All-Electron Ground-State and Time-Dependent Density Functional Theory: Fast Algorithms and Better Approximations

by

Bikash Kanungo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering and Scientific Computing)
in The University of Michigan
2019

Doctoral Committee:

Associate Professor Vikram Gavini (Chair)
Professor Krishna Garikipati
Associate Professor Emanuel Gull
Associate Professor Shravan Veerapaneni
Associate Professor Paul M. Zimmerman

Bikash Kanungo

bikash@umich.edu

ORCID iD: 0000-0002-3911-9354

Dedicated to *Bou* and *Baba* (mom and dad)

ACKNOWLEDGEMENTS

This work rests on the implicit and explicit support of many different people to whom I owe my deepest gratitude. First, to Vikram (Prof. Vikram Gavini) for being an immensely enterprising and exceedingly caring advisor, and for steadily and very patiently trying to make a mature researcher out of me. I can certainly vouch for the fact that he has lost as much sleep as me for some of the particularly challenging aspects of in this thesis.

I thank all my committee members for the few yet valuable discussions. I thank Prof. Paul Zimmerman for some of the crucial data used in this thesis as well as for his contribution to one of the manuscripts. I thank Prof. Krishna Garikipati, Prof. Emanuel Gull and Prof. Shravan Veerapaneni for not only being generous committee members but also for being wonderful teachers in some of my graduate courses, which provided me with a sound understanding of the key subjects involved in this work. I thank all my current and former labmates—Ian, Nelson, Nathan, Paavai, Krishnendu, Mrinal, Janakiraman, Bala—for the pleasure of an encouraging and joyous lab experience. I specially thank Sambit and Phani for the innumerable technical discussions, and, more importantly, for being two extraordinary human debuggers for my code.

It has been my good fortune to have many supportive and caring friends. With a terrified sense of unwarranted omissions, I would like to extend my gratefulness to some of them. To Ram, Srayan, Isha, and Ananda for keeping me delectably fed as well as for tolerating my idiosyncrasies. To Vikas, Anurag, Aahana, and Suren, whose

erudition has immensely shaped my philosophical perspective. To the friends at the Association for India's Development (AID) for all the parallel sociopolitical wisdom that has sharpened my overall critical outlook.

Lastly, to my parents and my sister for their unconditional love and affection.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Density functional theory (DFT), in its ground-state as well as time-dependent variant, have enjoyed incredible success in predicting a range of physical, chemical and materials properties. Although a formally exact theory, in practice DFT entails two key approximations—(a) the pseudopotential approximation, and (b) the exchange-correlation approximation. The pseudopotential approximation models the effect of sharply varying core-electrons along with the singular nuclear potential into a smooth effective potential called the pseudopotential, thereby mitigating the need for a highly refined spatial discretization. The exchange-correlation approximation, on the other hand, models the quantum many-electron interactions into an effective mean-field of the electron density ($\rho(\mathbf{r})$), and, remains an unavoidable approximation in DFT.

The overarching goal of this dissertation work is —(a) to develop efficient numerical methods for all-electron DFT and TDDFT calculations which can dispense with the pseudopotentials without incurring huge computational cost, and (b) to provide key insights into the nature of the exchange-correlation potential that can later constitute a route to systematic improvement of the exchange-correlation approximation through machine learning algorithms (i.e., which can learn these functionals using training data from wavefunction-based methods). This, in turn, involves—(a) obtaining training data mapping $\rho(\mathbf{r})$ to $v_{\text{xc}}(\mathbf{r})$, and (b) using machine learning on the training data ($\rho(\mathbf{r}) \Leftrightarrow v_{\text{xc}}(\mathbf{r})$ maps) to obtain the functional form of $v_{\text{xc}}[\rho(\mathbf{r})]$, with conformity to the known exact conditions.

The research efforts, in this thesis, constitute significant steps towards both the aforementioned goals. To begin with, we have developed a computationally effi-

cient approach to perform large-scale all-electron DFT calculations by augmenting the classical finite element basis with compactly supported atom-centered numerical basis functions. We term the resultant basis as enriched finite element basis. Our numerical investigations show an extraordinary $50-300$-fold and $5-8$-fold speedup afforded by the enriched finite element basis over classical finite element and Gaussian basis, respectively. In the case of TDDFT, we have developed an efficient *a priori* spatio-temporal discretization scheme guided by rigorous error estimates based on the time-dependent Kohn-Sham equations. Our numerical studies show a staggering 100-fold speedup afforded by higher-order finite elements over linear finite elements. Furthermore, for pseudopotential calculations, our approach achieve a $3-60$-fold speedup over finite difference based approaches. The aforementioned *a priori* spatio-temporal discretization strategy forms an important foundation for extending the key ideas of the enriched finite element basis to TDDFT. Lastly, as a first step towards the goal of machine-learned exchange-correlation functionals, we have addressed the challenge of obtaining the training data mapping $\rho(\mathbf{r})$ to $v_{\text{xc}}(\mathbf{r})$. This constitute generating accurate ground-state density, $\rho(\mathbf{r})$, from wavefunction-based calculations, and then inverting the Kohn-Sham eigenvalue problem to obtain the $v_{\text{xc}}(\mathbf{r})$ that yields the same $\rho(\mathbf{r})$. This is otherwise known as the *inverse* DFT problem. Heretofore, this remained an open challenge owing lack of accurate and systematically convergent numerical techniques. To this end, we have provided a robust and systematically convergent scheme to solve the inverse DFT problem, employing finite element basis. We obtained the exact v_{xc} corresponding to ground-state densities obtained from configuration interaction calculations, to unprecedented accuracy, for both weak and strongly correlated polyatomic systems ranging up to 40 electrons. This ability to evaluate exact v_{xc}'s from ground-state densities provides a powerful tool in the future testing and development of approximate exchange-correlation functionals.

CHAPTER I

Introduction

Density functional theory (DFT), in its both static (ground-state) and time-dependent variant, constitute an essential method for describing electronic states in all manner of nano-scale phenomena, including chemical bonds in molecules, band structures of materials, electron transport, optical properties, photo-chemical processes etc. The ground-state density functional theory (or simply density functional theory), quite justifiably, has enjoyed the distinction of the most widely used electronic structure method for over four decades. This incredible success is attributed to its low computational cost along with its reliable predictions for a wide range of materials properties [1, 2]. It relies on the Hohenberg-Kohn theorems [3] and the Kohn-Sham *ansatz* [4] to establish a one-to-one correspondence between the external potential and the ground-state electron density, $\rho(\mathbf{r})$, thereby allowing all ground-state properties to be defined as unique functionals of $\rho(\mathbf{r})$. In effect, it provides a formally exact reduction of the otherwise intractable time-independent many-electron Schrödinger equation to a set of single-electron equations, wherein each electron interacts with other electrons through an effective mean-field. The time-dependent density functional theory (TDDFT), on the other hand, extends the keys ideas of ground-state DFT to electronic excitations and time-dependent processes. It relies on the Runge-Gross [5] to establish, for a given initial state, a one-to-one correspon-

dence between the time-dependent external potential and the time-dependent electron density, $\rho(\mathbf{r}, t)$, thereby making the electron density the fundamental variable to define other physical quantities. Thus, when used in conjunction with the Kohn-Sham *anstaz*, TDDFT formally reduces the intractable time-dependent Schrödinger many-electron equation to single-electron equations. Despite the formal exactness of DFT (TDDFT), any typical DFT (TDDFT) calculation, entails two key approximations—the *pseudopotential* and the *exchange-correlation approximation*.

The pseudopotential approximation provides a balance of speed and accuracy by modeling the effect of the sharply varying core-electrons along with the singular nuclear potential into a smooth effective potential called the pseudopotential. This mitigates the need for a highly refined spatial discretization, otherwise required for describing the core-electrons and the singular nuclear potential. It relies on the fact that the core-electrons remain chemically inactive, and hence can allow for transferability of pseudopotentials from atomic calculations to molecules and solids. Although successful in predicting several material properties, pseudopotentials tend to oversimplify the description of core-electrons to be chemically inert for various systems and external conditions. For example, in systems under high pressure where the core and valence wavefunctions show increasing overlap with pressure, pseudopotentials tend to under-predict their phase transition pressures [6, 7, 8]; in systems at high temperature, where the contribution of core electrons to various thermodynamic potentials is non-negligible, pseudopotentials provide an inaccurate description of the equation of state [9]; in transition metals, where the penultimate d and f orbitals are not tightly bound, non-inclusion of these orbitals as valence electrons oftentimes lead to inaccurate bulk property prediction. More pronounced inaccuracies and sensitivity to core sizes are observed in prediction of ionization potentials [10], magnetizability [11], spectroscopic properties [12, 13] of heavier atoms wherein scalar relativistic pseudopotentials are widely employed, and in prediction of band-gap and excited state

properties [14].

Unlike the pseudopotential approximation, the second approximation of exchange-correlation (xc) is an unavoidable approximation in DFT, and the quest for an accurate exchange-correlation approximation remains the holy-grail in DFT. The exchange-correlation approximation models the quantum many-electron interactions into an effective mean-field, and hence, is indispensable to the reduction of the many-electron Schrödinger equation to effective single-electron equations in DFT. It describes an energy term, $E_{\rm xc}$, and a potential term, $v_{\rm xc}$, to characterize the interacting many-electron system in terms of a non-interacting one. Both $v_{\rm xc}$ and $E_{\rm xc}$ are known to be unique functionals of the electron density, $\rho(\mathbf{r})$, and are related to each other as $v_{\rm xc}[\rho(\mathbf{r})] = \frac{\delta E_{\rm xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}$. However, their exact forms are unknown, necessitating the use of approximations. The existing exchange-correlation approximations, despite their successes in prediction of various material properties, exhibit some notable failures—under-prediction of band-gaps, inaccurate bond-dissociation curves, inaccurate reaction barriers [15, 16], to name a few. The exchange-correlation approximation also features in TDDFT, albeit with an additional time-dependence for the exchange-correlation functionals. The existing approximations for TDDFT also exhibit certain notable failures in describing time-dependent processes—missing double excitations [17], inaccurate charge-transfer excitations [18], incorrect optical response in insulators [19], to name a few.

The aforementioned limitations in typical DFT (TDDFT) calculations beg two very ambitious questions—(i) can we dispense with the pseudopotential approximation and perform *all-electron* DFT (TDDFT) calculations without incurring heavy computational cost?, and (ii) can we provide key insights into the nature of the exchange-correlation potential and device a systematic way to improve the exchange-correlation approximation that can alleviate the shortcomings of the currently existing ones? These two questions are central to this thesis, and have been addressed in

substantial measures.

Traditionally, all-electron DFT (TDDFT) calculations have been done using atomic-orbital-type basis sets [20, 21, 22, 23, 24, 25, 26, 27, 28], wherein atom-specific basis, either analytic or numerical, are used with only a few basis functions per atom. However, owing to the incompleteness of the basis, systematic convergence for all materials systems remains a challenge. Additionally, the basis being extended in nature significantly limits the parallel scalability of their numerical implementations. Among the family of complete basis sets, several extensions to plane-waves, namely, Augmented Plane-wave (APW) [29, 30], Linearized Augmented Plane-wave (LAPW) [31, 32, 33] and APW+lo [34, 35, 36], provide an alternative basis for all-electron calculations. They achieve adaptive spatial resolution by describing the fields in terms of products of radial functions and spherical harmonics inside small muffin-tins (MTs) surrounding each atom, and in terms of plane waves in the interstitial regions between atoms. However, the basis functions within the MTs are constructed using trial energy parameters, and hence the quality of the basis remain sensitive to the choice of the trial energies. Additionally, since these augmentation schemes are based out of plane waves, they inherit some of the notable disadvantages associated with planewave, such as their restriction to periodic boundary conditions and the highly nonlocal communication associated with fast Fourier transform (FFT). Finite elements, on the other, owing to its completeness, locality, adaptive resolution, and ease of handling boundary conditions, constitute a lucrative alternative. However, the standard finite elements (hereafter termed as *classical* finite elements) suffer from two notable shortcomings in all-electron calculations—(i) the requirement of highly refined mesh ($\mathcal{O}(10^5)$ basis functions per atom) to capture the sharp oscillations of the electronic fields near the nuclei, and (ii) the widening of the spectral width of the discrete Hamiltonian with refinement, which in turn, negatively affects any eigensolver employed to solve the Kohn-Sham equations. In our case, where we employ the Chebyshev polynomial fil-

ter to selectively compute the occupied (bound) eigenstates, the use of highly refined classical finite element mesh results in the requirement of a high degree ($\mathcal{O}(10^3)$) of Chebyshev polynomial filter. We remedy both these shortcomings, by augmenting the classical finite elements with numerical atom-centered functions, termed as *enrichment* functions, which are evaluated from atomic solutions to the Kohn-Sham problem [37]. We term the resultant basis as *enriched* finite element basis. The motivation behind augmenting the classical finite element basis with these enrichment functions is that in a multi-atom materials system, the enrichment functions, being solutions to single atom Kohn-Sham orbitals and electrostatic potentials, can effectively capture the sharp variations in the electronic fields close to an atom, eliminating the need for a refined classical finite element mesh. The key ideas in the proposed method involve—(i) pre-computing the enrichment functions by solving radial Kohn-Sham equations and employing smooth cutoff functions to ensure the locality as well as control the conditioning of the enriched finite element basis; (ii) employing a divide and conquer strategy to construct an adaptive quadrature grid based on the nature of enrichment functions, so as to accurately and efficiently evaluate the integrals involving enrichment functions; (iii) implementing an efficient scheme to evaluate the inverse of the overlap matrix corresponding to the enriched finite element basis by using block-wise matrix inversion in conjunction with Gauss-Lobatto-Legendre reduced order quadrature rules; and (iv) in each self-consistent field iteration, using a Chebyshev polynomial based filter to compute the space spanned by the occupied eigenstates, and solving the Kohn-Sham eigenvalue problem by projecting the problem onto this Chebyshev-filtered space. Using this approach, we obtain a staggering $50-300$-fold and a $5-8$-fold reduction in computational cost when compared to classical finite elements and Gaussian basis, respectively, for various large-scale semiconducting and heavy-metallic system, with the largest system containing $\sim$9000 electrons. These results present the proposed enriched finite element basis based for-

mulation as a robust, systematically convergent, and efficient approach to large-scale all-electron DFT calculations.

The extension of the above enriched finite element basis to all-electron TDDFT requires some prior groundwork, given the relatively recent origins of TDDFT as compared to DFT. To elaborate, while one can rely on previously developed *a priori* mesh adaption techniques (most notably [38]) for DFT calculations, there exists no prior efforts at developing any efficient *a priori* spatio-temporal discretization scheme for TDDFT. Moreover, while the efficacy of the finite element basis in terms of its accuracy, efficiency, scalability and relative performance with other competing methods (e.g., planewaves, Gaussian basis, finite difference), have been thoroughly studied in the context of ground-state DFT [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 38, 52, 53, 37, 54, 55], a similarly comprehensive study on the efficacy of the finite element basis for TDDFT is, however, lacking. To this end, we present a computationally efficient approach to solve the time-dependent Kohn-Sham equations in real-time using higher-order finite element spatial discretization, applicable to both pseudopotential and all-electron calculations [56]. The keys ideas in this work can be summarized as: (i) developing an *a priori* mesh-adaption scheme, through a semi-discrete (discrete in space, continuous in time) error analysis of the time-dependent Kohn-Sham equations, and subsequently, obtaining a close to optimal finite element discretization for the problem; (ii) use of spectral finite elements in conjunction with Gauss-Legendre-Lobatto quadrature to render the overlap matrix diagonal, thereby simplifying the evaluation of the inverse of the overlap matrix that features in the discrete time-evolution operator; (iii) obtaining a suitable temporal discretization using a full-discrete error analysis, in the context of second-order Magnus propagator; and (iv) using an adaptive Lanczos iteration to efficiently compute the action of the Magnus propagator on the Kohn-Sham orbitals. We perform various numerical studies to assess the accuracy, rates of convergence, computational efficiency, and parallel scala-

bility of the above higher-order finite elements based formulation, for both pseudopotential and all-electron calculations. Moreover, for pseudopotential calculations, we compare the performance of higher-order finite elements with finite-difference based approach, as implemented in the Octopus software package. In particular, we observe an extraordinary 100-fold speedup in terms of the total computational time for the fourth-order finite elements over linear finite elements. In terms of relative performance with finite-difference, based on the benchmark system, the finite element discretization shows a 3 to 60-fold savings in computational time as compared to the finite-difference approach, for pseudopotential calculations. Furthermore, we demonstrate the capabilities of finite elements for efficiently handling systems subjected to strong perturbation by studying higher harmonic generation in Mg_2. Additionally, we showcase the competence of finite-elements for all-electron calculations on two benchmark systems—methane and benzene molecule. We observe good parallel scalability with $\sim$75% efficiency at 768 processors for a benchmark system of a Buckminsterfullerene molecule containing $\sim$3.5 million degrees of freedom. These results establish the higher-order finite elements based RT-TDDFT calculations as a computationally efficient alternative to the widely used RT-TDDFT approaches. More importantly, it presents the proposed finite element based approach as a unified treatment of both pseudopotential and all-electron RT-TDDFT calculations.

Lastly, we turn to the all-important challenge of developing better exchange-correlation approximations for DFT. Many of the notable failures of the existing exchange-correlation approximations can be traced to the violations of certain known exact conditions on $E_{\mathrm{xc}}[\rho(\mathbf{r})]$ and $v_{\mathrm{xc}}[\rho(\mathbf{r})]$. Typically, these approximations are constructed through semi-empirical parameter fitting in model systems, thereby making systematic improvement and conformity to the exact conditions difficult. To this end, we propose to construct these functionals in a data-driven fashion, using accurate ground-state densities, $\rho(\mathbf{r})$, from wavefunction-based methods (e.g., quantum

Monte-Carlo, full Configuration Interaction). The wavefunction-based methods, owing to their exponential computational complexity, remain untenable beyond a few tens of electrons. However, one can utilize the accurate ground-state densities from these small systems that are within the reach of the wavefunction-based methods to develop a machine-learned model for $v_{\mathrm{xc}}[\rho(\mathbf{r})]$ and $E_{\mathrm{xc}}[\rho(\mathbf{r})]$. This involves three distinct steps—(i) generation of training data set comprising of $\{\rho(\mathbf{r}), v_{\mathrm{xc}}(\mathbf{r})\}$ pairs, (ii) use of machine learning algorithms to learn the functional form of $v_{\mathrm{xc}}[\rho(\mathbf{r})]$, conforming to the exact conditions, and (iii) use of appropriate path-integrals on the learnt model of $v_{\mathrm{xc}}[\rho(\mathbf{r})]$ to obtain $E_{\mathrm{xc}}[\rho(\mathbf{r})]$. As a first step towards this goal of machine-learned exchange-correlation approximation, we address the challenge of generating the training data. That is, given an accurate ground-state density, $\rho(\mathbf{r})$, from a wavefunction based method, we are interested in evaluating $v_{\mathrm{xc}}(\mathbf{r})$ (as a spatial field) that yields the same $\rho(\mathbf{r})$. This is otherwise known as the *inverse DFT problem*. We remark that there have been several attempts at this inverse problem, using either iterative updates [57, 58, 59, 60] or constrained optimization approaches [61, 62, 63]. However, all of these approaches suffer from ill-conditioning, thereby resulting in non-unique solution or spurious oscillations in the resultant $v_{\mathrm{xc}}(\mathbf{r})$. The ill-conditioning in the above approaches has been, largely, attributed to the incompleteness of the Gaussian basis that were employed to solve them [64, 65, 66, 67, 68, 69]. Recently, Staroverov and co-workers [70, 71] have presented a different approach which utilizes the two-electron reduced density matrix (2-RDM) along with the density, both from wavefunction based methods, to remedy the non-uniqueness and the spurious oscillations in the obtained $v_{\mathrm{xc}}(\mathbf{r})$. However, this method is intended to only provide an approximate and physically meaningful $v_{\mathrm{xc}}(\mathbf{r})$ when the inverse problem is solved using a Gaussian basis. At convergence, it does not guarantee to yield the input ground-state density, and hence, does not constitute an inverse DFT calculation *per se*. In our approach [72], we mitigate these shortcomings by employing finite elements—a

complete basis— to discretize the inverse problem. In particular, we cast the inverse problem as a PDE-constrained optimization—with $v_{\text{xc}}(\mathbf{r})$ as the control variable and the Kohn-Sham eigenvalue problem as the PDE-constraint. The key ingredients in our approach are—(i) the effective use of finite element basis, which is a systematically convergent and complete basis, and, in turn, results in a well-posed inverse DFT problem; (ii) use of numerical corrections and appropriate far-field boundary conditions so as to alleviate the unphysical artifacts associated with the ubiquitous Gaussian basis-set densities that are obtained from any wavefunction-based calculation (e.g., configuration interaction (CI)). In our numerical studies, we, first, verify the accuracy of our method for densities generated from a known local density approximation based (LDA) functional. Next, we obtain the exact exchange-correlation potential for accurate ground-state densities obtained from incremental full configuration interaction (iFCI), for various weakly and strongly correlated molecular systems, with the largest system containing 40 electrons. Notably, for all the systems we obtain good agreement between the highest occupied molecular orbital (HOMO) level and the negative of the ionization potential, as mandated by the Koopmans' theorem [73, 74]. This ability to evaluate exact xc potentials from ground-state electron densities, enabled by our method, will provide a powerful tool in the future testing and development of approximate xc functionals. Further, it unlocks a systematic route to data-driven functional development, wherein one can employ machine-learning on the $\{\rho(\mathbf{r}), v_{\text{xc}}(\mathbf{r})\}$ data from inverse DFT calculations to model the functional dependence of v_{xc} on ρ, i.e., $v_{\text{xc}}[\rho]$, while honoring the known conditions on the exact functional.

The rest of this thesis is organized as follows. Chapter II presents an overview of the ground-state and time-dependent density functional theory. In Chapter III, we detail the enriched finite element formulation, and discuss its associated numerical and computational aspects. Furthermore, in the same chapter, we showcase the accuracy and efficiency of our enriched finite element implementation against that

of classical finite element and Gaussian basis. Chapter IV covers our proposed *a priori* mesh adaptation scheme for TDDFT calculations. In particular, we compare the efficacy of higher-order finite elements (in conjunction with our mesh adaption scheme) against lower-order finite elements as well as the more widely used finite-difference method. In Chapter V, we introduce the PDE-constrained optimization problem along with various numerical strategies to solve the inverse DFT problem, to unprecedented accuracy. In the same chapter, we showcase the exact exchange-correlation potentials for various benchmark molecular systems. Lastly, we summarize the major findings of this dissertation and provide an outlook of the future scope of this thesis in Chapter VI.

CHAPTER II

Density Functional Theory

In this chapter we present an overview of the fundamentals of both ground-state and time-dependent density functional theory. The rest of the chapter is organized as follows. In Sec. 2.1, we introduce the time-independent Schödinger equation. We provide the main results of ground-state density functional theory (or simply DFT) in Sec. 2.2. In Sec. 2.3, we introduce the time-dependent Schrödinger equation, and subsequently present the fundamentals of time-dependent density functional theory (TDDFT) in Sec. 2.4 .

2.1 Time-independent many-electron Schrödinger equation

The time-independent Schrödinger equation constitutes the most fundamental equation governing electronic structure of a materials system. Within the non-relativistic regime, it takes the form of an eigenvalue problem, given by

$$\hat{\mathcal{H}}\Psi = E\Psi\,, \tag{2.1}$$

where

$$\begin{aligned}\hat{\mathcal{H}} &= -\frac{1}{2}\sum_{\alpha=1}^{N_e}\nabla_\alpha^2 - \sum_{I=1}^{N_a}\frac{1}{2M_I}\nabla_I^2 - \sum_{\alpha=1}^{N_e}\sum_{I=1}^{N_a}\frac{Z_I}{|\mathbf{r}_\alpha - \mathbf{R}_I|} \\ &+ \frac{1}{2}\sum_{\alpha=1}^{N_e}\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{N_e}\frac{1}{|\mathbf{r}_\alpha - \mathbf{r}_\beta|} + \frac{1}{2}\sum_{I=1}^{N_a}\sum_{\substack{J=1\\ J\neq I}}^{N_a}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &\coloneqq \hat{\mathcal{T}} + \hat{\mathcal{T}}_N + \hat{\mathcal{V}}_{\text{ext}} + \hat{\mathcal{W}} + \hat{\mathcal{V}}_{NN}\,,\end{aligned} \tag{2.2}$$

$$\Psi \coloneqq \Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_e}; \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_a})\,. \tag{2.3}$$

In the above equations, $\hat{\mathcal{H}}$ is the many-body Hamiltonian (in a.u.) comprising of—(i) $\hat{\mathcal{T}}$: sum of the kinetic operators of N_e electrons, (ii) $\hat{\mathcal{T}}_N$: sum of kinetic operators of N_a nuclei, each with mass M_I and charge Z_I, (iii) $\hat{\mathcal{V}}_{\text{ext}}$: electrostatic interaction between electrons and the nuclei, (iv) $\hat{\mathcal{W}}$: electrostatic interaction between the electrons, and (v) $\hat{\mathcal{V}}_{NN}$: electrostatic repulsion between the nuclei. The variable $\mathbf{x}_\alpha = (\mathbf{r}_\alpha, s_\alpha)$ denotes the combined spatial coordinate and spin of the α^{th} electron, and $\mathbf{R}_I$ denotes the spatial coordinate of the I^{th} nucleus. Furthermore, Ψ is anti-symmetric with respect to interchange of any two electrons. In other words, Ψ changes sign when a pair of electronic coordinates (say $\mathbf{x}_\alpha$ and $\mathbf{x}_\beta$) are interchanged, i.e.,

$$\Psi\left(\mathbf{x}_1, \ldots, \mathbf{x}_\alpha, \ldots, \mathbf{x}_\beta, \ldots, \mathbf{x}_{N_e}; \mathbf{R}_1, \ldots, \mathbf{R}_{N_a}\right) = -\Psi\left(\mathbf{x}_1, \ldots, \mathbf{x}_\beta, \ldots, \mathbf{x}_\alpha, \ldots, \mathbf{x}_{N_e}; \mathbf{R}_1, \ldots, \mathbf{R}_{N_a}\right)\,. \tag{2.4}$$

We remark that the form of $\hat{\mathcal{H}}$ in Eq. 2.2 makes the electronic and nuclear motion coupled to each other. To this end, one can invoke the *Born-Oppenheimer* approximation to separate the electronic and nuclear motions. The *Born-Oppenheimer* approximation rests on the fact that the nuclei are massive compared to the electrons (e.g., a proton in more than 1800 times as heavy as an electron), and consequently, move much more slowly compared to the electrons. Therefore, the nuclei can be approximated to be fixed with respect to the electrons. In effect, we can neglect the $\hat{\mathcal{T}}_N$ term

(sum of kinetic operators for the nuclei) and take the electronic wavefunction, Ψ_e, to be only parametrically dependent on the nuclear positions $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_a}\}$. In other words, for a fixed position of nuclei, $\mathbf{R}$, we can write the electronic part of the Schrödinger equation as

$$\hat{\mathcal{H}}_e \Psi_e = E_e \Psi_e \,, \tag{2.5}$$

where $\hat{\mathcal{H}}_e = \hat{\mathcal{T}} + \hat{\mathcal{V}}_{\text{ext}} + \hat{\mathcal{W}} + \hat{\mathcal{V}}_{NN}$ and $\Psi_e \coloneqq \Psi_e(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_e}; \mathbf{R})$. We emphasize that although we have explicitly defined $\hat{\mathcal{V}}_{\text{ext}}$ in Eq. 2.2, for our subsequent discussion, it can be taken as any operator of the form,

$$\hat{\mathcal{V}}_{\text{ext}} = \sum_{\alpha=1}^{N_e} v_{\text{ext}}(\mathbf{r}_\alpha) \,, \tag{2.6}$$

where $v_{\text{ext}}(\mathbf{r}_\alpha)$ is a single-particle potential. This allows us to introduce the essentials of DFT in a generic manner, without having to express the specific form of v_{ext}. One may note that for the form of $\hat{\mathcal{V}}_{\text{ext}}$ in Eq. 2.2, $v_{\text{ext}}(\mathbf{r}_\alpha) = -\sum_{I=1}^{N^a} \frac{Z_I}{|\mathbf{r}_\alpha - \mathbf{R}_I|}$. Similarly, although we have explicitly defined $\hat{\mathcal{W}}$ in Eq. 2.2, for our discussion on DFT, $\hat{\mathcal{W}}$ can take the form of $\hat{\mathcal{W}} = \sum_{\alpha=1}^{N_e} \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{N_e} w(|\mathbf{r}_\alpha - \mathbf{r}_\beta|)$, where $w(|\mathbf{r}_\alpha - \mathbf{r}_\beta|)$ is a pair potential.

We remark that Ψ_e being an eigenfunction of $\hat{\mathcal{H}}_e$ is defined only up to a scaling constant. It is customary to choose the scaling constant such that Ψ_e is normalized to unity, i.e.,

$$\sum_{s_1, s_2, \ldots, s_{N_e}} \int d\mathbf{r}_1 \, d\mathbf{r}_2 \ldots \, d\mathbf{r}_{N_e} \, |\Psi_e|^2 = 1 \tag{2.7}$$

Additionally, Ψ_e is mandated to be anti-symmetric with respect to the exchange of any $\{\mathbf{x}_\alpha, \mathbf{x}_\beta\}$ pair.

The electronic ground-state is defined to be the eigenstate Ψ_e^0 (or a degenerate set of $\{\Psi_e\}$'s) that corresponds to the lowest eigenvalue of the Eq. 2.5, denoted as E_e^0.

Equivalently, the evaluation of E_e^0 can be defined as the following variational problem

$$E_e^0 = \inf_{\Psi_e} \langle \Psi_e | \, \hat{\mathcal{H}}_e \, | \Psi_e \rangle \ . \tag{2.8}$$

subject to the normalization and antisymmetry constraints on Ψ_e.

Given that Ψ_e is a multidimensional function from $\mathbb{R}^{3N_e} \times 2^{N_e}$ to $\mathbb{C}$, the problem of computing the electronic ground-state translates into a computationally intractable problem. To elaborate, if we are to neglect the spin variables and take K degrees of freedom (e.g., discrete points) on each of the $3N_e$ dimensions, the discrete $\hat{\mathcal{H}}_e$ matrix takes an astronomical dimension of $K^{3N_e} \times K^{3N_e}$. Consequently, diagonalizing such an enormous matrix remains computationally infeasible, beyond a few number of electrons. This is otherwise referred to as the *curse of dimensionality* in electronic structure theory. This calls for other means to evaluate electronic properties without having to resort to the solution of the many-electron Schrödinger equation. To this end, ground-state density functional theory presents a formal route to compute ground-state properties without having to explicitly solve for Ψ_e^0. We discuss the fundamentals of DFT in the following section.

2.2 Ground-state density functional theory

Ground-state density functional theory, or *simply* density functional theory (DFT), seeks to describe ground-state properties of a many-electron system through the electron density. The electron density, $\rho(\mathbf{r})$, is defined as

$$\rho(\mathbf{r}) = N_e \sum_{s_1} \int d\mathbf{r}_2 \, d\mathbf{r}_3 \ldots d\mathbf{r}_{N_e} \, |\Psi_e|^2 \ . \tag{2.9}$$

Physically, $\rho(\mathbf{r}) \, d\mathbf{r}$ gives us the probability of finding an electron in the infinitesimal volume $d\mathbf{r}$ centered at $\mathbf{r}$. Before proceeding to the details of DFT, a few things

need to be simplified. For brevity, we omit the label e which was used to denote the electronic part of the Schrödinger equation (within the Born-Oppenheimer approximation). Furthermore, we note that $\hat{\mathcal{V}}_{NN}$ operator in $\hat{\mathcal{H}}_e$ is independent of the electronic wavefunction and only contributes a constant nuclear-nuclear repulsive term to the energy, given by

$$E_{\rm NN} = \frac{1}{2}\sum_{I=1}^{N_a}\sum_{\substack{J=1\\J\neq I}}^{N_a}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{2.10}$$

Thus, we can omit $\hat{\mathcal{V}}_{NN}$ from the following discussion, without any loss of generality.

2.2.1 Existence and uniqueness of density functional

The foundation of DFT rests on two simple yet profound results from Hohenberg and Kohn [3]. The first Hohenberg-Kohn theorem establishes a one-to-one correspondence between the ground-state density $\rho_0(\mathbf{r})$ and the external potential $v_{\rm ext}(\mathbf{r})$ (defined up to an arbitrary constant). Below we present the proof for the non-degenerate case. Let us assume that two external potentials, $v^{(1)}_{\rm ext}(\mathbf{r})$ and $v^{(2)}_{\rm ext}(\mathbf{r})$, which differ by more than a constant (i.e., $v^{(1)}_{\rm ext} \neq v^{(2)}_{\rm ext} + C$), yield the same ground-state density $\rho_0(\mathbf{r})$. Obviously, $v^{(1)}_{\rm ext}$ and $v^{(2)}_{\rm ext}$ define two different $\hat{\mathcal{V}}_{\rm ext}$'s (i.e., $\hat{\mathcal{V}}^{(1)}_{\rm ext}$ and $\hat{\mathcal{V}}^{(2)}_{\rm ext}$), and hence, define two different $\hat{\mathcal{H}}$ (i.e., $\hat{\mathcal{H}}^{(1)}$ and $\hat{\mathcal{H}}^{(2)}$). Consequently, they generate two different ground-state wavefunctions Ψ_1 and Ψ_2. Now, let E_1 and E_2 denote the ground-state energies corresponding to $\hat{\mathcal{H}}^{(1)}$ and $\hat{\mathcal{H}}^{(2)}$, respectively. Then, using the definition of ground-state energy, and assuming non-degeneracy of the ground-state, we have

$$\begin{aligned} E_1 &= \langle\Psi_1|\,\hat{\mathcal{H}}^{(1)}\,|\Psi_1\rangle < \langle\Psi_2|\,\hat{\mathcal{H}}^{(1)}\,|\Psi_2\rangle \\ &= \langle\Psi_2|\,\hat{\mathcal{T}} + \hat{\mathcal{W}} + \hat{\mathcal{V}}^{(2)}_{\rm ext}\,|\Psi_2\rangle + \langle\Psi_2|\,(\hat{\mathcal{V}}^{(1)}_{\rm ext} - \hat{\mathcal{V}}^{(2)}_{\rm ext})\,|\Psi_2\rangle \\ &= E_2 + \int \rho_0(\mathbf{r})(v^{(1)}_{\rm ext}(\mathbf{r}) - v^{(2)}_{\rm ext}(\mathbf{r}))\,. \end{aligned} \tag{2.11}$$

Similarly, interchanging the labels 1 and 2, we obtain

$$E_2 < E_1 + \int \rho_0(\mathbf{r})(v_{\text{ext}}^{(2)} - v_{\text{ext}}^{(1)}(\mathbf{r}))\,. \tag{2.12}$$

Adding the above equation leads to $E_1 + E_2 < E_1 + E_2$—*a contradiction.* Hence, our assumption must be wrong. This concludes that the ground-state density $\rho_0(\mathbf{r})$ must uniquely determine (up to a constant) the external potential $v_{\text{ext}}(\mathbf{r})$.

We emphasize that the uniqueness of $v_{\text{ext}}(\mathbf{r})$ for a given ground-state $\rho(\mathbf{r})$ does not guarantee that $v_{\text{ext}}(\mathbf{r})$ exists for any given $\rho(\mathbf{r})$. A density for which a $v_{\text{ext}}(\mathbf{r})$ exists is called *v-representable*. We denote the set of *v-representable* densities as $\mathcal{A}$. The second Hohenberg-Kohn theorem assumes *v-representability* of densities to show that, for a given v_{ext}, a universal functional for the energy $E_{v_{\text{ext}}}[\rho]$ can be defined in terms of the density, and the ground-state energy can be determined as the global minimum of this functional with respect to ρ. The proof of this theorem, for non-degenerate case, follows from the first theorem along with the assumption of *v-representability*, i.e., $\rho(\mathbf{r}) \in \mathcal{A}$. To elaborate, owing to the *v-representability* assumption, for any $\tilde{\rho}(\mathbf{r}) \in \mathcal{A}$ we can uniquely determine the external potential $\tilde{v}(\mathbf{r})$ which yields $\tilde{\rho}(\mathbf{r})$ as its ground-state density. The $\tilde{v}(\mathbf{r})$, in turn, determines the ground-state wavefunction $\widetilde{\Psi}$. Thus, transitively, the ground-state wavefunction becomes a functional of $\tilde{\rho}$, i.e., $\widetilde{\Psi} = \widetilde{\Psi}\,[\tilde{\rho}]$. Thus, for any $\tilde{\rho} \in \mathcal{A}$, the functional $F_{\text{HK}}[\tilde{\rho}] = \left\langle \widetilde{\Psi}[\tilde{\rho}] \middle| \hat{\mathcal{T}} + \hat{\mathcal{W}} \middle| \widetilde{\Psi}[\tilde{\rho}] \right\rangle$ is well-defined and unique, since $\tilde{\rho}$ determines $\widetilde{\Psi}$ as well as N_e, and N_e determines both $\hat{\mathcal{T}}$ and $\hat{\mathcal{W}}$. The physical interpretation of $F_{\text{HK}}[\tilde{\rho}]$ is that of the internal electronic energy. We now define, for an arbitrary external potential v_{ext}, unrelated to the $\tilde{v}$ determined by $\tilde{\rho}$, the following energy functional

$$E_{v_{\text{ext}}}[\tilde{\rho}] = F_{\text{HK}}[\tilde{\rho}] + \int \tilde{\rho}(\mathbf{r}) v_{\text{ext}}(\mathbf{r})\, d\mathbf{r}\,. \tag{2.13}$$

The above definition proves the first part of the second Hohenberg-Kohn theorem

that a universal energy functional $E_{v_{\text{ext}}}[\tilde{\rho}]$ can be defined in terms of the density.

We now proceed to the second part of the second Hohenberg-Kohn theorem. We note that if E_0 denotes the ground-state energy for v_{ext}, then

$$\left\langle \widetilde{\Psi} \middle| \hat{\mathcal{T}}+\hat{\mathcal{W}}+\hat{\mathcal{V}}_{\text{ext}} \middle| \widetilde{\Psi} \right\rangle = \left\langle \widetilde{\Psi} \middle| \hat{\mathcal{T}}+\hat{\mathcal{W}} \middle| \widetilde{\Psi} \right\rangle + \left\langle \widetilde{\Psi} \middle| \hat{\mathcal{V}}_{\text{ext}} \middle| \widetilde{\Psi} \right\rangle = F_{\text{HK}}[\tilde{\rho}] + \int \tilde{\rho}(\mathbf{r}) v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} \geq E_0 \,. \tag{2.14}$$

The above inequality is a variational statement, i.e., for a given v_{ext}, its ground-state energy can be obtained as the global minimum of $E_{v_{\text{ext}}}[\rho]$ with respect to ρ. This concludes the proof of the second Hohenberg-Kohn theorem. While we have presented the proof for the non-degenerate case, the same result holds for degenerate case as well, and can be proven using density matrices. We refer the reader to [75] for a proof for the degenerate case.

We remark that Hohenberg-Kohn minimization requires the ρ to be *v-representable*. However, determining whether a given ρ is *v-representable* still remains an open question. To this end, one can relax the *v-representability* requirement by defining E_0 through the Levy-Lieb constrained search [76] approach. To begin with, it defines a new functional

$$F[\rho] = \min_{\Psi \to \rho(\mathbf{r})} \langle \Psi | \, \hat{\mathcal{T}} + \hat{\mathcal{W}} \, | \Psi \rangle \tag{2.15}$$

where $\Psi \to \rho(\mathbf{r})$ indicates that the $\rho(\mathbf{r})$ is obtained from an anti-symmetric wavefunction Ψ. A density which is obtained from an anti-symmetric wavefunction is known as *n-representable*. We denote the set of *n-representable* densities as $\mathcal{N}$. Thus, finding the ground-state energy E_0 for an external potential v_{ext} can be reformulated as

$$\begin{aligned} E_0 &= \min_{\rho(\mathbf{r}) \in \mathcal{N}} \left[\min_{\Psi \to \rho(\mathbf{r})} \langle \Psi | \, \hat{\mathcal{T}} + \hat{\mathcal{W}} \, | \Psi \rangle + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} \right] \\ &= \min_{\rho(\mathbf{r}) \in \mathcal{N}} \left[F[\rho] + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} \right] . \end{aligned} \tag{2.16}$$

The distinct advantage of the above minimization problem over the Hohenberg-Kohn

minimization is that the *n-representability* problem is a solved problem, as opposed to the *v-representability* problem. To elaborate, the sufficient conditions for a density to be *n-representable* are [77]—(i) $\rho(\mathbf{r}) \geq 0,\ \forall \mathbf{r} \in \mathbb{R}^3$; (ii) $\int \rho(\mathbf{r})\, d\mathbf{r} = N_e$, and (iii) $\int \left|\nabla \rho(\mathbf{r})^{1/2}\right|^2 d\mathbf{r} < \infty$.

We emphasize that although the above results prove the uniqueness of $F_{\mathrm{HK}}[\rho]$ or $E_{v_{\mathrm{ext}}}[\rho]$, the exact form of these functionals are unknown, and thus, finding E_0 by minimizing $E_{v_{\mathrm{ext}}}[\rho]$ remains as challenging as solving the original many-body problem of Eq. 2.8. To this end, the real practical breakthrough in DFT is offered by the approach of Kohn and Sham [4].

2.2.2 Kohn-Sham DFT

The first step in the Kohn-Sham approach is to split $E_{v_{\mathrm{ext}}}[\rho]$ as

$$E_{v_{\mathrm{ext}}}[\rho] = T_s[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] + \int \rho(\mathbf{r}) v_{\mathrm{ext}}(\mathbf{r})\, d\mathbf{r} + E_{\mathrm{NN}}\,. \tag{2.17}$$

In the above equation, $T_s[\rho]$ is the kinetic energy of a system of non-interacting electrons which, under the influence of an external potential $v_{\mathrm{KS}}[\rho]$, yields the same density as the interacting system. $E_{\mathrm{H}}[\rho]$ is the classical electron-electron electrostatic interaction energy, given by

$$E_{\mathrm{H}}[\rho] = \frac{1}{2} \int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}'\,. \tag{2.18}$$

E_{NN} is the constant nuclear-nuclear repulsive energy defined in Eq. 2.10. Finally, $E_{\mathrm{xc}}[\rho]$ encapsulates the remainder of $E_{v_{\mathrm{ext}}}[\rho]$. Thus, physically $E_{\mathrm{xc}}[\rho]$ encapsulates the quantum many-body interaction among the electrons. The underlying assumption of the Kohn-Sham approach is that the interacting density $\rho(\mathbf{r})$ is non-interacting *v-representable*. That is, there exists a potential $v_{\mathrm{KS}}(\mathbf{r})$ which yields the same density $\rho(\mathbf{r})$ as that of the interacting system. The ground-state of the non-interacting system,

Φ, corresponding to v_{KS} can be written as a Slater determinant of orthonormalized single-electron orbitals $\{\psi_1(\mathbf{r}), \psi_2(\mathbf{r}), \ldots, \psi_{N_e}(\mathbf{r})\}$. Thus, we can rewrite the energy as

$$E_{v_{\mathrm{ext}}}[\rho] = -\frac{1}{2}\sum_{\alpha=1}^{N_e}\int \psi_\alpha^\dagger(\mathbf{r})\nabla^2\psi_\alpha(\mathbf{r})\,d\mathbf{r} + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] + \int \rho(\mathbf{r})v_{\mathrm{ext}}(\mathbf{r})\,d\mathbf{r} + E_{\mathrm{NN}}\,, \tag{2.19}$$

where the superscript † denotes the complex conjugate of a function.

The above Kohn-Sham reformulation of the problem, now, reduces the evaluation of the ground-state energy to minimizing $E_{v_{\mathrm{ext}}}$ with respect to the ψ_α's, subject to the constraint $\int |\psi_\alpha|^2\ d\mathbf{r} = 1$. This leads to the following set of Euler-Lagrange equations known as the Kohn-Sham eigenvalue equations,

$$\left(-\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}[\rho]\right)\psi_\alpha(\mathbf{r}) = \epsilon_\alpha\psi_\alpha(\mathbf{r}) \quad \alpha = 1, 2, \ldots \tag{2.20}$$

In the above equation, ϵ_α and ψ_α are commonly referred to as the Kohn-Sham eigenvalue and eigenfunction (orbital). The density, ρ, is given by

$$\rho(\mathbf{r}) = \sum_\alpha h(\epsilon_\alpha, \mu)\,|\psi_\alpha(\mathbf{r})|^2\ , \tag{2.21}$$

where $h(\epsilon, \mu)$ is the Heavyside step function given by

$$h(\epsilon, \mu) = \begin{cases} 1, & \text{if } \epsilon \leq \mu \\ 0, & \text{otherwise}\,. \end{cases} \tag{2.22}$$

The scalar μ is the Fermi energy defined by the constraint on the number of electrons, i.e.,

$$\int \rho(\mathbf{r})\,d\mathbf{r} = \sum_\alpha h(\epsilon_\alpha, \mu) = N_e\,. \tag{2.23}$$

The potential $v_{\rm KS}[\rho]$ is defined as

$$v_{\rm KS}[\rho] = v_{\rm ext}(\mathbf{r}) + v_{\rm H}[\rho] + v_{\rm xc}[\rho]\,, \tag{2.24}$$

where $v_{\rm H}[\rho]$ and $v_{\rm xc}[\rho]$ denote the Hartree and exchange-correlation potentials, respectively and are given by

$$v_{\rm H}[\rho] = \frac{\delta E_{\rm H}}{\delta \rho} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}'\,, \tag{2.25}$$

$$v_{\rm xc}[\rho] = \frac{\delta E_{\rm xc}}{\delta \rho}\,. \tag{2.26}$$

We note that the above equations present a set of non-linear eigenvalue problem, given that the potential $v_{\rm KS}$ is dependent on the eigenfunction ψ_α's via $\rho(\mathbf{r})$. The practicality of the Kohn-Sham equations lies in the fact that if we can obtain a good approximation for $E_{\rm xc}[\rho]$ (and hence for $v_{\rm xc}[\rho]$), then we can solve for the ψ_α's self-consistently. Over the past four decades increasingly sophisticated approximations to $E_{\rm xc}$ have been developed ranging from local to non-local forms. The simplest of these forms is the local density approximation [78, 79], wherein $E_{\rm xc}[\rho] = \int \epsilon_{\rm xc}[\rho]\, d\mathbf{r}$ with $\epsilon_{\rm xc}[\rho]$ being a local functional in ρ. Other widely used forms include the generalized gradient approximation (GGA), wherein $E_{\rm xc} = \int \epsilon_{\rm xc}[\rho, \nabla\rho]\, d\mathbf{r}$ [80]. Recently, hybrid functionals which combine a portion of the non-local Hartree exchange operator with local approximations have resulted in improved predictions of a wide range of materials properties [81, 82, 83, 84, 85, 86, 87]. We refer to [88, 16, 1, 89, 2, 90] and references there-within for a comprehensive review of the existing $E_{\rm xc}$ approximations, their successes and failures.

