\rangle$, $\langle\psi\|\varphi\rangle$	bra vector, ket vector and bra–ket in Dirac notation
Coordinates	
$\mathbf{r}$, r_α	single electron spatial coordinate and its Cartesian components, $\alpha = x, y, z$ or $1, 2, 3$
$\mathbf{p}$, p_α	single electron momentum coordinate and its Cartesian components
$\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$	space-spin coordinates of the ith electron
Z_I	charge of the Ith nuclei
$\mathbf{R}_I$	spatial coordinate of the Ith nuclei
DFT-related	
Ψ or $\|\Psi\rangle$	N-electron wavefunction
P	single-particle density matrix
$\rho(\mathbf{r})$	electron density
$\psi_i(\mathbf{r})$ or $\varphi_i(\mathbf{r})$	ith single electron spatial orbital
ε_i	eigenvalue of the ith orbital
f_i	occupation number for the ith orbital
i, j	occupied eigenvalue index
a, b	unoccupied eigenvalue index
μ	chemical potential
$V(\mathbf{r})$	single-particle potential
$V_{\mathrm{ext}}(\mathbf{r})$	external potential
$V_{\mathrm{H}}, V_{\mathrm{x}}, V_{\mathrm{c}}, V_{\mathrm{xc}}, V_{\mathrm{Hxc}}$	Hartree, exchange, correlation, exchange-correlation and Hartree-exchange-correlation potentials
$\mathcal{F}_{\mathrm{KS}}[V]$	Kohn–Sham map from potential to density
Notation for matrix representation	
$A^\top$	transpose of A
A^* or $A^\dagger$	Hermitian transpose / adjoint of A
N_g	number of grid points / degrees of freedom
N_b	number of basis functions

Continued on next page

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Continued from previous page

$\Psi = [\psi_1, \ldots, \psi_N]$	a matrix collecting N single-particle orbitals
$\Phi = [\phi_1, \ldots, \phi_{N_b}]$	a matrix collecting N_b basis functions, usually of size $N_g \times N_b$
H, S	discretized Hamiltonian and overlap matrices
G	discretized Green's function
I	identity matrix
Other quantities	
$\mathbb{1}_{(-\infty,0)}$	indicator function
f_β	finite temperature Fermi–Dirac function
f_∞	zero temperature Fermi–Dirac function, the same as an indicator function
η	a small positive quantity approaching 0

REFERENCES[2]

H. M. Aktulga, L. Lin, C. Haine, E. G. Ng and C. Yang (2014), 'Parallel eigenvalue calculation based on multiple shift–invert Lanczos and contour integral based spectral projection method', *Parallel Comput.* **40**, 195–212.

P. Amestoy, I. Duff, J.-Y. L'Excellent and J. Koster (2001), 'A fully asynchronous multifrontal solver using distributed dynamic scheduling', *SIAM J. Matrix Anal. Appl.* **23**, 15–41.

O. K. Andersen (1975), 'Linear methods in band theory', *Phys. Rev.* B **12**, 3060–3083.

D. G. Anderson (1965), 'Iterative procedures for nonlinear integral equations', *J. Assoc. Comput. Mach.* **12**, 547–560.

E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen (1999), *LAPACK Users' Guide*, third edition, SIAM.

D. N. Arnold (1982), 'An interior penalty finite element method with discontinuous elements', *SIAM J. Numer. Anal.* **19**, 742–760.

D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini (2002), 'Unified analysis of discontinuous Galerkin methods for elliptic problems', *SIAM J. Numer. Anal.* **39**, 1749–1779.

C. Ashcraft and R. Grimes (1989), 'The influence of relaxed supernode partitions on the multifrontal method', *ACM Trans. Math. Software* **15**, 291–309.

I. Babuška and M. Zlámal (1973), 'Nonconforming elements in the finite element method with penalty', *SIAM J. Numer. Anal.* **10**, 863–875.

A. S. Banerjee, L. Lin, P. Suryanarayana, C. Yang and J. E. Pask (2018), 'Two-level Chebyshev filter based complementary subspace method for pushing the

[2] The URLs cited in this work were correct at the time of going to press, but the publisher and the authors make no undertaking that the citations remain live or are accurate or appropriate.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

envelope of large-scale electronic structure calculations', *J. Chem. Theory Comput.* **14**, 2930–2946.

G. Bao, G. Hu and D. Liu (2012), 'An h-adaptive finite element solver for the calculations of the electronic structures', *J. Comput. Phys.* **231**, 4967–4979.

S. Baroni and P. Giannozzi (1992), 'Towards very large-scale electronic-structure calculations', *Europhys. Lett.* **17**, 547–552.

M. Barrault, E. Cancès, W. Hager and C. Le Bris (2007), 'Multilevel domain decomposition for electronic structure calculations', *J. Comput. Phys.* **222**, 86–109.

A. P. Bartók, M. C. Payne and G. Csányi (2010), 'Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons', *Phys. Rev. Lett.* **104**, 1–4.

A. D. Becke (1988), 'Density-functional exchange-energy approximation with correct asymptotic behavior', *Phys. Rev.* A **38**, 3098–3100.

A. D. Becke (1993), 'Density functional thermochemistry, III: The role of exact exchange', *J. Chem. Phys.* **98**, 5648–5652.

L. Belpassi, F. Tarantelli, A. Sgamellotti and H. M. Quiney (2005), 'Computational strategies for a four-component Dirac–Kohn–Sham program: Implementation and first applications', *J. Chem. Phys.* **122**, 184109.

G. Bencteux, M. Barrault, E. Cancès, W. W. Hager and C. Le Bris (2008), Domain decomposition and electronic structure computations: A promising approach. In *Numerical Analysis and Scientific Computing for PDEs and Their Challenging Applications* (R. Glowinski and P. Neittaanmäki, eds), Vol. 16 of Computational Methods in Applied Sciences, Springer, pp. 147–164.

M. Benzi, P. Boito and N. Razouk (2013), 'Decay properties of spectral projectors with applications to electronic structure', *SIAM Rev.* **55**, 3–64.

L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker and R. C. Whaley (1997), *ScaLAPACK Users' Guide*, SIAM.

P. E. Blöchl (1994), 'Projector augmented-wave method', *Phys. Rev.* B **50**, 17953–17979.

E. I. Blount (1962), 'Formalisms of band theory', *Solid State Phys.* **13**, 305–373.

V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler (2009), '*Ab initio* molecular simulations with numeric atom-centered orbitals', *Comput. Phys. Commun.* **180**, 2175–2196.