2.3 Time-dependent many-electron Schrödinger equation

The time-dependent Schrödinger equation (TDSE) constitutes the most fundamental equation governing the dynamics of electrons and nuclei. While in its full glory TDSE involves coupled electronic and nuclear dynamics, one can invoke the *Born-Oppenheimer* approximation to separate the two motions. In other words, under the *Born-Oppenheimer* approximation, we can define the time-dependent electronic wavefunction ($\Psi(t)$) to be only parametrically dependent on the nuclei positions $\mathbf{R} = \{\mathbf{R}_I\}$. Thus, within the non-relativistic regime, the evolution of the electronic wavefunction $\Psi(t)$ is given by the following initial-value problem

$$i\frac{\partial \Psi(t)}{\partial t} = \hat{\mathcal{H}}(t)\Psi(t)\,, \quad \Psi(t_0) = \Psi_0\,, \quad \forall t \geq t_0\,. \tag{2.27}$$

In the above equation, analogous to the time-independent case, we have

$$\begin{aligned}\hat{\mathcal{H}}(t) &= -\frac{1}{2}\sum_{\alpha=1}^{N_e}\nabla_\alpha^2 + \sum_{\alpha=1}^{N_e} v_{\text{ext}}(\mathbf{r}_\alpha, t) + \frac{1}{2}\sum_{\alpha=1}^{N_e}\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{N_e}\frac{1}{|\mathbf{r}_\alpha - \mathbf{r}_\beta|} \\ &\coloneqq \hat{\mathcal{T}} + \hat{\mathcal{V}}_{\text{ext}}(t) + \hat{\mathcal{W}},\end{aligned} \tag{2.28}$$

$$\Psi(t) \coloneqq \Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_e}; \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_a}; t)\,. \tag{2.29}$$

Typically, the single-electron external potential $v_{\text{ext}}(\mathbf{r}, t)$ comprises of two parts

$$v_{\text{ext}}(\mathbf{r}, t) = v_{\text{ext}}^{(0)}(\mathbf{r}) + v_{\text{field}}(\mathbf{r}, t)\,, \tag{2.30}$$

where $v_{\text{ext}}^{(0)}(\mathbf{r})$ is the time-independent part (usually, the nuclear potential), and $v_{\text{field}}(\mathbf{r}, t)$ is the time-dependent potential. The time-dependent potential is, usually, given in terms of an electric field, i.e., $v_{\text{field}}(\mathbf{r}, t) = \mathbf{E}(t)\cdot\mathbf{r}$. We emphasize that in the above form of TDSE there remains a hidden assumption of the $v_{\text{field}}(\mathbf{r}, t)$ being a classical field. An accurate description warrants quantization of v_{field}, which, in turn, requires

solving the TDSE as a quantum electrodynamics (QED) problem. Nevertheless, when the photon density is high, as is typical of most of the light-matter interactions, one can safely approximate the v_{field} to be classical.

Needless to say, any *ab initio* approach to solve the TDSE suffers from the *curse of dimensionality*, owing to the high-dimensionality of $\Psi(t)$ (i.e., $\Psi(t)$ defines a function from $\mathbb{R}^{3N_e} \times 2^{N_e} \times [0,T]$ to $\mathbb{C} \times [0,T]$). Consequently, studying electronic dynamics through *ab initio* means remain computationally unviable, beyond a few number of electrons. This warrants a formal route to describe electron dynamics without resorting to the evaluation of $\Psi(t)$. To this end, the time-dependent density functional theory (TDDFT) offers a formal as well as practical solution to the problem.

2.4 Time-dependent density functional theory

2.4.1 Runge-Gross Theorem

The foundation of TDDFT rests on the Runge-Gross theorem [5] which, analogous to the Hohenberg-Kohn theorem, establishes a one-to-one correspondence between the time-dependent density $\rho(\mathbf{r},t)$ and the external potential $v_{\text{ext}}(\mathbf{r},t)$, for a given initial many-electron state Ψ_0. To elaborate, if there are two potentials $v^{(1)}_{\text{ext}}(\mathbf{r},t)$ and $v^{(2)}_{\text{ext}}(\mathbf{r},t)$ which differ by more than just a purely time-dependent constant (i.e., $v^{(1)}_{\text{ext}}(\mathbf{r},t) \neq v^{(2)}_{\text{ext}}(\mathbf{r},t) + C(t)$)), then their corresponding densities, $\rho^{(1)}(\mathbf{r},t)$ and $\rho^{(2)}(\mathbf{r},t)$, evolving from a common initial state Ψ_0, will start to differ infinitesimally later than t_0. The proof comprises of two steps. The first step establishes the one-to-one correspondence between $v_{\text{ext}}(\mathbf{r},t)$ and the current density $\mathbf{j}(\mathbf{r},t)$. The second step uses the continuity equation to show that if the current densities $\mathbf{j}(\mathbf{r},t)$ for two different external potentials are different, then the densities $\rho(\mathbf{r},t)$ must be different as well. We omit the details of the proof, in the interest of brevity. We refer an interested reader to the standard textbooks on TDDFT [91, 92] for the details of

the proof. We remark that there are two major assumptions that go into the Runge-Gross. The first assumption is that the external potentials ($v_{\text{ext}}^{(1)}(\mathbf{r}.t)$ and $v_{\text{ext}}^{(2)}(\mathbf{r},t)$) must be Taylor-expandable about t_0. For most practical applications this remains a valid assumption. The second assumption requires the external potential to decay faster than $1/r$ as well as the density to decay to zero at infinity. Alternatively, the second assumption warrants the external potential to be periodic. Thus, while the Runge-Gross theorem is straightforwadly applicable to finite (non-periodic systems), its applicability to periodic systems is restricted. The periodic case can be handled in a more generic and formal way through time-dependent current density functional theory (TDCDFT) [93], where one uses the current density $\mathbf{j}(\mathbf{r},t)$ as the fundamental quantity instead of $\rho(\mathbf{r},t)$.

In summary, the Runge-Gross theorem establishes the fact that, for a given initial state Ψ_0, the external potential can be expressed as a functional of the time-dependent density, i.e, $v_{\text{ext}}[\rho(\mathbf{r},t),\Psi_0]$. This implies that the many-body Hamiltonian $\hat{\mathcal{H}}(t)$, and thus, the many-body wavefunction $\Psi(t)$ are functionals of $\rho(\mathbf{r},t)$, i.e, $\hat{\mathcal{H}}(t)=\hat{\mathcal{H}}[\rho,\Psi_0]$ and $\Psi(t)=\Psi[\rho,\Psi_0]$. Consequently, the expectation value of all physical obervables becomes a functional of the density (i.e., $O(t)=\langle\Psi[\rho,\Psi_0]|\,\hat{O}\,|\Psi[\rho,\Psi_0]\rangle=O[\rho,\Psi_0]$). *Prima facie*, given that the functional dependence of $\Psi(t)$ on $\rho(\mathbf{r},t)$ is unknown, the above result is no more advantageous than solving the TDSE itself. However, the practical utility of Runge-Gross theorem is realized when one uses the Kohn-Sham approach.

2.4.2 Time-dependent Kohn-Sham approach

We remark that the Kohn-Sham approach in ground-state DFT assumes the non-interacting *v-representability* of the interacting density $\rho(\mathbf{r})$ (see Sec. 2.2.2). While the *v-representability* problem is still an open challenge in DFT, in TDDFT it has been resolved through the van Leeuwen theorem [94]. To elaborate, the van Leeuwen

theorem states that given a particle-particle interaction $w(|\mathbf{r}-\mathbf{r}'|)$, external potential $v_{\text{ext}}(\mathbf{r},t)$ and initial state Ψ_0 that yields the density $\rho(\mathbf{r},t)$, there exists a different many-body system with interaction $\tilde{w}(|\mathbf{r}-\mathbf{r}'|)$, initial state $\widetilde{\Psi}_0$ and a unique external potential $\tilde{v}_{\text{ext}}(\mathbf{r},t)$ (defined up to a purely time-depdendent $C(t)$) which yields the same $\rho(\mathbf{r},t)$. The initial state $\widetilde{\Psi}_0$ must yield the same density and time derivative as that of Ψ_0 at t_0. Further, both v_{ext} and $\tilde{v}_{\text{ext}}$ must be Taylor-expandable at t_0. To put it in simple terms, the van Leeuwen theorem allows us to set $\tilde{w}(\mathbf{r}-\mathbf{r}'|)=0$ and thereby establish an auxiliary system of non-interacting electrons that yields the same density as the interacting one. Thus, one can formally reduce the many-electron TDSE problem into a set of single-electron initial value equations given by

$$i\frac{\partial\psi_\alpha(\mathbf{r},t)}{\partial t}=\hat{\mathcal{H}}_{\text{KS}}\psi_\alpha(\mathbf{r},t)\,,\quad \alpha=1,2,\ldots,N_e,\quad \psi(t_0)=\psi_{\alpha,0}\,,\quad \forall t\geq t_0\,. \tag{2.31}$$

The above equations are known as the time-dependent Kohn-Sham (TDKS) equations. The $\psi_\alpha(\mathbf{r},t)$ denote the Kohn-Sham orbitals. The $\hat{\mathcal{H}}_{\text{KS}}$ is the Kohn-Sham Hamiltonian defined as

$$\hat{\mathcal{H}}_{\text{KS}}[\rho(\mathbf{r},t),\Psi_0,\Phi_0]=-\frac{1}{2}\nabla^2+v_{\text{KS}}[\rho(\mathbf{r},t),\Psi_0,\Phi_0]\,, \tag{2.32}$$

where

$$v_{\text{KS}}[\rho(\mathbf{r},t),\Psi_0,\Phi_0]=v_{\text{ext}}(\mathbf{r},t)+v_{\text{H}}[\rho(\mathbf{r},t)]+v_{\text{xc}}[\rho(\mathbf{r},t),\Psi_0,\Phi_0]\,. \tag{2.33}$$

In the above equation, Φ_0 denotes initial Kohn-Sham wavefunction defined as the Slater determinant composed of the initial Kohn-Sham orbitals $\{\psi_\alpha(\mathbf{r},t_0)\}$. $v_{\text{ext}}(\mathbf{r},t)$ is the time-dependent external potential. $v_{\text{H}}[\rho(\mathbf{r},t)]$, is the Hartree potential (same

as its ground-state counterpart) and is given by

$$v_{\mathrm{H}}[\rho(\mathbf{r},t)] = \int \frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' \,. \tag{2.34}$$

Finally, $v_{\mathrm{xc}}[\rho(\mathbf{r},t), \Psi_0, \Phi_0]$ is the exchange-correlation potential that encapsulates all the quantum many-electron interactions as a mean-field of $\rho(\mathbf{r},t)$, and the initial states Ψ_0 and Φ_0. Similar to ground-state DFT, the exact form of v_{xc} is unknown, thereby necessitating the use of approximations. The initial state dependence, naturally, increases the complexity of the problem as compared to the ground-state case. However, usually, one is interested in systems evolving from the ground-state. This allows us, owing to the Hohenberg-Kohn theorem, to eliminate the initial state dependence as the initial ground-states can be defined in terms of the ground-state density. The above simplification aside, v_{xc}, in general, is known to be nonlocal in both space and time. This presents a much more daunting challenge to model approximate v_{xc}'s in TDDFT as compared to the ground-state DFT. Typically, one uses the *adiabatic approximation* to neglect the time nonlocality (memory effect). This allows for direct use of the existing exchange-correlation approximations used in ground-state DFT, i.e,

$$v_{\mathrm{xc}}^{\mathrm{adia}}(\mathbf{r},t) = v_{\mathrm{xc}}^{\mathrm{gs}}[\rho_0(\mathbf{r})]|_{\rho_0(\mathbf{r})=\rho(\mathbf{r},t)} \,. \tag{2.35}$$

We remark that the above prescription of TDDFT, wherein we explicitly evolve the Kohn-Sham orbitals in time, is referred to as the real-time TDDFT (RT-TDDFT). There exists another framework for TDDFT known as the linear-response TDDFT (LR-TDDFT) which applies to the case of weak interaction between the external field and the system. Owing to the weak interaction, the system undergoes small perturbation from its ground-state. Consequently, one can forgo the real-time evolution and directly compute the linear density response from the ground-state. However, the validity of LR-TDDFT is restricted to handling only perturbative (linear response)

regimes. Thus, between the two frameworks, RT-TDDFT offers the more generic framework which captures the electronic dynamics in real-time, thereby, allowing to handle both perturbative and non-perturbative regimes (e.g., harmonic generation, electron transport) in a unified manner. In this thesis, we discuss the more general case of RT-TDDFT.

CHAPTER III

Large-scale All-electron Density Functional Theory Calculations Using an Enriched Finite Element Basis

In this chapter, we present an efficient and systematically convergent approach to large-scale all-electron DFT calculations by employing an *enriched finite element* basis [37]. We begin with an introduction of the concept of all-electron (AE) and pseudopotential (PSP) DFT and highlight the need for all-electron DFT calculations in Sec. 3.1. In Sec. 3.2, we provide a brief survey of existing numerical methods for all-electron DFT calculation, and motivate the need for an *enriched* finite element basis for all-electron DFT calculations. In Sec. 3.3, we revisit the Kohn-Sham eigenvalue problem, in the context of real space formulation of DFT. Sec. 3.4 briefly introduces the standard finite element discretization for the Kohn-Sham problem. Sec. 3.5 details the proposed enriched finite element discretization for the Kohn-Sham eigenvalue problem, and Sec. 3.6 discusses the key ideas based on Chebyshev polynomial filtering employed in the self-consistent field iteration (SCF) solution procedure. Sec. 3.7 presents the convergence, accuracy, performance and parallel scalability of the enriched finite element basis. In the same section, we provide the comparison of the enriched finite basis against classical finite element and Gaussian basis. Finally, we

summarize the findings from the present work and outline the future scope in Sec 3.8.

3.1 All-electron and Pseudopotenial DFT

The *all-electron* (AE) DFT pertains to the case of solving the Kohn-Sham eigenvalue problem (cf. Eq. 2.20) for all the occupied orbitals in presence of the full nuclear potential. To elaborate, in all-electron DFT, the external potential ($v_{\text{ext}}^{\text{ae}}$) takes the form of $v_{\text{ext}}^{\text{ae}}(\mathbf{r}) = -\sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$. Additionally, the set of Kohn-Sham orbitals that are solved for includes all the electrons—core, valence, and (possibly) few unoccupied electrons. Thus, within a given exchange-correlation approximation, all-electron DFT presents an accurate description of the physical system. Two key observations are to be made regarding all-electron DFT. First is the straightforward observation that $v_{\text{ext}}^{\text{ae}}(\mathbf{r})$ is singular at the positions of the nuclei (i.e., at $\mathbf{R}_I$). The second, less apparent, observation is that the orbitals (especially the valence ones) have a highly oscillatory profile near the nuclei. To elaborate, the core electrons are, usually, localized near the nuclei. Since the Kohn-Sham orbitals must be mutually orthogonal, the valence orbitals must oscillate rapidly in the core region to maintain its orthogonality with the core orbitals. As a consequence, an all-electron calculation requires a highly refined spatial discretization near the nuclei, so as to accurately describe the singular $v_{\text{ext}}^{\text{ae}}$ as well as the sharp variations in the Kohn-Sham orbitals. Thus, all-electron DFT calculations have a high computational cost. As an efficiency measure, most DFT calculations employ the *pseudopotential* approximation [95, 96, 97, 98], wherein the Kohn-Sham eigenvalue problem is reduced to the evaluation of smooth pseudo-orbitals corresponding to the valence electrons in the presence of a smooth effective external potential, namely the pseudopotential ($v_{\text{ext}}^{\text{psp}}(\mathbf{r})$). The pseudopotential approximation, in turn, rests on two observations—(i) the core-electrons are, typically, chemically inactive, and hence, hardly exhibit any change from their atomic states when they are brought in the vicinity of other atoms; (ii) the valence electrons, typically, ex-

hibit change only in a binding region that lies beyond a certain core region around the nuclei. Thus, it is reasonable to replace the singular $v_{\text{ext}}^{\text{ae}}$ along with the core-electrons with a smooth effective pseudopotential, and replace the oscillatory valence orbitals with smooth pseudo-orbitals. The construction of a pseudopotential, which is non-unique, entails, for a given atom, matching the pseudo-orbitals for the valence electrons to its corresponding all-electron orbitals, outside the user defined atomic core. Subsequently, they are transferred for use in molecules and crystals. Fig. 3.1 presents a schematic of the above process of constructing a pseudopotential. In the past few decades, pseudopotentials have seen a rapid evolution from norm-conserving potentials [99, 100, 101, 102] to ultrasoft potentials [103] to the state-of-the-art projector augmented wave (PAW) [104] method and have proven to be successful in predicting bulk, mechanical, electrical, magnetic, and chemical properties for a wide range of materials.

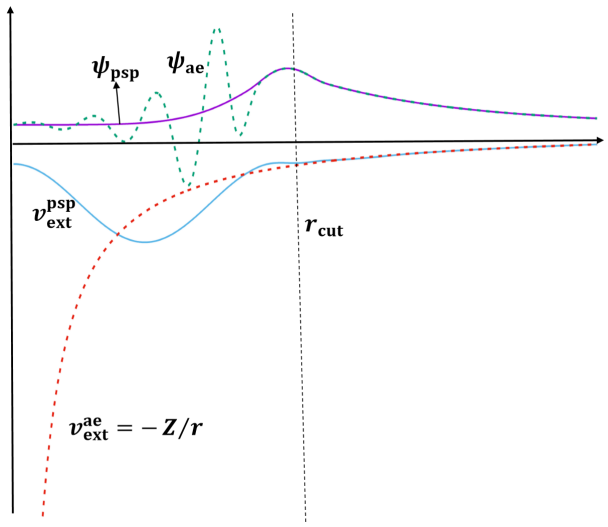


Figure 3.1: Schematic of a pseudopotential. The $v_{\text{ext}}^{\text{psp}}$ is constructed such that the ψ_{psp} matches its corresponding ψ_{ae}, beyond a cutoff radius (r_{cut}).

However, despite their success, pseudopotentials are often sensitive to the choice of core size used in their construction and tend to oversimplify the treatment of core electrons as chemically inert for various systems and external conditions. For

example, in systems under high pressure where the core and valence orbitals show increasing overlap with pressure, pseudopotentials tend to under-predict their phase transition pressures [6, 7, 8]; in systems at high temperature, where the contribution of core electrons to various thermodynamic potentials is non-negligible, pseudopotentials provide an inaccurate description of the equation of state [9]; in transition metals, where the penultimate d and f orbitals are not tightly bound, non-inclusion of these orbitals as valence electrons oftentimes lead to inaccurate bulk property prediction. More pronounced inaccuracies and sensitivity to core sizes are observed in prediction of ionization potentials [10], magnetizability [11], spectroscopic properties [12, 13] of heavier atoms wherein scalar relativistic pseudopotentials are widely employed, and in prediction of band-gap and excited state properties [14]. Thus, all-electron calculations are necessary for an accurate and more reliable description of such systems and conditions.

3.2 Enriched finite element basis—Motivation

The earliest and the most commonly employed method for all-electron calculations involves the use of atomic-orbital-type basis sets [20, 21, 22, 23, 24, 25, 26, 27, 28], wherein atom-specific basis, either analytic or numerical, are used with only a few basis functions per atom. However, owing to the incompleteness of the basis, systematic convergence for all materials systems remains a concern. Moreover, in many numerical implementations, their applicability is largely limited to isolated systems and are not easily amenable to arbitrary boundary conditions. Furthermore, the non-locality of the basis substantially limits parallel scalability of their numerical implementations. Among the family of complete basis sets, the plane-wave basis, owing to the straightforward evaluation of the Coulomb interactions in Fourier space and the exponential convergence afforded by the basis, has been the most popular choice for pseudopotential calculations. However, its applicability to all-electron calculations is

greatly hindered by its lack of adaptive spatial resolution, as any computationally efficient basis for all-electron calculations warrants finer resolution closer to nuclei, where the wavefunctions are rapidly oscillating, and coarser resolution elsewhere. This shortcoming has been, to a large extent, alleviated through the use of various augmentation schemes such as Augmented Plane-wave (APW) [29, 30], Linearized Augmented Plane-wave (LAPW) [31, 32, 33] and APW+lo [34, 35, 36]. All these methods involve the description of the wavefunctions in terms of products of radial functions and spherical harmonics inside muffin-tins (MTs) surrounding each atom, and in terms of plane-waves in the interstitial regions between atoms. Although these schemes attain adaptive spatial resolution, the basis functions within the MTs depend on the choice of trial energy parameters, typically based on atomic energies, for each azimuthal (l) quantum number. Owing to the lack of one-to-one correspondence between the Kohn-Sham eigenvalues and the trial energy parameters, the quality of the basis is sensitive to the choice of trial energy parameters, especially in systems where the chosen l quantum number based trial energies fail to describe all states with the same l-character, and in systems where the occupied bands show substantial deviation from their atomic counterparts [36]. Additionally, certain notable disadvantages of plane-waves such as their restriction to periodic boundary conditions, the highly non-local communication associated with Fast Fourier Transform (FFT), also persist in these augmentation schemes.

Blöchl, in his PAW formulation [104], generalized the notion of APW/LAPW and the pseudopotential approach to construct the all-electron orbitals through a linear transformation, $\hat{\mathscr{T}}$, of the smoothly varying pseudo-orbitals, thus providing a balance between accuracy and computational efficiency. However, typically, PAW is implemented within the frozen-core approximation, wherein, although the oscillatory behavior of the valence orbitals near the atomic centers is retrieved through $\hat{\mathscr{T}}$ acting on the pseudo valence orbitals, the core states are treated as frozen and do not

feature within the self-consistent field iteration. One can, in principle, relax the core states within the PAW framework, however, this involves achieving simultaneous self-consistency in core states, valence partial waves and the effective potential, which can severely affect the computational efficiency otherwise afforded by frozen-core approximation. Marsman *et al.* [105] presented a computationally efficient extension of PAW beyond the frozen-core approximation, wherein, first, the core states are updated self-consistently within a fixed valence charge density and a spherical approximation for the one-center potential. Subsequently, new valence partial waves are evaluated. However, as noted in that work, the spherical approximation of the one-center potential used in the core-state relaxation poses limitations in terms of accounting for core polarization effects and core-core interactions from neighboring atoms; capturing changes in valence-core interactions outside the augmentation spheres; preserving orthogonality of the valence partial waves with the core states under situations where the core charge density spills outside the augmentation spheres. Additionally, the construction of the valence all-electron and pseudo partial waves that feed into $\hat{\mathscr{T}}$, while using the actual one-center potential (crystal potential) in their construction, involves the use of trial energy parameters (analogous to APW/LAPW), thereby introducing sensitivity to the choice of these trial energies. Therefore, to account for these notable limitations, it is desirable to treat the core electrons on equal footing with the valence electrons while at the same time minimize the huge computational expense incurred by such explicit treatment of core electrons.

The limitations of plane-waves have, in the past two decades, led to the development of various real-space techniques for DFT calculations, of which the Finite Difference (FD) method [106, 107] remains the most prominent. The FD method can handle arbitrary boundary conditions, and exhibit improved parallel scalability in comparison to plane-wave basis. However, the FD method fails to retain the variational convergence of plane-waves. Moreover, a lack of basis in the FD method makes

an accurate treatment of singular potentials difficult, thereby, limiting its utility for all-electron calculations. Finite element basis [108, 109], on the other hand, being a local piecewise polynomial basis, retains the variational property of the plane-waves, and, in addition, has other desirable features such as locality of the basis that affords good parallel scalability, being easily amenable to adaptive spatial resolution, and the ease of handling arbitrary boundary conditions. While most studies employing the finite element basis in DFT calculations [39, 40, 41, 42, 44, 46, 49, 50, 51, 38] have shown its usefulness in pseudopotential calculations, some of the works [39, 47, 48, 38, 52, 53] have also demonstrated its promise for all-electron calculations. In particular, the work of Motamarri *et al.* [38] has combined the use of higher-order spectral finite elements along with Chebyshev polynomial based filtering technique to develop an efficient scheme for the computation of the occupied eigenstates. As detailed in the work, the aforementioned method outperforms the plane-wave basis in pseudopotential calculations for the benchmark systems considered. However, in the context of all-electron calculations, it remains an order of magnitude slower in comparison to the Gaussian basis. This relatively unsatisfactory performance of the finite element basis in all-electron calculations was attributed to the requirement of large number of basis functions ($\mathcal{O}(10^5)$ per atom, even for light atoms) as well as the high polynomial degree required in the Chebyshev filter ($\mathcal{O}(10^3)$) to accurately compute the occupied eigenstates. To elaborate, one requires a highly refined finite element mesh closer to the atomic cores in order to capture the sharp variations in the electronic wavefunctions and the singularity of the nuclear potential. This refinement, in turn, leads to an increase in the spectral width of the discrete Kohn-Sham Hamiltonian, thereby, requiring a very high polynomial degree Chebsyhev filter to compute the occupied eigenstates. This need for a high polynomial degree Chebyshev filter in all-electron calculations also negatively effects the computational complexity realized through reduced order scaling methods. As detailed in a recent work [53], which combines

Chebyshev filtered subspace projection with localization and Fermi-operator expansion, while pseudopotential calculations exhibited linear scaling for materials systems with a band-gap and subquadratic scaling for materials systems without a band gap, the overall scaling for all-electron calculations was close to quadratic even for materials with a band-gap.

In order to alleviate the aforementioned limitations of finite element basis in all-electron calculations, we propose employing a mixed basis comprising of finite element basis functions and compactly supported atomic-orbital-type basis functions. In particular, the atomic-orbital-type functions capture the essential features of the electronic fields near the nuclei, thereby, mitigating the need for high mesh refinement around atoms, while the finite element basis functions capture the smooth parts of the wavefunction away from the nuclei and also extend completeness to the basis. In this work, we formalize this idea of a mixed basis to develop, what we refer to as, the enriched finite element basis. The enriched finite element basis is generated by augmenting the piecewise continuous Lagrange polynomials in finite element basis, henceforth described as the classical finite element basis, with compactly supported atom-centered numerical basis functions that are obtained from the solutions of the Kohn-Sham problem (Kohn-Sham orbitals and electrostatic potentials) for single atoms. We term these atom-centered numerical basis functions as enrichment functions. We note that the proposed enriched finite element basis differs from other augmentation schemes in plane-waves like APW, LAPW, and APW+lo in the following ways: (i) unlike the plane-wave augmentation schemes, the enriched finite element basis does not partition the space into muffin tins (MTs) and interstitials, thereby eliminating the need of any matching or smoothness constraint for the augmenting basis functions; (ii) as opposed to the plane-wave augmentation schemes, the enrichment functions of our proposed method do not have any trial energy parameter dependence; and (iii) unlike the plane-wave augmentation scheme, wherein

the basis functions inside the MTs needs to be computed for every materials system separately, the enrichment functions, being atomic solutions to the electronic fields, are independent of the materials system and are computed *a priori*.

The key ideas in the proposed method involve: (i) pre-computing the enrichment functions by solving radial Kohn-Sham equations and employing smooth cutoff functions to ensure the locality as well as control the conditioning of the enriched finite element basis; (ii) employing a divide and conquer strategy to construct an adaptive quadrature grid based on the nature of enrichment functions, so as to accurately and efficiently evaluate the integrals involving enrichment functions; (iii) implementing an efficient scheme to evaluate the inverse of the overlap matrix corresponding to the enriched finite element basis by using block-wise matrix inversion in conjunction with Gauss-Lobatto-Legendre reduced order quadrature rules; and (iv) in each self-consistent field iteration, using a Chebyshev polynomial based filter to compute the space spanned by the occupied eigenstates, and solving the Kohn-Sham eigenvalue problem by projecting the problem onto this Chebyshev-filtered space. We have implemented the proposed method in a parallel computing framework using the Message Passing Interface (MPI) to enable large-scale all-electron calculations. To begin with, we demonstrate optimal convergence rates of the ground-state energies with respect to enriched finite element basis. Further, we investigate the accuracy and performance of the proposed method on benchmark semi-conducting (silicon nano-clusters) and heavy-metallic (gold nano-clusters) systems of various sizes, with the largest system containing 8694 electrons. The proposed formulation using the enriched finite element basis obtains close to 1 mHa accuracy in per-atom ground-state energies of the benchmark systems when compared to the reference ground-state energies obtained from classical finite element basis or Gaussian basis calculations. Furthermore, the proposed method achieves a staggering $50-300$ fold speedup relative to the classical finite element basis, and a significant speedup relative to the Gaussian basis even

for modest sized systems. Lastly, we observe good parallel efficiency of our implementation up to 384 processors for a silicon nano-cluster containing 3920 electrons discretized using $\sim$4 million basis functions.

3.3 Real space formulation of Kohn-Sham DFT

We recall that the ground-state properties of a materials system in the Kohn-Sham DFT framework can be computed by solving the non-linear Kohn-Sham eigenvalue problem [110], given by

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{KS}}[\rho, \mathbf{R}]\right)\psi_\alpha = \epsilon_\alpha \psi_\alpha, \quad \alpha = 1, 2, \ldots, \tag{3.1}$$

where ϵ_α and ψ_α denote the eigevalues and the corresponding eigenfunctions of the Kohn-Sham Hamiltonian, respectively, ρ is the electron charge density of the non-interacting system, $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_a}\}$ is the collective representation for all nuclear positions in the system, and $v_{\text{KS}}[\rho, \mathbf{R}]$ is the effective single-electron Kohn-Sham potential. In the present work, we limit our discussion to a non-periodic setting and spin-independent Hamiltonian. However, we note that all the ideas discussed subsequently can be generalized, in a straightforward manner, to periodic [44] or semi-periodic systems, and spin-dependent Hamiltonians [111].

The effective single-electron potential, $v_{\text{KS}}[\rho, \mathbf{R}]$, is same as that defined Eq. 2.24, i.e.,

$$v_{\text{KS}}[\rho, \mathbf{R}] = v_{\text{xc}}[\rho] + v_{\text{H}}[\rho] + v_{\text{ext}}(\mathbf{r}, \mathbf{R})\,. \tag{3.2}$$

In the present work, we have used the local density approximation (LDA) [111] for the exchange-correlation, specifically, the Ceperley-Alder [112] form. We recall that the Hartree potential, $v_{\text{H}}[\rho]$, and the external potential, $v_{\text{ext}}(\mathbf{r}, \mathbf{R})$, in Eq. 3.2 are the classical electrostatic potentials corresponding to the electron charge density and

nuclear charges, respectively, and are given by

$$v_{\mathrm{H}}[\rho] = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' \,, \tag{3.3}$$

$$v_{\mathrm{ext}}(\mathbf{r}, \mathbf{R}) = -\sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \,, \tag{3.4}$$

where Z_I denotes the atomic number of the Ith nucleus in the system.

We note that both the electrostatic potentials—Hartree (v_{H}) and external potential (v_{ext})—are extended in real space. However, noting that the $\frac{1}{|\mathbf{r}|}$ kernel in these extended interactions is the Green's function of the Laplace operator, one can reformulate their evaluation as solutions of the Poisson problems, given by

$$-\frac{1}{4\pi}\nabla^2 v_{\mathrm{H}}(\mathbf{r}) = \rho(\mathbf{r}) \,, \tag{3.5a}$$

$$-\frac{1}{4\pi}\nabla^2 v_{\mathrm{ext}}(\mathbf{r}, \mathbf{R}) = b(\mathbf{r}, \mathbf{R}) \,. \tag{3.5b}$$

In the above Eq. 3.5b, we define $b(\mathbf{r}, \mathbf{R}) = -\sum_I^{N_a} Z_I \widetilde{\delta}(\mathbf{r}, \mathbf{R}_I)$, where $\widetilde{\delta}(\mathbf{r}, \mathbf{R}_I)$ is a Dirac distribution centered at $\mathbf{R}_I$. We refer to previous works on finite element based DFT calculations [41, 44, 49, 113, 38] for a comprehensive treatment of the local reformulation of electrostatic potentials into Poisson problems.

The electron charge density, the central quantity of interest in DFT, is given in terms of the eigenfunctions in Eq. 3.1 as:

$$\rho(\mathbf{r}) = 2\sum_{\alpha} f(\epsilon_\alpha, \mu)|\psi_\alpha(\mathbf{r})|^2 \,, \tag{3.6}$$

where $f(\epsilon, \mu)$ is the orbital occupancy function and μ is the Fermi energy. Typically, in DFT calculations the orbital occupancy function f is chosen as the Fermi-Dirac

distribution [114, 115], given by

$$f(\epsilon, \mu) = \frac{1}{1 + \exp(\frac{\epsilon - \mu}{k_B T})}\,, \tag{3.7}$$

where k_B denotes the Boltzman constant and T is the temperature used for smearing the orbital occupancy function. The Fermi energy, μ, is evaluated from the constraint on the total number of electrons (N_e) in the system, given by

$$\int \rho(\mathbf{r})\, d\mathbf{r} = 2 \sum_{\alpha} f(\epsilon_\alpha, \mu) = N_e\,. \tag{3.8}$$

A reader might note the subtle difference in the expression of $\rho(\mathbf{r})$ given in Eq. 3.6 and Eq. 2.21. The difference of a factor 2 arises because in this chapter we are considering only spin unpolarized systems, i.e., systems where the Kohn-Sham orbitals are doubly occupied. The other difference is the use of a Fermi-Dirac distribution instead of a Heavyside function. The choice of a Fermi-Dirac distribution is made over a Heavyside function to avoid charge sloshing, wherein systems with degenerate energy levels at Fermi energy can exhibit large spatial deviation in electron charge density with SCF iterations on the account of different degenerate orbitals being occupied at different SCF iterations.

Finally, upon solving Eqs. 3.1, 3.6 and 3.8 self-consistently, the ground-state energy of the materials system is computed as

$$E_{\text{tot}} = E_{\text{band}} + E_{\text{xc}} - \int v_{\text{xc}}[\rho]\rho\, d\mathbf{r} - \frac{1}{2}\int \rho v_{\text{H}}[\rho]\, d\mathbf{r} + E_{\text{NN}}\,, \tag{3.9}$$

where E_{band} is the band energy, given by

$$E_{\text{band}} = 2 \sum_{\alpha} f(\epsilon_\alpha, \mu)\epsilon_\alpha\,, \tag{3.10}$$

and E_{NN} is the nuclear-nuclear repulsive energy (same as in Eq. 2.10), given by

$$E_{\text{NN}} = \frac{1}{2} \sum_{I=1}^{N_a} \sum_{\substack{J=1 \\ J \neq I}}^{N_a} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \,. \tag{3.11}$$

3.4 Classical finite element method

In this section, we briefly discuss the discretization of the Kohn-Sham eigenvalue problem using the classical finite element basis. In particular, we comment on the usefulness of higher-order spectral finite elements, employed in this work. We discuss how the use of spectra finite elements in conjunction with the reduced order Gauss-Lobatto-Legendre (GLL) quadrature rule enables an efficient inversion of the overlap matrix of the classical finite element basis functions.

3.4.1 Classical finite element discretization

In the finite element method, the spatial domain of interest is discretized into subdomains called finite elements. The finite element basis is constructed from piecewise polynomial functions that have a compact support on the finite elements, thus rendering locality to these basis functions. We note that there is an abundance of choice in terms of the form and order of the polynomial functions that can be used in constructing the finite element basis, and we refer to [109, 116] for a comprehensive discourse on the subject. Henceforth, we refer to the standard notion of finite element basis as the classical finite element basis in order to differentiate from the proposed enriched finite element basis in Sec 3.5, and refer to the corresponding discrete formulation as the classical finite element discretization.

Let X_h denote the finite element subspace of dimension n_h constructed from a finite element mesh with a characteristic mesh-size h. Let ψ_α^h and ϕ^h denote the classical finite element discretized fields corresponding to the Kohn-Sham orbitals

and the electrostatic potential (generically representing both Hartree and external potential), respectively, that are expressed as

$$\psi_\alpha^h(\mathbf{r}) = \sum_{j=1}^{n_h} N_j^C(\mathbf{r})\psi_{\alpha,j}^C \qquad \alpha = 1, 2, \dots , \tag{3.12}$$

$$\phi^h(\mathbf{r}) = \sum_{j=1}^{n_h} N_j^C(\mathbf{r})\phi_j^C \, . \tag{3.13}$$

The superscript C, in the above expressions and elsewhere in this chapter, is used to indicate the discretization based on classical finite element basis. Here $N_j^C : 1 \leq j \leq n_h$ denote the classical finite element basis functions spanning X_h, and $\psi_{\alpha,j}^C$ and ϕ_j^C are the coefficients corresponding to j^{th} basis function (N_j^C) in the expansion of the α^{th} Kohn-Sham orbital and electrostatic potential, respectively.

Using the classical finite element discretization in Eq. 3.12, the Kohn-Sham eigenvalue problem in Eq. 3.1 reduces to the following discrete form,

$$\mathbf{H}^C \boldsymbol{\psi}_\alpha^C = \epsilon_\alpha^C \mathbf{M}^C \boldsymbol{\psi}_\alpha^C \, , \tag{3.14}$$

where $\mathbf{H}^C$ denotes the discrete Kohn-Sham Hamiltonian, $\mathbf{M}^C$ denotes the overlap matrix of the classical finite element basis, ϵ_α^C denotes the α^{th} discrete Kohn-Sham eigenvalue, and $\boldsymbol{\psi}_\alpha^C$ denotes the corresponding eigenvector containing the expansion coefficients $\psi_{\alpha,j}^C$. For a non-periodic problem defined on a domain Ω with homogeneous Dirichlet boundary conditions, the discrete Hamiltonian matrix H_{jk}^C is given by

$$\begin{aligned} H_{jk}^C = \frac{1}{2} \int_\Omega \nabla N_j^C(\mathbf{r}) . \nabla N_k^C(\mathbf{r}) \, d\mathbf{r} \, + \\ \int_\Omega v_{\mathrm{KS}}^h(\mathbf{r}, \mathbf{R}) N_j^C(\mathbf{r}) N_k^C(\mathbf{r}) \, d\mathbf{r} \, . \end{aligned} \tag{3.15}$$

Although the above expression is for a non-periodic problem, it can be easily extended

to a periodic problem on a unit cell using the Bloch theorem [44]. We note that owing to the non-orthogonality of the classical finite element basis, the overlap matrix $\mathbf{M}^C$, defined by $M^C_{jk} = \int_\Omega N^C_j(\mathbf{r}) N^C_k(\mathbf{r})\, d\mathbf{r}$, is not an identity matrix, thereby, resulting in a generalized eigenvalue problem. However, utilizing the symmetric positive definiteness, and hence the invertibility of $\mathbf{M}^C$, we can transform the generalized eigenvalue problem in Eq. 3.14 to a standard eigenvalue problem, given by

$$(\mathbf{M}^C)^{-1}\mathbf{H}^C\boldsymbol{\psi}^C_\alpha = \epsilon^C_\alpha \boldsymbol{\psi}^C_\alpha \,. \tag{3.16}$$

We remark that this transformation of the generalized eigenvalue problem to a standard eigenvalue problem is essential for the use of Chebyshev polynomial based acceleration technique to compute the occupied eigenspace (to be discussed in the Section 3.6). Further, we note that this transformation to a standard eigenvalue problem relies on computationally efficient methods for computing $(\mathbf{M}^C)^{-1}$, which forms the basis for our use of spectral finite elements along with Gauss-Lobatto-Legendre quadrature rule, as will be discussed in Section 3.4.2.

Turning to the Poisson problems in Eq. 3.5, and using the discretization in Eq. 3.13, we obtain the following system of linear equations,

$$\mathbf{A}^C\boldsymbol{\phi}^C = \mathbf{b}^C \,, \tag{3.17}$$

where $\mathbf{A}^C$ represents the Laplace operator discretized in the classical finite element basis that is given by

$$A^C_{jk} = \frac{1}{4\pi}\int_\Omega \nabla N^C_j(\mathbf{r}).\nabla N^C_k(\mathbf{r})\, d\mathbf{r} \,, \tag{3.18}$$

$\boldsymbol{\phi}^C$ is the electrostatic potential vector containing the expansion coefficients ϕ^C_j, and

$\mathbf{b}^C$, the forcing vector, is given by

$$b_i^C = \int_\Omega g(\mathbf{r}) N_i^C(\mathbf{r})\, d\mathbf{r}\,, \tag{3.19}$$

where $g(\mathbf{r}) = \rho(\mathbf{r})$ or $g(\mathbf{r}) = b(\mathbf{r}, \mathbf{R})$ for the Hartree and external potential, respectively.

3.4.2 Spectral finite elements

As opposed to conventional classical finite element basis, which is typically constructed from a tensor product of Lagrange polynomials interpolated through equidistant nodal points in an element, spectral finite element basis employ a distribution of nodes obtained from the roots of the derivative of Legendre polynomials or the Chebyshev polynomials [117]. In the present work, we employ the Gauss-Lobatto-Legendre node distribution, where the nodes are located at the roots of the derivative of the Legendre polynomial and the boundary points. The resulting spectral finite element basis has been shown to provide better conditioning with increasing polynomial degree [117] and has been effective for electronic structure calculations using higher-order finite element discretization [113, 38]. However, the major advantage of this spectral finite element basis is realized when it is used in conjunction with Gauss-Lobatto-Legendre (GLL) quadrature rule [118] for evaluation of the integrals arising in the overlap matrix, wherein the quadrature points are coincident with the nodal points in the spectral finite element. Such a combination renders the overlap matrix $\mathbf{M}^C$ in the discrete Kohn-Sham eigenvalue problem diagonal, thereby making the transformation of the generalized eigenvalue problem in Eq. 3.14 to the standard eigenvalue problem in Eq. 3.16 to be trivial. We note that while an n point rule in the conventional Gauss quadrature rule integrates polynomials exactly up to degree $2n-1$, an n point GLL quadrature rule integrates polynomials exactly only up to

degree $2n-3$. Thus, we employ the GLL quadrature rule only in the construction of $\mathbf{M}^C$, while the more accurate Gauss quadrature rule is used for all other integrals featuring in the Kohn-Sham eigenvalue problem as well as the Poisson problems for the electrostatic potentials. We refer to Motamarri *et al.* [38] for a discussion on the accuracy and sufficiency of GLL quadrature in the evaluation of overlap matrix $\mathbf{M}^C$. Since we employ spectral finite elements all throughout the present work, any reference to classical finite elements, henceforth in this chapter, corresponds to spectral finite elements.

3.5 Enriched finite element method

In this section, we first discuss the proposed enriched finite element discretization for the Kohn-Sham eigenvalue problem. Then, we present various numerical and algorithmic strategies for efficient use of the enriched finite element basis.

3.5.1 Enriched finite element discretization

In enriched finite element discretization we augment the classical finite element basis by appending additional atom-centered basis functions called enrichment functions. We write the enriched finite element discretization for the Kohn-Sham orbitals, ψ_α^h, and the electrostatic potentials (both Hartree and external potential), ϕ^h, as follows:

$$\psi_\alpha^h(\mathbf{r}) = \underbrace{\sum_{j=1}^{n_h} N_j^C(\mathbf{r})\psi_{\alpha,j}^C}_{\text{Classical}} + \underbrace{\sum_{I=1}^{N_a}\sum_{k=1}^{n_I} N_{k,I}^{E,\psi}(\mathbf{r},\mathbf{R}_I)\psi_{\alpha,k,I}^E}_{\text{Enriched}}\,, \tag{3.20a}$$

$$\phi^h(\mathbf{r}) = \underbrace{\sum_{j=1}^{n_h} N_j^C(\mathbf{r})\phi_j^C}_{\text{Classical}} + \underbrace{\sum_{I=1}^{N_a} N_I^{E,\phi}(\mathbf{r},\mathbf{R}_I)\phi_I^E}_{\text{Enriched}}\,. \tag{3.20b}$$

In the above expressions, the superscripts C and E are used to distinguish between classical and enriched components, respectively. As with the classical finite element

discretization, N_j^C denotes the j^{th} classical finite element basis, and $\psi_{\alpha,j}^C$ and ϕ_j^C are the expansion coefficients corresponding to N_j^C for the α^{th} Kohn-Sham orbital and the electrostatic potential, respectively. In addition, we have the enrichment functions $N_{k,I}^{E,\psi}$ and $N_I^{E,\phi}$ for the Kohn-Sham orbitals and the electrostatic potentials, respectively, each centered on atom I located at $\mathbf{R}_I$. $\psi_{\alpha,k,I}^E$ denotes the expansion coefficient corresponding to $N_{k,I}^{E,\psi}$ for the α^{th} Kohn-Sham orbital, and ϕ_I^E denotes the expansion coefficient corresponding to $N_I^{E,\phi}$ for the electrostatic potential. The enrichment functions, $N_{k,I}^{E,\psi}$ and $N_I^{E,\phi}$, are the atom-centered numerical solutions to the Kohn-Sham orbitals and electrostatic potentials, respectively, for the atom type (chemical element) at $\mathbf{R}_I$. The index I runs over all the atoms, N_a, in the materials system, and the index k in Eq. 3.20a runs over the number of atomic Kohn-Sham orbitals, n_I, we would want to include for the atom I. Typically, we include all the occupied and a few lowest unoccupied atomic orbitals for a given atom I. We note that although we have represented the enrichment functions for both Hartree and external potential as $N_I^{E,\phi}$, they differ based on the electrostatic potential that is being discretized.

We now discuss the procedure to generate the enrichment functions. As aforementioned, the enrichment functions are chosen as the solutions to the Kohn-Sham orbitals and electrostatic potentials for any given single atom. Under the assumption of equal fractional occupancy for degenerate orbitals, the charge density for a single atom is spherically symmetric, which in turn, results in spherically symmetric $v_{\text{xc}}[\rho]$ and $v_{\text{H}}[\rho]$. Thus, rewriting the Eqs. 3.1 and 3.5 in spherical coordinates and using separation of variables, we obtain the following radial equations for any single atom with the atom type denoted by a superscript S:

$$-\frac{1}{4\pi}\frac{1}{r^2}\frac{d}{dr}\big(r^2\frac{d}{dr}\big)\phi^S(r) = g^S(r)\,, \tag{3.21a}$$

$$\Big[-\frac{1}{2}\frac{1}{r^2}\frac{d}{dr}\big(r^2\frac{d}{dr}\big) + \frac{l(l+1)}{r^2} + V^S_{\mathrm{KS}}(r)\Big] R^S_{nl}(r) = \epsilon^S_{nl} R^S_{nl}(r)\,, \tag{3.21b}$$

$$\rho^S(r) = 2\sum_n \sum_l \frac{2l+1}{4\pi} f(\epsilon^S_{nl}, \mu^S)\big(R^S_{nl}(r)\big)^2\,. \tag{3.21c}$$

In Eq. 3.21a, $\phi^S(r)$ denotes either the Hartree or the external potential; $g^S(r)$ denotes the charge density $\rho^S(r)$ or the nuclear charge $b^S(r) = Z_S\widetilde{\delta}(0)$ with Z_s denoting the atomic number, depending on whether $\phi^S(r)$ represents the Hartree or the external potential, respectively. In Eq. 3.21b, $R^S_{nl}(r)$ represents the radial part of the Kohn-Sham orbital corresponding to the principal quantum number n and azimuthal quantum number l. Equations in 3.21 are solved self-consistently until convergence in $\rho^S(r)$ is achieved. We note that these radial equations can be solved inexpensively using a 1D classical finite element mesh comprising of, typically, $1000 - 5000$ basis functions. Moreover, the radial atomic solutions can be pre-computed for all atom types spanning the periodic table and stored for later use in constructing the enrichment functions.

Having evaluated the radial part $R^S_{nl}(r)$, the full Kohn-Sham orbital is obtained by multiplying it with spherical harmonics as follows

$$\psi^S_{nlm}(r, \beta, \gamma) = R^S_{nl}(r) Y_{lm}(\beta, \gamma)\,, \tag{3.22}$$

where $Y_{lm}(\beta, \gamma)$ denotes the real form of spherical harmonics for the pair of azimuthal quantum number l and magnetic quantum number m, and β and γ represent the polar and azimuthal angles, respectively. Using the above atomic solutions, we write the orbital enrichment function $N^{E,\psi}_{k,I}$ centered at atom I of atom type S as

$$N^{E,\psi}_{k,I}(\mathbf{r}, \mathbf{R}_I) = \psi^S_{nlm}(|\mathbf{r} - \mathbf{R}_I|, \beta_{\mathbf{R}_I}, \gamma_{\mathbf{R}_I})\,, \tag{3.23}$$

where the index k represents a specific choice of n, l and m, and $\beta_{\mathbf{R}_I}$ and $\gamma_{\mathbf{R}_I}$ are the polar and azimuthal angles, respectively, for the point $\mathbf{r}$ with $\mathbf{R}_I$ as the origin.

Similarly, we define the electrostatic enrichment function $N_I^{E,\phi}(\mathbf{r})$ centered at atom I of atom type S as

$$N_I^{E,\phi}(\mathbf{r}, \mathbf{R}_I) = \phi^S(|\mathbf{r} - \mathbf{R}_I|)\,. \tag{3.24}$$

Henceforth in the paper, to make our notation of the enrichment functions more succinct, we combine the indices k and I into a single index denoted by q (or r) for the orbital enrichment functions and their coefficients, and drop the argument $\mathbf{R}_I$ in the enrichment functions. Furthermore, we define $n_E^{\psi} = \sum_{I=1}^{N_a} n_I$ to denote the total number of enrichment functions in the materials system used for discretization of any Kohn-Sham orbital ψ_α.

Discretizing the Kohn-Sham eigenvalue problem in the enriched finite element basis, we obtain a standard eigenvalue equation analogous to its classical counterpart (Eq. 3.16), and is given by

$$(\mathbf{M}^E)^{-1}\mathbf{H}^E\boldsymbol{\psi}_\alpha^E = \epsilon_\alpha^E\boldsymbol{\psi}_\alpha^E\,, \tag{3.25}$$

where $\mathbf{H}^E$ and $\mathbf{M}^E$ are the discrete Kohn-Sham Hamiltonian matrix and overlap matrix in the enriched finite element basis, ϵ_α^E denotes the α^{th} discrete Kohn-Sham eigenvalue and $\boldsymbol{\psi}_\alpha^E$ denotes the corresponding eigenvector containing the expansion coefficients $\psi_{\alpha,j}^C$ and $\psi_{\alpha,q}^E$ (defined in Eq. 3.20a). Both $\mathbf{H}^E$ and $\mathbf{M}^E$ matrices have a 2×2 block structure, given by

$$\mathbf{H}^E = \left[\begin{array}{c|c} \mathbf{H^{cc}} & (\mathbf{H^{ec}})^T \\ \hline \mathbf{H^{ec}} & \mathbf{H^{ee}} \end{array}\right] \tag{3.26}$$

$$\mathbf{M}^E = \left[\begin{array}{c|c} \mathbf{M^{cc}} & (\mathbf{M^{ec}})^T \\ \hline \mathbf{M^{ec}} & \mathbf{M^{ee}} \end{array}\right] \tag{3.27}$$

where $\mathbf{H^{cc}}$ and $\mathbf{M^{cc}}$ are the classical-classical blocks which comprise of matrix elements involving only the classical finite element basis functions and are same as the $\mathbf{H}^C$ and $\mathbf{M}^C$ matrices appearing in Eq. 3.14, respectively; $\mathbf{H^{ec}}$ and $\mathbf{M^{ec}}$ are the enriched-classical blocks containing the cross-term matrix elements involving both classical finite element basis functions and enrichment functions; and $\mathbf{H^{ee}}$ and $\mathbf{M^{ee}}$ are the enriched-enriched blocks comprising of matrix elements involving only the enrichment functions. Each of these blocks are given by

$$H^{cc}_{jk} = \frac{1}{2}\int_{\Omega} \nabla N^C_j(\mathbf{r}).\nabla N^C_k(\mathbf{r})\, d\mathbf{r} + \int_{\Omega} v^h_{\mathrm{KS}}(\mathbf{r},\mathbf{R}) N^C_j(\mathbf{r}) N^C_k(\mathbf{r})\, d\mathbf{r}\,, \tag{3.28a}$$

$$H^{ec}_{qj} = \frac{1}{2}\int_{\Omega} \nabla N^{E,\psi}_q(\mathbf{r}).\nabla N^C_j(\mathbf{r})\, d\mathbf{r} + \int_{\Omega} v^h_{\mathrm{KS}}(\mathbf{r},\mathbf{R}) N^{E,\psi}_q(\mathbf{r}) N^C_j(\mathbf{r})\, d\mathbf{r}\,, \tag{3.28b}$$

$$H^{ee}_{qr} = \frac{1}{2}\int_{\Omega} \nabla N^{E,\psi}_q(\mathbf{r}).\nabla N^{E,\psi}_r(\mathbf{r})\, d\mathbf{r} + \int_{\Omega} v^h_{\mathrm{KS}}(\mathbf{r},\mathbf{R}) N^{E,\psi}_q(\mathbf{r}) N^{E,\psi}_r(\mathbf{r})\, d\mathbf{r}\,; \tag{3.28c}$$

$$M^{cc}_{jk} = \int_{\Omega} N^C_j(\mathbf{r}) N^C_k(\mathbf{r})\, d\mathbf{r}\,, \tag{3.29a}$$

$$M^{ec}_{qj} = \int_{\Omega} N^{E,\psi}_q(\mathbf{r}) N^C_j(\mathbf{r})\, d\mathbf{r}\,, \tag{3.29b}$$

$$M^{ee}_{qr} = \int_{\Omega} N^{E,\psi}_q(\mathbf{r}) N^{E,\psi}_r(\mathbf{r})\, d\mathbf{r}\,, \tag{3.29c}$$

where $j,k = 1,2,\ldots,n_h$ and $q,r = 1,2,\ldots,n^{\psi}_E$.

Discretizing the Poisson problems (Eq. 3.5) in the enriched finite element basis, we obtain a system of linear equations analogous to its classical counterpart (Eq. 3.17), and is given by

$$\mathbf{A}^E \boldsymbol{\phi}^E = \mathbf{b}^E\,, \tag{3.30}$$

where $\mathbf{A}^E$ represents the discrete Laplace operator in the enriched finite element basis,

and $\boldsymbol{\phi}^E$ is the electrostatic potential vector containing the expansion coefficients ϕ_j^C and ϕ_I^E (defined in Eq. 3.20b). Similar to $\mathbf{H}^E$ and $\mathbf{M}^E$, the matrix $\mathbf{A}^E$ also assumes a 2×2 block structure containing classical-classical, enriched-classical and enriched-enriched blocks, given by

$$\mathbf{A}^E = \left[\begin{array}{c|c} \mathbf{A^{cc}} & (\mathbf{A^{ec}})^T \\ \hline \mathbf{A^{ec}} & \mathbf{A^{ee}} \end{array}\right] \tag{3.31}$$

with the individual blocks defined as

$$A_{jk}^{cc} = \int_\Omega \nabla N_j^C(\mathbf{r}).\nabla N_k^C(\mathbf{r})\, d\mathbf{r}\,, \tag{3.32a}$$

$$A_{Ij}^{ec} = \int_\Omega \nabla N_I^{E,\phi}(\mathbf{r}).\nabla N_j^C(\mathbf{r})\, d\mathbf{r}\,, \tag{3.32b}$$

$$A_{IJ}^{ee} = \int_\Omega \nabla N_I^{E,\phi}(\mathbf{r}).\nabla N_J^{E,\phi}(\mathbf{r})\, d\mathbf{r}\,, \tag{3.32c}$$

where $j, k = 1, 2, \ldots, n_h$; and $I, J = 1, 2, \ldots, N_a$.

The forcing vector $\mathbf{b}^E$, is also analogous to its classical counterpart, and is defined as the composite vector

$$\mathbf{b}^E = \left[\begin{array}{c} \mathbf{b^c} \\ \hline \mathbf{b^e} \end{array}\right] \tag{3.33}$$

where $\mathbf{b^c}$ is the classical part of $\mathbf{b}^E$ and is same as $\mathbf{b}^C$ (defined in Eq. 3.19). $\mathbf{b^e}$ is the enrichment part of $\mathbf{b}^E$ and is given by

$$b_I^e = \int_\Omega g(\mathbf{r}) N_I^{E,\phi}(\mathbf{r})\, d\mathbf{r}\,, \tag{3.34}$$

where $g(\mathbf{r}) = \rho(\mathbf{r})$ or $g(\mathbf{r}) = b(\mathbf{r}, \mathbf{R})$ for the Hartree and external potential, respec-

tively, and $I = 1, 2, \ldots, N_a$.

The key idea behind augmenting the classical finite element basis with these enrichment functions is that in a multi-atom materials system, the enrichment functions, being solutions to single atom Kohn-Sham orbitals and electrostatic potentials, can effectively capture the sharp variations in the orbitals and the electrostatic potentials close to an atom, thereby eliminating the need for a refined classical finite element mesh close to an atom. Loosely speaking, the role of the classical finite element basis is now to capture the deviation of an electronic field in a materials system from that of superposition of atomic solutions for the same field. Since these deviations are much smoother in nature compared to the actual field, we can use a coarse classical finite element mesh to accurately approximate them. As will be discussed in Section 3.6, the use of a coarse classical finite element mesh results in two-fold advantage: (i) a reduction in the total degrees of freedom, and (ii) a reduction in the polynomial degree of the Chebyshev filter required to compute the occupied Kohn-Sham eigenspace.

3.5.2 Conditioning of the enriched finite element basis

The enrichment functions, being solutions to the Kohn-Sham orbitals and electrostatic potentials for a single atom, have smooth tails away from their atomic cores. These smooth tails can cause linear dependency between the enrichment functions and the classical finite element basis, thereby resulting in an ill-conditioned basis. We avoid such ill-conditioning by multiplying the enrichment functions with a smooth radial cutoff function, which generates a compact support for each enrichment function. In the present work, we employ the smooth cutoff function given by

$$h(\widetilde{r}) = \frac{u(\widetilde{r})}{u(\widetilde{r}) + u(1-\widetilde{r})}\,, \tag{3.35}$$

where $u(\widetilde{r})$ is defined as

$$u(\widetilde{r}) = \begin{cases} e^{-\frac{1}{\widetilde{r}}} & \widetilde{r} > 0 \\ 0 & \widetilde{r} \leq 0 \end{cases}$$

and $\widetilde{r} = 1 - \frac{t(r-r_0)}{r_0}$ is a linear transformation of the variable r, which offers $h(\widetilde{r})$ the following properties

$$\begin{cases} h(\widetilde{r}) = 1 & 0 \leq r < r_0 \\ 0 \leq h(\widetilde{r}) < 1 & r_0 < r \leq r_0 + \frac{r_0}{t} \\ h(\widetilde{r}) = 0 & r > r_0 + \frac{r_0}{t} \ . \end{cases}$$

We multiply the radial part of each enrichment function, $N_q^{E,\psi}(\mathbf{r})$ or $N_I^{E,\phi}$, with $h(\widetilde{r})$ to smoothly truncate them to zero. In the above expression, the parameter r_0 is called the cutoff radius, beyond which the truncation begins, and t controls the width of the transition. In the present work, we employ different values of r_0 for different enrichment functions. In particular, for an orbital enrichment function, the value of r_0 is chosen to be the farthest turning point (extremum) in the radial part of the corresponding atomic orbital. One exception to this rule is the monotonically decreasing $1s$ radial function, $R_{10}(r)$, for which the r_0 is chosen such that $|\int_0^{r_0} \left(h(\widetilde{r})R_{10}(r)\right)^2 dr - 1| < 10^{-6}$, i.e., the density arising out of the truncated $R_{10}(r)$ must integrate to within 10^{-6} of unity. The maximum of the set of r_0's corresponding to orbital enrichment functions of a given atom is selected as the cutoff radius for the electrostatic enrichment functions of the atom. We use $t \in [0.5, 1]$ to avoid sharp truncation of the enrichment functions, which may otherwise require a very high density of quadrature points in the transition region in order to accurately compute any integrals involving the gradients of these truncated enrichment functions. Figure 3.2 presents a schematic of the radial part of the truncated atomic orbital. Henceforth, enrichment functions, $N_q^{E,\psi}(\mathbf{r})$ or $N_I^{E,\phi}(\mathbf{r})$, are assumed to be truncated with the aforementioned smooth cutoff function. We remark that, in addition to

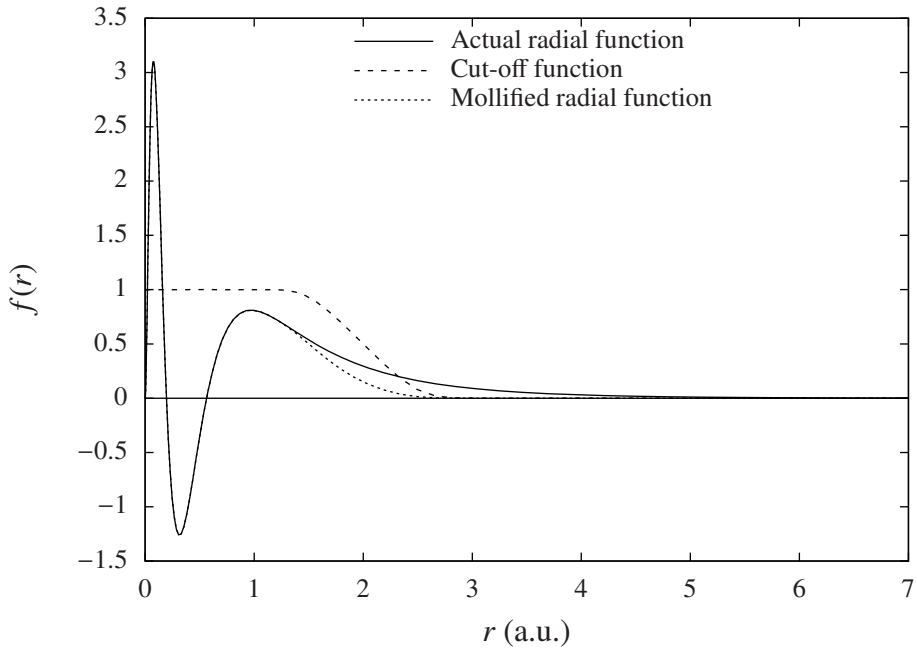


Figure 3.2: Schematic of truncated atomic orbital (radial part).

improving the conditioning of the basis, the truncation renders locality to the enrichment functions, which in turn renders sparsity to the discrete Hamiltonian, Laplacian and overlap matrices.

We note that several prior efforts have been made towards the generation of compactly supported (finite-range) atom-centered orbitals by employing different forms of confining potentials in the atomic Kohn-Sham equation, ranging from hard-wall potential [119] to polynomial [120, 121] to smooth exponential potential [25]. Other efforts [122, 123] were made to variationally optimize the parameters in the confining potential to strike a good balance between the locality and accuracy of the resultant basis. In our view, all these approaches can be adapted as an alternative to our approach of using smooth cutoff function.

3.5.3 Adaptive quadrature rule

We note that sharp gradients in regions close to atomic centers and cusps at atomic centers are characteristics of enrichment functions. Therefore, in order to accurately

compute any integral involving an enrichment function, we need a high quadrature density near the atomic centers, while a lower quadrature density may suffice in regions farther away from atomic centers. To this end, we employ an adaptive refinement of the quadrature grid on each finite element based on the characteristics of the enrichment functions. Specifically, we follow a divide and conquer strategy proposed in previous efforts [124, 125, 126], and outline here the main idea and specifics of our implementation for hexahedral finite elements employed in this work. For any given finite element, we begin by assigning it to be the parent element Ω^p. Further, we consider a trial function $f(\mathbf{r})$, an n-point Gauss quadrature rule, the 8 child elements $(\{\Omega_i^c\})_{i=1}^8$ that are obtained by sub-dividing Ω^p, a fixed tolerance τ, and an empty list labelled *points*. Next, we evaluate $I^p = \int_{\Omega^p} f(\mathbf{r})\, d\mathbf{r}$ and $I_i^c = \int_{\Omega_i^c} f(\mathbf{r})\, d\mathbf{r}$ for $i = 1, 2, \ldots, 8$, using their respective n-point Gauss quadrature rules. If the base condition, $|I^p - \sum_{i=1}^8 I_i^c| < \tau$, is satisfied, we append the Gauss quadrature points and weights of the parent element to the list *points* and terminate the algorithm. Otherwise, we treat each of the child elements as a parent element, and recursively sub-divide them until the base condition is satisfied. Whenever the base condition is satisfied, the Gauss quadrature points and weights corresponding to the parent element at the current recursion level are appended to the list *points*. Finally, the list *points* represents the quadrature points and weights for the given element. We repeat this process for each element present in the finite element mesh. Typically, instead of using a single trial function $f(\mathbf{r})$, we use n_t such trial functions, $\{f_\nu(\mathbf{r})\}_{\nu=1}^{n_t}$, which requires n_t base conditions corresponding to each $f_\nu(\mathbf{r})$ to be satisfied.

In the present work, we choose the following four trial functions to build the adaptive quadrature rule:

$$f_1(\mathbf{r}) = \sum_{I=1}^{N_a} \left(N_I^{E,\phi}(\mathbf{r})\right)^2, \tag{3.36a}$$

$$f_2(\mathbf{r}) = \sum_{I=1}^{N_a} \left|\nabla\left(N_I^{E,\phi}(\mathbf{r})\right)\right|^2, \tag{3.36b}$$

$$f_3(\mathbf{r}) = \sum_{q=1}^{n_E^\psi} \left(N_q^{E,\psi}(\mathbf{r})\right)^2, \tag{3.36c}$$

$$f_4(\mathbf{r}) = \sum_{q=1}^{n_E^\psi} \left|\nabla\left(N_q^{\psi}(\mathbf{r})\right)\right|^2. \tag{3.36d}$$

Although we have labeled just two trial functions, $f_1(\mathbf{r})$ and $f_2(\mathbf{r})$, defined by the electrostatic enrichment functions, these correspond to four trial functions—two each for enrichment functions corresponding to the Hartree potential and the external potential. We remark that the aforementioned adaptive quadrature construction is performed only on the finite elements which are within the compact support of the enrichment functions. Since only a small fraction of the total elements are within the compact support of any enrichment function, the adaptive quadrature construction is computationally inexpensive. Further, once constructed, the adaptive quadrature list remains fixed for a given finite element mesh, and only needs to be updated if the finite element mesh changes during the course of the calculation.

We now turn towards determining an economical choice for the tolerance parameter, τ, as a loose tolerance may result in an inadequate quadrature grid whereas an extremely tight tolerance will be computationally inefficient. In the present work, we employ the following heuristic to choose τ. For each atom type S of atomic number Z_S in the materials system, we obtain the atomic ground-state charge density, $\rho^S(r)$, its corresponding Hartree potential, $\phi_{\mathrm{H}}^S(r)$, and the atomic external potential, $\phi_{\mathrm{ext}}^S(r)$, by solving the the radial Kohn-Sham equations in Eqs. 3.21. Next, we evaluate the following two integrals $E_1^{S,1D} = \frac{1}{2}\int 4\pi r^2 \rho^S(r)\phi_{\mathrm{H}}^S(r)\,dr$ and $E_2^{S,1D} = \int 4\pi r^2 \rho^S(r)\phi_{\mathrm{ext}}^S(r)\,dr$, which correspond to the electrostatic interaction energies. We then construct a coarse 3D finite element mesh with atom S at the origin. In order to determine a judicious choice for τ^S corresponding to atom type S, we set its initial value as $\tau^S = 0.1$ and enter an iterative loop. For the current iterate of τ^S, we evaluate the 3D counterparts of $E_1^{S,1D}$ and $E_2^{S,1D}$, namely, $E_1^{S,3D}$ and $E_2^{S,3D}$, respectively, using the aforemen-

tioned adaptive quadrature rule. If the convergence criteria, $|E_1^{S,1D} - E_1^{S,3D}| < \gamma$ and $|E_2^{S,1D} - E_2^{S,3D}| < \gamma$, are satisfied for a pre-determined γ, we terminate the loop with the current value of τ^S. Else, the loop is repeated with τ^S set to half of its current value, until the above convergence criteria are met. We use the minimum of all such τ^S corresponding to the various atom types in the materials system as our τ for construction of the adaptive quadrature grid for the materials system calculation. In all our calculations, we have used $\gamma = 0.1$ mHa so as to ensure that the quadrature errors are an order lower than the desired discretization error ($\sim$1 mHa) that we are aiming in the ground-state energy per atom. We note that the above procedure to determine τ, is independent of the choice of 3D finite element mesh. Moreover, the τ^S for each S can be precomputed and stored for later use.

3.5.4 Inverse of overlap matrix

We now discuss a computationally efficient way of evaluating the inverse of the overlap matrix, $\mathbf{M}^E$, defined in Eq. 3.27, which is vital to the transformation of the generalized Kohn-Sham eigenvalue problem to a standard eigenvalue problem, and the subsequent use of Chebyshev polynomial based acceleration technique to compute the occupied eigenstates as will be discussed in Section 3.6. We make use of the block-wise matrix inversion theorem [127] (also known as Banachiewicz inversion formula), to obtain the following 2×2 block structure for $(\mathbf{M}^E)^{-1}$,

$$(\mathbf{M}^E)^{-1} = \left[\begin{array}{c|c} (\mathbf{M^{cc}})^{-1} + \mathbf{L^T S^{-1} L} & -\mathbf{L^T S^{-1}} \\ \hline -\mathbf{S^{-1} L} & \mathbf{S^{-1}} \end{array}\right] \tag{3.37}$$

where $\mathbf{L} = \mathbf{M^{ec}}(\mathbf{M^{cc}})^{-1}$, and $\mathbf{S} = \mathbf{M^{ee}} - \mathbf{M^{ec}}(\mathbf{M^{cc}})^{-1}(\mathbf{M^{ec}})^T$. Assuming that the enriched finite element basis is not ill-conditioned, we note that the overlap matrix

$\mathbf{M}^E$ is positive definite, and, hence invertible. Further, $\mathbf{M^{cc}}$ being the overlap matrix of the classical finite element basis functions, is also positive definite, and hence invertible. Subsequently, the positive definiteness, and hence invertibility, of $\mathbf{S}$ can be ascertained by noting that it is the Schur complement [127] of $\mathbf{M^{cc}}$ in $\mathbf{M}^E$. We note that the above expression for $(\mathbf{M}^E)^{-1}$ contains two matrix inverses, $(\mathbf{M^{cc}})^{-1}$ and $\mathbf{S^{-1}}$. As mentioned in Section 3.4.2, the matrix $\mathbf{M^{cc}}$ is rendered diagonal through the use of spectral finite elements along with Gauss-Lobatto-Legendre quadrature rule, which makes the evaluation of $(\mathbf{M^{cc}})^{-1}$ trivial. Regarding the evaluation of $\mathbf{S}^{-1}$, we note that $\mathbf{S}$ is a small matrix of the size of $n_E^\psi \times n_E^\psi$, where n_E^ψ is typically of the same order as the number of electrons in the system. Thus, $\mathbf{S}$ can be easily inverted through the use of direct solvers.

Further, we note that although the overlap matrix is sparse, its inverse is a dense matrix. However, the constituent matrices present in the 2×2 block structure of $(\mathbf{M}^E)^{-1}$ are either sparse or much smaller in size compared to $(\mathbf{M}^E)^{-1}$ itself. To elaborate, we note that $\mathbf{L}$ is of the size $n_E^\psi \times n_h$, and is hence much smaller than the size $(n_h + n_E^\psi) \times (n_h + n_E^\psi)$ of $(\mathbf{M}^E)^{-1}$. Furthermore, $\mathbf{L}$, owing to the locality of the enrichment functions, is sparse. As noted earlier, $\mathbf{S}^{-1}$ is a small $n_E^\psi \times n_E^\psi$ matrix and $(\mathbf{M^{cc}})^{-1}$, being diagonal, is sparse. Since we are only interested in the action of matrix $(\mathbf{M}^E)^{-1}$ on a vector (as will be discussed in Section 3.6), we perform the matrix-vector product using the constituent matrices without ever explicitly constructing the $(\mathbf{M}^E)^{-1}$ matrix. This matrix-free evaluation of any matrix-vector product presents a significant advantage for the above inversion technique over the Newton-Schultz [128, 129, 130] and Taylor expansion [131] based methods, wherein the construction of the $(\mathbf{M}^E)^{-1}$ matrix is explicit and hence have huge memory requirements owing to the dense structure of $(\mathbf{M}^E)^{-1}$.

Finally, we briefly compare the proposed enriched finite element method with the other existing methods which in a similar spirit seek to augment the classical

finite element basis with other basis functions that efficiently capture the known physics in regions of interest. One such approach is that of partition-of-unity finite element method (PUFEM) [132, 133], wherein a typical discretization can be defined as [134, 135]

$$\psi^h(\mathbf{r}) = \sum_{j=1}^{n_h} N_j^C(\mathbf{r})\psi_j^C + \sum_{q}^{n_E}\sum_{k=1}^{n_{PU}} N_k^{PU}(\mathbf{r})N_q^E(\mathbf{r})\psi_{q,k}^E\,, \tag{3.38}$$

where $N_j^C(\mathbf{r})$ are the classical finite element basis functions, and $N_k^{PU}(\mathbf{r})$ is a subset of the classical finite element basis functions used to modulate the enrichment functions, $N_q^{E,\psi}(\mathbf{r})$, thus providing a larger set of augmenting functions. Although PUFEM preserves the locality of the basis to be commensurate with conventional finite element basis, the effect of multiplying enrichment functions with a set of classical finite element basis functions results in smoother augmenting basis functions, thereby making it more prone to ill-conditioning (due to linear dependency of augmenting basis functions with classical finite element basis functions). A more serious limitation of PUFEM stems from the significant increase in the number of augmenting basis functions, which, in turn, significantly increases the size of the $\mathbf{M^{ee}}$ block of the overlap matrix $\mathbf{M}^E$, thereby making the evaluation of the $\mathbf{S^{-1}}$ in $(\mathbf{M}^E)^{-1}$ computationally prohibitive.

Another such approach is that of Gaussian finite element mixed basis [43], wherein a given choice of Gaussian basis is used to the augment the classical finite element basis instead of atomic solutions to the Kohn-Sham problem, as used in the present work. We note that compared to the Gaussian basis the atomic solutions provide a more natural choice for augmenting functions, given that the atomic functions that we construct preserve the cusp in the density at the nuclear positions, otherwise missing in the Gaussian basis. Also, the choice of atomic functions provide for better control over the conditioning of the basis through the use of smooth cutoff functions on the

radial part of the atomic orbitals. Further, in the work on Gaussian finite element mixed basis [43], the Kohn-Sham problem was solved as a generalized eigenvalue problem using preconditioned conjugate-gradient method [136] which is, in general, less efficient compared to the Chebyshev filtering method used in the present work, discussed subsequently.

3.6 Self-consistent field iteration and Chebyshev filtering

We begin this section with a brief outline of the well-known Kohn-Sham self-consistent field iteration (SCF) used to solve the nonlinear Kohn-Sham eigenvalue problem. This involves starting with an input guess for the charge density, ρ_{in}, which is used to construct the effective potential, $v_{\text{KS}}[\rho_{\text{in}}, \mathbf{R}]$. Having constructed $v_{\text{KS}}[\rho_{\text{in}}, \mathbf{R}]$, the Kohn-Sham eigenvalue problem is solved to obtain the eigenpairs (ϵ_α,ψ_α), which are in turn used to compute the output charge density, ρ_{out}. If the difference between ρ_{out} and ρ_{in}, in an appropriately chosen norm, is below a certain tolerance, then the charge density is deemed to have converged and ρ_{out} denotes the ground-state charge density. Otherwise, ρ_{in} is updated through a choice of mixing scheme [137, 138, 139, 140] involving ρ_{in} and ρ_{out} from the current as well as those from previous iterations, and the iteration continues until convergence in charge density is achieved.

The most computationally expensive step in every iterate of the SCF procedure is the solution of the discrete Kohn-Sham eigenvalue problem. Typically, one can use Krylov-subspace based methods such as Jacobi-Davidson [141] or Krylov-Schur [142] to evaluate the lowest few eigenpairs corresponding to the occupied eigenstates. However, benchmark studies presented in a recent work [38] have shown these Krylov-subspace based methods to be about ten-fold slower in comparison to the Chebyshev filtering technique [143] to compute the occupied eigenstates. Based on this relative merit, we have employed the Chebyshev filtering technique to compute the relevant eigenspectrum of the Kohn-Sham Hamiltonian.

The key idea involved in the Chebyshev filtering approach is to progressively improve the subspace V spanned by the eigenvectors of the previous SCF iteration through polynomial based power iteration to eventually compute the occupied eigenspectrum upon attaining self-consistency. It relies on two important properties of a Chebyshev polynomial $p_m(x)$ of degree m to magnify the relevant (occupied) spectrum of the discrete Kohn-Sham Hamiltonian: (i) $p_m(x)$ grows rapidly outside the interval $[-1,1]$, and (ii) $|p_m(x)| \leq 1$ for $x \in [-1,1]$. For the sake of notational simplicity, we denote the discrete Kohn-Sham Hamiltonian by $\widetilde{\mathbf{H}}$, which in the classical finite element basis is $(\mathbf{M}^C)^{-1}\mathbf{H}^C$ and in the enriched finite element basis is $(\mathbf{M}^E)^{-1}\mathbf{H}^E$. The filtering technique proceeds by first mapping the unoccupied eigenspectrum of $\widetilde{\mathbf{H}}$ to $[-1,1]$ through the affine transformation $t(x) = \frac{2x-a-b}{b-a}$, where a and b denote the upper bounds of the occupied and unoccupied eigenspectrum of $\widetilde{\mathbf{H}}$, respectively. The upper bound of the unoccupied spectrum, b, is obtained inexpensively through a few Arnoldi iterations on $\widetilde{\mathbf{H}}$. The upper bound of the occupied spectrum, a, is obtained as the highest Rayleigh quotient of $\widetilde{\mathbf{H}}$ in the subspace V of the previous SCF iteration. We denote the resultant transformed matrix as $\bar{\mathbf{H}}$. We then apply the m-degree Chebyshev polynomial filter $p_m(\bar{\mathbf{H}})$ on V to obtain $\widetilde{V} = p_m(\bar{\mathbf{H}})V$. Owing to the rapid growth property of Chebyshev polynomials outside $[-1,1]$, the aforementioned filtering of V amplifies, for each vector in V, the components along the eigenvectors corresponding to occupied states and damps the components along the eigenvectors corresponding to unoccupied states. The action of the Chebyshev filter on V can be achieved in an efficient manner by utilizing the recursive construction of the Chebyshev polynomial [144]: $p_{k+1}(x) = 2xp_k(x) - p_{k-1}(x)$. Next, we orthonormalize the Chebyshev-filtered vectors to obtain the orthonormal set of vectors Q spanning $\widetilde{V}$, and perform a Galerkin projection of $\widetilde{\mathbf{H}}$ onto $\widetilde{V}$ to obtain the following reduced

generalized eigenvalue problem,

$$Q^T \mathbf{H} Q \boldsymbol{v}_\alpha = \epsilon_\alpha Q^T \mathbf{M} Q \boldsymbol{v}_\alpha \,, \tag{3.39}$$

where $\{\mathbf{H},\, \mathbf{M},\, \epsilon_\alpha\}$ represent $\{\mathbf{H}^C,\, \mathbf{M}^C,\, \epsilon_\alpha^C\}$ or $\{\mathbf{H}^E,\, \mathbf{M}^E,\, \epsilon_\alpha^E\}$ corresponding to the classical or enriched finite element discretization, respectively. We can now solve the above generalized eigenvalue problem, whose dimension is commensurate with the number of electrons in the system, using direct solvers to obtain the eigenvalues ϵ_α and their corresponding projected eigenvectors $\boldsymbol{v}_\alpha$. We subsequently rotate the projected eigenvectors to the original space to obtain the eigenvectors $Q\boldsymbol{v}_\alpha$, which along with the eigenvalues ϵ_α are used to evaluate the charge density. Finally, the subspace V is updated to $\widetilde{V}$ for the next SCF iteration.

We remark that in order to gain computational efficiency, we exploit the elemental structure in $\mathbf{H}^E$ (or $\mathbf{H}^C$) and $(\mathbf{M}^E)^{-1}$ (or $(\mathbf{M}^C)^{-1}$) to perform the matrix-vector products involved in the evaluation of $\widetilde{V} = p_m(\bar{\mathbf{H}})V$. To elaborate, we consider the case of enriched finite element and note that all the blocks in the 2×2 block structure of $\mathbf{H}^E$ and all the constituent matrices (except $(\mathbf{M}^{\mathbf{cc}})^{-1}$ and $\mathbf{S}^{-1}$) can be constructed, owing to the locality of the basis, by assembling contributions from individual elements. However, since we are interested only in the action of these matrices on vectors, we perform the matrix-vector products by first evaluating elemental matrix-vector products and then assembling the resultant elemental vector, without explicitly assembling any global matrix. We also note that, the dimension of the subspace V, denoted by N, is chosen to be greater than the number of occupied orbitals so as to avoid numerical instabilities for systems with small band-gaps or degenerate energy levels close to the Fermi energy, and also to avoid missing out any occupied eigenstate between two successive SCF iterations. Typically, we choose $N \sim \frac{N_e}{2} + 20$. Further, we note that Kohn-Sham orbitals of single atoms represent a good initial guess for

the subspace V for the first SCF iteration, and is adopted in the present work.

We note that the degree m of the Chebyshev polynomial filter needed to obtain a good approximation to the occupied eigenspace of the Kohn-Sham Hamiltonian depends on: (i) the separation between eigenvalues in the occupied part of the eigenspectrum, and (ii) the ratio between the spectral widths of the occupied and unoccupied part of the eigenspectrum of $\widetilde{\mathbf{H}}$. While the separation between the occupied eigenvalues depends on the materials system, the ratio of the spectral widths of the occupied and unoccupied parts of the eigespectrum depends on the largest eigenvalue of $\widetilde{\mathbf{H}}$, which, in turn depends on the finite-element discretization—it increases with decreasing element size. Typically, in a pseudopotential calculation, where the orbitals and the electrostatic potentials vary smoothly, one can use a relatively coarse finite element mesh to achieve chemical accuracies of $\sim$1 meV per atom using the classical finite element method. For such coarse finite element discretizations, a Chebyshev polynomial degree between 10 to 50 is sufficient to compute the occupied eigenspace. However, in all-electron calculations, where the orbitals are characterized by sharp variations near atomic cores and the external potential has Coulomb-singularity, one requires a highly refined finite element mesh near the atomic cores to achieve chemical accuracies of $\sim$1 mHa per atom. In addition to the significant increase in the degrees of freedom, such mesh refinement also increases the upper bound of the unoccupied eigenspectrum, thereby requiring a very high Chebyshev polynomial degree, $\mathcal{O}(10^3)$, to effectively compute the occupied eigenspace. These shortcomings of the classical finite element discretization in the context of all-electron calculations are noted in [38], where comparisons were made against plane-wave basis for pseudopotential calculations and against Gaussian basis for all-electron. It was noted that while the classical finite elements basis outperforms the plane-wave basis in pseudopotential calculations on the benchmark systems studied, they were ten-fold slower in comparison to the Gaussian basis in all-electron calculations. These disadvantages of the

classical finite element basis for all-electron calculations are mitigated by using the proposed enriched finite element basis, as will be demonstrated in the subsequent section.

3.7 Results and Discussion

In this section, we discuss the rate of convergence, accuracy, performance, and parallel scalability of the proposed enriched finite method for all-electron calculations. We first study the rate of convergence of ground-state energy with respect to element size for methane and carbon monoxide molecules. We then demonstrate the accuracy and performance of the enriched finite element method using large-scale non-periodic semi-conducting and heavy metallic systems. We use non-periodic silicon nano-clusters of various sizes, with the largest one containing 621 atoms (8694 electrons), as our benchmark semi-conducting systems. For heavy metallic systems, we use gold nano-clusters, $\mathrm{Au}_n (n = 1 - 23)$, as our benchmark systems. In order to assess the accuracy, reduction in degrees of freedom, reduction in Chebyshev polynomial degree, and performance of the enriched finite element method, we use, wherever possible, the classical finite element method as a reference. Depending upon the system, we use spectral hexahedral finite elements of polynomial order 2 to 6, denoted as HEX27, HEX64SPEC, HEX125SPEC, HEX216SPEC, HEX343SPEC respectively. We also compare, wherever possible, the accuracy and performance of the enriched finite element method with that of Gaussian basis. All our calculations using the Gaussian basis are performed with the NWCHEM [26] package. We use n-stage Anderson mixing [137] for charge density mixing in all our enriched and classical finite element method based calculations. Finally, we present the parallel scalability of our implementation of the proposed enriched finite element method using Message Passing Interface (MPI). The scalability studies as well as the benchmark studies demonstrating the computational efficiency, reported subsequently, are conducted on

a parallel computing cluster with the following configuration: Intel Xeon E5-2680v3 CPU nodes with 24 processors (cores) per node, 128 GB memory per node, and Infiniband networking between all nodes for fast MPI communications.

3.7.1 Rate of convergence

In this section, we study the rate of convergence of the ground-state energy with element size, h, using quadratic (HEX27) and cubic (HEX64SPEC) spectral finite elements. To this end, we generate a sequence of finite element meshes with increasingly smaller element sizes by uniformly subdividing the coarsest mesh. The ground-state energy, E_h, obtained from each of the HEX64SPEC meshes are used in the expression

$$|E_h - E_0| = Ch^q \tag{3.40}$$

to compute the constants E_0, q and C through a least-square fit. In the above expression, E_0 is the extrapolated continuum ground-state energy obtained as $h \to 0$. We use the E_0 obtained from HEX64SPEC to compute the relative error $\frac{|E_h - E_0|}{|E_0|}$ for both HEX27 and HEX64SPEC meshes. To assess the accuracy of E_0, we also compare it against the ground-state energy obtained using the polarization consistent-4 (pc-4) [145] Gaussian basis.

For the benchmark systems in our convergence study, we consider two systems: (i) methane molecule with a C-H bond length of 2.0784 a.u. and H-C-H bond angle of 109.4712°, and (ii) carbon monoxide molecule with a C-O bond length of 2.1297 a.u. For both the systems, we use a Chebyshev filter of order 60 to compute the occupied eigenspace and Fermi-Dirac smearing with $T = 500K$. For methane, the value of E_0 is evaluated to be -40.11993 Ha and the ground-state energy from pc-4 Gaussian basis is -40.11992 Ha. For carbon monoxide, the value of E_0 is evaluated to be -112.47189 Ha and the ground-state energy from pc-4 Gaussian basis is

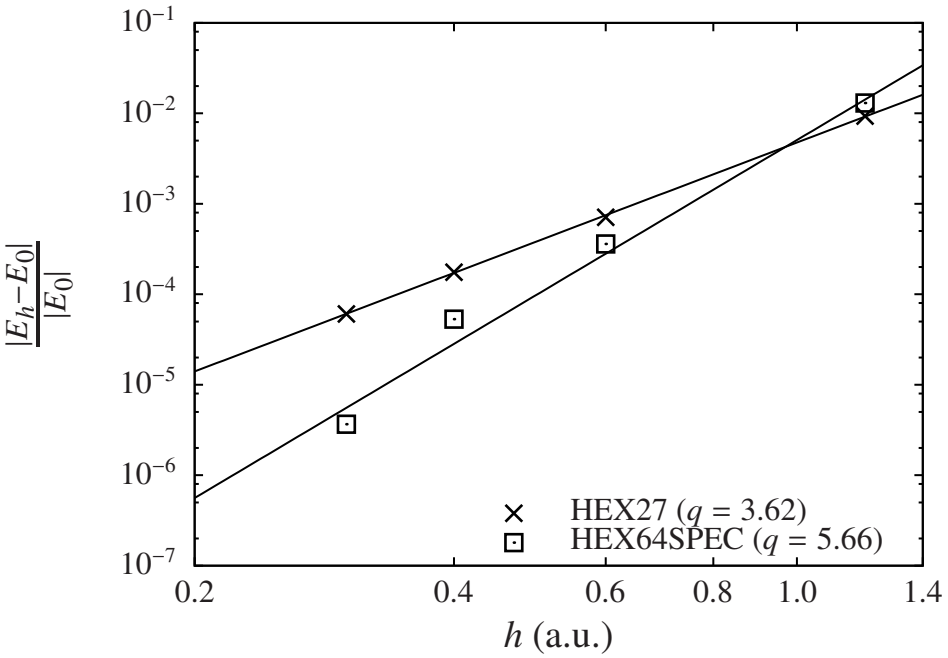


Figure 3.3: Convergence of energy with respect to element size for methane molecule

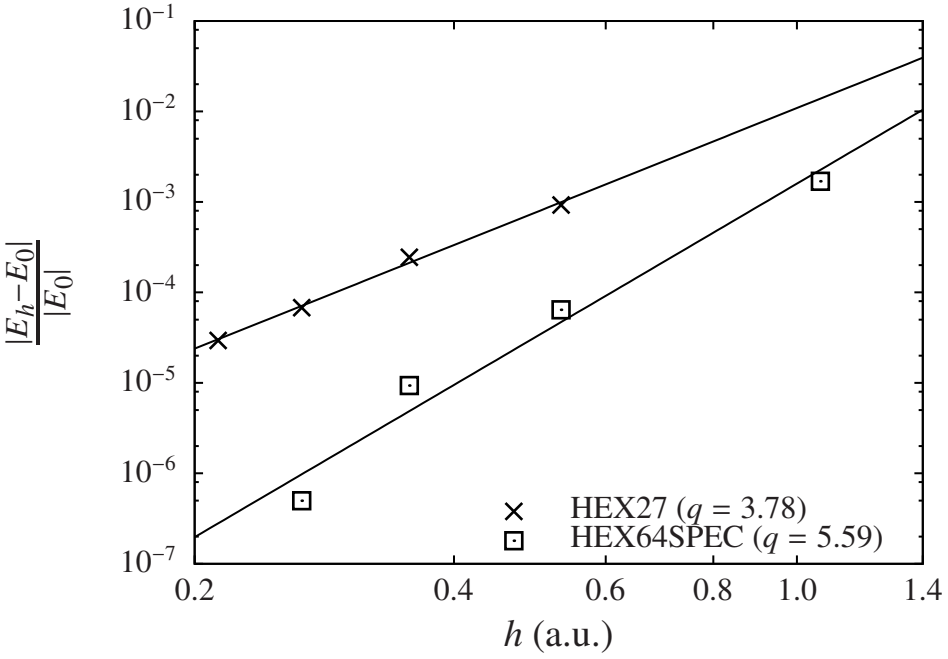


Figure 3.4: Convergence of energy with respect element size for carbon monoxide molecule

-112.47188 Ha. Next, we plot the relative error, $\frac{|E_h - E_0|}{|E_0|}$, against the smallest element size, and obtain the rates of convergence from the slopes of these plots. As evident from Figures 3.3 and 3.4, we obtain close to optimal rates of convergence in energy of $\mathcal{O}(h^{2k})$, where k is the polynomial order ($k = 2$ for HEX27 and $k = 3$ for HEX64SPEC). The results also suggest higher accuracies obtained with HEX64SPEC when compared to HEX27 for the same mesh size. We note that the numerically obtained rates of convergence deviate slightly from the theoretically optimal rates due to other numerical errors—beyond the discretization errors in the theoretical estimates—that are present in simulations, such as quadrature errors, errors due to stopping tolerance in the iterative solutions of the Poisson problem, diagonalization of the Hamiltonian and the self consistent field iteration.

3.7.2 Large-scale materials systems

We now discuss the accuracy and performance of the proposed enriched finite element method using large-scale semi-conducting and heavy metallic materials systems. We also compare, wherever possible, the proposed method against classical finite element and Gaussian basis based calculations.

3.7.2.1 Semi-conducting systems: Silicon nano-clusters

We consider silicon nano-clusters of various sizes, containing $1\times1\times1$ (252 electrons), $2 \times 1 \times 1$ (434 electrons), $2 \times 2 \times 2$ (1330 electrons), $3 \times 3 \times 3$ (3920 electrons); and $4 \times 4 \times 4$ (8694 electrons) diamond unit cells, as our benchmark semi-conducting systems. We employ a lattice constant of 10.26 a.u. in our calculations. These are isolated clusters in vacuum and we do not use any surface passivation. To obtain the characteristic element size to be used in the enriched finite element based calculations of the nano-clusters, we first obtain the reference ground-state energy, E_{ref}, for a single silicon atom by solving its 1D-radial Kohn-Sham eigenvalue problem as mentioned in

Section 3.5.1. Next, we choose a fourth-order (HEX125SPEC) finite element mesh for which the single atom ground-state energy obtained from the enriched finite element based calculation is within 1 mHa accuracy with respect to E_{ref}. Similarly, to obtain the characteristic element size for the classical finite element based calculations of the nano-clusters, we choose a fifth-order (HEX216SPEC) finite element mesh which is also within a 1 mHa accuracy for the single atom ground-state energy. We note that the smallest element size, thus obtained for the classical finite element based calculation is found to be an order of magnitude smaller than that of the smallest element size obtained in the enriched finite element based calculation. This, in turn, affects the largest eigenvalue of the Kohn-Sham Hamiltonian which is found to be $\mathcal{O}(10^6)$ in case of classical finite elements, thereby, requiring a Chebyshev polynomial filter of degree 1500 to compute the occupied eigenstates. Correspondingly, for the enriched finite element case, the largest eigenvalue is found to be $\mathcal{O}(10^3)$, thereby, requiring a $\sim$20-fold smaller Chebyshev polynomial degree of 80 to compute the occupied eigenstates. These choices for element sizes and Chebyshev polynomial degrees from single atom calculations are then carried forward to the nano-cluster calculations. We note that owing to the steep computational demand arising from large number of basis functions and high Chebyshev polynomial degree in the case of classical finite element based all-electron calculations, we could only perform studies up to $2 \times 2 \times 2$ nano-cluster size using the computational resources available to us. We also compare the accuracy and performance of the enriched finite element method with Gaussian basis. We use the polarization consistent (pc) family of Gaussian basis as it provides a hierarchy of increasingly larger basis sets. Specifically, we use pc-3 and pc-4 basis as they are both commensurate with the $\sim$1 mHa accuracy when compared with aforementioned E_{ref} for a single silicon atom. All the calculations with Gaussian basis are performed using Direct Inversion of Iterative Subspace (DIIS) [146] as well as the quadratically convergent minimization scheme [147], both available within the NWCHEM package,

and the computational time from the more efficient scheme is reported. For the DIIS scheme, extrapolation of up to 10 Fock matrices were used. Table 3.1 compares the ground-state energy, degrees of freedom (number of basis functions) per atom and the total computation CPU time (CPU time = number of cores $\times$ wall-clock time) for various cluster sizes using classical and enriched finite element basis. Similarly, Table 3.2 compares the ground-state energy and the total computation CPU time for various cluster sizes using enriched finite element, pc-3 and pc-4 basis. In all these calculations, we used a Fermi-Dirac smearing with $T = 500K$.