N. M. Boffi, M. Jain and A. Natan (2016), 'Efficient computation of the Hartree–Fock exchange in real-space with projection operators', *J. Chem. Theory Comput.* **12**, 3614–3622.

D. Bohm and D. Pines (1953), 'A collective description of electron interactions, III: Coulomb interactions in a degenerate electron gas', *Phys. Rev.* **92**, 609–625.

D. R. Bowler and T. Miyazaki (2012), '$O(N)$ methods in electronic structure calculations', *Rep. Prog. Phys.* **75**, 036503.

A. Brandt (1977), 'Multi-level adaptive solutions to boundary-value problems', *Math. Comp.* **31**, 333–390.

A. Brandt, S. McCormick and J. Ruge (1985), Algebraic multigrid (AMG) for sparse matrix equations. In *Sparsity and its Applications*, Cambridge University Press, pp. 257–284.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

N. P. Brawand, M. Vörös, M. Govoni and G. Galli (2016), 'Generalization of dielectric-dependent hybrid functionals to finite systems', *Phys. Rev.* X **6**, 041002.

W. Briggs, V. E. Henson and S. F. McCormick (2000), *A Multigrid Tutorial*, second edition, SIAM.

C. Brouder, G. Panati, M. Calandra, C. Mourougane and N. Marzari (2007), 'Exponential localization of Wannier functions in insulators', *Phys. Rev. Lett.* **98**, 046402.

K. Burke (2012), 'Perspective on density functional theory', *J. Chem. Phys.* **136**, 150901.

Y. Cai, Z. Bai, J. E. Pask and N. Sukumar (2013), 'Hybrid preconditioning for iterative diagonalization of ill-conditioned generalized eigenvalue problems in electronic structure calculations', *J. Comput. Phys.* **255**, 16–30.

E. Cancès and M. Lewin (2010), 'The dielectric permittivity of crystals in the reduced Hartree–Fock approximation', *Arch. Rational Mech. Anal.* **197**, 139–177.

E. Cancès and N. Mourad (2014), 'A mathematical perspective on density functional perturbation theory', *Nonlinearity* **27**, 1999.

E. Cancès and N. Mourad (2016), 'Existence of a type of optimal norm-conserving pseudopotentials for Kohn–Sham models', *Commun. Math. Sci.* **14**, 1315–1352.

E. Cancès, A. Deleurence and M. Lewin (2008), 'A new approach to the modeling of local defects in crystals: The reduced Hartree–Fock case', *Commun. Math. Phys.* **281**, 129–177.

E. Cancès, A. Levitt, G. Panati and G. Stoltz (2017), 'Robust determination of maximally localized Wannier functions', *Phys. Rev.* B **95**, 075114.

R. Car and M. Parrinello (1985), 'Unified approach for molecular dynamics and density-functional theory', *Phys. Rev. Lett.* **55**, 2471–2474.

D. M. Ceperley and B. J. Alder (1980), 'Ground state of the electron gas by a stochastic method', *Phys. Rev. Lett.* **45**, 566–569.

M. Ceriotti, T. Kühne and M. Parrinello (2008), 'An efficient and accurate decomposition of the Fermi operator', *J. Chem. Phys.* **129**, 024707.

S. Chandrasekaran, M. Gu and T. Pals (2006), 'A fast ULV decomposition solver for hierarchically semiseparable representations', *SIAM J. Matrix Anal. Appl.* **28**, 603–622.

J. Chelikowsky, N. Troullier and Y. Saad (1994), 'Finite-difference-pseudopotential method: Electronic structure calculations without a basis', *Phys. Rev. Lett.* **72**, 1240–1243.

G. P. Chen, V. K. Voora, M. M. Agee, S. G. Balasubramani and F. Furche (2017), 'Random-phase approximation method', *Ann. Rev. Phys. Chem.* **68**, 421–445.

H. Chen, X. Dai, X. Gong, L. He and A. Zhou (2014), 'Adaptive finite element approximations for Kohn–Sham models', *Multiscale Model. Simul.* **12**, 1828–1869.

J. Chen and J. Lu (2016), 'Analysis of the divide-and-conquer method for electronic structure calculations', *Math. Comp.* **85**, 2919–2938.

E. Chow, X. Liu, M. Smelyanskiy and J. R. Hammond (2015), 'Parallel scalability of Hartree–Fock calculations', *J. Chem. Phys.* **142**, 104103.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. C. Payne (2005), 'First principles methods using CASTEP', *Z. Kristallographie* **220**, 567–570.

B. Cockburn, G. Karniadakis and C.-W. Shu (2000), *Discontinuous Galerkin methods: Theory, Computation and Applications*, Vol. 11 of Lecture Notes in Computational Science and Engineering, Springer.

F. Corsetti (2014), 'The orbital minimization method for electronic structure calculations with finite-range atomic basis sets', *Comput. Phys. Commun.* **185**, 873–883.

A. Damle and L. Lin (2018), 'Disentanglement via entanglement: A unified method for Wannier localization', *Math. Model. Simul.* **16**, 1392–1410.

A. Damle, A. Levitt and L. Lin (2019), 'Variational formulation for Wannier functions with entangled band structure', *SIAM Multiscale Model. Simul.* **17**, 167–191.

A. Damle, L. Lin and L. Ying (2015), 'Compressed representation of Kohn–Sham orbitals via selected columns of the density matrix', *J. Chem. Theory Comput.* **11**, 1463–1469.

A. Damle, L. Lin and L. Ying (2017), 'SCDM-k: Localized orbitals for solids via selected columns of the density matrix', *J. Comput. Phys.* **334**, 1–15.

E. R. Davidson (1975), 'The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices', *J. Comput. Phys.* **17**, 87–94.

W. Dawson and F. Gygi (2015), 'Performance and accuracy of recursive subspace bisection for hybrid DFT calculations in inhomogeneous systems', *J. Chem. Theory Comput.* **11**, 4655–4663.

W. Dawson and T. Nakajima (2018), 'Massively parallel sparse matrix function calculations with NTPoly', *Comput. Phys. Commun.* **225**, 154–165.

M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist (2004), 'Van der Waals density functional for general geometries', *Phys. Rev. Lett.* **92**, 246401.