Table 3.1: Comparison of classical and enriched finite element (FE) basis: Energy per atom (E in Ha), degrees of freedom per atom (DoF), and total computation CPU time (in CPU hours) for various silicon nano-clusters.

Si $1 \times 1 \times 1$	Classical FE	Enriched FE
E	-288.320035	-288.319450
DoF	$402,112$	$14,728$
CPU Hrs	1599.15	24.81
Si $2 \times 1 \times 1$	Classical FE	Enriched FE
E	-288.334123	-288.333872
DoF	$386,205$	$13,557$
CPU Hrs	16441.43	57.10
Si $2 \times 2 \times 2$	Classical FE	Enriched FE
E	-288.359459	-288.359266
DoF	$360,467$	$10,642$
CPU Hrs	75936.4	553.13

As is evident from Tables 3.1 and 3.2, the enriched finite element basis achieves accuracies of within 1 mHa in the ground-state energies per atom when compared with classical finite element, pc-3 and pc-4 basis. We observe a staggering 60$-$ to 300$-$fold reduction in the total computation CPU time for the enriched finite element basis when compared with the classical finite element basis. This reduction in computation time stems from a $\sim$30$-$fold reduction in the degrees of freedom as

Table 3.2: Comparison of enriched finite element, pc-3 and pc-4 basis: Energy per atom (E in Ha) and total computation CPU time (in CPU hours) for various silicon nano-clusters.

Si $1 \times 1 \times 1$	Enriched FE	pc-3	pc-4
E	-288.319450	-288.318996	-288.319448
CPU Hrs	24.81	8.39	98.88
Si $2 \times 1 \times 1$	Enriched FE	pc-3	pc-4
E	-288.333872	-288.333447	-288.333898
CPU Hrs	57.10	151.74	1817.30
Si $2 \times 2 \times 2$	Enriched FE	pc-3	pc-4
E	-288.359266	-288.360045	FTC (Failed to converge)
CPU Hrs	553.13	4097.29	–
Si $3 \times 3 \times 3$	Enriched FE	pc-3	pc-4
E	-288.374721	FTC	FTC
CPU Hrs	6252.15	–	–
Si $4 \times 4 \times 4$	Enriched FE	pc-3	pc-4
E	-288.381425	FTC	FTC
CPU Hrs	45053.82	–	–

well as a $\sim$20$-$fold reduction in the Chebyshev polynomial degree as compared to the classical finite element basis. When compared with the pc-3 Gaussian basis, the enriched finite element is a factor $\sim$3 slower in the case of the smallest $(1 \times 1 \times 1)$ cluster. However, it outperforms the pc-3 basis, in total computation CPU time, by a factor of 2.5 for the $2 \times 1 \times 1$ cluster and by a factor of 7.5 for the $2 \times 2 \times 2$ cluster. Similarly, the enriched finite element basis outperforms the pc-4 Gaussian basis by factors 4 and 30 for the $1 \times 1 \times 1$ and $2 \times 1 \times 1$ clusters, respectively. We note that the pc-3 basis failed to converge for the $3 \times 3 \times 3$ and $4 \times 4 \times 4$ clusters, whereas the pc-4 basis failed to converge for $2 \times 2 \times 2$ and higher clusters. The failure of the pc-3 and pc-4 basis to converge for larger system sizes is primarily due to linear dependency of the Gaussian basis functions for larger system sizes. These results suggest that the enriched finite element basis offers a computationally efficient and robust basis

for all-electron calculations in semi-conducting systems as compared to both classical finite element and Gaussian basis.

3.7.2.2 Heavy metallic systems: Gold nano-clusters

Next, we consider gold nano-clusters, $Au_n (n = 1 - 23)$, to study the accuracy and performance of the enriched finite element basis. For $n = 2$ and $n = 6$, we use the stable geometries obtained in a previous work [148] wherein the Au_2 has a bond length of 4.818 a.u. and Au_6 has a planar triangle geometry with D_{3h} symmetry and bond length of 5.055 a.u. (refer Table A.1 in the Appendices for coordinates). The Au_{14} and Au_{23} nano-clusters were constructed as $1 \times 1 \times 1$ and $2 \times 1 \times 1$ face centered cubic (FCC) lattice, respectively, with a lattice constant of 6.812 a.u. We follow the same strategy as used for silicon nano-clusters to obtain the characteristic element sizes and Chebyshev polynomial degrees that are to be used in gold nano-cluster calculations, both using classical and enriched finite element basis. We use fifth-order (HEX216SPEC) and sixth-order (HEX343SPEC) finite elements for the enriched and classical finite element based calculations, respectively. We note that since gold is much heavier than silicon, it is characterized by more sharply oscillating orbitals and much steeper electrostatic potentials in comparison to silicon, thereby requiring smaller element sizes than those used in silicon to achieve similar accuracy. This, in turn, results in an increment in the largest eigenvalues of the Hamiltonian, which are found to be $\mathcal{O}(10^4)$ and $\mathcal{O}(10^8)$, for the enriched and classical finite element basis, respectively, thereby requiring higher Chebyshev polynomial degrees to accurately compute the occupied eigenstates. We note that the Chebyshev polynomial based filtering technique, being analogous to the power iteration method, can generate an ill-conditioned space for a very high polynomial degree, thereby resulting in numerical issues. To circumvent this, we employ, at each SCF iteration, multiple passes of a low polynomial degree Chebyshev filter and orthonormalize the filtered vectors between

two successive passes. For all our gold cluster calculations based on the enriched finite element basis we used 30 passes of a Chebyshev filter of degree 20, whereas 10 passes of a Chebyshev filter of degree 1200 have been employed for the classical finite element based calculations. We note that in the case of classical finite element based calculations, owing to the huge computational cost, we could perform calculations only up to Au_2 using the computational resources at our disposal. Further, we do not present a comparison with Gaussian basis owing to the lack of a good hierarchical non-relativistic basis for gold. Table 3.3 presents the comparison of the ground-state energies, degrees of freedom and total computation CPU times for the gold nano-clusters using classical and enriched finite element basis.

As is evident from Table 3.3, the enriched finite element basis obtains chemical accuracy in the ground-state energies per atom with far fewer degrees of freedom. In terms of computational efficiency, while the enriched finite element basis achieves $\sim$14$-$fold speedup over the classical finite Au_1, we observe $\sim$100$-$fold speedup for Au_2. Once again, these speedups for the enriched finite element basis are the result of a 40$-$fold reduction in the number of degrees of freedom and a 20$-$fold reduction in the Chebyshev polynomial degree as compared to that of the classical finite element basis. These numerical experiments demonstrate the accuracy and efficiency for all-electron calculations in heavy metallic systems.

3.7.3 Scalability

We demonstrate the parallel scalability (strong scaling) of the proposed enriched finite element basis in Figure 3.5. We choose the $3 \times 3 \times 3$ silicon nano-cluster containing $\sim$4 million degrees of freedom (number of basis functions) as our fixed benchmark system and report the relative speedup with respect to the wall time on 48 processors. The use of any number of processors below 48 was infeasible owing to the memory requirement posed by the system. As evident from the figure, the scaling

Table 3.3: Comparison of classical and enriched finite element (FE) basis: Energy per atom (E in Ha), degrees of freedom per atom (DoF), and total computation CPU time (in CPU hours) for various gold nano-clusters.

Au_1	Classical FE	Enriched FE
E	-17860.7623	-17860.7622
DoF	$5,040,409$	$120,361$
CPU Hrs	612.22	43.39
Au_2	Classical FE	Enriched FE
E	-17860.8001	-17860.8019
DoF	$4,659,399$	$122,300$
CPU Hrs	22950.25	220
Au_6	Classical FE	Enriched FE
E	–	-17860.8249
DoF	–	$178,906$
CPU Hrs	–	1924.42
Au_{14}	Classical FE	Enriched FE
E	–	-17860.8077
DoF	–	$88,657$
CPU Hrs	–	3740.29
Au_{23}	Classical FE	Enriched FE
E	–	-17860.8045
DoF	–	$80,397$
CPU Hrs	–	8171.40

is in good agreement with the ideal linear scaling behavior up to 384 processors, at which we observe a parallel efficiency of 87.8%. However, we observe a considerable deviation from linear scaling behavior at 768 processors with a parallel efficiency of 71.2%. This is attributed to the fact that at 768 processors the number of degrees of freedom possessed by each processor falls below 5000, which is too low to achieve good parallel scalability.

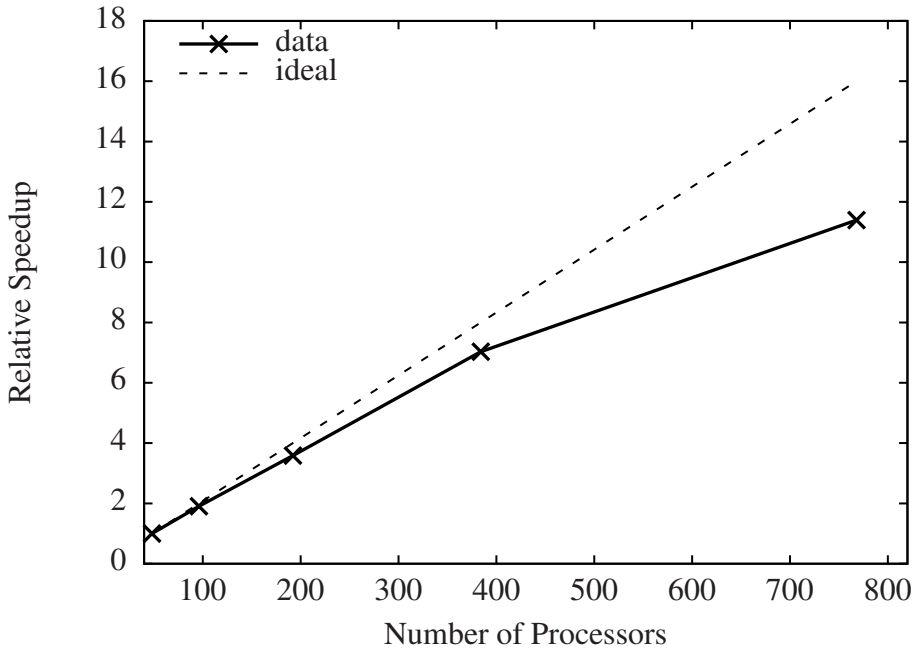


Figure 3.5: Parallel scalability of the enriched finite element implementation.

3.8 Summary

We have developed a computationally efficient mixed basis, termed as enriched finite element basis, for all-electron DFT calculations which combines the efficiency of atomic-orbitals-type basis to capture the sharp variations of the electronic fields closer to the atoms and the completeness of the classical finite element basis. This work demonstrates the marked computational advantage afforded by the enriched finite element basis over the classical finite element and the Gaussian basis for all-electron DFT calculations.

The proposed method is developed based on the following key ideas. Firstly, we augmented the classical spectral finite element basis with enrichment functions constructed from single-atom Kohn-Sham orbitals and electrostatic potentials. The enrichment functions are inexpensively pre-computed and stored by solving radial Kohn-Sham equations for all atoms in the periodic table. The enrichment functions are instrumental in capturing the sharp variations of the Kohn-Sham orbitals close to an atom, thereby, mitigating the need of high mesh refinement near the atomic cores.

Secondly, we used smooth cutoff functions to truncate the enrichment functions so as to ensure locality as well as better conditioning of the enriched finite element basis. Thirdly, we employ a divide and conquer strategy to construct an adaptive quadrature grid to efficiently evaluate the integrals involving the enrichment functions. Next, in order to convert the generalized Kohn-Sham eigenvalue problem to a standard eigenvalue problem, we employed a computationally efficient scheme to evaluate the inverse of the overlap matrix in the enriched finite element basis, by exploiting the block-wise matrix inversion. The use of spectral finite elements along with Gauss-Lobatto-Legendre quadrature rule is crucial in rendering the classical-classical block of the overlap matrix diagonal, whereas the use of the block-wise matrix inversion is crucial in utilizing the sparsity of the constituent matrices in the inverse of the overlap matrix for an efficient evaluation of the ensuing matrix-vector products. Finally, we employed the Chebyshev polynomial based filter to compute the occupied eigenstates. Here, we exploited the finite element structure in the Hamiltonian and the inverse overlap matrices to achieve an efficient and scalable implementation of the matrix-vector products involved in the action of the Chebyshev filter on a subspace.

In terms of the numerical convergence afforded by the enriched finite element basis, we demonstrated close to optimal rates of convergence for the ground-state energy with respect to the finite element discretization. We demonstrated the accuracy and performance of the proposed enriched finite element basis on: (i) silicon nano-clusters of various sizes, with the largest cluster containing 8694 electrons; and (ii) gold nano-clusters of various sizes, with the largest cluster containing 1817 electrons. We obtained good agreement in the ground-state energies when compared to classical finite element and Gaussian basis. In the larger clusters considered in this study, the enriched finite element basis provides a staggering $50 - 300$ fold speedup compared to the classical finite element basis, which is attributed to a $30-$fold reduction in the degrees of freedom as well as a $20-$fold reduction in the Chebyshev polynomial degree.

We also observed a significant outperformance by the enriched finite element basis relative to Gaussian basis (pc-3 and pc-4). Furthermore, we were able to perform ground-state energy calculations for silicon clusters containing 280 and 621 atoms ($\sim$9000 electrons), for which the Gaussian basis failed to converge owing to linear dependency of the basis. These large clusters underline the fact the proposed enriched finite element based approach can easily perform all-electron calculations ranging up to several thousands of electrons with moderate computational resources. In terms of parallel scalability, we obtained good parallel efficiency with almost linear scaling up to 384 processors for the benchmark system comprising of 280 atoms silicon nanocluster ($\sim$4 million basis functions).

The proposed method offers a computationally efficient, systematically improvable, and scalable basis for large scale all-electron DFT calculations, applicable to both light and heavy atoms. The use of the enrichment in developing linear-scaling DFT algorithms for all-electron calculations based on finite element basis [53, 54] or Tucker-tensor basis [149] holds good promise, and is currently being investigated. Furthermore, the use of enrichment ideas in conjunction with reduced-order scaling DFT algorithms can also be effectively utilized in the evaluation of the exact exchange operator, and forms a future direction of interest. Last, but not the least, the enriched finite element basis can be useful in a systematic study of the applicability and accuracy of various pseudopotential approximations on a wide range of materials and external conditions.

CHAPTER IV

Real Time Time-Dependent Density Functional Theory Using Higher-order Finite Element Methods

In this chapter, we present a computationally efficient approach to solve the time-dependent Kohn-Sham (TDDFT) equations in real-time using (classical) higher-order finite element spatial discretization, applicable to both pseudopotential and all-electron calculations [56]. In particular, we use error estimates to guide our spatial and temporal discretization. The temporal discretization related error analysis and numerics are presented in the context of second-order Magnus propagator. We underline that the numerical strategies developed in this work lays the essential groundwork for extending the key ideas of enriched finite element basis, explored in the ground-state DFT (refer Chapter III), to the TDDFT case.

The rest of the chapter is organized as follows. Sec. 4.1 provides a brief overview of real-time TDDFT (RT-TDDFT) and motivates the need for a higher-order finite elements for RT-TDDFT calculations. We revisit the TDKS equations in Sec. 4.2 and discuss the form of the exact time-evolution operator. We introduce the notion of semi- and full-discrete solution to the TDKS equation in Sec. 4.3. In Sec. 4.4, we provide formal spatial and time discretization error estimates in the Kohn-Sham

orbitals. Sec. 4.5 provides an efficient spatio-temporal discretization scheme guided by the error estimates. In Sec. 4.6, we describe the various numerical implementation aspects pertaining to spectral finite elements and the discrete second-order Magnus operator. Sec. 4.7 details the convergence, accuracy, efficiency and parallel scalability of the higher-order finite elements along with its relative performance against the finite-difference method. Finally, we summarize our findings and outline the future scope in Section 4.8.

4.1 Introduction

Time-dependent density functional theory (TDDFT), as we recall from Chapter II, extends the keys ideas of ground-state density functional theory (DFT) to electronic excitations and time-dependent processes. It relies on the Runge-Gross theorem [5] to establish, for a given initial state, a one-to-one correspondence between the time-dependent external potential and the time-dependent electronic density, thereby making the electronic density the fundamental variable to define other physical quantities. Subsequently, one invokes the Kohn-Sham *ansatz* [4] to reduce the many-electron time-dependent Schrödinger equation to a set of effective single electron equations, called the time-dependent Kohn-Sham (TDKS) equations. For all practical purposes, it requires the use of approximate exchange-correlation functionals, analogous to the ground-state case. However, TDDFT offers a great balance of accuracy and computational efficiency which have enabled the study of a wide-range of time-dependent phenomena—optical [150] and higher-order responses [151, 152], electron transport [153, 154], charge-transfer excitations [155, 156], dynamics of chemical bonds [157], multi-photon ionization [158, 159, 160], to name a few.

Given the practical significance of TDDFT calculations, there has been a growing interest in developing faster and more accurate numerical methods for solving the TDKS equations, over the past two decades. Broadly, these numerical methods can

be classified into two categories, characterized by the strength of the light-matter interaction, namely, linear-response time-dependent density functional theory (LR-TDDFT) [161, 162] and real-time time-dependent density functional theory (RT-TDDFT) [163, 164, 165]. As noted in Chapter I, the LR-TDDFT pertains to the case of weak interaction between the external field and the system, wherein the field induces a small perturbation from the ground-state. In such perturbative regime, one can compute the linear density response from the ground-state itself, which in turn can be used for the calculation of first-order response functions such as the absorption spectra. The RT-TDDFT, on the other hand, is a more generic framework which captures the electronic dynamics in real-time, thereby, allowing to handle both perturbative and non-perturbative regimes (e.g., harmonic generation, electron transport) in a unified manner. This involves propagating the TDKS equations in real-time without any restriction to the external field in terms of its frequency, shape or intensity. This work pertains to the more general RT-TDDFT.

Despite its generality in dealing with various time-dependent processes, there are two major challenges associated with RT-TDDFT. The first stems from the quality of the time-dependent exchange-correlation approximation used in the TDKS equations. The exact exchange-correlation functional is, in general, nonlocal in both space and time [166, 167, 91] and has an initial-state dependence [168]. However, the lack of insight into its time nonlocality and initial-state dependence has necessitated the use of the adiabatic approximation, wherein the exchange-correlation functional is defined in terms of the instantaneous electronic density. Although the applicability of the adiabatic approximation to various systems and materials properties are yet to be understood, they have shown remarkable agreement in estimating the transition frequencies [150], and, in most cases, is the underlying approximation in existing RT-TDDFT softwares. As with most of the numerical implementations in RT-TDDFT, this work is restricted to the adiabatic approximation. The second chal-

lenge stems from the huge computational cost associated with the non-linear TDKS equations. Numerical simulations for large length- and time-scales are still computationally challenging, and warrant systematically improvable, accurate, efficient and scalable spatio-temporal discretization. Addressing these numerical challenges constitutes the main subject of this work.

Significant efforts have been made towards efficient RT-TDDFT numerical schemes as extensions to popular ground-state DFT packages, borrowing from their respective spatial discretization. These include planewave basis in QBox [169, 170]; linear combination of atomic orbitals (LCAO) in Siesta [171, 172] and GPAW [173]; Gaussian basis in NWChem [26, 174]; and finite-difference based approaches in Octopus [175] and GPAW [176, 177]. The planewave basis, owing to its completeness, provides systematic convergence, and affords an efficient treatment of the electrostatic interactions through fast Fourier transforms. However, they remain restricted to only periodic geometries and boundary conditions, thereby ill-equipped to describe systems with defects, and non-periodic systems like isolated molecules and nano-clusters. Additionally, the nonlocality of the basis greatly hinders its parallel scalability. Atomic-type orbitals, such as LCAO and Gaussian basis, owing to their atom-specific basis, are well-suited to describe molecules and nano-clusters for both pseudopotential as well as all-electron calculations. However, owing to the incompleteness of such basis, systematic convergence for all materials systems remains a concern. The finite difference discretization (FD) provides systematic convergence, can handle a broad range of boundary conditions, and exhibits improved parallel scalability in comparison to planewave and atomic-type orbital basis. However, incorporating adaptive spatial resolution in FD through a non-uniform grid remains non-trivial. The lesser flexibility of FD to vary spatial resolution renders FD less straightforward to employ in the context of singular potentials (as in the case of all-electron calculations). On the other hand, the finite element basis [178, 109], being a local-piecewise polynomial

basis, offers several key advantages—it provides systematic convergence; is amenable to adaptive spatial resolution, and thereby suitable for both pseudopotential and all-electron calculations; exhibits excellent parallel scalability owing to the locality of the basis; and admits arbitrary geometries and boundary conditions. We add that many of these advantages of finite element basis are also shared by the wavelets basis [179]. While, at present, the use of wavelets basis has been restricted to LR-TDDFT [180], we expect them to be a competent basis for RT-TDDFT as well.

The efficacy of the finite element basis in terms of its accuracy, efficiency, scalability and relative performance with other competing methods (e.g., planewaves, Gaussian basis, FD), have been thoroughly studied in the context of ground-state DFT, for both pseudopotential [39, 40, 41, 42, 44, 45, 46, 49, 50, 51, 181, 38, 182, 55, 183] and all-electron calculations [39, 38, 55, 43, 47, 48, 52, 53, 37, 54, 183]. A similarly comprehensive study on the efficacy of the finite element basis for RT-TDDFT is, however, lacking. While two recent studies [184, 185] demonstrate the accuracy of finite elements for RT-TDDFT, they remain restricted to only linear and quadratic finite elements. As known from prior studies in ground-state DFT [186, 47, 38], lower-order (linear and quadratic) finite elements require a large number of basis functions ($50,000 - 500,000$ per atom for pseudopotential calculations) to achieve chemical accuracy, and hence, perform poorly in comparison to planewaves and other real-space based methods. However, this shortcoming of linear and quadratic finite elements for ground-state DFT calculations has been shown to be alleviated by the use of higher-order finite elements [38]. In this work, we extend the use of higher-order finite elements to RT-TDDFT calculations and demonstrate the resulting advantages over lower-order finite elements as well as finite-difference based methods.

The keys ideas in this work can be summarized as: (i) developing an *a priori* mesh-adaption based on semi-discrete (discrete in space, continuous in time) error analysis of the TDKS equations, and subsequently, obtaining an efficient finite element

discretization for the problem; (ii) use of spectral finite elements in conjunction with Gauss-Legendre-Lobatto quadrature to render the overlap matrix diagonal, thereby simplifying the evaluation of the inverse of the overlap matrix that features in the discrete time-evolution operator; (iii) obtaining an efficient temporal discretization using a full-discrete error analysis of the TDKS equations, in the context of second-order Magnus time-evolution operator; and (iv) using an adaptive Lanczos iteration to efficiently compute the action of the Magnus propagator on the Kohn-Sham orbitals. The *a priori* mesh-adaption in this work is performed by minimizing the discretization error in the observable of importance, subject to fixed number of elements in the finite element mesh. In particular, we minimize the semi-discrete error in the dipole moment of the system with respect to the mesh-size distribution, $h(\mathbf{r})$, to obtain an efficient *a priori* spatial discretization. Having obtained the spatial discretization, an efficient temporal discretization is obtained through a full-discrete error analysis, in the context of second-order Magnus time-evolution operator. This is, to the best of our knowledge, the first work that guides the spatio-temporal discretization for the RT-TDDFT problem using error estimates.

We study the key numerical aspects of the proposed higher-order finite element discretization for benchmark systems involving both nonlocal pseudopotential and all-electron calculations. To begin with, we study the numerical rates of convergence of the dipole moment with respect to spatial and temporal discretization. We use two benchmark systems: (i) a pseudopotential calculation on methane molecule; and (ii) an all-electron calculation on lithium hydride molecule, to demonstrate the rates of convergence for linear, quadratic and fourth-order finite elements. We observe numerical rates of convergence in the dipole moment close to the optimal rates obtained from our error analysis. Next, we assess the computational advantage afforded by higher-order finite elements over linear finite element, using the same benchmark systems. We observe an extraordinary 100-fold speedup in terms of the total computational

time for the fourth-order finite element over linear finite element, for calculations in the regime of chemical accuracy. We also compare the relative performance of the finite element discretization against finite-difference method for pseudopotential calculations. We use aluminum clusters (Al_2 and Al_{13}), and the Buckminsterfullerene (C_{60}) molecule as our benchmark pseudopotential systems. The finite-difference based calculations are done using the Octopus package [175]. Depending on the benchmark system, the finite element discretization shows a 3- to 60-fold savings in computational time as compared to the finite-difference approach, for pseudopotential calculations. We also demonstrate the efficacy of finite elements for systems subjected to strong perturbation by studying higher harmonic generation in Mg_2. Additionally, we demonstrate the competence of finite elements for all-electron calculations on two benchmark systems—methane and benzene molecule. Lastly, we study the strong scaling of our implementation and observe good parallel scalability with ∼75% efficiency at 768 processors for a benchmark system of a Buckminsterfullerene molecule containing 3.5 million degrees of freedom.

4.2 Real space time-dependent Kohn-Sham equations

We revisit the TDKS equations and elaborate on certain aspects of it, in the context of real-space discretization. The TDKS equations, in atomic units, are given as

$$\begin{aligned} i\frac{\partial \psi_\alpha(\mathbf{r},t)}{\partial t} &= H_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})\psi_\alpha(\mathbf{r},t) \\ &:= \left[-\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})\right]\psi_\alpha(\mathbf{r},t)\,, \end{aligned} \tag{4.1}$$

where $H_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})$, $v_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})$ and $\psi_\alpha(\mathbf{r},t)$ represent the time-dependent Kohn-Sham Hamiltonian, potential and orbitals, respectively, with the index α spanning over all the N_e electrons in the system. $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_{N_a}\}$ denotes the collective

representation for the positions of the N_a atoms in the system. The electron density, $\rho(\mathbf{r}, t)$, is given in terms of the Kohn-Sham orbitals as

$$\rho(\mathbf{r}, t) = \sum_{\alpha=1}^{N_e} |\psi_\alpha(\mathbf{r}, t)|^2 \,. \tag{4.2}$$

In the present work, we restrict ourselves to only non-periodic (clusters and molecules) as well as spin-unpolarized systems. However, we note that all the ideas discussed subsequently can be generalized to spin-polarized systems as well.

The time-dependent Kohn-Sham potential, $v_{\text{KS}}[\rho](\mathbf{r}, t; \mathbf{R})$ (same as that defined in Eq. 2.33), is given by

$$v_{\text{KS}}[\rho](\mathbf{r}, t; \mathbf{R}) = v_{\text{ext}}(\mathbf{r}, t; \mathbf{R}) + v_{\text{H}}[\rho](\mathbf{r}, t) + v_{\text{xc}}[\rho](\mathbf{r}, t) \,, \tag{4.3}$$

where $v_{\text{ext}}(\mathbf{r}, t; \mathbf{R})$ denotes the external potential, $v_{\text{H}}[\rho](\mathbf{r}, t)$ denotes the Hartree potential, and $v_{\text{xc}}[\rho](\mathbf{r}, t)$ represents the exchange-correlation potential. We recall from our discussion that the exchange-correlation potential, $v_{\text{xc}}[\rho](\mathbf{r}, t)$, in general, is non-local in both space and time [166, 167, 91], and has a dependence on the initial many-electron wavefunction [168]. However, in absence of the knowledge of its true form, most of the existing approximations use locality in time (adiabatic exchange-correlation) and non-dependence on the initial many-electron wavefunction. This allows for direct use of the existing exchange-correlation approximations used in ground-state DFT. In the present work, we use the adiabatic local-density approximation (ALDA) [187], which is local in both space and time. Specifically, we use the Ceperley-Alder form [112].

In Eq. 4.3, the Hartree potential is given by

$$v_{\text{H}}[\rho](\mathbf{r}, t) = \int \frac{\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}' \,. \tag{4.4}$$

The external potential comprises of the nuclear potential $v_{\mathrm{N}}(\mathbf{r};\mathbf{R})$ and the external field $v_{\mathrm{field}}(\mathbf{r},t)$. The nuclear potential is given by

$$v_{\mathrm{N}}(\mathbf{r};\mathbf{R}) = \begin{cases} v_{\mathrm{N}}^{\mathrm{ae}} = -\sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}, & \text{for all-electron}\,, \\ v_{\mathrm{N}}^{\mathrm{psp}}(\mathbf{R}), & \text{for pseudopotential}\,, \end{cases} \tag{4.5}$$

where Z_I and $\mathbf{R}_I$ represent the atomic charge and position of the I^{th} nucleus. For a typical pseudopotential calculation, $v_{\mathrm{N}}^{\mathrm{psp}}$ comprises of a local part, $v_{\mathrm{psp}}^{\mathrm{loc}}$, and a non-local part, $v_{\mathrm{psp}}^{\mathrm{nl}}$. For the nonlocal part, the action on a function $\phi(\mathbf{r})$, written in the Kleinman-Bylander form [188], is given by

$$v_{\mathrm{psp}}^{\mathrm{nl}}(\mathbf{R})\phi(\mathbf{r}) = \sum_{I=1}^{N_a}\sum_{l=0}^{L_I}\sum_{m=-l}^{l}\left(\frac{\int u_{lm}^I(\mathbf{r}')\delta v_l^I(\mathbf{r}')\phi(\mathbf{r}')\,d\mathbf{r}'}{\int u_{lm}^I(\mathbf{r}')\delta v_l^I(\mathbf{r}')u_{lm}^I(\mathbf{r}')\,d\mathbf{r}'}\right)\delta v_l^I(\mathbf{r})u_{lm}^I(\mathbf{r})\,, \tag{4.6}$$

where l and m denote the angular and magnetic quantum number, respectively. $u_{lm}^I(\mathbf{r})$ is a pseudo-atomic eigenfunction for the atom at $\mathbf{R}_I$, $\delta v_l^I(\mathbf{r})$ is the specified l angular component short-ranged potential for the atom at $\mathbf{R}_I$, and L_I is the maximum angular quantum number specified for the atom at $\mathbf{R}_I$. The external field, $v_{\mathrm{field}}(\mathbf{r},t)$, is typically provided as a laser pulse of the form

$$v_{\mathrm{field}}(\mathbf{r},t) = -\mathbf{E_0}(t)\cdot\mathbf{r}\,, \tag{4.7}$$

where $\mathbf{E_0}(t)$ represents the time-dependent electric field.

As with the ground-state DFT , we reformulate the evaluation of the electrostatic potentials as the solutions to the following Poisson equations (refer to Sec. 3.3)

$$-\frac{1}{4\pi}\nabla^2 v_{\mathrm{H}}(\mathbf{r},t) = \rho(\mathbf{r},t)\,, \quad v_{\mathrm{H}}(\mathbf{r},t)|_{\partial\Omega} = f(\mathbf{r},\mathbf{R})\,, \tag{4.8a}$$

$$-\frac{1}{4\pi}\nabla^2 v_{\mathrm{N}}^{\mathrm{ae}}(\mathbf{r};\mathbf{R}) = b(\mathbf{r},\mathbf{R})\,, \quad v_{\mathrm{N}}^{\mathrm{ae}}(\mathbf{r})|_{\partial\Omega} = -f(\mathbf{r},\mathbf{R})\,. \tag{4.8b}$$

In the above equation, $b(\mathbf{r};\mathbf{R}) = -\sum_{I=1}^{N_a} Z_I \delta(\mathbf{r};\mathbf{R}_I)$, where $\delta(\mathbf{r};\mathbf{R}_I)$ is a bounded regularization of the Dirac-delta distribution with compact support in a small ball around $\mathbf{R}_I$ and satisfies $\int \delta(\mathbf{r};\mathbf{R}_I)\, d\mathbf{r} = 1$; $f(\mathbf{r},R) = \sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$; and $\partial\Omega$ denotes the boundary of a sufficiently large bounded domain $\Omega \in \mathbb{R}^3$. We refer to previous works on finite elements based ground-state DFT calculations [41, 44, 49, 38, 113] for a detailed treatment of the local reformulation of the electrostatic potentials into Poisson equations.

Formally, the solution to Eq. 4.1 can be written as

$$\begin{aligned} \psi_\alpha(\mathbf{r},T) &= U(T,t_0)\psi_\alpha(\mathbf{r},t_0) \\ &= \mathcal{T}\exp\left\{-i\int_{t_0}^{T} H_{\mathrm{KS}}[\rho](\mathbf{r},\tau)d\tau\right\}\psi_\alpha(\mathbf{r},t_0)\,, \end{aligned} \tag{4.9}$$

where $U(T,t_0)$ represents the time-evolution operator (propagator) and $\mathcal{T}$ denotes the time-ordering operator. Although the above equation provides a formal way to directly evaluate the orbitals at any time, t, resolving the implicit time-dependence of the Kohn-Sham Hamiltonian through the density is too difficult. However, one can exploit the following composition property of the propagator,

$$U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0)\,, \qquad t_0 < t_1 < t_2\,, \tag{4.10}$$

to accurately resolve the implicit time-dependence in $H_{\mathrm{KS}}[\rho](\mathbf{r},t)$. To elaborate, the above property allows us to rewrite the propagator $U(T,t_0)$ as

$$U(T,t_0) = \prod_{i=0}^{N-1} U(t_{i+1},t_i)\,, \tag{4.11}$$

where $t_N = T$ and $t_{i+1} - t_i = \Delta t_i$, with Δt_i denoting the variable time step. Consequently, one can divide the evaluation of the orbitals at T into N short-time propa-

gation, given by

$$\begin{aligned}\psi_\alpha(\mathbf{r}, t+\Delta t) &= U(t+\Delta t)\psi_\alpha(\mathbf{r}, t) \\ &= \mathcal{T}\exp\left\{-i\int_t^{t+\Delta t} H_{\mathrm{KS}}[\rho](\mathbf{r}, \tau)d\tau\right\}\psi_\alpha(\mathbf{r}, t)\,.\end{aligned} \tag{4.12}$$

In addition to resolving the implicit time-dependence in $H_{\mathrm{KS}}[\rho](\mathbf{r}, t)$, the short time propagation provides the numerical advantage of containing the norm of the exponent in Eq. 4.9. To elaborate, any efficient numerical scheme to compute the action of the propagator on a wavefunction involves either a power series expansion or a subspace projection of the propagator, wherein the number of terms in the power series or the dimension of the subspace required for a given accuracy are dependent on norm of the exponent. Moreover, there is a physical upper bound imposed on the time step based on the maximum frequency, ω_{max}, that one wants to resolve in their calculations, i.e., $\Delta t_{max} = \frac{1}{\omega_{max}}$. Typically, ω_{max} is determined by the eigenspectrum of the ground-state Hamiltonian or by the frequency of the applied field, v_{field}. We note that, in practice, one uses a time step $\Delta t \ll \Delta t_{max}$ owing to the need of containing time-discretization errors that arise in approximating the continuous propagator, $\mathcal{T}\exp\left\{-i\int_t^{t+\Delta t} H_{\mathrm{KS}}[\rho](\mathbf{r}, \tau)d\tau\right\}$. We discuss these approximations and their associated time-discretization errors in greater detail in Sec. 4.3 and Sec. 4.4.

4.3 Semi- and full-discrete solutions

In this section, we introduce the notion of semi-discrete (discrete in space but continuous in time) and full-discrete solution to the TDKS equation. The full-discrete solution is provided in the context of second-order Magnus propagator.

To begin with, we provide some of the finite element essentials. In the finite element method, the spatial domain of interest ($\Omega \in \mathbb{R}^3$) is divided into non-overlapping

sub-domains, known as finite elements. Each finite element (e) is characterized by its spatial extent (Ω_e) and size (h_e). Subsequently, the finite element basis is constructed from piecewise Lagrange interpolating polynomials that have a compact support on the finite elements (i.e., on Ω_e), thus rendering locality to these basis functions. We note that there is an abundance of choice in terms of the form and order of the polynomial functions that can be used in constructing the finite element basis. We refer to Refs. [109] and [189] for a comprehensive discourse on the subject.

4.3.1 Semi-discrete solution

To begin with, we express the semi-discrete time-dependent Kohn-Sham orbitals, $\psi_\alpha^h(\mathbf{r},t)$, as

$$\psi_\alpha^h(\mathbf{r},t) = \sum_{j=1}^{n_h} N_j(\mathbf{r})\psi_\alpha^j(t)\,,\ \text{s.t.}\ \psi_\alpha^h(\mathbf{r},t)|_{\partial\Omega} = 0\ ,\forall t \geq 0\,, \tag{4.13}$$

where $\{N_j(\mathbf{r})\}$ represents the set of finite element basis functions, each of polynomial order p; and $\psi_\alpha^j(t)$ denote the time-dependent expansion coefficient corresponding to the N_j basis function. Using the discretization of Eq. 4.13 in the TDKS equation (Eq. 4.1) results in following discrete equation,

$$i\mathbf{M}\dot{\boldsymbol{\psi}}_\alpha(t) = \mathbf{H}\boldsymbol{\psi}_\alpha(t), \tag{4.14}$$

where $\mathbf{H}$ and $\mathbf{M}$ denote the discrete Hamiltonian and overlap matrix, respectively, and $\boldsymbol{\psi}_\alpha(t)$ denotes the vector containing the coefficients $\psi_\alpha^j(t)$. The solutions to the above equation are called the *semi-discrete* solutions to the TDKS equation. In the

above equation, the discrete Hamiltonian H_{jk} is given by

$$\begin{aligned} H_{jk} = \frac{1}{2} \int_{\Omega} \nabla N_j(\mathbf{r}) \cdot \nabla N_k(\mathbf{r}) \, d\mathbf{r} \\ + \int_{\Omega} v^h_{\mathrm{KS}}[\rho^h](\mathbf{r}, t) N_j(\mathbf{r}) N_k(\mathbf{r}) \, d\mathbf{r} \,, \end{aligned} \tag{4.15}$$

where $v^h_{\mathrm{KS}}[\rho^h](\mathbf{r}, t)$ is the discrete Kohn-Sham potential corresponding to the semi-discrete density, $\rho^h(\mathbf{r}, t)$ (i.e., evaluated from the solutions of Eq. 4.14). $v^h_{\mathrm{KS}}[\rho^h](t)$ is, in turn, given by

$$\begin{aligned} v^h_{\mathrm{KS}}[\rho^h](\mathbf{r}, t) = v^h_{\mathrm{H}}[\rho^h](\mathbf{r}, t) + v^h_{\mathrm{N}}(\mathbf{r}) \\ + v_{\mathrm{xc}}[\rho^h](\mathbf{r}, t) + v_{\mathrm{field}}(\mathbf{r}, t) \,, \end{aligned} \tag{4.16}$$

where $v^h_{\mathrm{H}}[\rho^h](\mathbf{r}, t)$ and $v^h_{\mathrm{N}}(\mathbf{r})$ denote the discrete Hartree and nuclear potential, respectively. We note that for the pseudopotential case, v^h_{N} is same as the continuous potential $v^{\mathrm{psp}}_{\mathrm{N}}$ and hence v^h_{N} is relevant only in the all-electron case. Similar to $\psi^h_\alpha(\mathbf{r}, t)$, the discrete electrostatic potentials ($v^h_{\mathrm{H}}[\rho^h](\mathbf{r}, t)$ and $v^h_{\mathrm{N}}(\mathbf{r})$) are obtained by discretizing them in the finite element basis, i.e.,

$$v^h_{\mathrm{H}}[\rho^h](\mathbf{r}, t) = \sum_{j=1}^{n_h} N_j(\mathbf{r}) \phi^j_H(t) \,, \tag{4.17}$$

$$v^{\mathrm{ae},h}_{\mathrm{N}}(\mathbf{r}) = \sum_{j=1}^{n_h} N_j(\mathbf{r}) \phi^j_N \,, \tag{4.18}$$

satisfying the boundary conditions presented in Eq. 3.5 in a discrete sense. Subsequently, the expansion coefficients $\phi^j_H(t)$ and ϕ^j_N can be obtained by using the above expressions in Eq. 4.8, and solving the resulting system of linear equations

$$\mathbf{K}\boldsymbol{\phi}_H(t) = 4\pi \mathbf{c}(t) \,, \text{ and} \tag{4.19}$$

$$\mathbf{K}\boldsymbol{\phi}_N = 4\pi\mathbf{d}\,, \tag{4.20}$$

where $K_{jk} = \int_\Omega \nabla N_j(\mathbf{r}).\nabla N_k(\mathbf{r})\, d\mathbf{r}$; $\boldsymbol{\phi}_H$ and $\boldsymbol{\phi}_N$ are the vectors containing the coefficients $\phi_H^j(t)$ and ϕ_N^j, respectively; $c_j(t) = \int_\Omega \rho^h(\mathbf{r},t)N_j(\mathbf{r})\, d\mathbf{r}$; and $d_j = \int_\Omega b(\mathbf{r},\mathbf{R})N_j(\mathbf{r})\, d\mathbf{r}$.

4.3.2 Full-discrete solution

We now discuss the full-discrete solution to the TDKS equations, in the context of second-order Magnus propagator. To begin with, we note that the overlap matrix $\mathbf{M}$, being positive definite, guarantees the existence of a unique positive definite square root, $\mathbf{M}^{1/2}$. This allows us to rewrite Eq. 4.14 as

$$i\dot{\bar{\boldsymbol{\psi}}}_\alpha(t) = \bar{\mathbf{H}}\bar{\boldsymbol{\psi}}_\alpha(t)\,, \tag{4.21}$$

where $\bar{\boldsymbol{\psi}}_\alpha(t) = \mathbf{M}^{1/2}\boldsymbol{\psi}_\alpha(t)$ and $\bar{\mathbf{H}} = \mathbf{M}^{-1/2}\mathbf{H}\mathbf{M}^{-1/2}$. To put it differently, $\bar{\boldsymbol{\psi}}_\alpha(t)$ is the representation of $\psi_\alpha^h(\mathbf{r},t)$ in a Löwdin orthonormalized basis [190]. We remark that the practicality of the above reformulation in terms of $\bar{\boldsymbol{\psi}}_\alpha$ is contingent upon the efficient evaluation of $\mathbf{M}^{-1/2}$. To that end, we present an efficient scheme for computing $\mathbf{M}^{-1/2}$ in Sec. 4.6.

We invoke the Magnus *ansatz* to write the solution of Eq. 4.21 as

$$\bar{\boldsymbol{\psi}}_\alpha(t) = \exp(\mathbf{A}(t))\bar{\boldsymbol{\psi}}_\alpha(0)\,, \qquad \forall t \geq 0\,. \tag{4.22}$$

The operator $\exp(\mathbf{A}(t))$ is termed as the *Magnus propagator*, where $\mathbf{A}(t)$ is given explicitly as [191, 192]

$$\begin{aligned}\mathbf{A}(t) = &\int_0^t -i\bar{\mathbf{H}}(\tau)d\tau \\ &-\frac{1}{2}\int_0^t \left[\int_0^\tau -i\bar{\mathbf{H}}(\sigma)d\sigma, -i\bar{\mathbf{H}}(\tau)\right] d\tau + \ldots \quad ,\end{aligned} \tag{4.23}$$

where $[\mathbf{X}, \mathbf{Y}] = \mathbf{XY} - \mathbf{YX}$ denotes the commutator. Although known explicitly, the above equation is not practically useful, given the difficulty in resolving the implicit dependence of $\bar{\mathbf{H}}(t)$ on $\rho^h(\mathbf{r}, t)$. As mentioned in Section 4.2, we resolve the implicit dependence by using the composition property of a propagator (cf. Eq. 4.11). This allows us to rewrite the Magnus propagator as

$$\exp(\mathbf{A}(t)) = \prod_{n=1}^{N} \exp(\mathbf{A}_n)\,, \tag{4.24}$$

where $\mathbf{A}_n$ is given by Eq. 4.23, albeit with the limits of integration modified to $[t_{n-1}, t_n]$.

In practice, one replaces the exact $\mathbf{A}_n$ with an approximate operator $\tilde{\mathbf{A}}_n$, which involves, first, a truncation of the Magnus expansion (defined in Eq. 4.23), and second, an approximation for the time integrals in the truncated Magnus expansion. Truncating the Magnus expansion after the first p terms results in a time-integration scheme of order $2p$. In this work, we restrict ourselves to the second-order Magnus propagator, i.e., obtained by truncating the Magnus expansion after the first term. Furthermore, we use a mid-point integration rule to evaluate $\int_{t_{n-1}}^{t_n} -i\bar{\mathbf{H}}(\tau)d\tau$. In particular, the action of the second-order Magnus propagator, with a mid-point integration rule, on the set of Kohn-Sham orbitals $\{\bar{\boldsymbol{\psi}}_1, \bar{\boldsymbol{\psi}}_2, \ldots, \bar{\boldsymbol{\psi}}_{N_e}\}$ which defines the density $\rho^h(\mathbf{r}, t)$, is given by

$$e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}}_\alpha(t) = e^{-i\bar{\mathbf{H}}\left[\rho^h\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]\Delta t}\bar{\boldsymbol{\psi}}_\alpha(t)\,, \tag{4.25}$$

where $\Delta t = t_n - t_{n-1}$ and $\bar{\mathbf{H}}\left[\rho^h\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]$ is the time-continuous Kohn-Sham Hamiltonian described by $\rho^h(\mathbf{r}, t)$ at the future time instance $t_{n-1} + \Delta t/2$. We remark that $\bar{\mathbf{H}}\left[\rho^h\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]$, being dependent on a future instance of the density, is evaluated either by an extrapolation of $\bar{\mathbf{H}}$ using $m(>2)$ previous steps or by a second (or higher) order predictor-corrector scheme.