K. Dong, W. Hu and L. Lin (2018), 'Interpolative separable density fitting through centroidal Voronoi tessellation with applications to hybrid functional electronic structure calculations', *J. Chem. Theory Comput.* **14**, 1311–1320.

R. M. Dreizler and E. K. U. Gross (1990), *Density Functional Theory*, Springer.

I. Duchemin and F. Gygi (2010), 'A scalable and accurate algorithm for the computation of Hartree–Fock exchange', *Comput. Phys. Commun.* **181**, 855–860.

T. H. Dunning (1989), 'Gaussian basis sets for use in correlated molecular calculations, I: The atoms boron through neon and hydrogen', *J. Chem. Phys.* **90**, 1007–1023.

W. E and J. Lu (2011), 'The electronic structure of smoothly deformed crystals: Wannier functions and the Cauchy–Born rule', *Arch. Ration. Mech. Anal.* **199**, 407–433.

W. E, T. Li and J. Lu (2010), 'Localized bases of eigensubspaces and operator compression', *Proc. Nat. Acad. Sci.* **107**, 1273–1278.

A. Edelman, T. A. Arias and S. T. Smith (1998), 'The geometry of algorithms with orthogonality constraints', *SIAM J. Matrix Anal. Appl.* **20**, 303–353.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

A. Erisman and W. Tinney (1975), ‘On computing certain elements of the inverse of a sparse matrix’, *Comm. Assoc. Comput. Mach.* **18**, 177–179.

H. Eschrig (1996), *The Fundamentals of Density Functional Theory*, Springer.

H.-R. Fang and Y. Saad (2009), ‘Two classes of multisecant methods for nonlinear acceleration’, *Numer. Linear Algebra Appl.* **16**, 197–221.

J. L. Fattebert and J. Bernholc (2000), ‘Towards grid-based $O(N)$ density-functional theory methods: Optimized nonorthogonal orbitals and multigrid acceleration’, *Phys. Rev.* B **62**, 1713–1722.

E. Fermi (1927), ‘Un metodo statistico per la determinazione di alcune prioprietà dell’atomo’, *Rend. Accad. Naz. Lincei.* **6**, 602–607.

M. Feyereisen, G. Fitzgerald and A. Komornicki (1993), ‘Use of approximate integrals in *ab initio* theory: An application in MP2 energy calculations’, *Chem. Phys. Lett.* **208**, 359–363.

B. Fornberg (1998), *A Practical Guide to Pseudospectral Methods*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press.

J. M. Foster and S. F. Boys (1960), ‘Canonical configurational interaction procedure’, *Rev. Mod. Phys.* **32**, 300–302.

T. Fukazawa and H. Akai (2015), ‘Optimized effective potential method and application to static RPA correlation’, *J. Phys. Condens. Matter* **27**, 115502.

W. Gao and W. E (2009), ‘Orbital minimization with localization’, *Discrete Contin. Dyn. Syst.* **23**, 249–264.

C. J. Garcia-Cervera, J. Lu, Y. Xuan and W. E (2009), ‘Linear-scaling subspace-iteration algorithm with optimally localized nonorthogonal wave functions for Kohn–Sham density functional theory’, *Phys. Rev.* B **79**, 115110.

M. Gell-Mann and K. A. Brueckner (1957), ‘Correlation energy of an electron gas at high density’, *Phys. Rev.* **106**, 364–368.

L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman and R. Schneider (2008), ‘Daubechies wavelets as a basis set for density functional pseudopotential calculations’, *J. Chem. Phys.* **129**, 014109.

A. Georges, G. Kotliar, W. Krauth and M. J. Rozenberg (1996), ‘Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions’, *Rev. Mod. Phys.* **68**, 13–125.

P. Ghosez, X. Gonze and R. W. Godby (1997), ‘Long-wavelength behavior of the exchange-correlation kernel in the Kohn–Sham theory of periodic systems’, *Phys. Rev.* B **56**, 12811–12817.

P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni (2017), ‘Advanced capabilities for

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

materials modelling with QUANTUM ESPRESSO', *J. Phys. Condens. Matter* **29**, 465901.

R. W. Godby, M. Schlüter and L. J. Sham (1986), 'Accurate exchange-correlation potential for Silicon and its discontinuity on addition of an electron', *Phys. Rev. Lett.* **56**, 2415–2418.

R. W. Godby, M. Schlüter and L. J. Sham (1988), 'Self-energy operators and exchange-correlation potentials in semiconductors', *Phys. Rev.* B **37**, 10159–10175.

S. Goedecker (1999), 'Linear scaling electronic structure methods', *Rev. Mod. Phys.* **71**, 1085–1123.

S. Goedecker and L. Colombo (1994), 'Efficient linear scaling algorithm for tight-binding molecular dynamics', *Phys. Rev. Lett.* **73**, 122–125.

L. Goerigk and S. Grimme (2014), 'Double-hybrid density functionals', *WIREs Comput. Mol. Sci.* **4**, 576–600.

G. H. Golub and C. F. Van Loan (2013), *Matrix Computations*, fourth edition, Johns Hopkins University Press.

X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D. Hamann, L. He, G. Jomard, J. Laflamme Janssen, S. Le Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukačević, A. Martin, C. Martins, M. Oliveira, S. Poncé, Y. Pouillon, T. Rangel, G.-M. Rignanese, A. Romero, B. Rousseau, O. Rubel, A. Shukri, M. Stankovski, M. Torrent, M. Van Setten, B. Van Troeye, M. Verstraete, D. Waroquiers, J. Wiktor, B. Xu, A. Zhou and J. Zwanziger (2016), 'Recent developments in the ABINIT software package', *Comput. Phys. Commun.* **205**, 106–131.

L. Greengard and V. Rokhlin (1987), 'A fast algorithm for particle simulations', *J. Comput. Phys.* **73**, 325–348.

S. Grimme (2006), 'Semiempirical hybrid density functional with perturbative second-order correlation', *J. Chem. Phys.* **124**, 034108.

O. Gunnarsson and B. I. Lundqvist (1976), 'Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism', *Phys. Rev.* B **13**, 4274–4298.

F. Gygi (2008), 'Architecture of Qbox: A scalable first-principles molecular dynamics code', *IBM J. Res. Dev.* **52**, 137–144.

F. Gygi (2009), 'Compact representations of Kohn–Sham invariant subspaces', *Phys. Rev. Lett.* **102**, 166406.

W. Hackbusch (1999), 'A sparse matrix arithmetic based on $\mathcal{H}$-matrices, I: Introduction to $\mathcal{H}$-matrices', *Computing* **62**, 89–108.

D. R. Hamann (2013), 'Optimized norm-conserving Vanderbilt pseudopotentials', *Phys. Rev.* B **88**, 085117.

D. R. Hamann, M. Schlüter and C. Chiang (1979), 'Norm-conserving pseudopotentials', *Phys. Rev. Lett.* **43**, 1494–1497.

C. Hartwigsen, S. Goedecker and J. Hutter (1998), 'Relativistic separable dual-space Gaussian pseudopotentials from H to Rn', *Phys. Rev.* B **58**, 3641–3662.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

L. Hedin (1965), 'New method for calculating the one-particle Green's function with application to the electron-gas problem', *Phys. Rev.* A **139**, 796–823.

J. Heyd, G. E. Scuseria and M. Ernzerhof (2003), 'Hybrid functionals based on a screened Coulomb potential', *J. Chem. Phys.* **118**, 8207–8215.

N. Higham (2008), *Functions of Matrices: Theory and Computation*, SIAM.

P. Hohenberg and W. Kohn (1964), 'Inhomogeneous electron gas', *Phys. Rev.* B **136**, 864–871.

J. Hu, B. Jiang, L. Lin, Z. Wen and Y. Yuan (2018), Structured quasi-Newton methods for optimization with orthogonality constraints. arXiv:1809.00452

W. Hu, L. Lin and C. Yang (2015), 'DGDFT: A massively parallel method for large scale density functional theory calculations', *J. Chem. Phys.* **143**, 124110.

W. Hu, L. Lin and C. Yang (2017*a*), 'Interpolative separable density fitting decomposition for accelerating hybrid density functional calculations with applications to defects in silicon', *J. Chem. Theory Comput.* **13**, 5420–5431.

W. Hu, L. Lin and C. Yang (2017*b*), 'Projected commutator DIIS method for accelerating hybrid functional electronic structure calculations', *J. Chem. Theory Comput.* **13**, 5458–5467.

W. Hu, L. Lin, A. Banerjee, E. Vecharynski and C. Yang (2017*c*), 'Adaptively compressed exchange operator for large scale hybrid density functional calculations with applications to the adsorption of water on silicene', *J. Chem. Theory Comput.* **13**, 1188–1198.

M. Jacquelin, L. Lin and C. Yang (2016), 'PSelInv: A distributed memory parallel algorithm for selected inversion: The symmetric case', *ACM Trans. Math. Software* **43**, 21.

M. Jacquelin, L. Lin and C. Yang (2018), 'PSelInv: A distributed memory parallel algorithm for selected inversion: The non-symmetric case', *Parallel Comput.* **74**, 84–98.

F. Jensen (2013), 'Atomic orbital basis sets', *WIREs Comput. Mol. Sci.* **3**, 273–295.

W. Jia and L. Lin (2017), 'Robust determination of the chemical potential in the pole expansion and selected inversion method for solving Kohn–Sham density functional theory', *J. Chem. Phys.* **147**, 144107.

Y. Jin, D. Zhang, Z. Chen, N. Q. Su and W. Yang (2017), 'Generalized optimized effective potential for orbital functionals and self-consistent calculation of random phase approximation', *J. Phys. Chem. Lett.* **8**, 4746–4751.

M. Kaltak, J. Klimeš and G. Kresse (2014*a*), 'Cubic scaling algorithm for the random phase approximation: Self-interstitials and vacancies in Si', *Phys. Rev.* B **90**, 054115.

M. Kaltak, J. Klimeš and G. Kresse (2014*b*), 'Low scaling algorithms for the random phase approximation: Imaginary time and Laplace transformations', *J. Chem. Theory Comput.* **10**, 2498–2507.

E. Kaxiras (2003), *Atomic and Electronic Structure of Solids*, Cambridge University Press.

J. Kaye, L. Lin and C. Yang (2015), '*A posteriori* error estimator for adaptive local basis functions to solve Kohn–Sham density functional theory', *Commun. Math. Sci.* **13**, 1741–1773.

G. P. Kerker (1981), 'Efficient iteration scheme for self-consistent pseudopotential calculations', *Phys. Rev.* B **23**, 3082–3084.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

J. Kim, F. Mauri and G. Galli (1995), 'Total-energy global optimization using nonorthogonal localized orbitals', *Phys. Rev.* B **52**, 1640–1648.

S. Kivelson (1982), 'Wannier functions in one-dimensional disordered systems: Application to fractionally charged solitons', *Phys. Rev.* B **26**, 4269–4277.

L. Kleinman and D. M. Bylander (1982), 'Efficacious form for model pseudopotentials', *Phys. Rev. Lett.* **48**, 1425–1428.

G. Knizia and G. Chan (2012), 'Density matrix embedding: A simple alternative to dynamical mean-field theory', *Phys. Rev. Lett.* **109**, 186404.

A. V. Knyazev (2001), 'Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method', *SIAM J. Sci. Comput.* **23**, 517–541.

M. Kobayashi and H. Nakai (2009), 'Divide-and-conquer-based linear-scaling approach for traditional and renormalized coupled cluster methods with single, double, and noniterative triple excitations', *J. Chem. Phys.* **131**, 114108.

E. Koch and S. Goedecker (2001), 'Locality properties and Wannier functions for interacting systems', *Solid State Commun.* **119**, 105–109.

W. Kohn (1959), 'Analytic properties of Bloch waves and Wannier functions', *Phys. Rev.* **115**, 809–821.

W. Kohn (1996), 'Density functional and density matrix method scaling linearly with the number of atoms', *Phys. Rev. Lett.* **76**, 3168–3171.

W. Kohn and L. Sham (1965), 'Self-consistent equations including exchange and correlation effects', *Phys. Rev.* A **140**, 1133–1138.

G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet and C. A. Marianetti (2006), 'Electronic structure calculations with dynamical mean-field theory', *Rev. Mod. Phys.* **78**, 865–951.

G. Kresse and J. Furthmüller (1996), 'Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set', *Phys. Rev.* B **54**, 11169–11186.

R. Lai and J. Lu (2016), 'Localized density matrix minimization and linear scaling algorithms', *J. Comput. Phys.* **315**, 194–210.