Thus, time-discrete approximation to $\bar{\boldsymbol{\psi}}_\alpha(t_n)$, denoted by $\bar{\boldsymbol{\psi}}_\alpha^n$, is given by

$$\bar{\boldsymbol{\psi}}_\alpha^n = \exp(\tilde{\mathbf{A}}_n)\bar{\boldsymbol{\psi}}_\alpha^{n-1}. \tag{4.26}$$

Consequently, the orbitals $\psi_\alpha^{h,n}(\mathbf{r})$ defined by the coefficient vectors $\boldsymbol{\psi}_\alpha^n = \mathbf{M}^{-1/2}\bar{\boldsymbol{\psi}}_\alpha^n$ represent the *full-discrete* solution to the TDKS equation.

4.4 Discretization Errors

In this section, we provide the discretization error in the Kohn-Sham orbitals which will later on form the basis of our efficient spatio-temporal discretization. To begin with, we decompose the discretization error in the Kohn-Sham orbitals into two parts, one arising due to spatial discretization and the other due to temporal discretization. To elaborate, if $\psi_\alpha^h(\mathbf{r}, t_n)$ and $\psi_\alpha^{h,n}(\mathbf{r})$ represent the semi-discrete (discrete in space but continuous in time) and full-discrete solution to $\psi_\alpha(\mathbf{r}, t_n)$, respectively, then one decompose the discretization error in $\psi_\alpha(\mathbf{r}, t_n)$ as

$$\begin{aligned}\psi_\alpha(\mathbf{r}, t_n) - \psi_\alpha^{h,n}(\mathbf{r}) = &\big(\psi_\alpha(\mathbf{r}, t_n) - \psi_\alpha^h(\mathbf{r}, t_n)\big) \\ &+ \big(\psi_\alpha^h(\mathbf{r}, t_n) - \psi_\alpha^{h,n}(\mathbf{r})\big)\end{aligned} \tag{4.27}$$

We present the spatial and temporal error estimates (i.e., error estimates for the two right-hand terms in the above equation in Sec. 4.4.1 and Sec. 4.4.2), respectively. The detailed derivation for these estimates are presented in Sec. 4.4.4 and in Sec. 4.4.5.

4.4.1 Spatial discretization error

If $\psi_\alpha^h(\mathbf{r},t)$ denotes the semi-discrete solution to $\psi_\alpha(\mathbf{r},t)$, then the following bound holds true (see Sec. 4.4.4 for the derivation)

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq C(t)\sum_e h_e^p \sum_{\alpha=1}^{N_e}\Big(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha})\Big) \\ + C(t)\sum_e h_e^p \sum_{\alpha=1}^{N_e}\Big(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\Big) \\ + C(t)\sum_e h_e^p \Big(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e}\Big)\,,\end{aligned} \tag{4.28}$$

$$\text{for some } \{s_{1,\alpha}\}, \{s_{2,\alpha}\}, \text{and } s_3 \in [0,t]\,.$$

In the above equations, h_e and Ω_e denote the size and spatial-extent of the $e-$th finite element, respectively. $C(t)$ is a time-dependent constant independent of the finite element mesh. $|.|_{p,\Omega_e}$ is the semi-norm in $H^p(\Omega_e)$. The importance of the above equations lies in relating the semi-discrete error to the mesh parameters (i.e., h_e and p), and hence, is instrumental in obtaining an efficient spatial discretization (discussed in Sec. 4.5). In particular, the above equation informs that the semi-discrete error in $\left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}$ decays as $\mathcal{O}(h_e^p)$.

4.4.2 Time discretization error

We now present the formal bounds on the time discretization error in $\psi_\alpha(\mathbf{r},t)$. Assuming each time interval $[t_{n-1},t_n]$ to be of length Δt, we obtain the following bound for the time-discretization error for a second-order Magnus propagator with a mid-point integration rule (see Sec. 4.4.5 for the derivation)

$$\left\|\psi_\alpha^h(t_n) - \psi_\alpha^{h,n}\right\|_{L^2(\Omega)} \leq \quad C(\Delta t)^2 t_n \max_{0\leq t\leq t_n}\left\|\psi_\alpha^h(t)\right\|_{H^1(\Omega)}\,. \tag{4.29}$$

The essence of the above relation lies in relating the time-discretization error to $\left\|\psi_\alpha^h(t)\right\|_{H^1(\Omega)}$, which in turn is related to the spatial discretization. Thus, the above equation, is crucial in selecting an efficient Δt, for a given finite element mesh (see Sec. 4.5).

We now present the derivation for the estimates presented in Eq. 4.28 and Eq. 4.29. We remark that the proof is rather involved and a reader may choose to skip it and proceed directly to Sec. 4.5, without any loss of flow or generality. Before proceeding to the proofs, we introduce some notations, key assumptions and preliminaries.

4.4.3 Notations, assumptions and preliminaries

For a bounded closed domain Ω and bounded time interval $[0, T]$, we denote $\Omega_T = \Omega \times [0, T]$. For any two complex-valued functions $f(\mathbf{r}, t), g(\mathbf{r}, t) : \Omega_T \to \mathbb{C}$, the inner product $(f, g)(t) = \int_\Omega f(\mathbf{r}, t) g^\dagger(\mathbf{r}, t)\, d\mathbf{r}$, where $g^\dagger(\mathbf{r}, t)$ denotes the complex conjugate of $g(\mathbf{r}, t)$. Correspondingly, the norm $||f||_{L^2(\Omega)}(t) = \sqrt{(f, f)(t)}$. Thus, we extend the definition of the standard $L^2(\Omega)$ and $H^1(\Omega)$ spaces to define

$$L^2(\Omega_T) = \left\{ f(\mathbf{r}, t) \,\middle|\, ||f||_{L^2(\Omega)}(t) \leq \infty, \quad \forall t \in [0, T] \right\}, \tag{4.30a}$$

$$H^1(\Omega_T) = \left\{ f(\mathbf{r}, t) \,\middle|\, f, \frac{\partial f}{\partial t}, Df \in L^2(\Omega_T) \right\}, \tag{4.30b}$$

$$H_0^1(\Omega_T) = \left\{ f(\mathbf{r}, t) \,\middle|\, f \in H^1(\Omega_T), f(\mathbf{r}, t)|_{\partial\Omega} = 0, \quad \forall t \in [0, T] \right\}, \tag{4.30c}$$

where Df denotes the spatial partial derivatives of f, and $\partial\Omega$ denotes the boundary of Ω. Additionally, we define two more spaces relevant to the Poisson problem (Eq. 4.8),

$$H_Z^1(\Omega_T) = \left\{ f(\mathbf{r}, t) \,\middle|\, f \in H^1(\Omega_T), f(\mathbf{r}, t)|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \quad \forall t \in [0, T] \right\} \tag{4.31a}$$

$$H_{-Z}^1(\Omega) = \left\{ f(\mathbf{r}) \,\middle|\, f \in H^1(\Omega), f(\mathbf{r})|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{-Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right\}. \tag{4.31b}$$

For conciseness of notation, in all our subsequent discussion, we drop the argument t from the inner product as well as all the L^p and H^1 norms. Thus, any occurrence of $(.\,,.)$, $||.||_{L^p(\Omega)}$, and $||.||_{H^1(\Omega)}$ are to treated as time-dependent, unless otherwise specified.

We list certain weak assumptions that we invoke throughout our error-estimates.

$\mathcal{A}1$ The time-dependent Kohn-Sham orbitals and their spatial derivatives are bounded and have a compact support on Ω, which, in turn, is a large but a bounded subset of $\mathbb{R}^3$. To elaborate, $\psi_\alpha \in H_0^1(\Omega_T) \cap L^\infty(\Omega_T)$.

$\mathcal{A}2$ The nuclear potential (in the all-electron case), due to the use of regularized nuclear charge distribution $b(\mathbf{r};\mathbf{R})$ (defined in Eq. 4.8b), is bounded, i.e., $v_{\mathrm{N}}^{\mathrm{ae}} \in L^\infty(\mathbb{R}^3)$.

$\mathcal{A}3$ The local part of the pseudopotential is bounded, i.e., $v_{\mathrm{psp}}^{\mathrm{loc}} \in L^\infty(\mathbb{R}^3)$.

$\mathcal{A}4$ The short-ranged potentials appearing in the nonlocal part of the pseudopotential are bounded, i.e., $\delta v_l^I \in L^\infty(\Omega)$.

$\mathcal{A}5$ The exchange-correlation potential and its derivative with respect to density are both bounded, i.e., $v_{\mathrm{xc}}[\rho], v'_{\mathrm{xc}}[\rho] \in L^\infty(\mathbb{R}^3), \forall t \in [0,T]$.

$\mathcal{A}6$ The external field is bounded, i.e., $v_{\mathrm{field}} \in L^\infty(\mathbb{R}^3), \forall t \in [0,T]$.

$\mathcal{A}7$ The induced operator (or matrix) norm of the Kohn-Sham Hamiltonian and the Laplace operator are equivalent, i.e., $\exists$ time-independent bounded constants C_1, C_2 such that:
$C_1 \left\|\nabla^2\phi\right\|_{L^2(\Omega)} \leq \left\|H_{\mathrm{KS}}\phi\right\|_{L^2(\Omega)} \leq C_2 \left\|\nabla^2\phi\right\|_{L^2(\Omega)}$, $\forall \phi \in H_0^1(\Omega)$, $\forall t \in [0,T]$.

$\mathcal{A}8$ The first and second time-derivatives of the Kohn-Sham potential are bounded, i.e., $\left\|\frac{d}{dt}v_{\mathrm{KS}}(t)\right\|_{L^2(\Omega)} \leq C_1$ and $\left\|\frac{d^2}{dt^2}v_{\mathrm{KS}}(t)\right\|_{L^2(\Omega)} \leq C_2\,, \forall t \in [0,T]$, where C_1, C_2 are time-independent bounded constants.

We remark that while the validity of $\mathcal{A}1$ and $\mathcal{A}7$ are apparent in the case of pseudopotential calculations, for the all-electron case, it is reasonable to assume the same owing to the use of regularized nuclear charge distribution $b(\mathbf{r};\mathbf{R})$. Using these assumptions, we derive certain formal bounds that will subsequently be used in deriving the error estimates. To this end, given two different densities $\rho_{\Psi_1}(\mathbf{r},t)$ and $\rho_{\Psi_2}(\mathbf{r},t)$ defined by the set of orbitals $\Psi_1 = \{\psi_{1,1}, \psi_{1,2}, \ldots, \psi_{1,N_e}\}$ and $\Psi_2 = \{\psi_{2,1}, \psi_{2,2}, \ldots, \psi_{2,N_e}\}$, respectively, we seek to bound $\|v_{\mathrm{KS}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{KS}}[\rho_{\Psi_2}]\psi_{2,\alpha}\|_{L^2(\Omega)}$ in terms of $(\psi_{1,\alpha} - \psi_{2,\alpha})$ and $(\rho_{\Psi_1} - \rho_{\Psi_2})$. We remark that all the subsequent results hold $\forall \alpha \in \{1,2,\ldots,N_e\}$, unless otherwise specified. Moreover, the constants C, its subscripted forms (i.e., C_1, C_2, etc.), and primed forms (C'), that appear subsequently, are positive and bounded.

To begin with, we note, through straightforward use of Cauchy-Schwarz and Sobolev inequalities, that

$$\|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^1(\Omega)} \leq C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\,, \tag{4.32a}$$

$$\|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^2(\Omega)} \leq C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{H^1(\Omega)}\,. \tag{4.32b}$$

Furthermore, for the convolution integral of ρ and $\frac{1}{|\mathbf{r}|}$, denoted by $|\mathbf{r}|^{-1} * \rho = \int_\Omega \frac{\rho(\mathbf{x})}{|\mathbf{r}-\mathbf{x}|} d\mathbf{x}$, we have

$$\left\||\mathbf{r}|^{-1} * \rho\right\|_{L^\infty(\Omega)} \leq C \left\||\mathbf{r}|^{-1}\right\|_{L^2(\Omega)} \|\rho\|_{L^2(\Omega)}\,, \tag{4.33a}$$

$$\left\||\mathbf{r}|^{-1} * \rho\right\|_{L^2(\Omega)} \leq C \left\||\mathbf{r}|^{-1}\right\|_{L^2(\Omega)} \|\rho\|_{L^1(\Omega)}\,, \tag{4.33b}$$

where we have used the Young's inequality along with the fact that $|\mathbf{r}|^{-1} \in L^2(\Omega)$.

We now bound $\|v_{\mathrm{KS}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{KS}}[\rho_{\Psi_2}]\psi_{2,\alpha}\|_{L^2(\Omega)}$, by decomposing v_{KS} into its

Hartree (v_{H}), nuclear (v_{N}), exchange-correlation (v_{xc}) and field (v_{field}) components, and bounding each of the components. For the Hartree potential, we have, for $\forall v \in H_0^1(\Omega_T)$,

$$(v_{\mathrm{H}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{H}}[\rho_{\Psi_2}]\psi_{2,\alpha}, v) = (v_{\mathrm{H}}[\rho_{\Psi_1}](\psi_{1,\alpha} - \psi_{2,\alpha}), v) + (v_{\mathrm{H}}[\rho_{\Psi_1} - \rho_{\Psi_2}]\psi_{2,\alpha}, v) \ . \tag{4.34}$$

Thus, using result of Eq. 4.33b along with the fact that $\psi_{2,\alpha} \in L^\infty(\Omega_T)$ (from $\mathcal{A}1$) and $v_{\mathrm{H}}[\rho_{\Psi_1}] \in L^\infty(\Omega_T)$ (from Eq. 4.33a), it follows that

$$\begin{aligned} |(v_{\mathrm{H}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{H}}[\rho_{\Psi_2}]\psi_{2,\alpha}, v)| \leq\ & C\left(\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}\right) \\ & + C\left(\|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^1(\Omega)} \|v\|_{L^2(\Omega)}\right) \ . \end{aligned} \tag{4.35}$$

Next, for the exchange-correlation potential, we use the mean value theorem to note that

$$v_{\mathrm{xc}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{xc}}[\rho_{\Psi_2}]\psi_{2,\alpha} = (v_{\mathrm{xc}}[\rho_X] + 2\chi_\alpha^2 v'_{\mathrm{xc}}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}) \, , \tag{4.36}$$

where ρ_X is defined by the orbitals $\chi_\alpha = \lambda_\alpha\psi_{1,\alpha} + (1-\lambda_\alpha)\psi_{2,\alpha}$, for some $\lambda_\alpha \in [0,1]$. Using the above relation, we have, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned} |(v_{\mathrm{xc}}[\rho_{\Psi_1}]\psi_{1,\alpha} - v_{\mathrm{xc}}[\rho_{\Psi_2}]\psi_{2,\alpha}, v)| &= \left|\left((v_{\mathrm{xc}}[\rho_X] + 2\chi_\alpha^2 v'_{\mathrm{xc}}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}), v\right)\right| \\ &\leq \left\|v_{\mathrm{xc}}[\rho_X] + 2\chi_\alpha^2 v'_{\mathrm{xc}}[\rho_X]\right\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \ , \end{aligned} \tag{4.37}$$

where we have used the boundedness assumption on v_{xc} and v'_{xc} (assumption $\mathcal{A}5$).

Similarly, using the boundedness assumptions on $v_{\mathrm{N}}^{\mathrm{ae}}$ ($\mathcal{A}2$), $v_{\mathrm{N}}^{\mathrm{psp}}$($\mathcal{A}3$, $\mathcal{A}4$), and

v_{field} ($\mathcal{A}6$) it is easy to observe, $\forall v \in H_0^1(\Omega_T)$,

$$\left|(v_{\text{N}}^{\text{ae}}\psi_{1,\alpha} - v_{\text{N}}^{\text{ae}}\psi_{2,\alpha}, v)\right| \leq C \left\|\psi_{1,\alpha} - \psi_{2,\alpha}\right\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,. \tag{4.38}$$

$$\left|(v_{\text{N}}^{\text{psp}}\psi_{1,\alpha} - v_{\text{N}}^{\text{psp}}\psi_{2,\alpha}, v)\right| \leq C \left\|\psi_{1,\alpha} - \psi_{2,\alpha}\right\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,. \tag{4.39}$$

$$\left|(v_{\text{field}}\psi_{1,\alpha} - v_{\text{field}}\psi_{2,\alpha}, v)\right| \leq C \left\|\psi_{1,\alpha} - \psi_{2,\alpha}\right\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,. \tag{4.40}$$

We now define the weak solution of the TDKS equation (Eq. 4.1) as follows: given an initial state $\psi_\alpha(\mathbf{r}, 0) \in H_0^1(\Omega)$, we seek $\psi_\alpha(\mathbf{r}, t) \in H_0^1(\Omega_T)$ such that

$$i\left(\frac{\partial \psi_\alpha}{\partial t}, v\right) = \frac{1}{2}\left(\nabla\psi_\alpha, \nabla v\right) + \left(v_{\text{KS}}[\rho]\psi_\alpha, v\right), \quad \forall v \in H_0^1(\Omega_T)\,, \text{ and } \forall t \in [0, T]\,. \tag{4.41}$$

Similarly, the weak solutions to the Poisson problems defined in Eq. 4.8 are defined to be $v_{\text{H}}(\mathbf{r}, t) \in H_Z^1(\Omega_T)$, and $v_{\text{N}}^{\text{ae}}(\mathbf{r}, \mathbf{R}) \in H_{-Z}^1(\Omega)$, satisfying,

$$\left(\nabla v_{\text{H}}, \nabla v\right) = 4\pi(\rho, v), \quad \forall v \in H_0^1(\Omega_T), \text{ and } \forall t \in [0, T] \tag{4.42a}$$

$$\left(\nabla v_{\text{N}}^{\text{ae}}, \nabla v\right) = 4\pi\left(b, v\right), \quad \forall v \in H_0^1(\Omega)\,. \tag{4.42b}$$

4.4.4 Derivation of spatial discretization error

We denote $X^{h,p} \in H^1(\Omega)$ to be the finite-dimensional space of dimension n^h, spanned by finite element basis functions of order p. Further, we denote $X_0^{h,p} = X^{h,p} \cap H_0^1(\Omega)$. We now define the semi-discrete solution, $\psi_\alpha^h(\mathbf{r}, t)$, to Eq. 4.41 as follows: given an initial state $\psi_\alpha^h(\mathbf{r}, 0) \in X_0^{h,p}$, we seek $\psi_\alpha^h(\mathbf{r}, t) \in X_0^{h,p} \times [0, T]$ such that

$$i\left(\frac{\partial \psi_\alpha^h}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla\psi_\alpha^h, \nabla v^h\right) + \left(v_{\text{KS}}^h[\rho^h]\psi_\alpha^h, v^h\right), \quad \forall v^h \in X_0^{h,p} \times [0, T], \text{ and } \forall t \in [0, T]\,, \tag{4.43}$$

where $\rho^h(\mathbf{r},t) = \sum_{\alpha=1}^{N_e} \left|\psi_\alpha^h(\mathbf{r},t)\right|^2$ and $v_{\mathrm{KS}}^h[\rho^h](\mathbf{r},t) = v_{\mathrm{H}}^h[\rho^h](\mathbf{r},t) + v_{\mathrm{N}}^h(\mathbf{r};\mathbf{R}) + v_{\mathrm{xc}}[\rho^h](\mathbf{r},t) + v_{\mathrm{field}}(\mathbf{r},t)$.

We now elaborate on the different terms appearing in the expression for $v_{\mathrm{KS}}^h[\rho^h](\mathbf{r},t)$. First, to define appropriate boundary conditions for $v_{\mathrm{H}}^h[\rho^h](\mathbf{r},t)$ and $v_{\mathrm{N}}^h(\mathbf{r})$, we introduce the function $f^h(\mathbf{r};\mathbf{R}) = \sum_{j=1}^{n_h} q_j N_j(\mathbf{r})$, with

$$q_j = \begin{cases} \sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}_j - \mathbf{R}_I|}, & \text{if } j^{th} \text{ node (positioned at } \mathbf{r}_j) \text{ is a boundary node} \\ 0, & \text{otherwise}\,, \end{cases}$$

as an interpolation of the boundary conditions of Eqs. 4.8 into $X^{h,p}$. This allows us to define the discrete counterpart of the weak solution described in Eq. 4.42a as $v_{\mathrm{H}}^h[\rho^h](\mathbf{r},t) = v_{\mathrm{H},0}^h[\rho^h](\mathbf{r},t) + f^h(\mathbf{r};\mathbf{R})$, with $v_{\mathrm{H},0}^h[\rho^h](\mathbf{r},t) \in X_0^{h,p} \times [0,T]$, such that

$$(\nabla v_{\mathrm{H},0}^h, \nabla v^h) = 4\pi(\rho^h, v^h) - (\nabla f^h, \nabla v^h), \quad \forall v^h \in X_0^{h,p} \times [0,T]\,, \text{and } \forall t \in [0,T]\,. \tag{4.44}$$

Similarly, we define the discrete analog of the weak solution defined in Eq. 4.42b as $v_{\mathrm{N}}^{\mathrm{ae},h}(\mathbf{r};\mathbf{R}) = v_{\mathrm{N},0}^{\mathrm{ae},h}(\mathbf{r};\mathbf{R}) - f^h(\mathbf{r};\mathbf{R})$, with $v_{\mathrm{N},0}^{\mathrm{ae},h}(\mathbf{r};\mathbf{R}) \in X_0^{h,p}$, such that

$$(\nabla v_{\mathrm{N},0}^{\mathrm{ae},h}, \nabla v^h) = 4\pi(b, v^h) + (\nabla f^h, \nabla v^h), \quad \forall v^h \in X_0^{h,p}\,. \tag{4.45}$$

For the pseudopotential case, $v_{\mathrm{N}}^h(\mathbf{r};\mathbf{R})$ is same as the continuous function $v_{\mathrm{N}}^{\mathrm{psp}}(\mathbf{r};\mathbf{R})$.

We now introduce the concept of Ritz projection, $\mathcal{P}_h$, which will be used in subsequent error estimates. The Ritz projection $\mathcal{P}_h : H_0^1(\Omega_T) \to X_0^{h,p} \times [0,T]$ is defined through the following Galerkin orthogonality condition,

$$\left(\nabla(\psi - \mathcal{P}_h\psi), \nabla v^h\right) = 0, \quad \forall \psi \in H_0^1(\Omega_T),\, \forall v^h \in X_0^{h,p} \times [0,T], \text{ and } \forall t \in [0,T]\,. \tag{4.46}$$

This allows us to use some standard finite element error estimates [193] to bound

$\|\psi - \mathcal{P}_h\psi\|_{L^2(\Omega)}$.

In order to prove the bound of Eq. 4.28, we, first, present a general case with no assumptions on the initial orbitals $\psi_\alpha(\mathbf{r}, 0)$. We then present the special case of the initial orbitals being ground-state Kohn-Sham orbitals, as a corollary to the general case. Furthermore, we note that an error estimate for $\left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}$, in turn, requires an estimate for $\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}$. Therefore, in our subsequent analysis we report estimates for both $\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}$ and $\left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}$. We emphasize that, although the numerical studies presented in this work have used hexagonal elements, the following results apply to other shapes of finite element, and hence, in our analysis we denote the mesh using the generic term 'triangulation' [193]. In particular, we take a triangulation $\mathcal{T}^{h,p}$ of p^{th} order finite elements covering the domain Ω.

Proposition IV.1. *Assuming uniqueness and existence of the solution to Eqs. 4.41 and 4.43, we obtain the following bounds on the finite element semi-discrete approximation error to the Kohn-Sham orbitals:*

$$\begin{aligned}\sum_{\alpha=1}^{N_e}\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq\ & C_1 e^{C_2 t}(t+1)\sum_e h_e^{p+1}\sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha})\right)\\ &+ C_1 e^{C_2 t}(t+1)\sum_e h_e^{p+1}\sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)\\ &+ C_1 e^{C_2 t} t\sum_e h_e^{p+1}\left(\left|v_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_N|_{p+1,\Omega_e}\right)\\ &+ e^{C_2 t}\sum_{\alpha=1}^{N_e}\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0)\,,\end{aligned} \tag{4.47a}$$

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq C_3 e^{C_2 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right)$$
$$+ C_3 e^{C_2 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right)$$
$$+ C_3 e^{C_2 t} t \sum_e h_e^p \left(\left|v_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_N|_{p+1,\Omega_e} \right)$$
$$+ C_3 e^{C_2 t} h_{min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0), \tag{4.47b}$$

where e denotes a finite element of mesh size h_e and cover Ω_e in the triangulation $\mathcal{T}^{h,p}$, h_{min} represents the smallest element in the triangulation $\mathcal{T}^{h,p}$, and $|.|_{p,\Omega_e}$ is the semi-norm in $H^p(\Omega_e)$. The arguments $s_{1,\alpha}, s_{2,\alpha},$ and s_3 are defined as

$$s_{1,\alpha} = \arg\max_{0\leq s\leq t} \|\psi_\alpha - \mathcal{P}_h\psi_\alpha\|_{L^2(\Omega)}(s), \quad s_{2,\alpha} = \arg\max_{0\leq s\leq t} \left\|\frac{\partial\psi_\alpha}{\partial t} - \mathcal{P}_h\frac{\partial\psi_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s), \text{ and}$$
$$s_3 = \arg\max_{0\leq s\leq t} \left\|v_H[\rho^h] - v_H^h[\rho^h]\right\|_{L^2(\Omega)}(s). \tag{4.48}$$

Proof. Taking $v = v^h \in X_0^{h,p}\times[0,T]$ in Eq. 4.41 (continuous solution) and subtracting it from Eq. 4.43 (semi-discrete solution), we get $\forall v^h \in X_0^{h,p}\times[0,T]$,

$$i\left(\frac{\partial\left(\psi_\alpha - \psi_\alpha^h\right)}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla\left(\psi_\alpha - \psi_\alpha^h\right), \nabla v^h\right) + \left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v_{\mathrm{KS}}^h[\rho^h]\psi_\alpha^h, v^h\right). \tag{4.49}$$

We rewrite $\psi_\alpha - \psi_\alpha^h = (\psi_\alpha - \mathcal{P}_h\psi_\alpha) + (\mathcal{P}_h\psi_\alpha - \psi_\alpha^h)$ and derive bounds on each of the terms. For simpler notation, we use $u_\alpha = \psi_\alpha - \mathcal{P}_h\psi_\alpha$ and $w_\alpha = (\mathcal{P}_h\psi_\alpha - \psi_\alpha^h)$. Thus, using $\psi_\alpha - \psi_\alpha^h = u_\alpha + w_\alpha$, we rewrite Eq. 4.49 as

$$i\left(\frac{\partial w_\alpha}{\partial t}, v^h\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, v^h\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla v^h\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla v^h\right)$$
$$+ \left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v_{\mathrm{KS}}^h[\rho^h]\psi_\alpha^h, v^h\right). \tag{4.50}$$

Taking $v^h = w_\alpha$, we have

$$\begin{aligned} i\left(\frac{\partial w_\alpha}{\partial t}, w_\alpha\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla w_\alpha\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla w_\alpha\right) \\ + \left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right) . \end{aligned} \tag{4.51}$$

Noting that

$$\frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|^2_{L^2(\Omega)} = \mathrm{Re}\left\{\left(\frac{\partial}{\partial t}w_\alpha, w_\alpha\right)\right\} , \tag{4.52}$$

and comparing the imaginary parts of Eq. 4.51, we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|^2_{L^2(\Omega)} = -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla u_\alpha, \nabla w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla w_\alpha, \nabla w_\alpha\right)\right\} \\ + \mathrm{Im}\left\{\left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right\} \end{aligned} \tag{4.53}$$

In the above equation, we note that $\left(\nabla u_\alpha, \nabla w_\alpha\right) = 0$, as a consequence of Eq. 4.46. Furthermore, $\left(\nabla w_\alpha, \nabla w_\alpha\right)$ is real. Thus, Eq. 4.53 simplifies to,

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|^2_{L^2(\Omega)} &= -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \mathrm{Im}\left\{\left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right\} \\ &\leq \left|\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right| + \left|\left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right| . \end{aligned} \tag{4.54}$$

We now decompose v_{KS} into its components to rewrite the second term on the right of the above equation as

$$\begin{aligned} \left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right) = \left(v_{\mathrm{xc}}[\rho]\psi_\alpha - v_{\mathrm{xc}}[\rho^h]\psi^h_\alpha, w_\alpha\right) + \left(v_{\mathrm{H}}[\rho]\psi_\alpha - v_{\mathrm{H}}[\rho^h]\psi^h_\alpha, w_\alpha\right) \\ + \left(\left(v_{\mathrm{H}}[\rho^h] - v^h_{\mathrm{H}}[\rho^h]\right)\psi^h_\alpha, w_\alpha\right) + \left(v_{\mathrm{N}}\psi_\alpha - v_{\mathrm{N}}\psi^h_\alpha, w_\alpha\right) \\ + \left(v_{\mathrm{N}}\psi^h_\alpha - v^h_{\mathrm{N}}\psi^h_\alpha, w_\alpha\right) + \left(v_{\mathrm{field}}\psi_\alpha - v_{\mathrm{field}}\psi^h_\alpha, w_\alpha\right) . \end{aligned} \tag{4.55}$$

We note that the term $\left(v_{\mathrm{N}}\psi^h_\alpha - v^h_{\mathrm{N}}\psi^h_\alpha, w_\alpha\right)$, on the right side of the above equation,

is relevant only in the all-electron case (i.e., zero for the pseudopotential case as $v_{\mathrm{N}} = v_{\mathrm{N}}^h$). Combining the results from Eqs. 4.35, 4.37, 4.38, 4.39, and 4.40, with $v = w_\alpha$, and using the fact that $\psi_\alpha^h \in L^\infty(\Omega)$, it is straightforward to show that

$$\begin{aligned}\left|\left(v_{\mathrm{KS}}[\rho]\psi_\alpha - v_{\mathrm{KS}}^h[\rho^h]\psi_\alpha^h, w_\alpha\right)\right| \leq\ & C_0 \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)} \\ & + C_1 \left(\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)} + \left\|\rho - \rho^h\right\|_{L^1(\Omega)}\right) \|w_\alpha\|_{L^2(\Omega)} \\ & + C_2 \left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)} \\ & + C_3 \left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)}\ .\end{aligned} \tag{4.56}$$

Using the above result in Eq. 4.54, we obtain

$$\begin{aligned}\frac{d}{dt}\|w_\alpha\|_{L^2(\Omega)} &\leq \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0 \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)} \\ &\quad + C_1 \left(\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)} + \left\|\rho - \rho^h\right\|_{L^1(\Omega)}\right) \\ &\quad + C_2 \left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)} + C_3 \left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)} \\ &\leq \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0 \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)} \\ &\quad + C_2 \left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)} + C_3 \left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)} \\ &\quad + C_4 \sum_{\beta=1}^{N_e} \left\|\psi_\beta - \psi_\beta^h\right\|_{L^2(\Omega)}\ ,\end{aligned} \tag{4.57}$$

where we have used Eq. 4.32a in the second line to simplify the term involving

$\left\|\rho - \rho^h\right\|_{L^1(\Omega)}$. Summing the above equation over all index α, we have

$$\begin{aligned}
\frac{d}{dt}\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)} &\leq \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}\right) + C_6\left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)} \\
&\quad + C_7\left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)} \\
&\leq \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5\left\|u_\alpha\right\|_{L^2(\Omega)} + C_5\left\|w_\alpha\right\|_{L^2(\Omega)}\right) \\
&\quad + C_6\left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)} + C_7\left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)}\,,
\end{aligned} \tag{4.58}$$

where in the second line we have split $\psi_\alpha - \psi_\alpha^h$ into u_α and w_α. Now, integrating the above equation, gives

$$\begin{aligned}
\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) &\leq \sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(0) + C_5\int_0^t\sum_{\alpha=1}^{N_e}\|w_\alpha\|(s)\,ds \\
&\quad + C_5\int_0^t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s) + \|u_\alpha\|_{L^2(\Omega)}(s)\right)ds \\
&\quad + C_8\int_0^t\left(\left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)}(s) + \left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)}\right)ds\,.
\end{aligned} \tag{4.59}$$

Noting that $u_\alpha = \psi_\alpha - \mathcal{P}_h\psi_\alpha$, $\frac{\partial u_\alpha}{\partial t} = \frac{\partial \psi_\alpha}{\partial t} - \mathcal{P}_h\frac{\partial \psi_\alpha}{\partial t}$, and using the definitions of $s_{1,\alpha}$, $s_{2,\alpha}$, and s_3 (cf. Eq. 4.48), we can simplify the above equation as

$$\begin{aligned}
\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) &\leq \sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(0) + C_5\int_0^t\sum_{\alpha=1}^{N_e}\|w_\alpha\|(s)\,ds \\
&\quad + C_5 t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha})\right) \\
&\quad + C_8 t\left(\left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)}(s_3) + \left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)}\right).
\end{aligned} \tag{4.60}$$

Invoking the Grönwall's inequality on the above equation yields

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq e^{C_5 t} \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(0) \\ + e^{C_5 t} C_5 t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} (s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha}) \right) \\ + C_8 e^{C_5 t} t \left(\left\| v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h] \right\|_{L^2(\Omega)} (s_3) + \left\| v_{\mathrm{N}} - v_{\mathrm{N}}^h \right\|_{L^2(\Omega)} \right).\end{aligned} \tag{4.61}$$

Noting that $\|w_\alpha\|_{L^2(\Omega)}(0) \leq \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0)$, we rewrite the above equation as

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq C_5 e^{C_5 t} t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} (s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha}) \right) \\ + C_8 e^{C_5 t} t \left(\left\| v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h] \right\|_{L^2(\Omega)} (s_3) + \left\| v_{\mathrm{N}} - v_{\mathrm{N}}^h \right\|_{L^2(\Omega)} \right) \\ + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\| \psi_\alpha - \psi_\alpha^h \right\|_{L^2(\Omega)}(0).\end{aligned} \tag{4.62}$$

Bounds on the terms involving $\|u_\alpha\|_{L^2(\Omega)}$, $\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}$, $\left\|v_{\mathrm{H}}[\rho^h] - v_{\mathrm{H}}^h[\rho^h]\right\|_{L^2(\Omega)}$, and $\left\|v_{\mathrm{N}} - v_{\mathrm{N}}^h\right\|_{L^2(\Omega)}$, can now be obtained using the Ceá's lemma [193]— a standard finite element error estimates. The Ceá's lemma, in simple terms, is stated as follows. Let $\phi \in H^1(\Omega_T)$ and $\phi^h \in V^h \subseteq X^{h,p}$. If $y = \phi - \phi^h$ satisfies the following Galerkin orthogonality condition,

$$(\nabla y, \nabla v^h)(t) = 0, \quad \forall v^h \in V^h \text{ and } \forall t \in [0, T], \tag{4.63}$$

then

$$\|y\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} |\phi|_{p+1,\Omega_e}, \quad \text{and} \tag{4.64a}$$

$$\|y\|_{H^1(\Omega)} \leq C \sum_e h_e^{p} |\phi|_{p+1,\Omega_e}. \tag{4.64b}$$

By definition of Ritz projection (Eq. 4.46), $y = u_\alpha = \psi_\alpha - \mathcal{P}_h \psi_\alpha$ satisfies the Eq. 4.63.

Further, taking the time-derivative of Eq. 4.46, it is easy to verify that $y = \frac{\partial u_\alpha}{\partial t} = \frac{\partial \psi_\alpha}{\partial t} - \mathcal{P}_h \frac{\partial \psi_\alpha}{\partial t}$ also satisfies the Eq. 4.63. Thus, applying the Ceá's lemma (Eq. 4.64a) to u_α and $\frac{\partial u_\alpha}{\partial t}$ yields

$$\|u_\alpha\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left|\psi_\alpha\right|_{p+1,\Omega_e}, \quad \text{and} \tag{4.65}$$

$$\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left| \frac{\partial \psi_\alpha}{\partial t} \right|_{p+1,\Omega_e}. \tag{4.66}$$

We further simplify the above inequality, by using Eq. 4.1

$$\begin{aligned}
\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} &\leq C \sum_e h_e^{p+1} \left| \frac{\partial \psi_\alpha}{\partial t} \right|_{p+1,\Omega_e} = C \sum_e h_e^{p+1} \left| -\frac{1}{2}\nabla^2 \psi_\alpha + v_{\text{KS}}[\rho]\psi_\alpha \right|_{p+1,\Omega_e} \\
&\leq C \sum_e h_e^{p+1} \left(\left|\psi_\alpha\right|_{p+3,\Omega_e} + \left|(v_{\text{H}} + v_{\text{N}} + v_{\text{xc}} + v_{\text{field}})\psi_\alpha\right|_{p+1,\Omega_e} \right) \\
&\leq C \sum_e h_e^{p+1} \left(\left|\psi_\alpha\right|_{p+3,\Omega_e} + \left|\psi_\alpha\right|_{p+1,\Omega_e} \right),
\end{aligned} \tag{4.67}$$

which follows from the definition of the $|.|_{p+3}$ semi-norm and the boundedness assumptions on v_{N}, v_{xc}, and v_{field} (assumptions $\mathcal{A}2$–$\mathcal{A}6$). Lastly, it is straightforward to observe that both $y = v_{\text{H}}[\rho^h] - v_{\text{H}}^h[\rho^h]$ and $y = v_{\text{N}} - v_{\text{N}}^h$ satisfy Eq. 4.63 (take the difference of Eqs. 4.44 and 4.42a; and Eqs. 4.45 and 4.42b, respectively). Thus, once again, applying the Ceá's lemma (Eq. 4.64a), we get

$$\left\| v_{\text{H}}[\rho^h] - v_{\text{H}}^h[\rho^h] \right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left| v_{\text{H}}[\rho^h] \right|_{p+1,\Omega_e}. \tag{4.68a}$$

$$\left\| v_{\text{N}} - v_{\text{N}}^h \right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left| v_{\text{N}} \right|_{p+1,\Omega_e}. \tag{4.68b}$$

Using Eqs. 4.65, 4.67, and 4.68 in Eq. 4.62, we have

$$\begin{aligned}
\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq\ & C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right) \\
& + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) .
\end{aligned} \tag{4.69}$$

Finally, expressing $\psi_\alpha - \psi_\alpha^h = w_\alpha + u_\alpha$ and using the result of Eq. 4.65 in the above equation, we obtain

$$\begin{aligned}
\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq\ & C \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} |\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) \\
& + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right) \\
& + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \\
\leq\ & C_{10} e^{C_5 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{10} e^{C_5 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{10} e^{C_5 t} t \sum_e h_e^{p+1} \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right) \\
& + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) .
\end{aligned} \tag{4.70}$$

This concludes the proof of Eq. 4.47a.

In order to derive estimates for $\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t)$, we use the inverse estimate [178] for $w_\alpha = (\mathcal{P}_h\psi_\alpha - \psi_\alpha^h) \in X_0^{h,p}$ to obtain

$$\|w_\alpha\|_{H^1(\Omega)}(t) \leq C h_{min}^{-1} \|w_\alpha\|_{L^2(\Omega)}(t). \tag{4.71}$$

Additionally, applying the Ceá's lemma (Eq. 4.64b) on $u_\alpha = (\psi_\alpha - \mathcal{P}_h\psi_\alpha)$, we have

$$\|u_\alpha\|_{H^1(\Omega)}(t) \leq C \sum_e h_e^p \left|\psi_\alpha\right|_{p+1,\Omega_e}(t). \tag{4.72}$$

Combining Eqs. 4.71 and 4.72, we get

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) &\leq \sum_{\alpha=1}^{N_e} \Big(\|u_\alpha\|_{H^1(\Omega)}(t) + \|w_\alpha\|_{H^1(\Omega)}(t)\Big) \\ &\leq C_{11} \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left|\psi_\alpha\right|_{p+1,\Omega_e}(t) + C_{12} h_{min}^{-1} \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t).\end{aligned} \tag{4.73}$$

Finally, using the inequality obtained in Eq. 4.69 in the above equation and using the

fact that $h_e/h_{min} \leq C$ for all the elements in $\mathcal{T}^{h,p}$, yields

$$
\begin{aligned}
\sum_{\alpha}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq\; & C_{11} \sum_e h_e^p \sum_{\alpha=1}^{N_e} |\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) \\
& + C_{13} e^{C_5 t} t \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{13} e^{C_5 t} t \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{13} e^{C_5 t} t \sum_e h_e^p \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right) \\
& + C_{12} e^{C_5 t} h_{min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \\
\leq\; & C_{14} e^{C_5 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{14} e^{C_5 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\
& + C_{14} e^{C_5 t} t \sum_e h_e^p \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right) \\
& + C_{14} e^{C_5 t} h_{min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) .
\end{aligned}
\tag{4.74}
$$

This concludes the proof for Eq. 4.47b. □

Corollary IV.2. *If the initial orbitals $\psi_\alpha(\boldsymbol{r}, 0)$ are obtained from a ground-state DFT calculation, wherein [38]*

$$
\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \leq C \sum_e h_e^{p+1} \left(|\psi_\alpha|_{p+1,\Omega_e} + \left|v_H[\rho^h]\right|_{p+1,\Omega_e} + |v_N|_{p+1,\Omega_e} \right) , \tag{4.75}
$$

the results of Proposition IV.1 can be simplified, $\forall t \in [0,T]$, to

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) + C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) + C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \left(\left|v_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_N|_{p+1,\Omega_e} \right) . \tag{4.76a}$$

$$\sum_{\alpha}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq C_3' e^{C_2 t}(t+1) \sum_e h_e^{p} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) \right) + C_3' e^{C_2 t}(t+1) \sum_e h_e^{p} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) + C_3' e^{C_2 t}(t+1) \sum_e h_e^{p} \left(\left|v_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_N|_{p+1,\Omega_e} \right) \tag{4.76b}$$

The last equation concludes the proof of Eq. 4.28.

4.4.5 Derivation of time discretization error

Before proceeding to the proof for Eq. 4.29, we note that for an exponential operator of the form $e^{\mathbf{L}(t)}$, the partial derivative with respect to t is given by [191],

$$\frac{\partial}{\partial t} e^{\mathbf{L}(t)} = \mathrm{dexp}_{\mathbf{L}(t)} \left(\dot{\mathbf{L}}(t) \right) e^{\mathbf{L}(t)}, \tag{4.77}$$

where $\mathrm{dexp}_{\mathbf{X}}(\mathbf{Y}) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_{\mathbf{X}}^k(\mathbf{Y})$. The operator $\mathrm{ad}_{\mathbf{X}}^k(\mathbf{Y})$ is defined recursively as

$$\mathrm{ad}_{\mathbf{X}}^k(\mathbf{Y}) = \mathrm{ad}_{\mathbf{X}} \left(\mathrm{ad}_{\mathbf{X}}^{k-1}(\mathbf{Y}) \right) , \tag{4.78}$$

with $\mathrm{ad}_{\mathbf{X}}^1(\mathbf{Y}) = \mathbf{XY} - \mathbf{YX}$, and $\mathrm{ad}_{\mathbf{X}}^0(\mathbf{Y}) = \mathbf{Y}$.

We now present the proof for Eq. 4.29 in the following Proposition. In the follow-

ing analysis, we assume each time interval $[t_{n-1}, t_n]$ to be of length Δt. Moreover, for simpler terminology, we term $e^{\tilde{\mathbf{A}}_n}$ (cf. Eq. 4.25) as the second-order Magnus propagator without explicitly spelling out the mid-point integration rule invoked in it.

Proposition IV.3. *For a second-order Magnus propagator with a mid-point integration rule, we obtain the following bound for the time-discretization error in* ψ_α^h

$$\left\|\psi_\alpha^h(t_n) - \psi_\alpha^{h,n}\right\|_{L^2(\Omega)} \leq C(\Delta t)^2 t_n \max_{0\leq t\leq t_n} \left\|\psi_\alpha^h(t)\right\|_{H^1(\Omega)} , \tag{4.79}$$

Proof. To begin with, we introduce the following operators

$$\mathbf{S}_0^k = e^{\mathbf{A}_k} e^{\mathbf{A}_{k-1}} \dots e^{\mathbf{A}_1} = \prod_{l=0}^{k-1} e^{\mathbf{A}_{k-l}} \quad \text{for } 0 < k \leq n\,, \quad \mathbf{S}_0^0 = I \tag{4.80a}$$

$$\mathbf{R}_k^n = e^{\tilde{\mathbf{A}}_n} e^{\tilde{\mathbf{A}}_{n-1}} \dots e^{\tilde{\mathbf{A}}_{k+1}} = \prod_{l=0}^{n-k-1} e^{\tilde{\mathbf{A}}_{n-l}} \quad \text{for } 0 \leq k < n\,, \quad \mathbf{R}_n^n = I\,. \tag{4.80b}$$

To elaborate, S_0^k denotes the exact Magnus propagator from t_0 to t_k, and R_k^n denotes the second-order Magnus propagator from t_k to t_n. Let $\boldsymbol{\psi}_\alpha^h(t_n)$ and $\boldsymbol{\psi}_\alpha^{h,n}$ denote the vectors containing the finite element expansion coefficients for $\psi_\alpha^h(t_n)$ and $\psi_\alpha^{h,n}$, respectively. Further, let $\bar{\boldsymbol{\psi}}_\alpha(t_n) = \mathbf{M}^{1/2}\boldsymbol{\psi}_\alpha(t_n)$ and $\bar{\boldsymbol{\psi}}_\alpha^n = \mathbf{M}^{1/2}\boldsymbol{\psi}_\alpha^n$. Thus, we can rewrite the time-discretization error in $\bar{\boldsymbol{\psi}}_\alpha(t_n)$ in terms of the following telescopic series,

$$\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n = \left(\mathbf{R}_n^n\mathbf{S}_0^n - \mathbf{R}_0^n\mathbf{S}_0^0\right)\bar{\boldsymbol{\psi}}_\alpha(0) = \sum_{k=1}^{n}\left(\mathbf{R}_k^n\mathbf{S}_0^k - \mathbf{R}_{k-1}^n\mathbf{S}_0^{k-1}\right)\bar{\boldsymbol{\psi}}_\alpha(0)\,. \tag{4.81}$$

Noting that $\mathbf{S}_0^k = e^{\mathbf{A}_k}\mathbf{S}_0^{k-1}$ and $\mathbf{R}_{k-1}^n = \mathbf{R}_k^n e^{\tilde{\mathbf{A}}_k}$, we rewrite the above equation as

$$\begin{aligned}\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n &= \sum_{k=1}^{n}\left(\mathbf{R}_k^n\mathbf{S}_0^k - \mathbf{R}_{k-1}^n\mathbf{S}_0^{k-1}\right)\bar{\boldsymbol{\psi}}_\alpha(0) = \sum_{k=1}^{n}\left(\mathbf{R}_k^n e^{\mathbf{A}_k}\mathbf{S}_0^{k-1} - \mathbf{R}_k^n e^{\tilde{\mathbf{A}}_k}\mathbf{S}_0^{k-1}\right)\bar{\psi}_\alpha(0) \\ &= \sum_{k=1}^{n}\mathbf{R}_k^n\left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right)\mathbf{S}_0^{k-1}\bar{\boldsymbol{\psi}}_\alpha(0) \\ &= \sum_{k=1}^{n}\mathbf{R}_k^n\left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\,.\end{aligned} \tag{4.82}$$

Since $\mathbf{R}_k^n$ is a unitary operator, bounding $\left(\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n\right)$ reduces to finding the bound on $\left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$. To this end, we extend the proof presented in [192] to the non-linear case of the TDKS equations. To begin with, we split $\left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ as

$$\left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) = \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) + \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\,, \tag{4.83}$$

where $\bar{\mathbf{A}}_k = \int_{t_{k-1}}^{t_k} -i\bar{\mathbf{H}}[\rho(t)]dt$. The two terms on the right hand side of the above equation denote the error due to truncation of the Magnus expansion and the time integral approximation, respectively.