R. Lai, J. Lu and S. Osher (2015), 'Density matrix minimization with ℓ_1 regularization', *Commun. Math. Sci.* **13**, 2097–2117.

L. Landau and E. Lifshitz (1991), *Quantum Mechanics: Non-Relativistic Theory*, Butterworth-Heinemann.

D. C. Langreth and J. P. Perdew (1975), 'The exchange-correlation energy of a metallic surface', *Solid State Commun.* **17**, 1425–1429.

C. Lee, W. Yang and R. G. Parr (1988), 'Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density', *Phys. Rev.* B **37**, 785–789.

M. Levy (1979), 'Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem', *Proc. Nat. Acad. Sci.* **76**, 6062–6065.

S. Li, S. Ahmed, G. Klimeck and E. Darve (2008), 'Computing entries of the inverse of a sparse matrix using the FIND algorithm', *J. Comput. Phys.* **227**, 9408–9427.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Y. Li and L. Lin (2019), 'Globally constructed adaptive local basis set for spectral projectors of second order differential operators', *Multiscale Model. Simul.* **17**, 92–116.

E. H. Lieb (1983), 'Density functionals for Coulomb systems', *Int. J. Quantum Chem.* **24**, 243–277.

E. H. Lieb and M. Loss (2001), *Analysis*, Vol. 14 of Graduate Studies in Mathematics, AMS.

L. Lin (2016), 'Adaptively compressed exchange operator', *J. Chem. Theory Comput.* **12**, 2242–2249.

L. Lin (2017), 'Localized spectrum slicing', *Math. Comp.* **86**, 2345–2371.

L. Lin and M. Lindsey (2019), 'Convergence of adaptive compression methods for Hartree–Fock-like equations', *Commun. Pure Appl. Math.* **72**, 451–499.

L. Lin and J. Lu (2016), 'Decay estimates of discretized Green's functions for Schrödinger type operators', *Sci. China Math.* **59**, 1561–1578.

L. Lin and J. Lu (2019), *A Mathematical Introduction to Electronic Structure Theory*, SIAM, to appear.

L. Lin and B. Stamm (2016), '*A posteriori* error estimates for discontinuous Galerkin methods using non-polynomial basis functions, I: Second order linear PDE', *Math. Model. Numer. Anal.* **50**, 1193–1222.

L. Lin and B. Stamm (2017), '*A posteriori* error estimates for discontinuous Galerkin methods using non-polynomial basis functions, II: Eigenvalue problems', *Math. Model. Numer. Anal.* **51**, 1733–1753.

L. Lin and C. Yang (2013), 'Elliptic preconditioner for accelerating self consistent field iteration in Kohn–Sham density functional theory', *SIAM J. Sci. Comp.* **35**, S277–S298.

L. Lin, M. Chen, C. Yang and L. He (2013), 'Accelerating atomic orbital-based electronic structure calculation via pole expansion and selected inversion', *J. Phys. Condens. Matter* **25**, 295501.

L. Lin, J. Lu, L. Ying and W. E (2009*a*), 'Pole-based approximation of the Fermi–Dirac function', *Chin. Ann. Math.* B **30**, 729–742.

L. Lin, J. Lu, L. Ying and W. E (2012*a*), 'Adaptive local basis set for Kohn–Sham density functional theory in a discontinuous Galerkin framework, I: Total energy calculation', *J. Comput. Phys.* **231**, 2140–2154.

L. Lin, J. Lu, L. Ying and W. E (2012*b*), 'Optimized local basis function for Kohn–Sham density functional theory', *J. Comput. Phys.* **231**, 4515–4529.

L. Lin, J. Lu, L. Ying, R. Car and W. E (2009*b*), 'Fast algorithm for extracting the diagonal of the inverse matrix with application to the electronic structure analysis of metallic systems', *Commun. Math. Sci.* **7**, 755–777.

L. Lin, Z. Xu and L. Ying (2017), 'Adaptively compressed polarizability operator for accelerating large scale *ab initio* phonon calculations', *Multiscale Model. Simul.* **15**, 29–55.

L. Lin, C. Yang, J. Meza, J. Lu, L. Ying and W. E (2011), 'SelInv: An algorithm for selected inversion of a sparse symmetric matrix', *ACM. Trans. Math. Software* **37**, 40.

B. Liu (1978), The simultaneous expansion method for the iterative solution of several of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. Report LBL-8158, Lawrence Berkeley Laboratory, University of California, Berkeley.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

P.-O. Löwdin (1950), 'On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals', *J. Chem. Phys.* **18**, 365–375.

J. Lu and K. Thicke (2017*a*), 'Cubic scaling algorithm for RPA correlation using interpolative separable density fitting', *J. Comput. Phys.* **351**, 187–202.

J. Lu and K. Thicke (2017*b*), 'Orbital minimization method with ℓ^1 regularization', *J. Comput. Phys.* **336**, 87–103.

J. Lu and L. Ying (2015), 'Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost', *J. Comput. Phys.* **302**, 329–335.

J. Lu and L. Ying (2016), 'Fast algorithm for periodic density fitting for Bloch waves', *Ann. Math. Sci. Appl.* **1**, 321–339.

J. Lu, C. D. Sogge and S. Steinerberger (2018), Approximating pointwise products of Laplacian eigenfunctions. arXiv:1811.10447

T. Lu, W. Cai, J. Xin and Y. Guo (2013), 'Linear scaling discontinuous Galerkin density matrix minimization method with local orbital enriched finite element basis: 1-D lattice model system', *Commun. Comput. Phys.* **14**, 276–300.

A. Luenser, H. F. Schurkus and C. Ochsenfeld (2017), 'Vanishing-overhead linear-scaling Random Phase Approximation by Cholesky decomposition and an attenuated Coulomb-metric', *J. Chem. Theory Comput.* **13**, 1647–1655.

J. MacQueen (1967), Some methods for classification and analysis of multivariate observations. In *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 281–297, University of California Press.

G. Mahan (2000), *Many-Particle Physics*, Plenum.

G. Makov and M. C. Payne (1995), 'Periodic boundary conditions in *ab initio* calculations', *Phys. Rev.* B **51**, 4014–4022.

F. Malet and P. Gori-Giorgi (2012), 'Strong correlation in Kohn–Sham density functional theory', *Phys. Rev. Lett.* **109**, 246402.