In order to bound the error in the first term on the right side of Eq. 4.83, we introduce the following auxiliary function

$$\boldsymbol{\xi}_\alpha^k(t) = e^{\mathbf{B}_k(t)}\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}), \qquad \forall t \in [t_{k-1}, t_k]\,, \tag{4.84}$$

where $\mathbf{B}_k(t) = \int_{t_{k-1}}^{t} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau$. We remark that $\boldsymbol{\xi}_\alpha^k(t)$ denotes the time-evolution of $\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ using the truncated the Magnus expansion, in the time interval $[t_{k-1}, t_k]$. Differentiating the above equation and using the result of Eq. 4.77 gives

$$\dot{\boldsymbol{\xi}}_\alpha^k(t) = \mathrm{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))e^{\mathbf{B}_k(t)}\boldsymbol{\xi}_\alpha^k(t_{k-1}) = -i\bar{\mathbf{G}}_k(t)\boldsymbol{\xi}_\alpha^k(t)\,, \qquad \forall t \in [t_{k-1}, t_k], \tag{4.85}$$

where $\bar{\mathbf{G}}_k(t) = i\text{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))$. We observe that $\bar{\mathbf{G}}_k$ is Hermitian. This can be proven as follows. First, note that for two Hermitian (or skew-Hermitian) matrices $\mathbf{X}$, $\mathbf{Y}$, the operator $\text{ad}_{\mathbf{X}}(\mathbf{Y})$ is skew-Hermitian. Second, owing to the Hermiticity of $\bar{\mathbf{H}}$, both $\mathbf{B}_k(t) = \int_{t_{k-1}}^{t} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau$ and $\dot{\mathbf{B}}_k(t) = -i\bar{\mathbf{H}}(t)$ ($\forall t \in [t_{k-1}, t_k]$) are skew-Hermitian. Thus, by expanding $\text{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))$ and using the above two arguments, it can be shown that $\bar{\mathbf{G}}_k$ is Hermitian. We now introduce the function $\boldsymbol{\gamma}_\alpha^k(t) = \bar{\boldsymbol{\psi}}_\alpha(t) - \boldsymbol{\xi}_\alpha^k(t),\ \forall t \in [t_{k-1}, t_k]$. It is important to note that

$$\boldsymbol{\gamma}_\alpha^k(t_k) = \bar{\boldsymbol{\psi}}_\alpha(t_k) - \boldsymbol{\xi}_\alpha^k(t_k) = \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\,, \tag{4.86}$$

where the second equality follows from the definition of $\boldsymbol{\xi}_\alpha^k$ (Eq. 4.84) and the fact that $\mathbf{B}_k(t_k) = \int_{t_{k-1}}^{t_k} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau = \bar{\mathbf{A}}_k$. Thus, the problem of bounding $\left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ (the first term in Eq. 4.83) reduces to bounding $\boldsymbol{\gamma}_\alpha^k(t_k)$. To this end, we proceed, by first expressing the time-derivative of $\boldsymbol{\gamma}_\alpha^k$ as

$$\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t))\bar{\boldsymbol{\psi}}_\alpha(t), \qquad \forall t \in [t_{k-1}, t_k]\,, \tag{4.87}$$

which follows from Eqs. 4.21 and 4.85. Now, taking the dot product with $\boldsymbol{\gamma}_\alpha^k(t)^\dagger$ on both sides yields

$$\boldsymbol{\gamma}_\alpha^k(t)^\dagger\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\boldsymbol{\gamma}_\alpha^k(t)^\dagger\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i\boldsymbol{\gamma}_\alpha^k(t)^\dagger\left(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\,. \tag{4.88}$$

We note that $2\text{Re}\{\boldsymbol{\gamma}_\alpha^k(t)^\dagger\dot{\boldsymbol{\gamma}}_\alpha^k(t)\} = \frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\|^2$, where $\|.\|$ represents the Euclidean norm of a vector. Further, we note $\boldsymbol{\gamma}_\alpha^k(t)^\dagger\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t)$ is real, owing to the Hermiticity of $\bar{\mathbf{G}}$. Thus, comparing the real parts of the above equation results in

$$\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\|^2 = \text{Im}\{\boldsymbol{\gamma}_\alpha^k(t)^\dagger\left(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\}\,. \tag{4.89}$$

Consequently,

$$\frac{d}{dt}\left\|\boldsymbol{\gamma}_{\alpha}^{k}\right\| \leq \left\|\left(\bar{\mathbf{H}}(t)-\bar{\mathbf{G}}_{k}(t)\right)\bar{\boldsymbol{\psi}}_{\alpha}(t)\right\| . \tag{4.90}$$

Time integrating the above equation yields

$$\begin{aligned}\left\|\boldsymbol{\gamma}_{\alpha}^{k}\right\|(t_k) &= \left\|\bar{\boldsymbol{\psi}}_{\alpha}(t_k)-\boldsymbol{\xi}_{\alpha}^{k}(t_k)\right\| = \left\|\left(e^{\mathbf{A}_k}-e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_{\alpha}(t_{k-1})\right\| \\ &\leq \int_{t_{k-1}}^{t_k}\left\|\left(\bar{\mathbf{H}}(\tau)-\bar{\mathbf{G}}_{k}(\tau)\right)\bar{\boldsymbol{\psi}}_{\alpha}(\tau)\right\| d\tau ,\end{aligned} \tag{4.91}$$

where we have used the result of Eq. 4.86 along with the fact that $\left\|\boldsymbol{\gamma}_{\alpha}^{k}\right\|(t_{k-1}) = \bar{\boldsymbol{\psi}}_{\alpha}(t_{k-1})-\boldsymbol{\xi}_{\alpha}^{k}(t_{k-1}) = 0$ (by the definition of $\boldsymbol{\xi}_{\alpha}^{k}(t)$, cf. Eq. 4.84). Thus, the problem of bounding $\left\|\left(e^{\mathbf{A}_k}-e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_{\alpha}(t_{k-1})\right\|$ further simplifies to finding a bound for $\int_{t_{k-1}}^{t_k}\left\|\left(\bar{\mathbf{H}}(\tau)-\bar{\mathbf{G}}_{k}(\tau)\right)\bar{\boldsymbol{\psi}}_{\alpha}(\tau)\right\| d\tau$. To this end, we use the fact that $\bar{\mathbf{G}}_{k}(\tau) = i\mathrm{dexp}_{\mathbf{B}_k(\tau)}(\dot{\mathbf{B}}_{k}(\tau))$ and the definition of the operator $\mathrm{dexp}_{\mathbf{X}}(\mathbf{Y})$, to obtain

$$\bar{\mathbf{H}}(\tau)-\bar{\mathbf{G}}_{k}(\tau) = -\frac{i}{2}[\mathbf{B}_{k}(\tau),\dot{\mathbf{B}}_{k}(\tau)]+h.o.t = -\frac{i}{2}\int_{t_{k-1}}^{\tau}[\bar{\mathbf{H}}(\tau),\bar{\mathbf{H}}(\sigma)]d\sigma + h.o.t. , \tag{4.92}$$

where *h.o.t.* stands for higher-order terms. In order to bound $[\bar{\mathbf{H}}(\tau),\bar{\mathbf{H}}(\sigma)]$, we begin by rewriting $\bar{\mathbf{H}}$ in terms of $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$, i.e., its kinetic and Kohn-Sham potential components. To elaborate, $\bar{\mathbf{U}} = \mathbf{M}^{-1/2}\mathbf{U}\mathbf{M}^{-1/2}$ and $\bar{\mathbf{V}} = \mathbf{M}^{-1/2}\mathbf{V}\mathbf{M}^{-1/2}$, with

$$U_{jk} = \frac{1}{2}\int_{\Omega}\nabla N_j(\mathbf{r}).\nabla N_k(\mathbf{r})\, d\mathbf{r} , \tag{4.93}$$

$$V_{jk} = \int_{\Omega} v_{\mathrm{KS}}^{h}[\rho^{h}](\mathbf{r},t)N_j(\mathbf{r})N_k(\mathbf{r})\, d\mathbf{r} . \tag{4.94}$$

Noting that $\bar{\mathbf{U}}$ is time-independent, we Taylor expand $\bar{\mathbf{H}}(\sigma)$ about τ to rewrite

$[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)]$ as

$$[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)] = [\bar{\mathbf{H}}(\tau), \bar{\mathbf{V}}'(\tau)](\sigma - \tau) + \mathcal{O}((\sigma - \tau)^2)\,, \tag{4.95}$$

where $\bar{\mathbf{V}}'(\tau) = \frac{d}{dt}(\bar{\mathbf{V}}(t))|_{\sigma=\tau}$. Thus, using the above relation in Eq. 4.92 we get

$$(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau))\bar{\boldsymbol{\psi}}_\alpha(\tau) = \frac{i}{4}\left([\bar{\mathbf{H}}, \bar{\mathbf{V}}'(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau)\right)(\tau - t_{k-1})^2 + \mathcal{O}((\tau - t_{k-1})^3)\,. \tag{4.96}$$

We now invoke the boundedness assumption on $\bar{\mathbf{V}}'$ (assumption $\mathcal{A}8$), and the norm equivalence of $\bar{\mathbf{U}}$ and $\bar{\mathbf{H}}$ (assumption $\mathcal{A}7$), to obtain

$$\left\|(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau))\bar{\boldsymbol{\psi}}_\alpha(\tau)\right\| \leq C(\tau - t_{k-1})^2 \left\|\bar{\boldsymbol{\psi}}_\alpha(\tau)\right\|_{\bar{\mathbf{U}}} + \mathcal{O}((\tau - t_{k-1})^3). \tag{4.97}$$

Thus, substituting the above result into Eq. 4.91 provides the following bound

$$\left\|\left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right)\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\right\| \leq C(\Delta t)^3 \max_{t_{k-1}\leq t\leq t_k} \left\|\bar{\boldsymbol{\psi}}_\alpha(t)\right\|_{\bar{\mathbf{U}}}\,. \tag{4.98}$$

This provides a bound for the first term (truncation error) on the right side of Eq. 4.83. In order to bound the second term on the right side of Eq. 4.83, i.e., the error due to mid-point quadrature rule, we begin with the following identity

$$e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k} = \int_0^1 \frac{d}{dx}\left(e^{(1-x)\tilde{\mathbf{A}}_k} e^{x\bar{\mathbf{A}}_k}\right)\,dx = \int_0^1 e^{(1-x)\tilde{\mathbf{A}}_k}(\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k)e^{x\bar{\mathbf{A}}_k}\,dx. \tag{4.99}$$

Furthermore, we note that for a function $f(x)$ if $F_{1/2}$ denotes the midpoint approximation to $F = \int_a^b f(x)\,dx$, then $\left|F - F_{1/2}\right| \leq C(b-a)^3 f''(\eta)$, for some $\eta \in [a, b]$. Thus, for the mid-point integration rule, $\left\|\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k\right\| \leq C(\Delta t)^3 \left\|\frac{d^2}{dt^2}(\bar{\mathbf{H}})|_{t'}\right\|$ for some $t' \in [t_{k-1}, t_k]$. Using this result along with the unitarity of the operators $e^{(1-x)\tilde{\mathbf{A}}_k}$ and

$e^{x\bar{\mathbf{A}}_k}$, we obtain

$$\left\| \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k} \right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| \leq \left\| (\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k) \right\| \leq C(\Delta t)^3 \left\| \frac{d^2}{dt^2}(\bar{\mathbf{H}})|_{t'} \right\|, \text{ for some } t' \in [t_{k-1}, t_k]\,. \tag{4.100}$$

Noting that $\frac{d^2}{dt^2}\bar{\mathbf{H}} = \frac{d^2}{dt^2}\bar{\mathbf{V}}$, $\forall t \in [t_{k-1}, t_k]$, and invoking the boundedness assumption on $\frac{d^2}{dt^2}\bar{\mathbf{V}}$ (assumption $\mathcal{A}8$), we get

$$\left\| \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k} \right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| \leq C(\Delta t)^3\,. \tag{4.101}$$

Thus, using the results of Eqs. 4.98 and 4.101 in Eq. 4.82 along with the unitarity of the operators $\mathbf{R}_k^n$, yields

$$\left\| \bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n \right\| \leq C(\Delta t)^2 t_n \max_{t_0 \leq t \leq t_n} \left\| \bar{\boldsymbol{\psi}}_\alpha(t) \right\|_{\bar{\mathbf{U}}}\,. \tag{4.102}$$

Finally, noting that the coefficient vectors for the spatial fields $\psi_\alpha^h(\mathbf{r}, t_n)$ and $\psi_\alpha^{h,n}(\mathbf{r})$ are given by $\mathbf{M}^{-1/2}\bar{\boldsymbol{\psi}}_\alpha^h(t_n)$ and $\mathbf{M}^{-1/2}\bar{\boldsymbol{\psi}}_\alpha^{h,n}$, respectively, it is now trivial to arrive at Eq. 4.29 from the above equation. □

4.5 Efficient spatio-temporal discretization

We now utilize our spatial and temporal discretization error estimates (Eqs. 4.28 and 4.29) to obtain an efficient spatio-temporal discretization. We follow along the lines of [194, 113, 38] to obtain an efficient spatial discretization by minimizing the semi-discrete error in the dipole moment at a given time, subject to a fixed number of finite elements. We remark that the choice of dipole moment as an observable for this exercise is solely a matter of convenience, and any observable which can be inexpensively evaluated in terms of the density or the Kohn-Sham orbitals can be used instead. Representing the x-component of the continuous and the semi-discrete

dipole moment at time t as $\mu_x(t)$ and $\mu_x^h(t)$, respectively, we have

$$\begin{aligned} |\mu_x(t) - \mu_x^h(t)| &\leq \|x\|_{L^2(\Omega)} \left\|\rho(t) - \rho^h(t)\right\|_{L^2(\Omega)} \\ &\leq C \left\|\rho(t) - \rho^h(t)\right\|_{L^2(\Omega)} \\ &\leq C \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t)\,. \end{aligned} \tag{4.103}$$

Now using Eq. 4.28 in the above equation, results in

$$\begin{aligned} |\mu_x(t) - \mu_x^h(t)| \leq C_1(t) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \Big(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \Big) \\ + C_1(t) \sum_e h_e^p \left(\left|v_{\mathrm{H}}[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |v_{\mathrm{N}}|_{p+1,\Omega_e} \right)\,, \end{aligned} \tag{4.104}$$

for some $\{s_{1,\alpha}\}, \{s_{2,\alpha}\},$ and $s_3 \in [0,t]$. We now use the definition of the semi-norm (in terms of partial spatial derivative) and introduce an element size distribution, $h(\mathbf{r})$, to rewrite the above equation as

$$\begin{aligned} |\mu_x(t) - \mu_x^h(t)| \leq C_1(t) \int_\Omega h^p(\mathbf{r}) \left[\sum_{\alpha=1}^{N_e} \left(\bar{D}^{p+1}\psi_\alpha(s_{1,\alpha}) + \bar{D}^{p+1}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+3}\psi_\alpha(s_{2,\alpha}) \right) \right] d\mathbf{r} \\ + C_1(t) \int_\Omega h^p(\mathbf{r}) \left(\bar{D}^{p+1} v_{\mathrm{H}}[\rho^h] + \bar{D}^{p+1} v_{\mathrm{N}} \right)\, d\mathbf{r}\,, \end{aligned} \tag{4.105}$$

where $\bar{D}^k f = \sum_{|l|=k} |D^l f|$, with D^l being the $l-$th order spatial derivative (i.e., $D^l f = \frac{\partial^l f}{\partial x^{l_1} \partial y^{l_2} \partial z^{l_3}}$, with $l = \{l_1, l_2, l_3\}$ and $|l| = l_1 + l_2 + l_3$). Thus, obtaining the optimal element size distribution, for a fixed number of elements (N_{elem}), reduces to the following minimization problem,

$$\min_{h(\mathbf{r})} \int_\Omega h^p(\mathbf{r}) \sum_{\alpha=1}^{N_e} \mathcal{L}_\alpha(\mathbf{r})\, d\mathbf{r}\,; \quad \text{subject to}: \int_\Omega \frac{d\mathbf{r}}{h^3(\mathbf{r})} = N_{elem}\,, \tag{4.106}$$

where

$$\mathcal{L}_\alpha(\mathbf{r}) = \bar{D}^{p+1}\psi_\alpha(s_{1,\alpha}) + \bar{D}^{p+1}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+3}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+1}v_{\mathrm{H}}[\rho^h](s_3) + \bar{D}^{p+1}v_{\mathrm{N}} \tag{4.107}$$

Solving the Euler-Lagrange equation corresponding to the above optimization problem yields the following element size distribution,

$$h(\mathbf{r}) = E\left[\sum_{\alpha=1}^{N_e}\mathcal{L}_\alpha(\mathbf{r})\right]^{\frac{-1}{p+3}} \tag{4.108}$$

where the constant E is evaluated from the constraint on the number of elements. We remark that although the above discretization approach requires *a priori* knowledge of the unknown $\psi_\alpha(s_{1,\alpha})$, $\psi_\alpha(s_{2,\alpha})$, and $v_{\mathrm{H}}[\rho^h](s_3)$, we use the ground-state atomic solutions to the Kohn-Sham orbitals and the electrostatic potentials to construct the adaptive finite element mesh for all calculations. Although, this does not inform us about the optimal discretization required near the nuclei, this affords an efficient strategy to accurately handle the regions away from the nuclei, wherein the time-dependent Kohn-Sham orbitals, typically, have similar decay properties as their ground-state counterparts. We note that, in practice, the finite element mesh obtained can deviate from $h(\mathbf{r})$ due to conformity and quality requirements, especially in the context of hexahedral elements that are employed in this work. However, the resulting finite element mesh broadly captures the optimal coarse-graining rate, and has the general adaptive characteristics that significantly enhance the computational efficiency, as demonstrated in Sec. 4.7. Figure 4.1 shows an adaptive mesh for all-electron Al_2 generated using this approach.

Having determined the required spatial discretization, a suitable temporal discretization for the given finite element mesh can be estimated by using the time-discrete error bound for the dipole moment. To elaborate, if $\mu_x^{h,n}$ denotes the x-

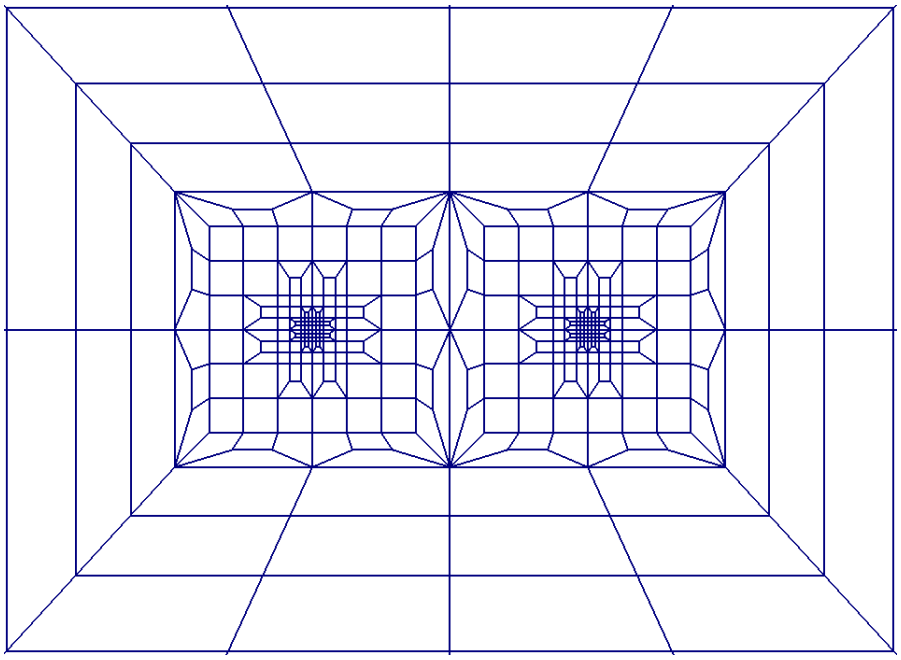


Figure 4.1: Adaptive finite element mesh for all-electron Al_2 (slice shown on the plane of the molecule)

component of the full-discrete dipole moment, then using the result in Eq. 4.29 it is straightforward to show,

$$|\mu_x^h(t_n) - \mu_x^{h,n}| \leq C t_n (\Delta t)^2 \sum_{\alpha=1}^{N_e} \max_{0 \leq t \leq t_n} \left\| \psi_\alpha^h(t) \right\|_{H^1(\Omega)} . \tag{4.109}$$

As is evident from the above relation, our choice of Δt is intrinsically tied to the spatial discretization through $\left\| \psi_\alpha^h(t) \right\|_{H^1(\Omega)}$. Furthermore, we remark that although the presence of t_n in the above inequality indicates increasing time-discretization error with time, it does not pose a limitation in containing the errors due to the fact that $t_n \leq T$, where, typically, T lies between $10-30$ fs. Now, in order to evaluate a suitable Δt that can contain the above error bound to a fixed tolerance, we need to estimate the values of $\left\| \psi_\alpha^h(t) \right\|_{H^1(\Omega)}$ and the value of the constant, C, featuring in it. The value of $\left\| \psi_\alpha^h(t) \right\|_{H^1(\Omega)}$ can be reliably approximated from its ground-state value, i.e., $\left\| \psi_\alpha^h(0) \right\|_{H^1(\Omega)}$. The characteristic value of the constant, C, is determined from atomic RT-TDDFT calculations at different Δt. To elaborate, for a given atom, the constant

can be evaluated from a linear fit to the log-log plot of the error $|\mu_x^h(t_n) - \mu_x^{h,n}|$ with respect to Δt. For a multi-atom system, we use the least Δt obtained for each of its constituent atomic species.

4.6 Numerical Implementation

We now discuss some of the key numerical aspects involved in our implementation of the finite element discretization of the TDKS equations.

4.6.1 Higher-order spectral finite elements

Finite-elements, with their varied choices of forms and orders [189, 109], have been widely used in several engineering applications. While the use of linear finite elements remains popular in engineering applications that warrant moderate levels of accuracy, it remains computationally inefficient to attain chemical accuracy in electronic structure calculations. To highlight, the use of linear finite elements have been shown to require large number of basis functions per atom ($\sim$100,000) to achieve chemical accuracy in ground-state DFT calculations [186, 47]. However, this shortcoming has been, demonstrably, mitigated by the use of higher-order finite elements [113, 38]. In this work, we explore the possibility of similar gains from using higher-order finite elements for RT-TDDFT calculations. Unlike conventional finite elements, we employ spectral finite elements for the spatial discretization of the TDKS equations. We refer to Sec. 3.4.2 for a discussion on spectral finite elements and its key advantages. A major advantage of the spectral finite elements is realized when used in conjunction with the Gauss-Legendre-Lobatto (GLL) quadrature rule for evaluating the integrals involved in the overlap matrix $\mathbf{M}$, wherein the quadrature points are coincident with the nodal points in the spectral finite elements. Such a combination renders $\mathbf{M}$ diagonal, thereby greatly simplifying the evaluation of $\mathbf{M}^{-1/2}$ (or $\mathbf{M}^{1/2}$) that features in the discrete TDKS equations (cf. Eq. 4.21). We note that while an n point rule in

the conventional Gauss quadrature rule integrates polynomials exactly up to degree $2n-1$, an n point GLL quadrature rule integrates polynomials exactly only up to degree $2n-3$. Therefore, we employ the GLL quadrature rule only in the construction of $\mathbf{M}$, while the more accurate Gauss quadrature rule is used for all other integrals featuring in the discrete TDKS equations. We refer to Motamarri *et al.* [38] for a discussion on the accuracy and sufficiency of GLL quadrature in the evaluation of overlap matrix $\mathbf{M}$. For the sake of brevity, we use the term finite elements instead of spectral finite elements in all subsequent discussions in this chapter.

4.6.2 Approximating the second-order Magnus operator

The form of the Magnus operator, as shown in Eq. 4.26, calls for efficient means of evaluating $\exp(\tilde{\mathbf{A}}_n)\bar{\boldsymbol{\psi}}_\alpha^{n-1}$. Direct means of evaluating the matrix $\exp(\tilde{\mathbf{A}}_n)$ remain computationally prohibitive beyond small sizes of $\tilde{\mathbf{A}}_n$. Thus, one needs to evaluate the action of $\exp(\tilde{\mathbf{A}}_n)$ on $\bar{\boldsymbol{\psi}}_\alpha^{n-1}$ in an iterative fashion. Several such schemes are in use in RT-TDDFT calculations, namely, Taylor series expansion, Chebyshev polynomial expansion, split-operator techniques, and Krylov subspace projection method. We refer to Castro *et al.* [195] and references there-within for a detailed discussion on each of these schemes. In this work, we adopt the Krylov subspace projection method for its superior efficiency and robustness compared to the other methods. To elaborate, the Krylov subspace projection allows for an *a posteriori* error control mechanism based on error estimates that will be described below. On the other hand, polynomial expansion methods such as Taylor series or Chebyshev polynomial expansion offer no such *a posteriori* mechanism. While one can use *a priori* estimates, based on the spectral radius of $\tilde{\mathbf{A}}_n$, the number of terms required in the polynomial expansion, for a desired accuracy, remain highly over-estimated. Furthermore, in the case of split-operator, the efficacy of it rests on operating back and forth between Fourier and real space, so as to diagonalize, in succession, the kinetic and the potential part

of the Kohn-Sham Hamiltonian. Thus, it involves the use of fast Fourier transforms (FFTs) which pose serious challenges to the parallel scalability of the code. The Krylov subspace projection, on the other hand involves operations only in real space and affords good parallel scalability (as will be shown in Sec. 4.7.4).

We now discuss the details of the Krylov subspace projection approach for the second-order Magnus operator. To begin with, a k-dimensional Krylov subspace for the matrix $\tilde{\mathbf{A}}_n$ and the vector $\bar{\boldsymbol{\psi}}$ (a generic representation for the Kohn-Sham vectors $\bar{\boldsymbol{\psi}}_\alpha^{n-1}$) is given by

$$\mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}}) = span\{\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n^{\,2}\bar{\boldsymbol{\psi}}, \ldots, \tilde{\mathbf{A}}_n^{\,k-1}\bar{\boldsymbol{\psi}}\}. \tag{4.110}$$

The Lanczos iteration provides a recipe for generating an orthonormal set of vectors $\mathbf{Q}_k = \{q_1, q_2, \ldots, q_k\}$, with $q_1 = \bar{\boldsymbol{\psi}}/\left\|\bar{\boldsymbol{\psi}}\right\|$, that spans the same space as $\mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}})$. In particular, the Lanczos iteration, allows for the following approximation to $e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}}$, denoted by $\boldsymbol{z}_k \in \mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}})$, given by

$$\boldsymbol{z}_k = \left\|\bar{\boldsymbol{\psi}}\right\| \mathbf{Q}_k e^{\mathbf{Q}_k^T \tilde{\mathbf{A}}_n \mathbf{Q}_k} e_1 = \left\|\bar{\boldsymbol{\psi}}\right\| \mathbf{Q}_k e^{\mathbf{T}_k} e_1\,, \tag{4.111}$$

where $\mathbf{T}_k = \mathbf{Q}_k^T \tilde{\mathbf{A}}_n \mathbf{Q}_k$ is a tridiagonal matrix, and e_1 is the i-th unit vector in $\mathbb{C}^k$. As is evident from the above form, the problem is now reduced to the evaluation of $\exp(\mathbf{T}_k)$, wherein $\mathbf{T}_k$ is a small matrix of size $k \times k$, and hence, $\exp(\mathbf{T}_k)$ can be evaluated inexpensively either through Taylor series expansion or exact eigenvalue decomposition of $\mathbf{T}_k$. The error, ϵ_k, incurred in the above approximation is given by [196]

$$\epsilon_k = \left\| e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}} - \left\|\bar{\boldsymbol{\psi}}\right\| \mathbf{Q}_k e^{\mathbf{T}_k} e_1 \right\| \approx \beta_{k+1,k} \left\|\bar{\boldsymbol{\psi}}\right\| \left| \left[e^{\mathbf{T}_k}\right]_{k,1} \right|\,, \tag{4.112}$$

where $\beta_{k+1,k}$ is the $(k+1,k)$ entry of $\mathbf{T}_{k+1} = \mathbf{Q}_{k+1}^T \tilde{\mathbf{A}}_n \mathbf{Q}_{k+1}$. Thus, the above relation provides a robust and inexpensive scheme to adaptively determine the dimension of

the Krylov subspace by checking if ϵ_k is below a set tolerance. An economic choice for the tolerance for ϵ_k is determined from atomic RT-TDDFT calculations, such that it achieves < 10 meV accuracy in the excitation energies. For a multi-atom system, we employ the lowest such tolerance obtained for each of the constituent atomic species.

Finally, we comment upon the numerical details of the second-order Magnus propagator with midpoint integration rule. As discussed in Sec. 4.3.2, the use of second-order Magnus propagator with midpoint integration rule, i.e., $e^{\tilde{\mathbf{A}}_n}$, requires the knowledge of $\bar{\mathbf{H}}$ at a future time instant i.e., $\bar{\mathbf{H}}[t_{n-1}+\Delta t/2]$, which is *a priori* unknown. A fully consistent approach involves, for a given $\bar{\boldsymbol{\psi}}^{n-1}$, the following steps: (i) approximate $\bar{\mathbf{H}}[t_{n-1}+\Delta t/2]$ through extrapolation over previous instants of $\bar{\mathbf{H}}$; (ii) use it to obtain $\bar{\boldsymbol{\psi}}^{n}$, and then evaluate $\bar{\mathbf{H}}[t_n]$; (iii) re-evaluate $\bar{\mathbf{H}}[t_{n-1}+\Delta t/2]$ by interpolating between $\bar{\mathbf{H}}[t_{n-1}]$ and $\bar{\mathbf{H}}[t_n]$; and (iv) repeat steps (ii)–(iii) until convergence. Although robust and accurate, this approach comes at a huge computational cost arising out of the Lanczos procedure at each iterate of the self-consistent iteration. An efficient and sufficiently accurate approach is to use a predictor-corrector method to, first, predict $\bar{\mathbf{H}}[t_{n-1}+\Delta t/4]$ through an extrapolation (linear or higher-order) from previous instants of $\bar{\mathbf{H}}$, use it to propagate $\bar{\boldsymbol{\psi}}^{n-1}$ to $\bar{\boldsymbol{\psi}}^{n-1/2}$, which is then used to evaluate $\bar{\mathbf{H}}[t_{n-1}+\Delta t/2]$. We refer to [197] for the details of the predictor-corrector scheme. We remark that this predictor-corrector scheme is accurate to $\mathcal{O}(\Delta t^2)$, and hence, does not affect the results of our time-discretization error estimates.

4.7 Results and discussion

In this section, we discuss the accuracy, rate of convergence, computational efficiency and the parallel scalability of higher-order finite element discretization in conjunction with second-order Magnus propagator, for both pseudopotential and all-electron RT-TDDFT calculations. Based on the system, we use hexahedral spectral finite elements of polynomial order 1 to 5, denoted as HEX8, HEX27, HEX64SPEC,

HEX125SPEC, and HEX216SPEC, respectively. For the pseudopotential calculations, we provide comparison, in terms of accuracy and performance, of the higher-order finite elements against the finite-difference method. The finite-difference based calculations are performed using the Octopus [175] software package. In all our finite-difference based calculations, we have used a stencil of order 4 in each direction (default stencil order in Octopus). All the pseudopotential calculations are done using the norm-conserving Troullier-Martins pseudopotentials [102]. For all calculations, the ground-state Kohn-Sham orbitals are used as the initial states. We use the Chebyshev filter acceleration technique (refer to Sec. 3.6 or [198, 143, 38]) to efficiently compute the ground-state, for all the calculations done using finite elements. All our scalability as well as benchmark studies demonstrating the computational efficiency are conducted on a parallel computing cluster with the following configuration: Intel Xeon Platinum 8160 (Skylake) CPU nodes with 48 processors (cores) per node, 192 GB memory per node, and Infiniband networking between all nodes for fast MPI communications.

4.7.1 Rates of Convergence

In this section, we study the rates of convergence of the dipole moment with respect to both finite element mesh-size, h, as well as time-step, Δt. We use methane and lithium hydride molecules as our benchmark systems for studying the rates of convergence, for pseudopotential and all-electron calculations, respectively.

We note that in order to study the convergence with respect to mesh-size, the dominant error must arise from spatial-discretization. To this end, we contain other sources of error, namely, time-discretization error and Krylov subspace projection error, by choosing a very small time-step of $\Delta t = 10^{-4}$, and using a small tolerance of 10^{-12} for the Krylov subspace error (cf. Eq. 4.112). In effect, we mimic a semi-discrete (discrete in space but continuous in time) error analysis. We employ finite elements

of three different orders (p)—HEX8, HEX27, and HEX125SPEC—in all our convergence studies. For each p, we start with a given N_{elem} and use the superposition of ground-state atomic orbitals and electrostatic potentials, to determine the coarsening rate of the mesh, as per Eq. 4.108. The resultant mesh forms the coarsest mesh. Subsequently, we increase the value of N_{elem} to obtain a sequence of increasingly refined meshes. We remark that although we propose for using the ground-state electronic fields to determine the mesh coarsening rate, it, nevertheless, forms a reliable and cost effective way for discretizing the mesh as opposed to any ad-hoc coarsening or using uniform discretization. Furthermore, it allows us to use large computational domain sizes without significantly increasing the number of elements. This, in turn, allows us to circumvent the need for artificial absorbing boundary conditions, otherwise essential to tackle wave reflection effects that are observed while dealing with small computational domains.

In order to perform the convergence study with respect to mesh-size, and compare to the semi-discrete error estimate obtained in Eq. 4.104, we require the knowledge of the continuous value of the dipole moment, $\mu_x(t)$. To this end, we use the discrete dipole moment $\mu_x^h(t)$, obtained from a sequence of increasingly refined HEX125SPEC finite element meshes to obtain a least-square fit of the form

$$\frac{\left|\mu_x(t) - \mu_x^h(t)\right|}{|\mu_x(t)|} = C h_{min}^q \,, \tag{4.113}$$

to determine $\mu_x(t)$, C, and q. In the above equation h_{min} represents the minimum element size, h_e, present in the mesh. The obtained $\mu_x(t)$ represents the extrapolated continuum limit (continuous in space) for the dipole moment computed using HEX125SPEC element, and is used as the reference value to compute $\frac{\left|\mu_x(t)-\mu_x^h(t)\right|}{|\mu_x(t)|}$ for HEX8 and HEX27 finite elements.

Next, we consider the convergence with respect to temporal discretization, i.e.,

Δt. To this end, we use a sufficiently refined HEX125SPEC finite element mesh and use increasingly refined Δt to obtain a least-square fit of the form

$$\frac{\left|\mu_x^h(t_n) - \mu_x^{h,n}\right|}{|\mu_x^h(t_n)|} = C(\Delta t)^q\,, \tag{4.114}$$

to determine $\mu_x^h(t_n)$, C, and q. The value of $\mu_x^h(t_n)$ obtained from the above equation represents the extrapolated continuum limit (continuous in time) for the dipole moment at t_n.

4.7.1.1 All-electron calculations: Lithium Hydride

In this example, we conduct all-electron RT-TDDFT study on a lithium hydride molecule (LiH) with Li-H bond-length of 3.014 a.u. A large cubical domain of length of 50 a.u. is chosen to ensure that the electron density decays to zero on the domain boundary, thereby, allowing us to impose Dirichlet boundary condition on the time-dependent Kohn-Sham orbitals and the Hartree potential. We use the aforementioned adaptive mesh generation strategy to construct a sequence of HEX8, HEX27 and HEX125SPEC meshes. For all the meshes under consideration, we first obtain the ground-state and employ a weak delta-kick to excite the system. To elaborate, we use an electric field of the form $\mathbf{E}_0(t) = \kappa\delta(t)\hat{x}$, with $\kappa = 10^{-3}$ a.u., where $\delta(t)$ is the Dirac-delta distribution and $\hat{x}$ is the unit vector along x-axis. This amounts to perturbing the ground-state Kohn-Sham orbitals, ψ_α^{GS}, by a factor $e^{-i\kappa x}$. Thus, our initial-states are defined as $\psi_\alpha(0) = e^{-i\kappa x}\psi_\alpha^{GS}$. Figure 4.2 depicts the rates of convergence for the dipole moment at $t = 1.0$ a.u. for different orders of finite elements. For all the three types of finite elements under consideration, we observe close to optimal rates of convergence, $\mathcal{O}(h^p)$, where p is the degree of the finite element interpolating polynomial. As is evident from Figure 4.2, much higher accuracies are obtained with HEX125SPEC when compared to HEX8 and HEX27 of the same mesh size.

In particular, 200,000 HEX8 elements (210,644 degrees of freedom) are required to achieve relative errors of 10^{-2}, whereas we achieve relative error of 10^{-3} with just 3,000 HEX125SPEC elements (83,156 degrees of freedom). Next, we study the rate of convergence of the dipole moment with respect to temporal discretization. To this end, we used a sufficiently refined HEX125SPEC mesh which affords 10^{-4} relative error with respect to spatial discretization. We then propagate the initial-states using second-order Magnus propagator with different Δt. Figure 4.3 depicts the rate of convergence of the dipole moment with respect to Δt at $t_n = 1.0$ a.u. We obtain a rate of convergence of $q = 1.96$ (defined in Eq. 4.114), which agrees remarkably well with the quadratic rate of convergence for second-order Magnus propagator (cf. Eq. 4.109).

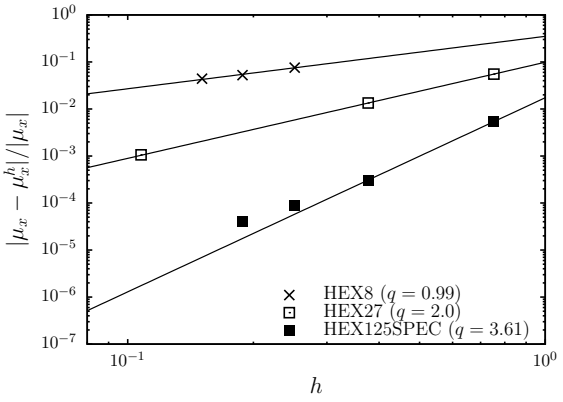


Figure 4.2: Rates of convergence with respect to spatial discretization for LiH (all-electron)

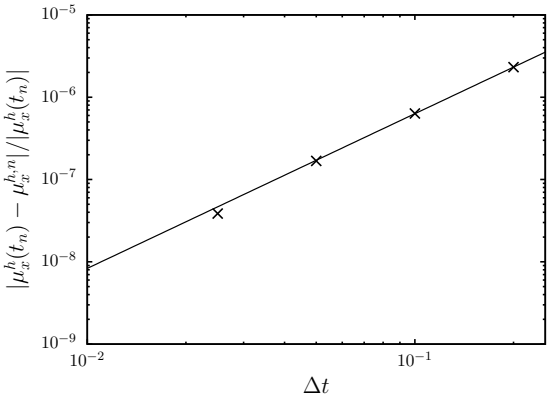


Figure 4.3: Rate of convergence with respect to temporal discretization for LiH (all-electron)

4.7.1.2 Pseudopotential calculations: Methane (CH_4)

We now turn to examining rates of convergence for the pseudopotential case. We use a methane molecule with C-H bond-length of 2.07846 a.u. and a H-C-H tetrahedral angle of 109.4712° as our benchmark system. Similar to lithium hydride, we use the ground-state single-atom electronic fields to obtain a sequence of adaptively refined

HEX8, HEX27, and HEX125SPEC meshes. We, once again, make use of a large cubical domain of length 50 a.u. to mimic simulations in $\mathbb{R}^3$. For all the meshes, we first, obtain the ground-state and then excite the system using a Gaussian electric field of the form $\mathbf{E}_0(t) = \kappa e^{(t-t_0)^2/w^2}\hat{x}$, with $\kappa = 2 \times 10^{-5}$ a.u., $t_0 = 3.0$ a.u., and $w = 0.2$ a.u. Figure 4.4 illustrates the rates of convergence of the dipole moment at $t = 5.0$ a.u. for different orders of finite elements. As in the case of lithium hydride, we obtain close to optimal rates of convergence, and observe significantly higher accuracies for HEX125SPEC over HEX8 and HEX27. Next, we study the rate of convergence afforded by the second-order Magnus propagator with respect to the time-step using a sufficiently refined HEX125SPEC mesh. We propagate the ground-state Kohn-Sham orbitals under the influence of the same Gaussian electric field using different Δt. Figure 4.5 shows the rate of convergence of the dipole moment with respect to Δt at $t_n = 5.0$ a.u. As was the case with lithium hydride, we obtain a convergence rate of $q = 1.98$, which is remarkably close to the optimal (i.e., quadratic) rate of convergence (cf. Eq. 4.109).

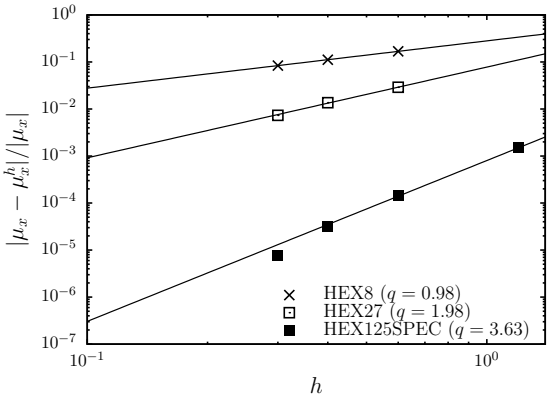


Figure 4.4: Rates of convergence with respect to spatial discretization for CH_4 (pseudopotential)

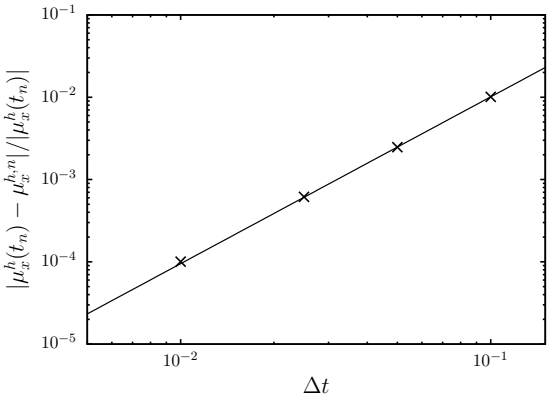


Figure 4.5: Rate of convergence with respect to temporal discretization for CH_4 (pseudopotential)

4.7.2 Computational Cost

In this section, we investigate the relative computational efficiency afforded by higher-order finite elements over linear finite element. We consider the previous two systems, lithium hydride and methane, for all-electron and pseudopotential calculations, respectively. We use the same mesh adaption strategy as detailed in Sec. 4.7.1. Since the objective of this study is to compare the relative performance of various orders of finite elements, we eliminate any time-discretization effect by setting $\Delta t = 10^{-4}$. Furthermore, we use a tolerance of 10^{-12} for the adaptive Lanczos (cf. Eq. 4.112) in order to eliminate any Krylov subspace projection error influencing the spatial discretization error. We repeat the previous numerical studies by exciting the lithium hydride molecule with a delta-kick (see Sec. 4.7.1.1), and the methane molecule with a Gaussian electric field (see Sec. 4.7.1.2). Figures 4.6 and 4.7 show the relative error in the dipole moment against the normalized computational time, for three different orders of finite elements. The normalization of the computational time is done with respect to the longest time among the various meshes under consideration. As is evident, the relative computational efficiency afforded by higher-order finite elements improves as the desired accuracy is increased. In particular, for a relative accuracy of 10^{-3}, HEX125SPEC outperforms HEX8 and HEX27 by factor $150 - 200$ and $10 - 18$, respectively. This underscores the efficacy of higher-order finite elements for RT-TDDFT calculations, an aspect which had, heretofore, remained unexplored for RT-TDDFT.

4.7.3 Other materials systems

In this section, we investigate the accuracy and computational efficiency afforded by higher-order finite elements for other materials systems, in both pseudopotential and all-electron RT-TDDFT calculations. We use Al_2, Al_{13}, and Mg_2 as the benchmark metallic systems for pseudopotential calculations. Furthermore, we use

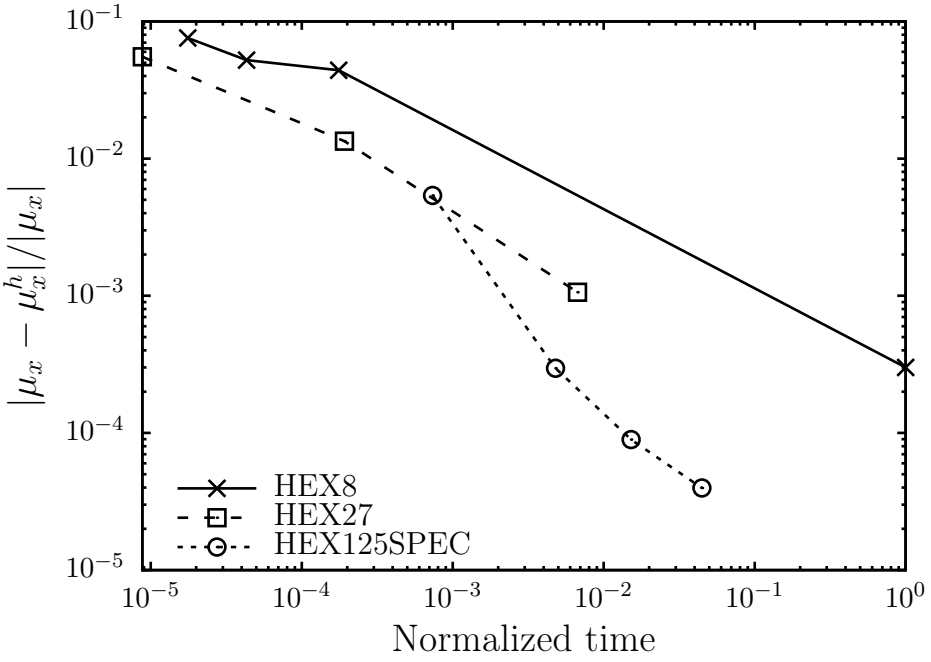


Figure 4.6: Computational efficiency of various orders finite elements for LiH (all-electron)

Buckminsterfullerene (C_{60}) as our benchmark insulating system for pseudopotential calculations. For the all-electron case, we use methane and benzene as our benchmark systems. Additionally, for the all-electron calculations we provide comparison, in the absorption spectrum, with their pseudopotential counterparts. For all the above systems under consideration, except Mg_2, we use weak electric fields to excite them. For Mg_2, we use a strong laser pulse to study the efficacy of higher-order finite elements for nonlinear response. Table 4.7.3 lists the important simulation parameters, for all the benchmark systems under consideration. We remark that the Δt as well as the tolerance for Krylov subspace projection error, listed in Table 4.7.3, are chosen such that they achieve < 10 meV accuracy in the excitation energies. For pseudopotential systems, we also provide comparison, wherever possible, against calculations based on a finite-difference discretization, by employing the same propagator (i.e., second-order Magnus) and simulation details (as listed in Table 4.7.3). To this end, we use the Octopus [175] software package to perform the finite-difference based calculations.

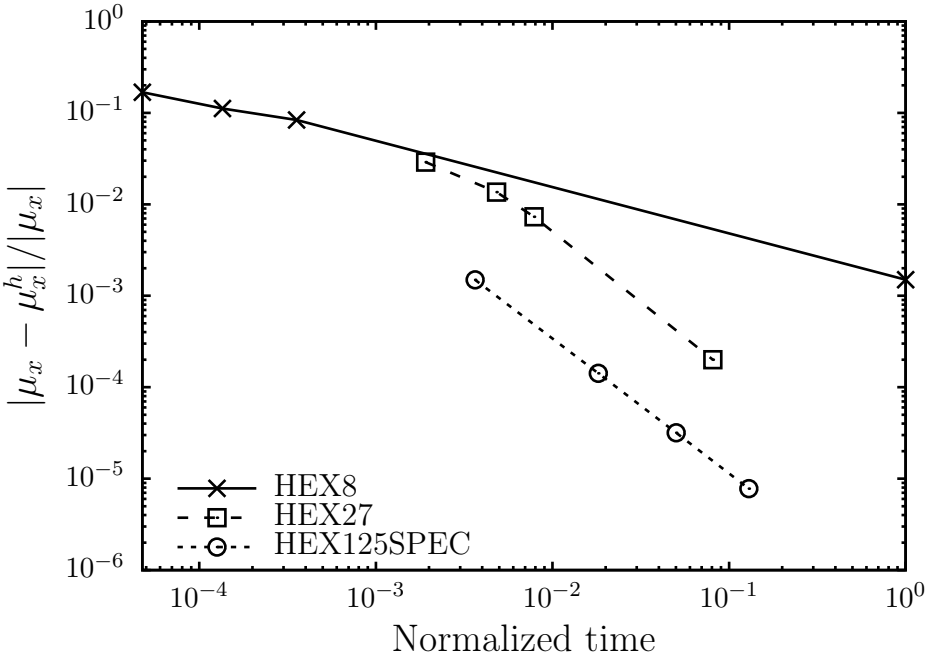


Figure 4.7: Computational efficiency of various orders finite elements for CH_4 (pseudopotential)

We now briefly discuss about the choice of spatial discretization and domain sizes in our calculations. For both finite elements and finite-difference based calculations, the spatial discretization and the domain sizes are chosen such that it attains < 10 meV accuracy in the excitation energies. Typically, one needs a larger domain for RT-TDDFT calculations than ground-state calculations, so as to avoid reflection at the domain boundaries. For finite elements, owing to adaptive meshing capability, choosing a large enough domain has little bearing on its computational expense. However, for finite-difference, wherein Octopus uses a uniform mesh, the use of large domain sizes can significantly effect its computational cost. In order to obtain a suitable grid in Octopus, we first obtain the optimal grid-spacing and domain size that achieves an accuracy of 10 meV in the ground-state energy per atom, commensurate with the accuracy targeted in the finite element discretization. We then increase the domain size until it achieves < 10 meV accuracy in the excitation energies (defined

in Sec 4.7.3.1). The calculation based on the resulting Octopus mesh is considered as the point of comparison (for both accuracy and efficiency) against the corresponding finite elements based calculation. We add that, while dealing with uniform mesh, a typical workaround to the large domain requirement is to use a smaller domain with absorbing boundary conditions. Hence, to better assess the effects of absorbing boundary conditions, we employ them in finite-difference based calculations for some of the benchmark systems discussed below.

Table 4.1: Simulation details for both pseudopotential (PSP) and all-electron (AE) benchmark systems: Type of the electric field $\mathbf{E}_0(t)$; time-step (Δt in a.u.); tolerance for Krylov subspace projection error (ϵ , cf. Eq. 4.112); total duration of simulation (T in fs)

System	Field type	Δt	ϵ	T
Al_2 (PSP)	Weak-Gaussian[1]	0.05	10^{-8}	10
Al_{13} (PSP)	Weak-Gaussian[1]	0.05	10^{-8}	10
C_{60} (PSP)	Weak-Gaussian[1]	0.05	10^{-8}	10
Mg_2(PSP)	Strong-Sinusoidal[2]	0.025	10^{-8}	25.33
CH_4 (PSP)	Weak-Gaussian[1]	0.05	10^{-8}	10
CH_4 (AE)	Weak-Gaussian[1]	0.025	10^{-8}	10
C_6H_6 (PSP)	Weak-Gaussian[1]	0.05	10^{-8}	10
C_6H_6 (AE)	Weak-Gaussian[1]	0.025	10^{-8}	10

[1] $\mathbf{E}_0(t) = \kappa e^{(t-t_0)^2/\omega^2}\hat{x}$, with $\kappa = 2 \times 10^{-5}$, $t_0 = 3.0$, and $\omega = 0.2$ (all in a.u.).

[2] $\mathbf{E}_0(t) = \kappa \sin^2(\pi/T)\sin(\omega t)\hat{x}$, with $\kappa = 0.01$, $\omega = 0.03$, $T = 5 \times (2\pi/\omega)$ (all in a.u.).

4.7.3.1 Pseudopotential calculations: Al_2

We consider an aluminum dimer (Al_2) of bond-length 4.74 a.u. In order to generate a suitable mesh, we use an adaptive HEX64SPEC finite elements discretization that follows the coarsening rate obtained from Eq. 4.108 and is commensurate with an accuracy of 10 meV in the ground-state energy per atom. We use a cubical domain of

length 60 a.u. to ensure that the wavefunctions decay to zero, and thereby, avoid any reflection effects. We excite the ground-state using the simulation parameters listed in Table 4.7.3. We use the Fourier transform of the dipole moment to obtain the dynamic polarizability, $\alpha_{a,b}(\omega)$, where a is the index of the electric field's polarization direction and b is the index of the measurement direction of the dipole. Subsequently, we obtain the absorption spectrum (dipole strength function), $S(\omega)$, given by $S(\omega) = \frac{2\omega}{3\pi}\mathrm{Tr}\left[\mathrm{Im}[\boldsymbol{\alpha}(\omega)]\right]$. The peaks in the absorption spectrum correspond to the excitation energies. We also assess the performance of higher-order finite elements by comparing against the finite-difference scheme of Octopus [175]. In order to highlight the effects of domain size for the finite-difference mesh, we use three cubical domains of sizes 38 a.u, 46 a.u, and 52 a.u., all with a grid-spacing of 0.2 a.u. Furthermore, to understand the effect of absorbing boundary conditions, we perform an additional finite-difference calculation on the 38 a.u. mesh with a negative imaginary potential (NIP) near the boundaries. In particular, we use a potential of the following form

$$V_{NIP}(\mathbf{x}) = \begin{cases} 0 & |x| \leq L \\ -i\,\eta \sin^2\left(\frac{2\pi(x-L)}{L}\right) & L < |x| \leq L + \Delta L \end{cases}$$

with $\eta = 0.4$, $L = 18.0$ and $\Delta L = 1.0$ (all in a.u.). For clarity, we refer to the four finite-difference calculations, namely, with domain size 52 a.u., with domain size 46 a.u., with domain size 38 a.u., and with domain size 38 a.u. along with NIP absorbing boundary condition as FD-52, FD-46, FD-38, and FD-38-ABS, respectively. We use the same simulation details, namely, time-step, duration of propagation, choice of propagator, and tolerance for Krylov subspace, as used for the finite element case. Fig. 4.8 compares absorption spectrum obtained from finite elements against finite-difference. We have used a Gaussian window of the form $g(t) = e^{-\alpha t^2}$, with $\alpha = 0.005$ a.u., in the Fourier transform of the dipole moment to artificially broaden the peaks.

As is evident from the figure, we get good agreement with the finite-difference based results for FD-46 and FD-52. The finite-difference calculation with domain size 38 a.u., with and without the absorbing boundary condition, provides qualitatively different results with two peaks around 5 eV. We attribute this discrepancy to possible reflection effects from the boundary, as a domain size of 38 a.u. may not be sufficient to avoid finite-domain size effects. Furthermore, comparing FD-38 and FD-38-ABS curves, we observe that the use of NIP based absorbing boundary condition, on its own, hardly improves the answer. This suggests that, for the system under consideration, one cannot rely, solely, on absorbing boundary conditions to avoid reflection effects, and hence, must use a larger domain. Table 4.2 compares the first two excitation peaks, the degrees of freedom and the total computational time for the finite element and the finite-difference (46 a.u. domain size) based calculations. As is evident from the table, both finite element and finite-difference based results agree to within 10 meV in the excitation energies. Furthermore, in terms of computational efficiency, we observe a ∼65-fold speedup for finite elements over finite-difference. We remark that this superior efficiency for the finite elements is largely attributed to fewer degrees of freedom that one can afford in finite elements due to adaptive resolution of the mesh, as opposed to a uniform mesh in finite-difference. We underline this by noting that while finite-difference requires over 12 million degrees of freedom, the finite elements require only 31,411 degrees of freedom to attain similar accuracies. Finally, comparing FD-46 and FD-52 vis-á-vis the FE results, we observe that although we achieve convergence in the excitation energies by 46 a.u., the convergence of the peak values requires a larger domain.

4.7.3.2 Pseudopotential calculations: Al_{13}

We now consider a 13 atom aluminum cluster with an icosahedral symmetry (refer Table A.2 in Appendices for coordinates). We use the same characteristic finite

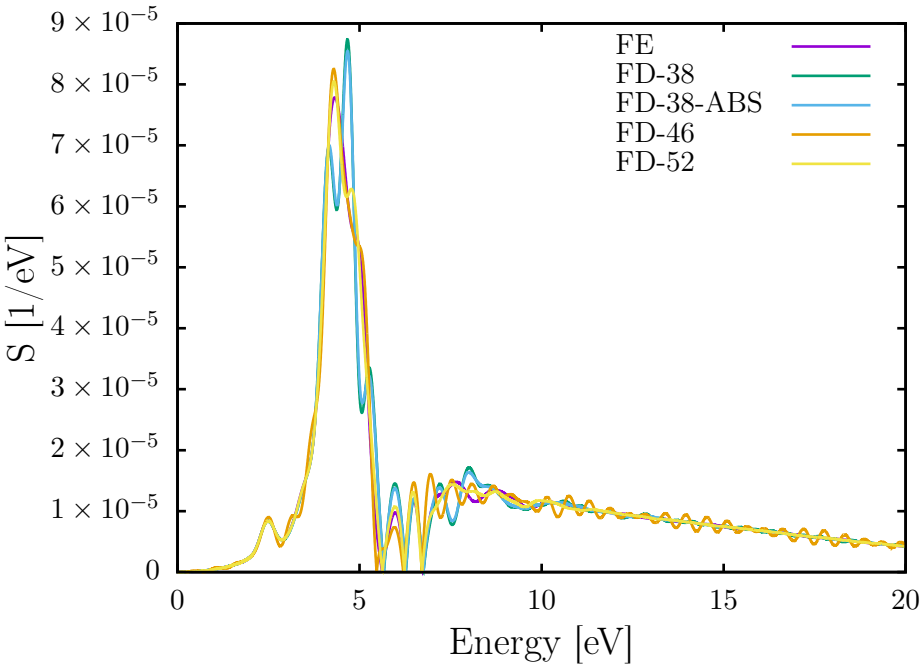


Figure 4.8: Absorption spectra for Al_2

Table 4.2: Comparison of finite element (FE) and finite-difference (FD) for Al_2: First and second excitation energies (E_1, E_2, respectively, in eV), degrees of freedom (DoF), and total computation CPU time (in CPU hours).

Method	E_1	E_2	DoF	CPU Hrs
FE	2.477	4.325	$31,411$	2.11
FD	2.486	4.332	$12,326,391$	138.8

element mesh as that of Al_2 but with a cubical domain of length 70 a.u., to avoid reflection effects. We excite the system from its ground-state using the parameters listed in Table 4.7.3. We, once again, provide a comparative study against finite-difference based calculation by using a uniform cubical mesh of size 56 a.u. and grid-spacing 0.2 a.u. Fig. 4.9 compares absorption spectrum obtained from finite elements against finite-difference. We have used the same Gaussian window as in the case of Al_2. As is evident from the figure, the peaks for both finite element and finite-difference are in good agreement. Table 4.3 compares the first two excitation peaks, degrees of freedom, and the total computational time for the finite element and the

finite-difference based calculations. Both the methods agree to within 10 meV in the first two excitation energies. In terms of computational efficiency, the finite elements attain an ∼8-fold savings in the computational time against finite-difference, once again, attributed to the fewer degrees of freedom in finite elements owing to adaptive resolution of the mesh. In particular, the finite elements afford ∼30-fold fewer degrees of freedom as compared to finite-difference.

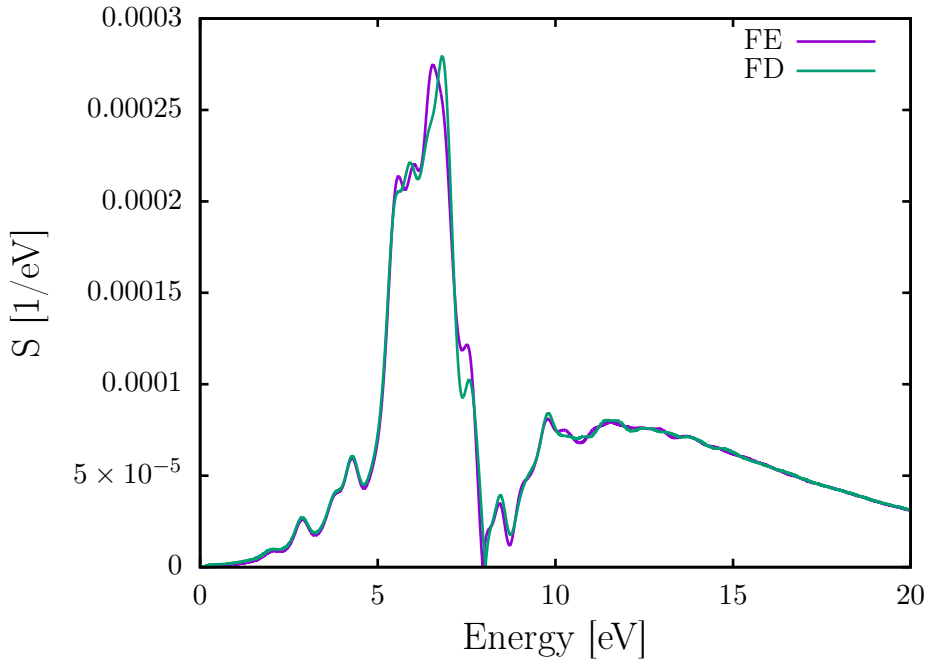


Figure 4.9: Absorption spectra for Al_{13}.

Table 4.3: Comparison of finite element (FE) and finite-difference (FD) for Al_{13}: First and second excitation energies (E_1, E_2, respectively, in eV), degrees of freedom (DoF), and total computation CPU time (in CPU hours).

Method	E_1	E_2	DoF	CPU Hrs
FE	2.876	4.280	698, 782	82.2
FD	2.880	4.282	22, 188, 041	624.6

4.7.3.3 Pseudopotential calculations: Buckminsterfullerene

In this example, we consider the Buckminsterfullerene molecule comprising of 60 carbon atoms (240 electrons) packed into the shape of a buckyball (refer Table A.2 in Appendices for coordinates). As with Al_2, we use an adaptive HEX64SPEC finite elements discretization that follows the coarsening rate obtained from Eq. 4.108 and is commensurate with an accuracy of 10 meV in the ground-state energy per atom. We use a cubical domain of length 50 a.u. to eliminate any reflection effects from the boundaries. We use the simulation parameters listed in Table 4.7.3 to excite the system from its ground-state. As with previous cases, we also assess the performance of higher-order finite elements by comparing against finite-difference based method, as implemented in the Octopus package. We assess the effects of domain size for the finite-difference mesh, by using two cubical domains of sizes 30 and 36 a.u., both with a grid-spacing of 0.15 a.u. Furthermore, we study the effect of absorbing boundary conditions by performing an additional finite-difference calculation on the 30 a.u. mesh with a negative imaginary potential (NIP) near the boundaries. We use an NIP of the same form as used in Al_2, albeit with $L = 14.0$ a.u. We denote these three finite-difference calculations, namely, with domain size 36 a.u., with domain size 30 a.u., and with domain size 30 a.u. along with NIP absorbing boundary condition as FD-36, FD-30, and FD-30-ABS, respectively. Figure 4.10 shows the absorption spectrum obtained from finite element and the three different finite-difference based calculations. We have used the same Gaussian window of the form $g(t) = e^{-\alpha t^2}$, with $\alpha = 0.01$ a.u., to artificially broaden the peaks. As is evident from the figure, there is good agreement between the finite element and FD-36 for all the excitation peaks. On the other hand, while FD-30 and FD-30-ABS have good agreement with finite elements for the first two peaks, they differ for the rest, possibly because of reflection effects. Furthermore, comparing FD-30 and FD-30-ABS, we remark that the use of NIP based absorbing boundary condition did not improve the absorption spectrum.

This, once again, indicates that one cannot always dispense with the need for a larger domain by, solely, using absorbing boundary conditions. Table 4.4 compares the first two excitation peaks, degrees of freedom, and the computational time for finite elements against that of FD-36. Both finite element and FD-36 based results match within 30 meV in the first two peaks. Furthermore, the excitation energies are also in good agreement with results presented in [199] (the first two excitation peaks, as we estimate from the absorption spectrum reported in [199], are ∼5.6 eV and ∼11.5 eV, respectively.). In terms of computational efficiency, finite elements attain a ∼3-fold speedup over FD-36. This higher efficiency of the finite elements, is once again, attributed to a ∼9-fold fewer degrees of freedom required by the finite elements against that of finite-difference.

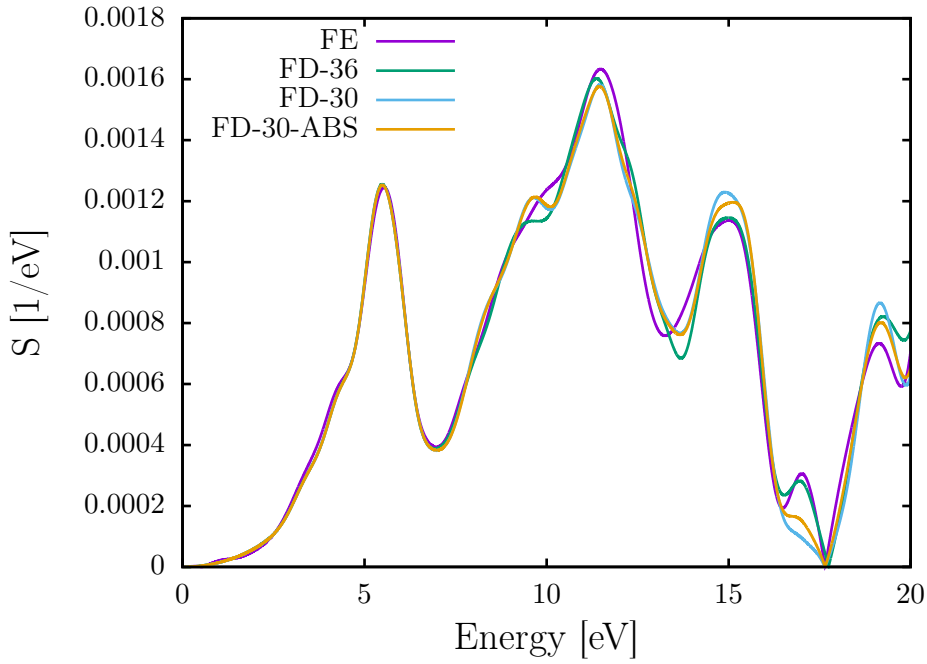


Figure 4.10: Absorption spectra of Buckminsterfullerene

Table 4.4: Comparison of finite element (FE) and finite-difference (FD) for C_{60}: First and second excitation energies (E_1, E_2, respectively, in eV), degrees of freedom (DoF), and total computation CPU time (in CPU hours).

Method	E_1	E_2	DoF	CPU Hrs
FE	5.499	11.412	$1,548,073$	$5,200$
FD	5.476	11.439	$13,997,521$	$15,361$

4.7.3.4 Pseudopotential calculations: Mg_2

In this example, we study the higher harmonic generation in a magnesium dimer with bond-length of 4.74 a.u. Unlike the previous examples, we use a strong laser pulse to excite the system from its ground-state (see Table 4.7.3 for the simulation details). We use an adaptive HEX125SPEC mesh with the coarsening rate determined by Eq. 4.108. Furthermore, we use a cubical domain of length 100 a.u. to eliminate any reflection effects from the boundaries. We obtain the dipole power spectrum, $P(\omega)$, of the system by taking the imaginary part of the Fourier transform of the acceleration of the dipole moment, $\mu(t)$. To elaborate, $P(w) = \text{Im}\left(\int_0^T e^{-i\omega t}\frac{d^2}{dt^2}\mu(t)\,dt\right)$. Theoretically, for a system with spatial inversion symmetry, only odd multiples of the frequency of the exciting laser pulse must be emitted. We verify this in Figure 4.11 wherein the peaks in the power spectrum coincide with odd harmonics. Furthermore, we observe that the decay of the intensity of the peaks flattens beyond the 13-th harmonic, which corroborates well with the plateau phenomenon, typically observed in experiments [200]. We emphasize that despite the large domain size used in this calculation, we require only $\sim 60,000$ basis functions. This underlines the efficacy of higher-order finite elements for even nonlinear regime in RT-TDDFT.

4.7.3.5 All-electron calculations: Methane (CH_4)

We now examine the competence of higher-order finite elements for all-electron RT-TDDFT calculations by providing a comparative study with its pseudopotential counterpart. In this example, we consider a methane molecule with the same geometry

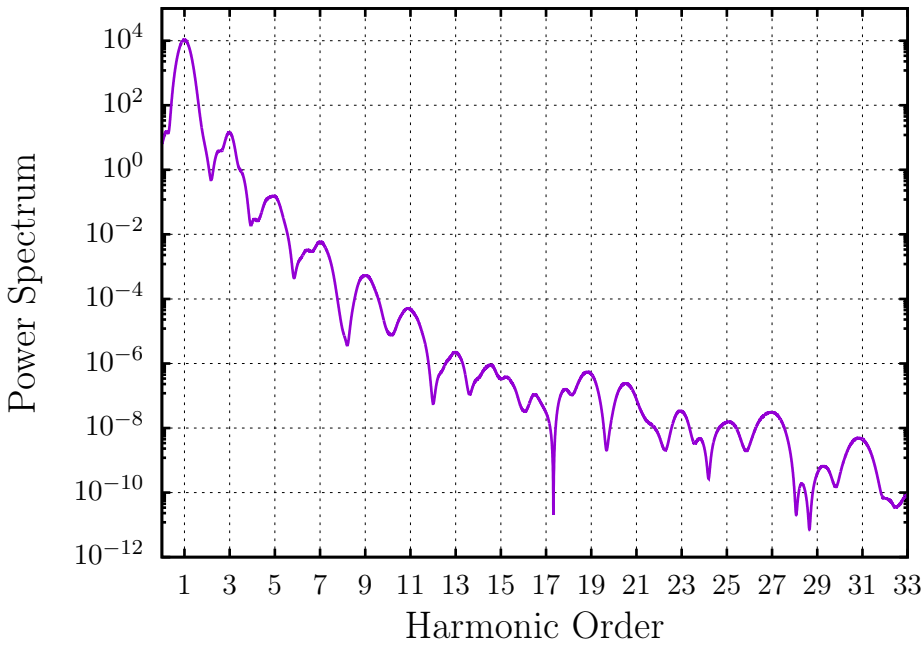


Figure 4.11: Dipole power spectrum of Mg_2

as described in Sec. 4.7.1.2. We use HEX64SPEC and HEX125SPEC elements for the pseudopotential and all-electron case, respectively. For both all-electron and pseudopotential cases, we use the same mesh adaption strategy as used in all previous examples. For both the meshes, we use a large cubical domain of length 40 a.u., so as to eliminate reflection from the boundaries. Both the systems are excited from their respective ground-states using the simulation details listed in Table 4.7.3. The absorption spectra for both the calculations are shown in Figure 4.12. We used the same Gaussian window as in the case of Buckminsterfullerene (see Sec. 4.7.3.3) to artificially broaden the peaks. As evident from the figure, we obtain remarkable agreement between the all-electron and pseudopotential results, i.e., the two curves are almost identical. Table 4.5 we list the first two excitation peaks, degrees of freedom, and total computational time for both the calculations. The first two excitation peaks agree to within 10 meV. We remark that the all-electron calculation requires $\sim$100x more computational time as compared to the pseudopotential case. This

large computational expense for the all-electron calculation stems primarily from the need of a highly refined mesh near the nuclei, so as to accurately capture the sharp variations in the electronic fields near the nuclei. This refinement has two major consequences: (i) an increase in the degrees of freedom; and (ii) increase in $\left\|\psi_\alpha^h\right\|_{H^1(\Omega)}$, which in turn, warrants a smaller time-step (cf. Eq. 4.109) as well as a larger Krylov subspace to achieve the prescribed accuracy. In particular, for the case of all-electron methane, we required $\sim$4x degrees of freedom and $\sim$10x the size of the Krylov subspace as compared to that of the pseudopotential case. We emphasize that while finite elements are expensive for the all-electron calculation, they provide the desired accuracy and offer systematic convergence (see Sec. 4.7.1.1). Moreover, one can mitigate the need of a refined mesh for the all-electron calculation by using an enriched finite element basis [37] (i.e., augmenting the standard (classical) finite element basis with numerical atom-centered basis). This enrichment idea, for ground-state DFT, has been detailed in Chapter III, wherein we have successfully attained $100-300$x speedup over the standard (classical) finite elements. Similar ideas can be extended to RT-TDDFT to further the capabilities of finite elements for all-electron RT-TDDFT calculations.

Table 4.5: Comparison of all-electron (AE) and pseudopotential (PSP) calculations for methane: First and second excitation energies (E_1, E_2, respectively, in eV), degrees of freedom (DoF), and total computation CPU time (in CPU hours).

Method	E_1	E_2	DoF	CPU Hrs
AE	8.898	11.238	$348,289$	$13,653$
PSP	8.907	11.244	$80,185$	145

4.7.3.6 All-electron calculations: Benzene

In this example, we perform similar comparative studies between all-electron and pseudopotential calculations for benzene molecule (refer Table A.2 in Appendices for coordinates). As with the methane molecule, we use HEX64SPEC and HEX125SPEC

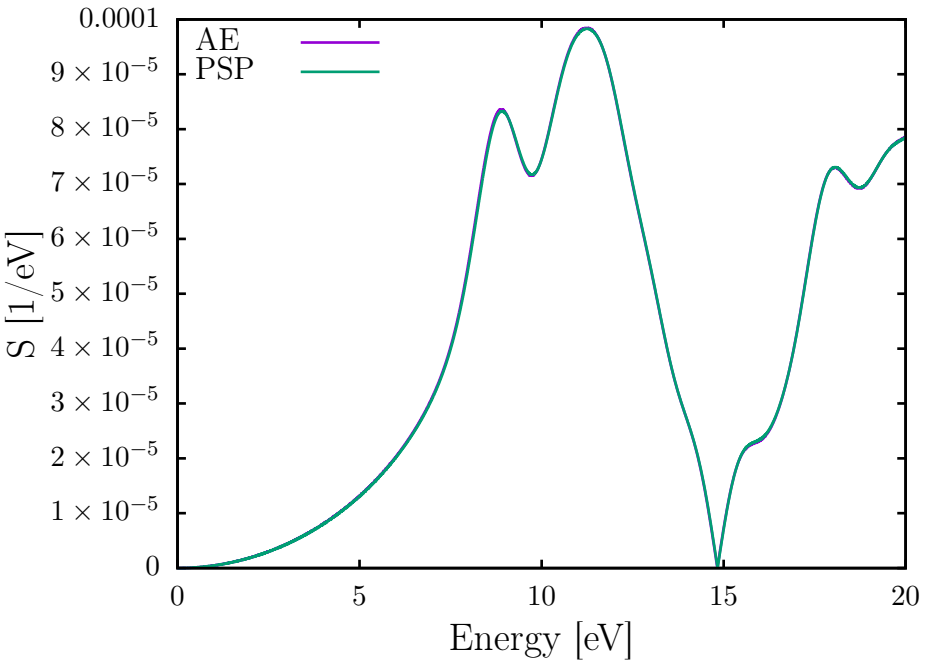


Figure 4.12: Absorption spectra of methane

finite elements for the pseudopotential and all-electron calculation, respectively. Furthermore, we use the same characteristic mesh features (i.e., refinement near the nuclei, coarsening rate, simulation domain), in both the meshes, as their counterparts in the methane calculation. The simulation details, for both the cases, are listed in Table 4.7.3. Figure 4.13 compares the absorption spectra from the all-electron and pseudopotential calculations. Both the spectra compares well with the results presented in [184], in terms of first two excitation peaks (the first two excitation peaks, as we estimate from the absorption spectrum reported in [184], are ∼6.6 eV and ∼10 eV, respectively). We remark that while there is qualitative agreement between the pseudopotential and all-electron calculations, quantitatively the predictions from all-electron and pseudopotential calculations differ. In particular, the first two excitation peaks (see Table 4.6) differ up to ∼0.2 eV. This suggests that one ought to carefully test for the transferability of the pseudopotential approximation used, to provide reliable quantitative predictions from RT-TDDFT calculations. We take note that a

more careful comparison of pseudopotential and all-electron calculations warrants a scan through a range of pseudopotential approximation. Nevertheless, the objective of this exercise is to highlight the fact that finite elements, by treating both pseudopotential and all-electron calculations on an equal footing, allows for a robust tool for such transferability studies.

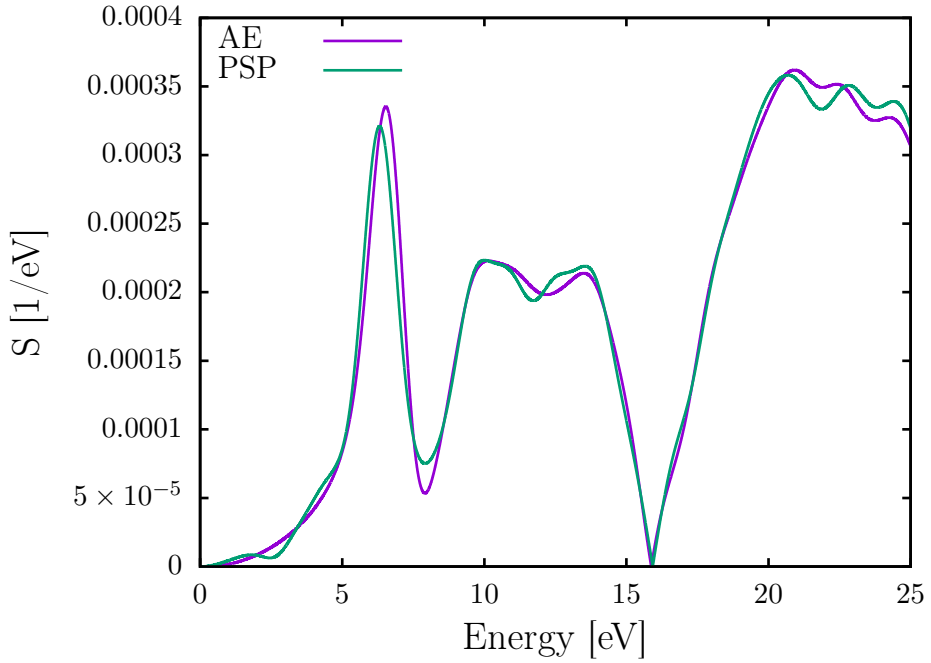


Figure 4.13: Absorption spectra of benzene

Table 4.6: Comparison of all-electron (AE) and pseudopotential (PSP) calculations for benzene: First and second excitation energies (E_1, E_2, respectively, in eV), degrees of freedom (DoF), and total computation CPU time (in CPU hours).

Method	E_1	E_2	DoF	CPU Hrs
AE	6.521	10.131	989,649	153,600
PSP	6.316	10.007	257,473	1,574

4.7.4 Scalability

Lastly, we demonstrate the parallel scalability (strong scaling) of the proposed finite element basis in Figure 4.14. We choose the Buckminsterfullerene molecule

containing ∼3.5 million degrees of freedom (number of basis functions) as our fixed benchmark system and report the relative speedup with respect to the wall time on 24 processors. The use of any number of processors below 24 was unfeasible owing to the memory requirement posed by the system. As is evident from the figure, the scaling is in good agreement with the ideal linear scaling behavior up to 384 processors, at which we observe a parallel efficiency of 86.2%. However, we observe a deviation from linear scaling behavior at 768 processors with a parallel efficiency of 74.2%. This is attributed to the fact that, at 768 processors, the number of degrees of freedom possessed by each processor falls below 5000, which is low to achieve good parallel scalability.

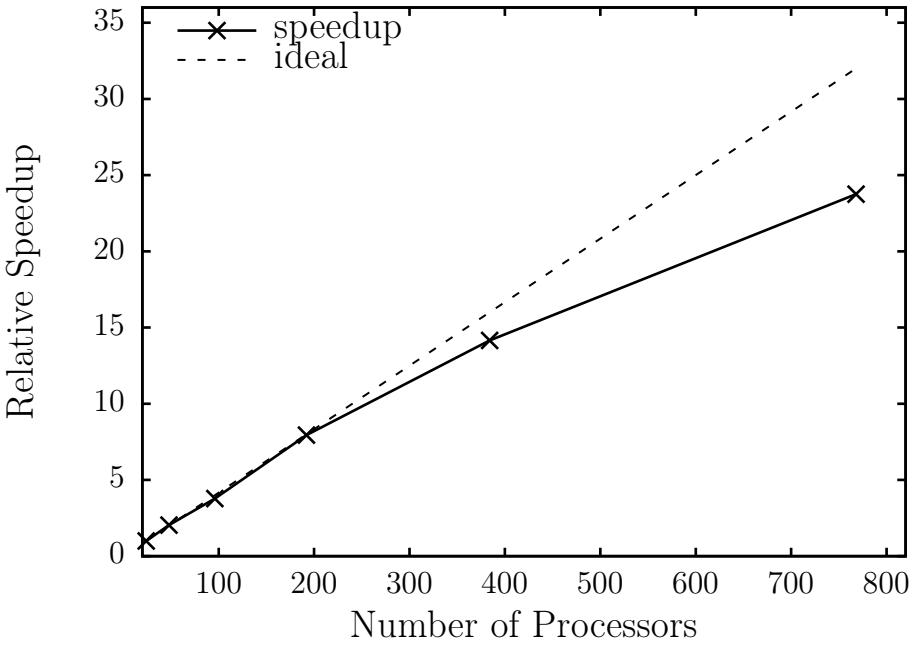


Figure 4.14: Parallel scalability of the higher-order finite element implementation.

4.8 Summary

In summary, we have investigated the accuracy, computational efficiency and scalability of higher-order finite elements for the RT-TDDFT problem, for both pseu-

dopotential and all-electron calculations. We presented an efficient *a priori* spatio-temporal scheme guided by the discretization errors in the time-dependent Kohn-Sham orbitals, in the context of second-order Magnus propagator. In particular, we used the knowledge of the ground-state electronic fields to determine an efficient adaptively resolved finite element mesh. This adaptive resolution is crucial in affording the use of large simulation domains without significant increase in the number of basis functions, and hence, allows us to circumvent the use of any artificial absorbing boundary conditions. A key aspect of the finite element discretization in this work is the use of higher-order spectral finite elements, which while providing a better conditioned basis also renders the overlap matrix diagonal when combined with special quadrature rules for numerical integration. This, in turn, enabled an efficient construction of the Magnus propagator (or any exponential time-integrator) for finite element discretization. Furthermore, we employed an adaptive Lanczos subspace projection to evaluate the action of the Magnus propagator, defined as exponential of a matrix, on the Kohn-Sham orbitals.

We demonstrated the accuracy of the proposed approach through numerical convergence studies on both pseudopotential and all-electron systems, where we obtained close to optimal rates of convergence with respect to both spatial and temporal discretization, as determined by our error estimates. The computational efficiency afforded by using higher-order finite element discretization was established, where a staggering $10 - 100$ fold speedup was obtained on benchmark systems by using a fourth-order finite element in comparison to linear and quadratic finite elements. Furthermore, we assessed the accuracy and efficiency afforded by our approach against the finite-difference based method of Octopus software package, for pseudopotential calculations. Across all the benchmark systems considered, we obtained good agreement between the absorption spectrum evaluated using the finite elements and finite-difference (Octopus) based calculations. In terms of computational efficiency,

we obtained $3-60$ fold speedup over finite-difference, which is largely attributed to the adaptive spatial resolution afforded by our approach. We also demonstrated the efficacy of finite elements, especially its efficient handling of large domains, for nonlinear response by studying the higher harmonic generation under a strong electric field. We also demonstrated the competence of higher-order finite elements for the all-electron RT-TDDFT calculations. This underscores the versatility of finite elements in handling both pseudopotential and all-electron calculations on an equal footing. Lastly, in terms of parallel scalability, we obtained good parallel efficiency up to 768 processors for a benchmark system comprising of the Buckminsterfullerene molecule containing $\sim$3.5 million basis functions.

Thus, the proposed approach offers a computationally efficient, systematically improvable, and scalable basis for RT-TDDFT calculations, applicable to both pseudopotential and all-electron cases. We remark that, for the all-electron case, the need for a highly refined mesh near the nuclei increases the computational cost, as observed from the numerical studies reported in this chapter. For systems with heavier atoms, the mesh requirements are expected to become even more exacting. However, this can be alleviated by augmenting the finite element basis with numerical atom-centered basis along with efficient numerical strategies to evaluate the inverse of the resultant overlap matrix as well handle the increased quadrature costs [124]. These ideas have been successfully used for ground-state DFT [201, 202, 135, 37] (also see Chapter III), and its extension to RT-TDDFT is currently being investigated. Further, assessing the transferability of pseudopotentials for electron dynamics, enabled by the unified treatment of all-electron and pseudoptential calculations, is another interesting direction for future investigation.

CHAPTER V

Exact Exchange-correlation Potentials from Ground-state Electron Densities

The success of DFT rests fundamentally on the quality of the approximation to the exchange-correlation (xc) functional, as it describes the many-electron quantum mechanical behavior through a computationally tractable quantity—the electron density—without resorting to multi-electron wave functions. Naturally, the quest for accurate exchange-correlation functionals has long remained the grand challenge in DFT. To this end, the *inverse* DFT problem of mapping the ground-state density to its exchange-correlation potential is a powerful tool to aid in functional development in DFT. However, the lack of an accurate and systematically convergent approach has left the problem unresolved, heretofore. This work presents a numerically robust and accurate scheme to evaluate the exact exchange-correlation potentials from correlated *ab initio* densities [72].

The rest of the chapter is organized as follows. Sec. 5.1 introduces the subject and motivates the need for an accurate numerical method for inverse DFT calculations. In Sec. 5.2, we formalize inverse DFT problem as a PDE-constrained optimization problem. We provide the numerical implementation aspects of our approach in Sec. 5.3. Sec. 5.4 provides the verification of our approach against densities obtained from DFT calculations performed using known exchange-correlation approximations (e.g.,

LDA). In Sec. 5.5 we present two important numerical strategies that are essential to circumvent the Gaussian basis-set density induced oscillations in v_{xc}. In Sec. 5.6, we present the exact v_{xc} for various molecular systems, ranging up to 40 electrons, using ground-state densities from configuration-interaction (CI) based calculations. In Sec. 5.7, we compare the accuracy of our results with that of other existing methods, using data from published literature. Finally, we present the summary of this work and lay out its future scope in Sec. 5.8.

5.1 Introduction

The inverse DFT problem [61, 58, 203, 63, 204] rests on the fact that the v_{xc} is a unique functional of the electron density ($\rho(\mathbf{r})$), so there exists a one-to-one relationship from $v_{\text{xc}}(\mathbf{r})$ to $\rho(\mathbf{r})$, and vice-versa. This observation presents a possible route forward to construct accurate exchange-correlation (xc) functionals, via the transformation of the electron density into $v_{\text{xc}}(\mathbf{r})$ (refer to the schematic in Figure 5.1). The inverse problem not only provides a route for finding the sole unknown quantity in DFT, it is also central for describing quantum mechanics without resorting to complicated multi-electron wave functions.

Given the large importance of this problem, there have been several attempts to solve the inverse DFT problem, employing either iterative updates [57, 58, 59, 203, 60] or constrained optimization approaches [61, 62, 63, 205]. However, these approaches have suffered from ill-conditioning, thereby resulting in non-unique solutions or causing spurious oscillations in the resultant $v_{\text{xc}}(\mathbf{r})$. This ill-conditioning has been largely attributed to the incompleteness of the Gaussian basis sets that were employed to solve the inverse DFT problem [66, 67, 205]. Recent efforts [70, 206, 71] have presented a different approach, which utilizes the two-electron reduced density matrix (2-RDM) to remedy the non-uniqueness and the spurious oscillations in the obtained $v_{\text{xc}}(\mathbf{r})$. However, this does not represent the solution of the inverse DFT problem,

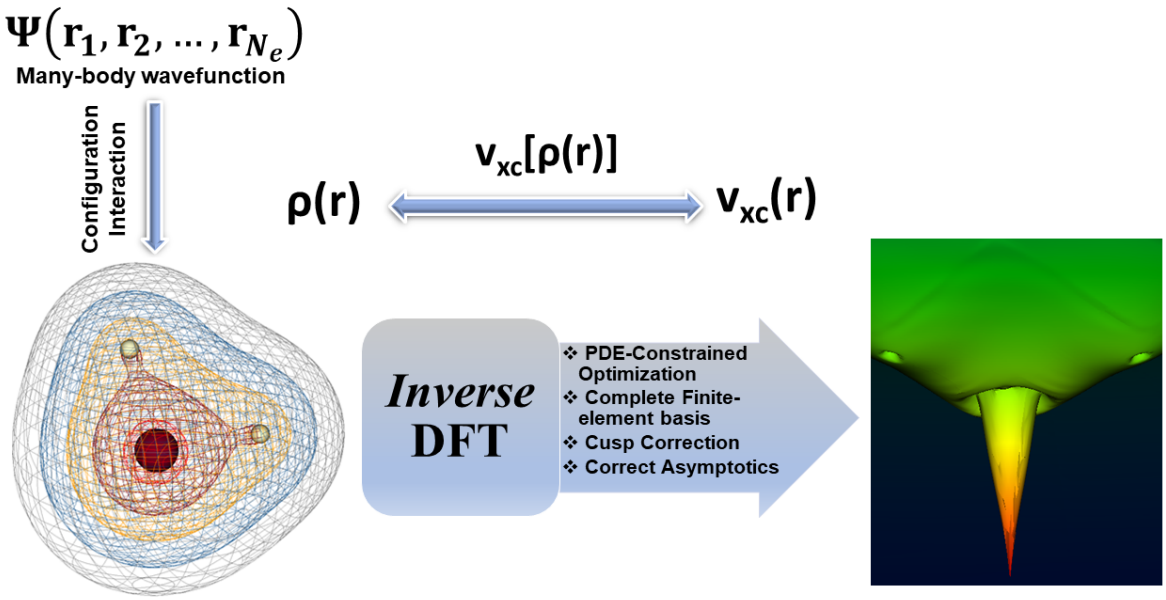


Figure 5.1: Schematic of the inverse DFT problem. The exact ground-state many-body wavefunction ($\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e})$), and hence, the ground-state electron density ($\rho(\mathbf{r})$) is obtained from configuration interaction calculation. The inverse DFT calculation evaluates the exact exchange-correlation potential ($v_{\text{xc}}(\mathbf{r})$) that yields the given $\rho(\mathbf{r})$. The ability to accurately solve the inverse DFT problem, presented in this work, presents a powerful tool to construct accurate density functionals ($v_{\text{xc}}[\rho(\mathbf{r})]$), either through conventional approaches or via machine-learning. The schematic shows the ground-state density and the exact exchange-correlation potential for H_2O obtained in this work.

i.e., the v_{xc} obtained from this approach is not guaranteed to yield the input electron density [71]. Thus, the inverse DFT problem has, heretofore, remained an open challenge.

In this work, we present an advance that provides an accurate solution to the inverse DFT problem, enabling the evaluation of exact v_{xc} from an ab initio density. Specifically, the approach uses a finite element basis that is systematically convergent and complete, thereby eliminating ill-conditioning in the discrete solution of the inverse DFT problem. Our approach is tested on a range of molecular systems, both weakly and strongly correlated, showing robustness and efficacy in treating realistic

polyatomic molecules. The proposed approach therefore unlocks the door to constructing accurate exchange-correlation functionals that provide precise energies and electronic properties of a huge range of chemical, materials, and biological systems. To elaborate, we envisage the inverse DFT problem to be instrumental in generating $\{\rho^{(i)}, v_{\text{xc}}^{(i)}\}$ pairs, using $\rho^{(i)}$'s from correlated ab-initio calculations. Subsequently, these can be used as training data to model $v_{\text{xc}}[\rho]$ through machine-learning algorithms [207, 208] that are designed to preserve the functional derivative requirement on $v_{\text{xc}}[\rho]$ [209]. Further, the exchange-correlation energy ($E_{\text{xc}}[\rho]$) can be directly evaluated through line-integration on $v_{\text{xc}}[\rho]$.