F. R. Manby, M. Stella, J. D. Goodpaster and T. F. Miller III (2012), 'A simple, exact density-functional-theory embedding scheme', *J. Chem. Theory Comput.* **8**, 2564–2568.

N. Mardirossian, J. D. McClain and G. Chan (2018), 'Lowering of the complexity of quantum chemistry methods by choice of representation', *J. Chem. Phys.* **148**, 044106.

A. Marek, V. Blum, R. Johanni, V. Havu, B. Lang, T. Auckenthaler, A. Heinecke, H.-J. Bungartz and H. Lederer (2014), 'The ELPA library: Scalable parallel eigenvalue solutions for electronic structure theory and computational science', *J. Phys. Condens. Matter* **26**, 213201.

L. D. Marks and D. R. Luke (2008), 'Robust mixing for *ab initio* quantum mechanical calculations', *Phys. Rev.* B **78**, 075114–075125.

R. Martin (2008), *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press.

D. Marx and J. Hutter (2009), *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press.

N. Marzari and D. Vanderbilt (1997), 'Maximally localized generalized Wannier functions for composite energy bands', *Phys. Rev.* B **56**, 12847–12865.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza and D. Vanderbilt (2012), 'Maximally localized Wannier functions: Theory and applications', *Rev. Mod. Phys.* **84**, 1419–1475.

F. Mauri and G. Galli (1994), 'Electronic-structure calculations and molecular-dynamics simulations with linear system-size scaling', *Phys. Rev.* B **50**, 4316–4326.

F. Mauri, G. Galli and R. Car (1993), 'Orbital formulation for electronic-structure calculations with linear system-size scaling', *Phys. Rev.* B **47**, 9973–9976.

R. McWeeny (1960), 'Some recent advances in density matrix theory', *Rev. Mod. Phys.* **32**, 335–369.

N. Mermin (1965), 'Thermal properties of the inhomogeneous electron gas', *Phys. Rev.* A **137**, 1441–1443.

S. Mohr, L. E. Ratcliff, P. Boulanger, L. Genovese, D. Caliste, T. Deutsch and S. Goedecker (2014), 'Daubechies wavelets for linear scaling density functional theory', *J. Chem. Phys.* **140**, 204110.

P. Mori-Sánchez, Q. Wu and W. Yang (2005), 'Orbital-dependent correlation energy in density-functional theory based on a second-order perturbation approach: Success and failure', *J. Chem. Phys.* **123**, 062204.

J. E. Moussa (2014), 'Cubic-scaling algorithm and self-consistent field for the random-phase approximation with second-order screened exchange', *J. Chem. Phys.* **140**, 014107.

J. E. Moussa (2016), 'Minimax rational approximation of the Fermi–Dirac distribution', *J. Chem. Phys.* **145**, 164108.

J. I. Mustafa, S. Coh, M. L. Cohen and S. G. Louie (2015), 'Automated construction of maximally localized Wannier functions: Optimized projection functions method', *Phys. Rev.* B **92**, 165134.

G. Nenciu (1983), 'Existence of the exponentially localised Wannier functions', *Comm. Math. Phys.* **91**, 81–85.

A. M. N. Niklasson (2002), 'Expansion algorithm for the density matrix', *Phys. Rev.* B. **66**, 155115.

A. M. N. Niklasson (2011), Linear-scaling techniques in computational chemistry and physics. In *Challenges and Advances in Computational Chemistry and Physics* (R. Zalesny *et al.*, eds), Springer, pp. 439–473.

A. M. N. Niklasson, C. J. Tymczak and M. Challacombe (2003), 'Trace resetting density matrix purification in $\mathcal{O}(N)$ self-consistent-field theory', *J. Chem. Phys.* **118**, 8611–8620.

J. Nocedal and S. J. Wright (1999), *Numerical Optimization*, Springer.

N. Ohba, S. Ogata, T. Kouno, T. Tamura and R. Kobayashi (2012), 'Linear scaling algorithm of real-space density functional theory of electrons with correlated overlapping domains', *Comput. Phys. Commun.* **183**, 1664–1673.

G. Onida, L. Reining and A. Rubio (2002), 'Electronic excitations: Density-functional versus many-body Green's-function approaches', *Rev. Mod. Phys.* **74**, 601–659.

P. Ordejón, D. A. Drabold, M. P. Grumbach and R. M. Martin (1993), 'Unconstrained minimization approach for electronic computations that scales linearly with system size', *Phys. Rev.* B **48**, 14646–14649.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

P. Ordejón, D. A. Drabold, R. M. Martin and M. P. Grumbach (1995), 'Linear system-size scaling methods for electronic-structure calculations', *Phys. Rev.* B **51**, 1456–1476.

T. Ozaki (2007), 'Continued fraction representation of the Fermi–Dirac function for large-scale electronic structure calculations', *Phys. Rev.* B **75**, 035123.

A. H. R. Paler and D. E. Manolopoulos (1998), 'Canonical purification of the density matrix in electronic-structure theory', *Phys. Rev.* B **58**, 12704–12711.

G. Panati and A. Pisante (2013), 'Bloch bundles, Marzari–Vanderbilt functional and maximally localized Wannier functions', *Commun. Math. Phys.* **322**, 835–875.

R. Parr and W. Yang (1989), *Density Functional Theory of Atoms and Molecules*, Oxford University Press.

R. M. Parrish, E. G. Hohenstein, T. J. Martínez and C. D. Sherrill (2012), 'Tensor hypercontraction, II: Least-squares renormalization', *J. Chem. Phys.* **137**, 224106.

R. M. Parrish, E. G. Hohenstein, T. J. Martínez and C. D. Sherrill (2013), 'Discrete variable representation in electronic structure theory: Quadrature grids for least-squares tensor hypercontraction', *J. Chem. Phys.* **138**, 194107.

M. C. Payne, M. P. Teter, D. C. Allen, T. A. Arias and J. D. Joannopoulos (1992), 'Iterative minimization techniques for *ab initio* total energy calculation: Molecular dynamics and conjugate gradients', *Rev. Mod. Phys.* **64**, 1045–1097.

J. P. Perdew (2013), 'Climbing the ladder of density functional approximations', *MRS Bull.* **38**, 743–750.

J. P. Perdew and K. Schmidt (2001), Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings*, Vol. 577, pp. 1–20.

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

J. P. Perdew, K. Burke and M. Ernzerhof (1996*a*), 'Generalized gradient approximation made simple', *Phys. Rev. Lett.* **77**, 3865–3868.

J. P. Perdew, M. Ernzerhof and K. Burke (1996*b*), 'Rationale for mixing exact exchange with density functional approximations', *J. Chem. Phys.* **105**, 9982–9985.

D. E. Petersen, S. Li, K. Stokbro, H. H. B. Sørensen, P. C. Hansen, S. Skelboe and E. Darve (2009), 'A hybrid method for the parallel computation of Green's functions', *J. Comput. Phys.* **228**, 5020–5039.

B. Pfrommer, J. Demmel and H. Simon (1999), 'Unconstrained energy functionals for electronic structure calculations', *J. Comput. Phys.* **150**, 287–298.

R. Pick, M. Cohen and R. Martin (1970), 'Microscopic theory of force constants in the adiabatic approximation', *Phys. Rev.* B **1**, 910–920.

E. Polizzi (2009), 'Density-matrix-based algorithm for solving eigenvalue problems', *Phys. Rev.* B **79**, 115112–115117.

E. Prodan and W. Kohn (2005), 'Nearsightedness of electronic matter', *Proc. Nat. Acad. Sci.* **102**, 11635–11638.

P. Pulay (1969), '*Ab initio* calculation of force constants and equilibrium geometries in polyatomic molecules, I: Theory', *Mol. Phys.* **17**, 197–204.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

P. Pulay (1980), 'Convergence acceleration of iterative sequences: The case of SCF iteration', *Chem. Phys. Lett.* **73**, 393–398.

P. Pulay (1982), 'Improved SCF convergence acceleration', *J. Comput. Chem.* **3**, 54–69.

M. J. Rayson and P. R. Briddon (2009), 'Highly efficient method for Kohn–Sham density functional calculations of 500–10 000 atom systems', *Phys. Rev.* B **80**, 205104.

S. Reine, T. Helgaker and R. Lindh (2012), 'Multi-electron integrals', *WIREs Comput. Mol. Sci.* **2**, 290–303.

X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter and M. Scheffler (2012*a*), 'Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions', *New J. Phys.* **14**, 053020.

X. Ren, P. Rinke, C. Joas and M. Scheffler (2012*b*), 'Random-phase approximation and its applications in computational chemistry and materials science', *J. Mater. Sci.* **47**, 7447–7471.

X. Ren, P. Rinke, G. E. Scuseria and M. Scheffler (2013), 'Renormalized second-order perturbation theory for the electron correlation energy: Concept, implementation, and benchmarks', *Phys. Rev.* B **88**, 035120.

Y. Saad and M. H. Schultz (1986), 'GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems', *SIAM J. Sci. Statist. Comput.* **7**, 856–869.

O. Schenk and K. Gartner (2006), 'On fast factorization pivoting methods for symmetric indefinite systems', *Elec. Trans. Numer. Anal.* **23**, 158–179.

G. Schofield, J. R. Chelikowsky and Y. Saad (2012), 'A spectrum slicing method for the Kohn–Sham problem', *Comput. Phys. Commun.* **183**, 497–505.

H. F. Schurkus and C. Ochsenfeld (2016), 'Communication: An effective linear-scaling atomic-orbital reformulation of the random-phase approximation using a contracted double-Laplace transformation', *J. Chem. Phys.* **144**, 031101.

M. Seidl, P. Gori-Giorgi and A. Savin (2007), 'Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities', *Phys. Rev.* A **75**, 042511.

Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser,

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill and M. Head-Gordon (2015), 'Advances in molecular quantum chemistry contained in the Q-Chem 4 program package', *Mol. Phys.* **113**, 184–215.

F. Shimojo, R. K. Kalia, A. Nakano and P. Vashishta (2008), 'Divide-and-conquer density functional theory on hierarchical real-space grids: Parallel implementation and applications', *Phys. Rev.* B **77**, 085103.

F. Shimojo, S. Ohmura, A. Nakano, R. Kalia and P. Vashishta (2011), 'Large-scale atomistic simulations of nanostructured materials based on divide-and-conquer density functional theory', *Eur. Phys. J. Spec. Top.* **196**, 53–63.

C. Skylaris, P. Haynes, A. Mostofi and M. Payne (2005), 'Introducing ONETEP: Linear-scaling density functional simulations on parallel computers', *J. Chem. Phys.* **122**, 084119.

J. C. Slater (1937), 'Wave functions in a periodic potential', *Phys. Rev.* **51**, 846–851.

J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón and D. Sánchez-Portal (2002), 'The SIESTA method for *ab initio* order-N materials simulation', *J. Phys. Condens. Matter* **14**, 2745–2779.

I. Souza, N. Marzari and D. Vanderbilt (2001), 'Maximally localized Wannier functions for entangled energy bands', *Phys. Rev.* B **65**, 035109.

V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew (2003), 'Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes', *J. Chem. Phys.* **119**, 12129–12137.

E. M. Stoudenmire and S. R. White (2017), 'Sliced basis density matrix renormalization group for electronic structure', *Phys. Rev. Lett.* **119**, 046401.

J. Sun, A. Ruzsinszky and J. P. Perdew (2015), 'Strongly constrained and appropriately normed semilocal density functional', *Phys. Rev. Lett.* **115**, 036402.

Q. Sun and G. K.-L. Chan (2016), 'Quantum embedding theories', *Acc. Chem. Res.* **49**, 2705–2712.

Q. Sun, T. C. Berkelbach, J. D. McClain and G. Chan (2017), 'Gaussian and plane-wave mixed density fitting for periodic systems', *J. Chem. Phys.* **147**, 164119.

P. Suryanarayana, V. Gavani, T. Blesgen, K. Bhattacharya and M. Ortiz (2010), 'Non-periodic finite-element formulation of Kohn–Sham density functional theory', *J. Mech. Phys. Solids* **58**, 258–280.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

J. J. Sylvester (1852), 'A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares', *Philos. Mag.* **4**, 138–142.

A. Szabo and N. Ostlund (1989), *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill.

M. Teter, M. Payne and D. Allan (1989), 'Solution of Schrödinger's equation for large systems', *Phys. Rev.* B **40**, 12255–12263.

B. Thaller (1992), *The Dirac Equation*, Springer.

L. H. Thomas (1927), 'The calculation of atomic fields', *Proc. Camb. Phil. Soc.* **23**, 542–548.

L. N. Trefethen (2008), 'Is Gauss quadrature better than Clenshaw–Curtis?', *SIAM Rev.* **50**, 67–87.

N. Troullier and J. L. Martins (1991), 'Efficient pseudopotentials for plane-wave calculations', *Phys. Rev.* B **43**, 1993–2006.

L. A. Truflandier, R. M. Dianzinga and D. R. Bowler (2016), 'Communication: Generalized canonical for density matrix minimization', *J. Chem. Phys.* **144**, 091102.

E. Tsuchida (2007), 'Augmented orbital minimization method for linear scaling electronic structure calculations', *J. Phys. Soc. Japan* **76**, 034708.

E. Tsuchida and M. Tsukada (1995), 'Electronic-structure calculations based on the finite-element method', *Phys. Rev.* B **52**, 5573–5578.

M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. De Jong (2010), 'NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations', *Comput. Phys. Commun.* **181**, 1477–1489.

D. Vanderbilt (1990), 'Soft self-consistent pseudopotentials in a generalized eigenvalue formalism', *Phys. Rev.* B **41**, 7892–7895.

E. Vecharynski, C. Yang and J. E. Pask (2015), 'A projected preconditioned conjugate gradient algorithm for computing many extreme eigenpairs of a Hermitian matrix', *J. Comput. Phys.* **290**, 73–89.

C. Vömel (2010), 'ScaLAPACK's MRRR algorithm', *ACM Trans. Math. Software* **37**, 1.

U. von Barth and L. Hedin (1972), 'A local exchange-correlation potential for the spin polarized case', *J. Phys.* C *Solid State Phys.* **5**, 1629–1642.

L.-W. Wang, Z. Zhao and J. Meza (2008), 'Linear-scaling three-dimensional fragment method for large-scale electronic structure calculations', *Phys. Rev.* B **77**, 165113.

G. H. Wannier (1937), 'The structure of electronic excitation levels in insulating crystals', *Phys. Rev.* **52**, 191–197.

A. Warshel and M. Levitt (1976), 'Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme', *J. Mol. Biol.* **103**, 227–249.

F. Weigend (2002), 'A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency', *Phys. Chem. Chem. Phys.* **4**, 4285–4291.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs (1998), ‘RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency’, *Chem. Phys. Lett.* **294**, 143–152.

Z. Wen and W. Yin (2013), ‘A feasible method for optimization with orthogonality constraints’, *Math. Program.* **142**, 397–434.

H. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz (2012), ‘Molpro: A general-purpose quantum chemistry program package’, *WIREs Comput. Mol. Sci.* **2**, 242–253.

S. R. White (2017), ‘Hybrid grid/basis set discretizations of the Schrödinger equation’, *J. Chem. Phys.* **147**, 244102.

J. Wilhelm, P. Seewald, M. Del Ben and J. Hutter (2016), ‘Large-scale cubic-scaling random phase approximation correlation energy calculations using a Gaussian basis’, *J. Chem. Theory Comput.* **12**, 5851–5859.

X. Wu, A. Selloni and R. Car (2009), ‘Order-N implementation of exact exchange in extended insulating systems’, *Phys. Rev.* B **79**, 085102.

Q. Xu, P. Suryanarayana and J. E. Pask (2018), ‘Discrete discontinuous basis projection method for large-scale electronic structure calculations’, *J. Chem. Phys.* **149**, 094104.

C. Yang, J. Meza and L. Wang (2006), ‘A constrained optimization algorithm for total energy minimization in electronic structure calculations’, *J. Comput. Phys.* **217**, 709–721.

W. Yang (1991*a*), ‘Direct calculation of electron density in density-functional theory’, *Phys. Rev. Lett.* **66**, 1438–1441.

W. Yang (1991*b*), ‘Direct calculation of electron density in density-functional theory: Implementation for benzene and a tetrapeptide’, *Phys. Rev.* A **44**, 7823–7826.

W. Yang and T.-S. Lee (1995), ‘A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules’, *J. Chem. Phys.* **103**, 5674–5678.

V. W.-z. Yu, F. Corsetti, A. García, W. P. Huhn, M. Jacquelin, W. Jia, B. Lange, L. Lin, J. Lu, W. Mi, A. Seifitokaldani, A. Vazquez-Mayagoitia, C. Yang, H. Yang and V. Blum (2018), ‘ELSI: A unified software interface for Kohn–Sham electronic structure solvers’, *Comput. Phys. Commun.* **222**, 267–285.

G. Zhang, L. Lin, W. Hu, C. Yang and J. E. Pask (2017), ‘Adaptive local basis set for Kohn–Sham density functional theory in a discontinuous Galerkin framework, II: Force, vibration, and molecular dynamics calculations’, *J. Comput. Phys.* **335**, 426–443.

H. Zhang, B. Smith, M. Sternberg and P. Zapol (2007), ‘SIPs: Shift-and-invert parallel spectral transformations’, *ACM Trans. Math. Software* **33**, 9–19.

I. Y. Zhang, P. Rinke and M. Scheffler (2016), ‘Wave-function inspired density functional applied to the H_2 / H_2^+ challenge’, *New J. Phys.* **18**, 073026.

L. Zhang, J. Han, H. Wang, R. Car and W. E (2018), ‘Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics’, *Phys. Rev. Lett.* **120**, 143001.

Y. Zhang, X. Xu and W. A. Goddard III (2009), ‘Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics’, *Proc. Nat. Acad. Sci. USA* **106**, 4963–4968.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047

Z. Zhao, J. Meza and L.-W. Wang (2008), ‘A divide-and-conquer linear scaling three-dimensional fragment method for large scale electronic structure calculations’, *J. Phys. Condens. Matter* **20**, 294203.

Y. Zhou, J. R. Chelikowsky and Y. Saad (2014), ‘Chebyshev-filtered subspace iteration method free of sparse diagonalization for solving the Kohn–Sham equation’, *J. Comput. Phys.* **274**, 770–782.

Y. Zhou, Y. Saad, M. L. Tiago and J. R. Chelikowsky (2006), ‘Self-consistent-field calculations using Chebyshev-filtered subspace iteration’, *J. Comput. Phys.* **219**, 172–184.

J. M. Ziman (1979), *Principles of the Theory of Solids*, Cambridge University Press.

Downloaded from https://www.cambridge.org/core. Access paid by the UC Berkeley Library, on 15 Jun 2019 at 01:21:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0962492919000047