5.2 PDE-constrained optimization for inverse DFT

We cast the inverse DFT problem of finding the $v_{\text{xc}}(\mathbf{r})$ that yields a given density $\rho_{\text{data}}(\mathbf{r})$ as the following partial differential equation (PDE) constrained optimization:

$$\arg\min_{v_{\text{xc}}(\mathbf{r})} \int w(\mathbf{r}) \left(\rho_{\text{data}}(\mathbf{r}) - \rho(\mathbf{r})\right)^2 d\mathbf{r}\,, \tag{5.1}$$

subject to

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})\right)\psi_i = \epsilon_i \psi_i\,, \tag{5.2}$$

$$\int |\psi_i(\mathbf{r})|^2\, d\mathbf{r} = 1\,. \tag{5.3}$$

In the above equation, $w(\mathbf{r})$ is an appropriately chosen weight to expedite convergence; $v_{\text{ext}}(\mathbf{r})$ represents the nuclear potential; $v_{\text{H}}(\mathbf{r})$ is the Hartree potential corresponding to $\rho_{\text{data}}(\mathbf{r})$; and ψ_i and ϵ_i denote the Kohn-Sham orbitals and eigenvalues, respectively. For simplicity, we restrict ourselves to only closed-shell systems, and hence, the Kohn-Sham density $\rho(\mathbf{r}) = 2\sum_{i=1}^{N_e/2} |\psi_i(\mathbf{r})|^2$. Equivalently, the above

PDE-constrained optimization can be solved by minimizing the following Lagrangian,

$$\mathcal{L}\left(v_{\text{xc}}, \{\psi_i\}, \{p_i\}, \{\epsilon_i\}, \{\mu_i\}\right) = \int w(\mathbf{r}) \left(\rho_{\text{data}}(\mathbf{r}) - \rho(\mathbf{r})\right)^2 \, d\mathbf{r} + \sum_{i=1}^{N_e/2} \int p_i(\mathbf{r}) \left(\hat{H} - \epsilon_i\right) \psi_i \, d\mathbf{r} + \sum_{i=1}^{N_e/2} \mu_i \left(\int |\psi_i(\mathbf{r})|^2 \, d\mathbf{r} - 1 \right), \tag{5.4}$$

with respect to all its constituent variables—p_i, μ_i, ψ_i, ϵ_i and v_{xc}. In the above equation, $\hat{H} = -\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})$ is the Kohn-Sham Hamiltonian; p_i is the adjoint function which enforces the Kohn-Sham eigenvalue equation corresponding to ψ_i; and μ_i is the Lagrange multiplier corresponding to the normality condition of ψ_i. The optimality of $\mathcal{L}$ with respect to p_i, μ_i, ψ_i and ϵ_i are given by:

$$\hat{H}\psi_i = \epsilon_i \psi_i \,, \tag{5.5}$$

$$\int |\psi_i(\mathbf{r})|^2 \, d\mathbf{r} = 1 \,, \tag{5.6}$$

$$(\hat{H} - \epsilon_i) p_i(\mathbf{r}) = g_i(\mathbf{r}) \,, \tag{5.7}$$

$$\int p_i(\mathbf{r}) \psi_i(\mathbf{r}) \, d\mathbf{r} = 0 \,, \tag{5.8}$$

where $g_i(\mathbf{r}) = 8w(\mathbf{r})(\rho_{\text{data}}(\mathbf{r}) - \rho(\mathbf{r}))\psi_i - 2\mu_i\psi_i$. We remark that the operator $\hat{H} - \epsilon_i$ in Eq. 5.7 is singular with ψ_i as its null-vector. However, the orthogonality of $g_i(\mathbf{r})$ and ψ_i (consequence of Eq. 5.7) along with the orthogonality of p_i and ψ_i (Eq. 5.8) guarantee a unique solution for p_i. Having solved the above optimality conditions in Eqs. 5.5–5.8, the variation (gradient) of $\mathcal{L}$ with respect to v_{xc} is given by

$$\frac{\delta \mathcal{L}}{\delta v_{\text{xc}}} = \sum_{i=1}^{N_e/2} p_i \psi_i \,. \tag{5.9}$$

This constitutes the central equation for updating $v_{\text{xc}}(\mathbf{r})$ via any gradient based optimization technique.

Summing up, the proposed approach involves: (i) obtaining $\rho_{\text{data}}(\mathbf{r})$ from correlated ab initio calculations (i.e., configuration interaction calculations); (ii) using an initial guess for $v_{\text{xc}}(\mathbf{r})$; (iii) solving Eqs. 5.5–5.8 using the current iterate of v_{xc}; (iv) updating v_{xc} using Eq. 5.9 as the gradient; (v) repeating (iii)-(iv) until $\rho(\mathbf{r})$ converges to $\rho_{\text{data}}(\mathbf{r})$. We note that the general idea of PDE-constrained optimization has been explored recently in [204]. However, its utility had only been demonstrated on non-interacting model systems in 1D.

5.3 Numerical Implementation

5.3.1 Finite-element discretization

We employ spectral finite element (FE) basis to discretize all the spatial fields—v_{xc}, $\{\psi_i\}$, $\{p_i\}$. The FE basis is constructed from piecewise Lagrange interpolating polynomials on non-overlapping subdomains called elements. The basis, thus constructed, can be systematically improved to completeness by reducing the element size and/or increasing the polynomial order [183]. We remark that the spectral FE basis are not orthogonal, and hence, result in a generalized eigenvalue problem as opposed to the more desirable case of standard eigenvalue problem. To this end, we use special reduced order quadrature (Gauss-Legendre-Lobatto quadrature rule) to render the overlap matrix diagonal, and thereby, trivially transform the generalized eigenvalue problem into a standard one (see Sec. 3.4.2 for more details). For all the H_2 molecules used in our numerical studies (refer Sec. 5.5 and Sec. 5.6), we used adaptively-refined quadratic finite elements to discretize the $\{\psi_i\}$ and $\{p_i\}$, whereas for all other systems we used adaptively-refined fourth-order finite elements. The v_{xc}, in all the calculations, is discretized using linear finite elements. Most importantly,

the form of the FE basis is chosen carefully, so as to guarantee the cusp in ψ_i's (and hence in ρ) at the nuclei, which in turn is critical to obtaining accurate v_{xc}'s near the nuclei (as will be discussed in Sec. 5.5). To elaborate, in our calculations, we construct an adaptive finite element mesh such that the nuclei are positioned on corner nodes of elements. Since Lagrange interpolating functions have C^0 continuity at element boundaries, the resultant finite element basis admits cusps in the Kohn-Sham orbitals ($\psi_i(\mathbf{r})$) (and hence in $\rho(\mathbf{r})$) at the nuclear positions. We illustrate this through a 1D cubic finite element basis, as shown in Figure 5.2. In this figure, the black circles mark the C^0 continuity at the element boundary. This C^0 continuity of the basis at the element boundaries is vital to ensuring the cusp in $\rho(\mathbf{r})$ at the nuclei.

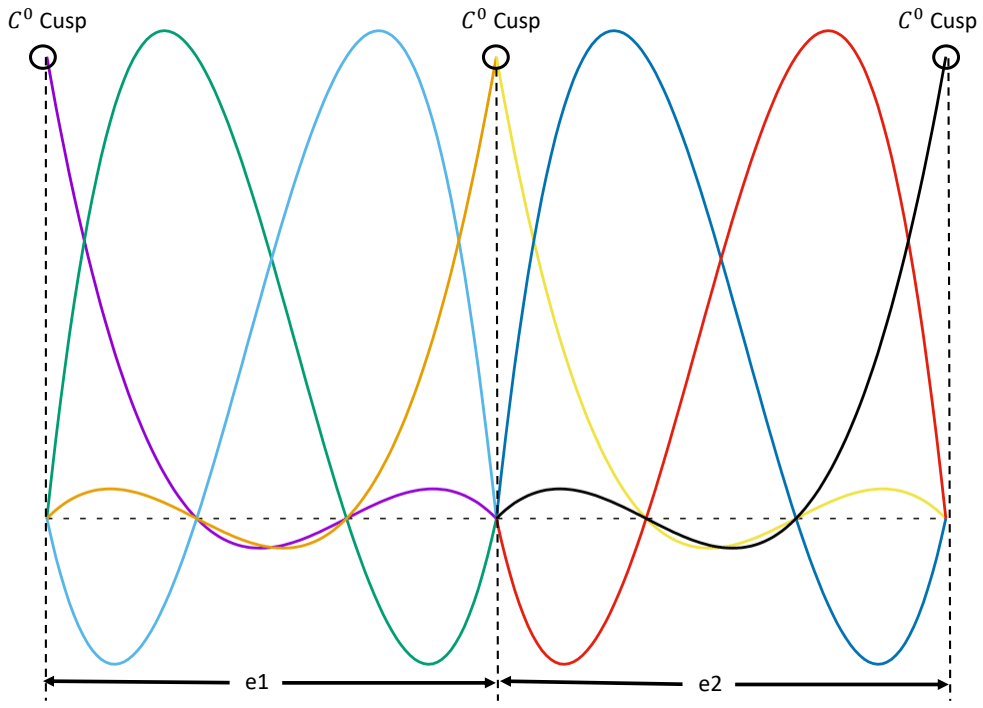


Figure 5.2: Illustration of two adjacent 1D cubic finite elements—e1 and e2. The vertical dashed lines denote the boundary between adjacent elements. The black circles highlight the C^0 continuity (cusp) of the basis at the element boundary.

5.3.2 Solvers

In order to efficiently solve the Kohn-Sham eigenvalue problem in Eq. 5.5, we employ the Chebyshev polynomial based filtering technique (refer Sec. 3.6 or [198, 38, 183]). Compared to a forward ground-state DFT calculation, the inverse DFT calculation warrants much tighter accuracy in solving the Kohn-Sham eigenvalue equation(s). However, the use of a very high polynomial degree Chebyshev filter can generate an ill-conditioned subspace, akin to any power iteration based eigen-solver. To circumvent the ill-conditioning and attain higher accuracy, we employ multiple passes of a low polynomial degree Chebyshev filter (polynomial order $\sim$1000) and orthonormalize the Chebyshev filtered vectors between two successive passes. The number of passes is determined adaptively so as to guarantee an accuracy of 10^{-9} in $||\hat{H}\psi_i - \epsilon_i\psi_i||_{L^2}$.

The discrete adjoint function (p_i) is solved by, first, projecting Eq. 5.7 onto a space orthogonal to the corresponding ψ_i (or degenerate ψ_i's), and then employing the conjugate-gradient method to compute the solution. The discrete adjoint problem is solved to an accuracy of 10^{-12} in $||(\hat{H} - \epsilon_i)p_i - g_i||_{L^2}$.

The update for v_{xc} is computed using limited-memory BFGS (L-BFGS), a memory-efficient quasi-Newton solver which constructs approximate Hessian matrices using the history of the gradients [210]. In all the calculations, we used a history of size 100 to construct the approximate Hessian.

5.3.3 Ab Initio Densities

Accurate electron densities were generated using the Incremental Full Configuration Interaction (iFCI) method [211] in the Q-Chem software package [24]. This method solves the electronic Schrodinger equation via a many-body expansion, and asymptotically produces the exact electronic energy and density as the number of bodies in the expansion approaches the all-electron limit. For this study, electron

densities were provided in the all-valence-electron limit of iFCI, that is, where the full valence set is fully correlated and the core orbitals of H_2O and C_6H_4 (as will be discussed in Sec. 5.6) are treated as an uncorrelated electron pair. Additionally, we also evaluated the ionization energies at the same level of theory, for each system with one less electron.

Lastly, to expedite the convergence of the non-linear solver, we make use of two different weights, $w(\mathbf{r}) = 1$ and $w(\mathbf{r}) = 1/\rho_{\text{data}}^{\alpha}$ $(1 \leq \alpha \leq 2)$, in sequence. The latter penalizes the objective function in the low density region.

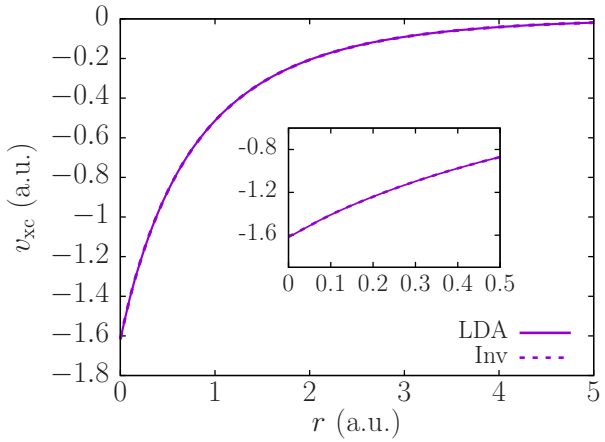
5.4 Verification with LDA-based densities

To assess the accuracy and robustness of the proposed approach, we use ρ_{data} obtained from local density approximation (LDA) [78, 79] based DFT calculations, discretized using the finite element (FE) basis—a systematically improvable and complete basis constructed from piecewise polynomials. This verification test allows us to compare the v_{xc} obtained from the inverse DFT calculation against $v_{\text{xc}}^{\text{LDA}}[\rho_{\text{data}}]$. As remarked earlier, most of the previous attempts at this verification test have suffered from either non-unique solutions or had resulted in unphysical oscillations in v_{xc}, owing to the incompleteness of the Gaussian basis employed in these works. Figure 5.3 presents the comparison of $v_{\text{xc}}^{\text{LDA}}[\rho_{\text{data}}]$ against the v_{xc} obtained from the inverse calculation, for various atomic systems. We also provide, in Figure 5.4, the v_{xc} for 1,3-dimethylbenzene (C_8H_{10}) obtained from the inverse calculation with LDA based ρ_{data}, highlighting the efficacy of our approach in accurately treating large systems. We note that all the inverse DFT calculations have been performed in 3D, and the L^2 norm error in the density, $||\rho_{\text{data}} - \rho||_{L^2}$, is driven below 10^{-5}. We remark that the L^2 error norm is a natural convergence criterion, given the form of the objective function in Eq. 5.1. However, given that previous works on this inverse problem have reported the L^1 error, we provide the same in the Table 5.1, for all the systems considered

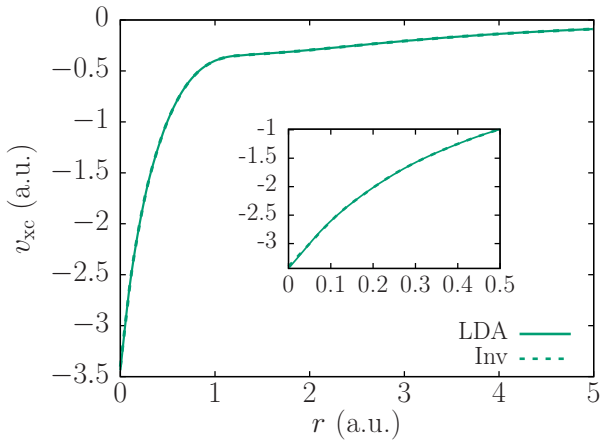
in our verification studies. As evident from these figures, the v_{xc} determined from the inverse DFT calculation is devoid of any spurious oscillations, and is in excellent agreement with $v_{\text{xc}}^{\text{LDA}}[\rho_{\text{data}}]$. Additionally, the Kohn-Sham eigenvalues computed using the inverted v_{xc} are in excellent agreement (i.e., $|\epsilon_i^{\text{LDA}} - \epsilon_i| < 1$ mHa), further validating the accuracy of the method. While we have reported the verification of our method for LDA based densities, similar accuracy was obtained using generalized gradient approximation (GGA) based densities.

Table 5.1: $\|\rho(\mathbf{r}) - \rho_{\text{data}}(\mathbf{r})\|_{L^1} / N_e$ for the verification tests against the local density approximation (LDA) based ρ_{data}.

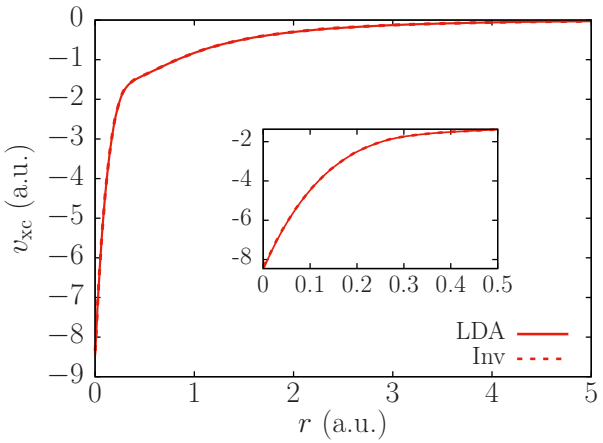
He	Be	Ne	C_8H_{10}
8.1×10^{-6}	8.4×10^{-6}	7.2×10^{-6}	6.8×10^{-6}



(a)

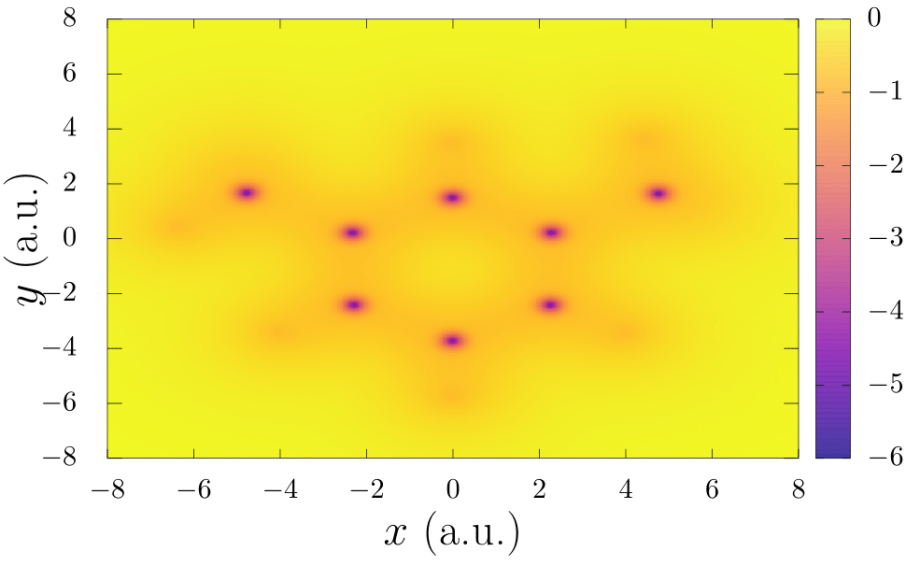


(b)

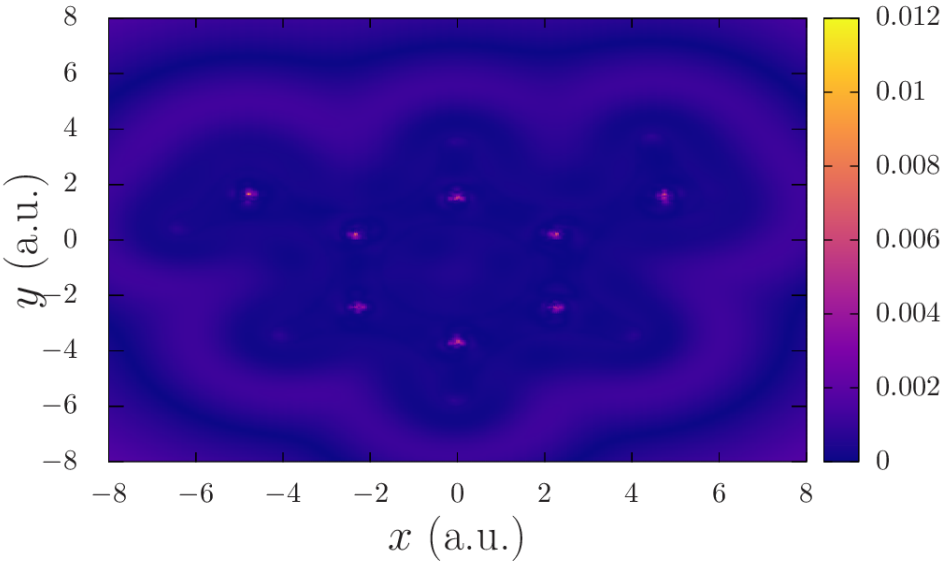


(c)

Figure 5.3: The exchange-correlation potential (v_{xc}) for various atomic systems, each corresponding to the local density approximation (LDA) based density (ρ_{data}). The dashed line corresponds to the exchange-correlation potential obtained from the inverse DFT calculation, and the solid line corresponds to the LDA exchange-correlation potential. The atomic systems considered are: **(a)** He; **(b)** Be; **(c)** Ne.



(a)



(b)

Figure 5.4: **(a)**: The $v_{\text{xc}}(\mathbf{r})$ (in a.u.) for 1,3-dimethylbenzene (C_8H_{10}) obtained from the inverse DFT calculation with the local density approximation (LDA) based ρ_{data}, displayed on the plane of the benzene ring. **(b)**: Relative error in the v_{xc} for 1,3-dimethylbenzene (C_8H_{10}) obtained from the inverse DFT calculation with the local density approximation (LDA) based ρ_{data} (refer Table A.3 in Appendices for coordinates).

5.5 Removing Gaussian basis-set artifacts

We next turn to employing the proposed method with input densities generated from configuration interaction (CI) calculations. All the CI calculations reported in this work are performed using the incremental full-CI approach presented in [211], and discretized using the universal Gaussian basis set (UGBS) [212] or cc-PVTZ Gaussian basis set [213]. It is known that Gaussian basis-set densities, owing to their lack of cusp at the nuclei as well as incorrect far-field decay, induce highly unphysical features in the v_{xc}'s obtained from inverse calculations. To this end, we provide two numerical strategies which for all practical purposes remedy the Gaussian basis-set artifacts, and thereby, allow for accurate evaluation of exact v_{xc}'s from CI densities. It is to be noted that the following numerical strategies are only necessitated due to the unphysical asymptotics in the Gaussian basis-set densities, and not due to any inadequacy of the proposed inverse DFT algorithm.

To begin with, the CI density obtained from a Gaussian basis has wrong decay characteristics away from the nuclei (i.e., Gaussian decay instead of exponential decay). This, in turn, results in incorrect asymptotics in the v_{xc} obtained from an inverse DFT calculation. Thus, to ensure the correct asymptotics in v_{xc}, we employ the following approach. First, we use an initial guess for v_{xc} that satisfies the correct $-1/r$ decay. In particular, we use the Fermi-Amaldi potential (v_{FA}) [214]. Next, we enforce homogeneous Dirichlet boundary condition on the adjoint function (p_i) in the low density region (i.e., $\rho_{\text{data}} < 10^{-6}$), while solving Eq. 5.7. In effect, this fixes the v_{xc} to its initial value in the low density region, thereby ensuring correct far-field asymptotics in the v_{xc}. This approach is also crucial to obtaining an agreement between the highest occupied Kohn-Sham eigenvalue (ϵ_{H}) and the negative of the ionization potential (I_{p}), as mandated by the Koopmans' theorem [73, 74].

Furthermore, the Gaussian basis-set based CI densities lack the cusp at the nuclei, which, in turn, leads to undesirable oscillations in the v_{xc} near the nuclei in any

inverse DFT calculation [65, 64, 68]. We demonstrate this in the case of equilibrium H_2 molecule (bond-length $R_{\text{H-H}} = 1.4$ a.u.), henceforth denoted as $H_2(eq)$. Figure 5.5 shows the v_{xc} profile for $H_2(eq)$ corresponding to the $\rho_{\text{data}}(\mathbf{r})$ obtained from a CI calculation, discretized using UGBS. As evident, we observe large unphysical oscillations in the v_{xc} near the nuclei. We remedy these oscillations by adding a small correction ($\Delta\rho(\mathbf{r})$) to $\rho_{\text{data}}(\mathbf{r})$, so as to correct for the missing cusp at the nuclei. The $\Delta\rho(\mathbf{r})$ is given by

$$\Delta\rho(\mathbf{r}) = \rho_{\text{FE}}^{\text{DFT}}(\mathbf{r}) - \rho_{\text{G}}^{\text{DFT}}(\mathbf{r})\,, \tag{5.10}$$

where $\rho_{\text{FE}}^{\text{DFT}}(\mathbf{r})$ is the ground-state density obtained from a forward DFT calculation using a known xc functional (e.g., LDA, GGA) and discretized using the FE basis; and $\rho_{\text{G}}^{\text{DFT}}(\mathbf{r})$ denotes the same, albeit obtained using the Gaussian basis employed in the CI calculation. The key idea here is that $\rho_{\text{FE}}^{\text{DFT}}(\mathbf{r})$, obtained from the FE basis, contains the cusp. Thus, one can expect $\Delta\rho$ to reasonably capture the Gaussian basis set error near the nuclei. Additionally, $\int \Delta\rho(\mathbf{r})\, d\mathbf{r} = 0$, preserving the number of electrons. A conceptually similar approach has been explored in [68], wherein one post-processes the v_{xc} instead of pre-processing the ρ_{data}, to remove the oscillations arising from the lack of cusp in ρ_{data}. We illustrate the efficacy of the $\Delta\rho$ correction with the $H_2(eq)$ molecule as an example. Figure 5.6 presents the v_{xc} corresponding to the cusp-corrected density (i.e., $\rho_{\text{data}} + \Delta\rho$) for $H_2(eq)$, with two different $\Delta\rho$: $\Delta\rho_{\text{LDA}}$ evaluated using an LDA functional [78, 79], and $\Delta\rho_{\text{GGA}}$ evaluated using a GGA functional [80]. As evident, both $\Delta\rho_{\text{LDA}}$ and $\Delta\rho_{GGA}$ based cusp-correction generate smooth v_{xc} profiles. More importantly, both the profiles are nearly identical, except for small differences in the bonding region between the H atoms. Further, a comparison of both these v_{xc}'s against the LDA based xc potential ($v_{\text{xc}}^{\text{LDA}}$) elucidates the significant difference between the exact v_{xc} and $v_{\text{xc}}^{\text{LDA}}$ even for a simple system that is not strongly correlated. Lastly, for both the v_{xc}'s we obtain the same HOMO level of -0.601 Ha, which, in turn, is in excellent agreement with the $-I_{\text{p}}$ (listed in

Table 5.2). We remark that the agreement of $\epsilon_{\rm H}$ with $-I_{\rm p}$ is a stringent test of the accuracy of the inversion, and is particularly sensitive to the $v_{\rm xc}$ in the far-field.

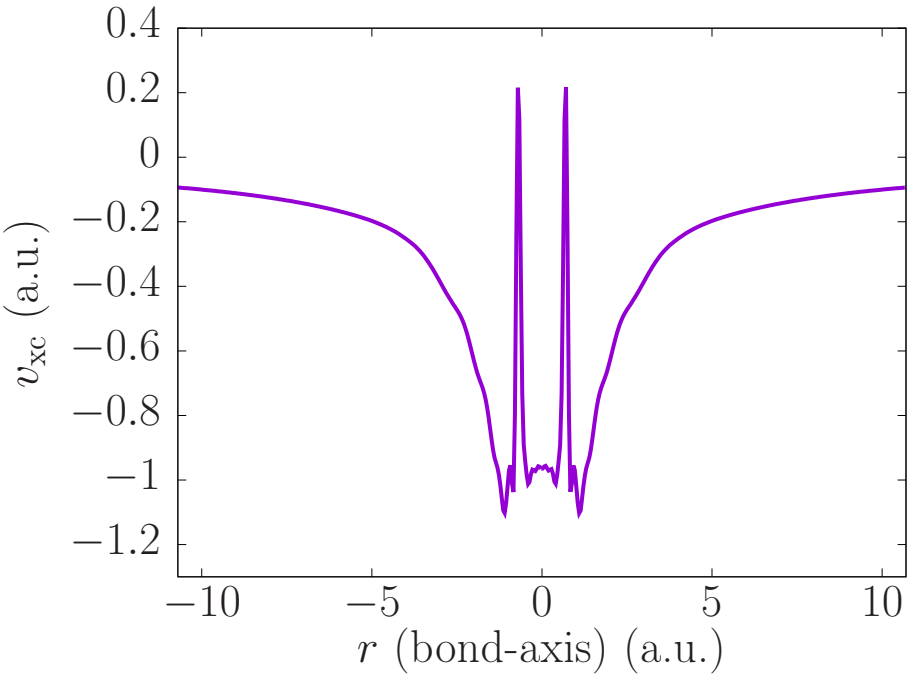


Figure 5.5: Artifact of Gaussian basis-set based density. The exchange-correlation potential ($v_{\rm xc}$) is evaluated from inverse DFT, using $\rho_{\rm data}$ obtained from a Gaussian basis-set based configuration interaction (CI) calculation for the equilibrium hydrogen molecule ($H_2(eq)$). The lack of cusp in $\rho_{\rm data}$ at the nuclei induces wild oscillations in the $v_{\rm xc}$ obtained through inversion. The two atoms are located at $r = \pm 0.7$ a.u.

Table 5.2: Comparison of the Kohn-Sham HOMO level ($\epsilon_{\rm H}$) and the negative of the ionization potential ($I_{\rm p}$) (all in Ha).

	$H_2(eq)$	$H_2(2eq)$	$H_2(d)$	H_2O	C_6H_4
$\epsilon_{\rm H}$	−0.601	−0.482	−0.479	−0.452	−0.354
$-I_{\rm p}$	−0.604	−0.484	−0.498	−0.454	−0.355

5.6 Exact $v_{\rm xc}$ from CI densities for molecules

We now combine the above numerical strategies to evaluate the exact $v_{\rm xc}$ for four other benchmark systems—two stretched H_2 molecules, and two polyatomic systems

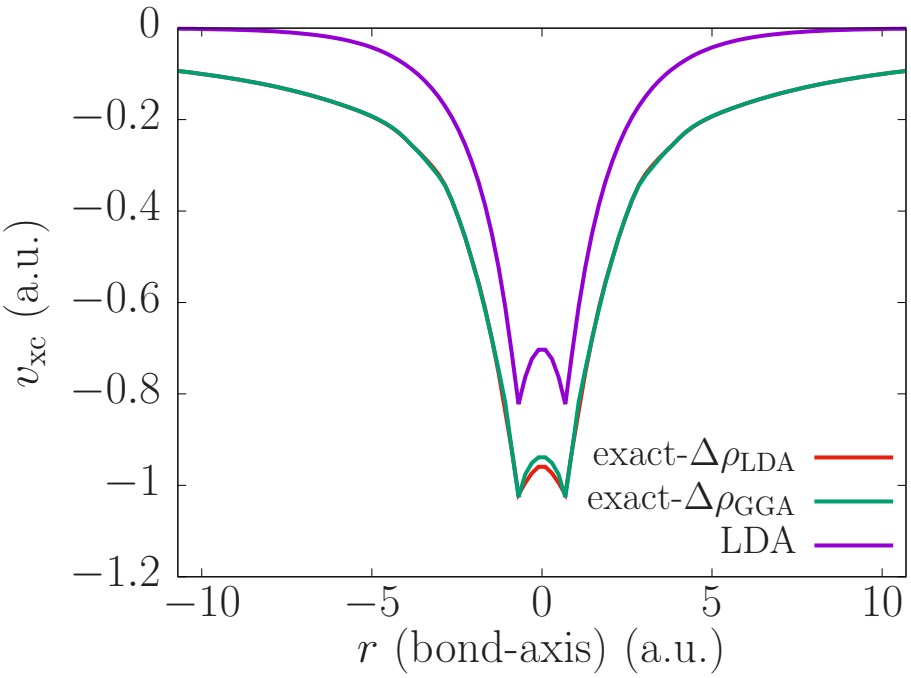


Figure 5.6: Exchange-correlation potentials (v_{xc}) for equilibrium hydrogen molecule ($H_2(eq)$). A comparison is provided between the exact and the LDA based v_{xc} potential. The exact exchange-correlation potential is evaluated using the cusp-corrected configuration interaction (CI) density. The effect of the choice of the functional used in evaluating the cusp correction is demonstrated using two different functionals—LDA (exact-$\Delta\rho_{LDA}$) and GGA (exact-$\Delta\rho_{GGA}$).

(water and ortho-benzyne molecules). The CI calculations for all the molecules, excepting ortho-benzyne, are performed using the UGBS. For ortho-benzyne, we used the cc-PVTZ basis. Given the weak sensitivity of the inverted v_{xc} to the choice of xc functional used in $\Delta\rho$, we employ $\Delta\rho_{LDA}$ for performing the cusp-correction in all our calculations. Further, for all the systems, the inverse problem is deemed to have converged when $||\rho_{data} - \rho||_{L^2} < 10^{-4}$. Given that previous works on this inverse problem have reported the L^1 error, we provide the same in the Table 5.3, for all the benchmark systems considered.

Figure 5.7 compares the v_{xc} for two stretched H_2 molecules—$H_2(2eq)$ ($R_{H\text{-}H} = 2.83$ a.u., roughly twice the equilibrium bond-length), and $H_2(d)$ ($R_{H\text{-}H} = 7.56$ a.u., at dissociation). We emphasize that these are prototypical systems where all existing xc approximations perform poorly, owing to their failure in handling strong correlations.

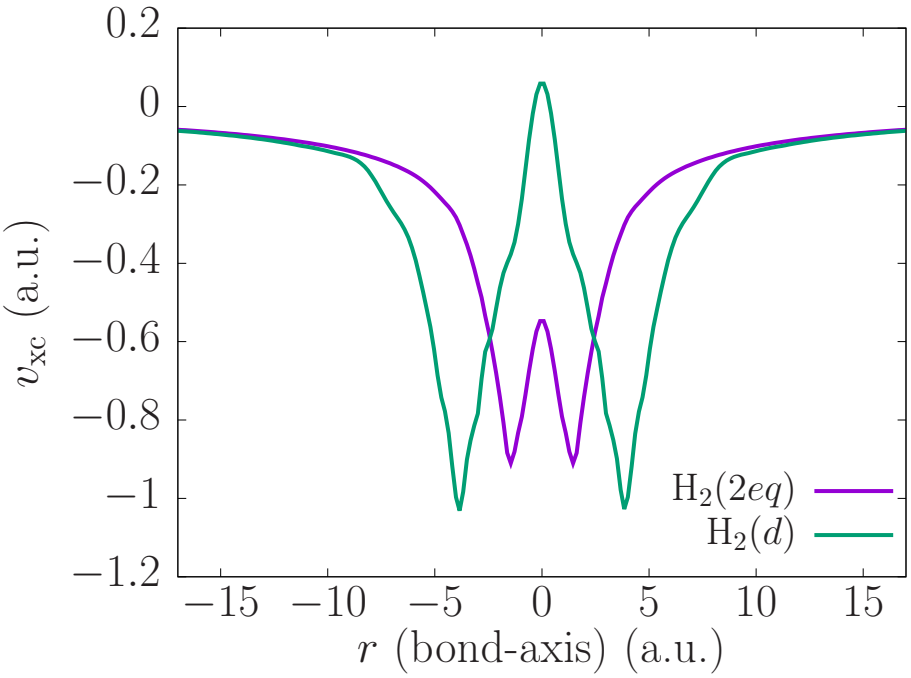


Figure 5.7: Exact v_{xc} for stretched H_2 molecules. The exact v_{xc} is provided for two stretched hydrogen molecules: one at twice the equilibrium bond-length ($H_2(2eq)$) and the other at dissociation ($H_2(d)$). The H atoms for $H_2(2eq)$ and $H_2(d)$ are located at $r = \pm 1.415$ a.u. and $r = \pm 3.78$ a.u., respectively.

We could successfully solve the inverse DFT problem for these systems ($||\rho_{\text{data}} - \rho||_{L^2} \sim 8 \times 10^{-5}$), thereby, suggesting that our approach works equally well for strongly correlated systems. As indicated in Table 5.2, we get remarkable agreement between ϵ_{H} and $-I_{\text{p}}$ for $H_2(2eq)$. However, for $H_2(d)$, we obtain ϵ_{H} within 19 mHa of $-I_{\text{p}}$. We attribute this larger difference between ϵ_{H} and $-I_{\text{p}}$ (as compared to $H_2(eq)$ and $H_2(2eq)$) to the use of v_{FA} as the boundary condition for v_{xc} in the low density region. To elaborate, for a single-orbital system, v_{FA} is the exact v_x (exchange-only potential), and, hence, represents the exact v_{xc} in regions where the correlations are negligible. While for the $H_2(eq)$ and $H_2(2eq)$ molecules the correlations are short-ranged, they are relatively longer-ranged for $H_2(d)$. We highlight this in Figure 5.8 by comparing the v_{xc} against v_x for $H_2(eq)$, $H_2(2eq)$, and $H_2(d)$. As evident, $H_2(d)$ has strong correlations extending to a significantly larger domain (in the far-field) in comparison to $H_2(eq)$ and $H_2(2eq)$. Thus, for $H_2(d)$, the use of v_{FA} is warranted only in regions

of much lower density (i.e., $\rho_{\text{data}} \ll 10^{-6}$) than considered here. However, at such low densities, the wrong far-field asymptotics of a Gaussian basis-set density produces spurious oscillations in the far-field v_{xc}. Thus, for the want of more accurate densities, we are restricted to using v_{FA} in regions where $\rho_{\text{data}} < 10^{-6}$, at the cost of incurring some error in ϵ_{H}.

We now turn to a polyatomic system—the H_2O molecule. Figure 5.9 compares the exact v_{xc} against $v_{\text{xc}}^{\text{LDA}}$, on the plane of the H_2O molecule. Additionally, Figure 5.10 provides the comparison along the O-H bond. For the exact v_{xc}, we observe an atomic inter-shell structure—marked by a yellow ring around the O atom in Figure 5.9(b) (as well as the local maxima and minima at around $r = \pm 0.4$ a.u. in Figure 5.10). This atomic inter-shell structure is a distinctive feature of the exact v_{xc} [215, 216], and is absent in the standard xc approximations, as evident from $v_{\text{xc}}^{\text{LDA}}$. Further, we observe a deeper potential around the O atom, as compared to $v_{\text{xc}}^{\text{LDA}}$, thereby suggesting a higher electronegativity on the O atom than that predicted by LDA. Moreover, we observe a distinct local maximum at the H atom, as opposed to a local minimum in $v_{\text{xc}}^{\text{LDA}}$. Lastly, as indicated in Table 5.2, we obtain striking agreement between ϵ_{H} and $-I_{\text{p}}$ for this polyatomic system.

Finally, we evaluate the exact v_{xc} for the singlet state of the ortho-benzyne radical (C_6H_4)—a strongly correlated species that has previously served as a test for accurate wave function theories [217]. Figure 5.11 compares the exact v_{xc} against $v_{\text{xc}}^{\text{LDA}}$, on the plane of the benzyne molecule. This example underscores the efficacy of our approach in handling both large as well as strongly correlated systems. As expected for the exact v_{xc}, we observe an atomic inter-shell structure—marked by a yellow ring around the C atoms, which, on the other hand, are absent in the case of $v_{\text{xc}}^{\text{LDA}}$. As is the case with H_2O, we observe a deeper potential around the C atom, as compared to $v_{\text{xc}}^{\text{LDA}}$, suggesting a higher electronegativity on the C atom than that predicted by LDA. Furthermore, as indicated in Table 5.2, we obtain remarkable agreement between ϵ_{H}

and $-I_{\mathrm{p}}$.

Table 5.3: $\|\rho(\mathbf{r}) - \rho_{\mathrm{data}}(\mathbf{r})\|_{L^1} / N_e$ for ρ_{data} obtained from CI calculations.

$H_2(eq)$	$H_2(2eq)$	$H_2(d)$	H_2O	C_6H_4
3.5×10^{-5}	4.7×10^{-5}	4.2×10^{-4}	3.4×10^{-5}	8.2×10^{-5}

5.7 Comparison with existing methods

The proposed approach to the inverse DFT problem is now compared to prior approaches in terms of accuracy, robustness and computational viability, using results from existing literature. Given the importance of the inverse DFT problem, several attempts have been made at solving this over the past two decades. Broadly speaking, the approaches proposed can be classified into three major categories—(a) iterative update based algorithms [57, 58, 59, 203, 218, 60]; (b) constrained optimization approaches [61, 62, 63, 66, 67, 205]; (c) linear response based approaches [64]. Irrespective of the underlying approach, the two major factors that determine the accuracy of obtained $v_{\mathrm{xc}}(\mathbf{r})$ are—(i) the completeness of the basis in which the Kohn-Sham orbitals ($\psi_i(\mathbf{r})$) and the $v_{\mathrm{xc}}(\mathbf{r})$ are discretized, and (ii) the accuracy of the input density $\rho_{\mathrm{data}}(\mathbf{r})$, including the correct near-field (cusp at the nuclear positions) and far-field asymptotics. The major criteria for judging the accuracy of the resulting $v_{\mathrm{xc}}(\mathbf{r})$ are the L^1 and L^2 errors in the Kohn-Sham density (i.e. the output density) that is produced by this $v_{\mathrm{xc}}(\mathbf{r})$, as well as errors in the $v_{\mathrm{xc}}(\mathbf{r})$ potential itself (as in the case of the verification studies presented in Sec. 5.4). These metrics will form the basis for our discussion that follows.

To begin with, we discuss the importance of using a complete basis to discretize the inverse DFT problem, which in this work is achieved by using a systematically convergent finite element basis. To this end, we compare the results of our verification studies using $\rho_{\mathrm{data}}^{\mathrm{LDA}}$ with similar results from published literature. This verification test

allows for a direct assessment of the accuracy of the v_{xc} obtained using inversion, by comparing it against $v_{\text{xc}}^{\text{LDA}}[\rho_{\text{data}}]$. As noted earlier, several attempts at this verification test have suffered from either non-unique solutions or had resulted in unphysical oscillations in v_{xc}, owing to the incompleteness of the Gaussian basis employed in these works. In particular, as demonstrated in [66, 67], the use of Gaussian basis results in errors in $\left|v_{\text{xc}}(\mathbf{r}) - v_{\text{xc}}^{\text{LDA}}[\rho_{\text{data}}]\right|$ in the range of $\mathcal{O}(10^0 - 10^1)$ (cf. Figure 2 in [66] and Figure 2 in [67]). A workaround to suppress this incomplete basis induced oscillation is to either introduce a regularization constraint on the v_{xc}, or to construct a balanced potential basis (for the corresponding orbital basis). Although these techniques alleviate the wild oscillations, the resulting v_{xc} still exhibits $\mathcal{O}(10^0)$ error (cf. Figures 4 and 6 in [66]; Figures 6 and 7 in [67]), and is sensitive to the choice of regularization parameters/balanced basis. On the other hand, we demonstrate that the use of finite element basis, owing to its completeness, results in smooth v_{xc} with a tight accuracy of $\mathcal{O}(10^{-2})$, for similar LDA-based verification studies (refer Fig. 5.3 and Fig. 5.4(b)). A recent effort [205] employs additional constraints to obtain the most optimal v_{xc} in a given Slater basis. Although this approach results in better v_{xc} for similar verification studies, the normalized error in the density (i.e., $\|\rho(\mathbf{r}) - \rho_{\text{data}}(\mathbf{r})\|_{L^1} / N_e$) remains high ($\sim 3 \times 10^{-3} - 6 \times 10^{-4}$). In comparison, we obtain $\|\rho(\mathbf{r}) - \rho_{\text{data}}(\mathbf{r})\|_{L^1} / N_e$ in the range of $6 \times 10^{-6} - 8 \times 10^{-6}$ (cf. Table 5.1), for similar verification studies. We remark that, in addition to these quantitative errors, the use of Gaussian basis has also resulted in qualitatively incorrect v_{xc}'s. For example, in [63], the v_{xc} for H_2, corresponding to CCSD based ρ_{data}, does not have a local maximum between the two H atoms, a feature present in both the exact as well approximate v_{xc}'s (such as LDA, GGA)—cf. Fig. 5.5. Thus, the above comparisons show that the finite element basis simultaneously provides accurate potentials and densities in solving the inverse DFT problem.

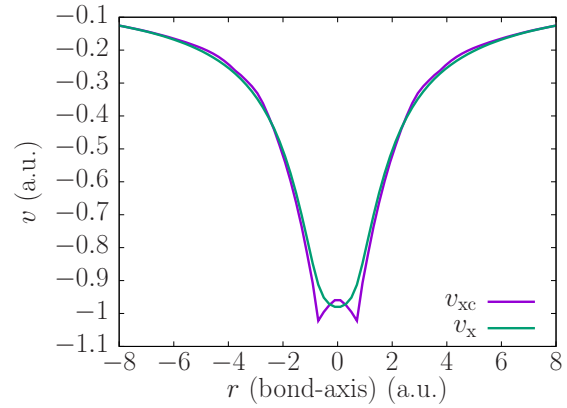
Using accurate densities (ρ_{data}) is also important in the inverse DFT problem,

and the presence of cusp in ρ_{data} at nuclei is particularly important. The vitality of the nuclear cusp has been highlighted in Fig. 5.5, wherein the lack of cusp in the Gaussian basis-set density induces wild oscillations in the v_{xc} near the nuclei. Similar Gaussian-density induced oscillations have also been observed in [65, 64, 218]. Thus, it is desirable to use ρ_{data} generated using Slater or other complete basis sets. However, the difficulty of performing CI calculations in Slater basis (owing to the large computational cost involved in the four-center integrals associated with Slater basis) restricts the availability of physically meaningful densities to atomic and small molecular systems. To this end, we provide a practically useful solution by adding a small $\Delta\rho$ correction to ρ_{data} (cf. Eq. 5.10), so as to correct for the missing cusp at the nuclei. This has enabled us to obtain exact v_{xc}'s of remarkable quality using Gaussian densities, with an accuracy of $\mathcal{O}(10^{-5} - 10^{-4})$ in $\|\rho(\mathbf{r}) - \rho_{\text{data}}(\mathbf{r})\|_{L^1} / N_e$ (cf. Table 5.3), for systems comprising up to 40 electrons. Similar accuracy has been obtained in 1D (atomic) systems [58, 64, 203, 60], wherein Slater densities (or densities obtained on a radial grid) are used in conjunction with solution of 1D (radial) Kohn-Sham equations via numerical integration. An extension of these techniques to 3D systems remains computationally challenging, due to the difficulty in obtaining accurate Slater densities for large molecular systems as well as adapting the numerical approach used in these works for the solution of Kohn-Sham equations to 3D setting. To this end, the use of finite elements, as employed in this work, affords both the availability of accurate densities and the computational viability to handle large 3D molecular systems.

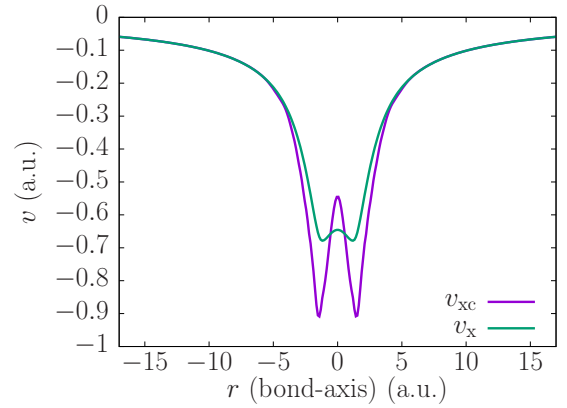
5.8 Summary

We have presented an accurate and robust method to evaluate the exact v_{xc}, solely from the ground-state electron density. The key ingredients in our approach are—(i) the effective use of FE basis, which is a systematically convergent and complete basis,

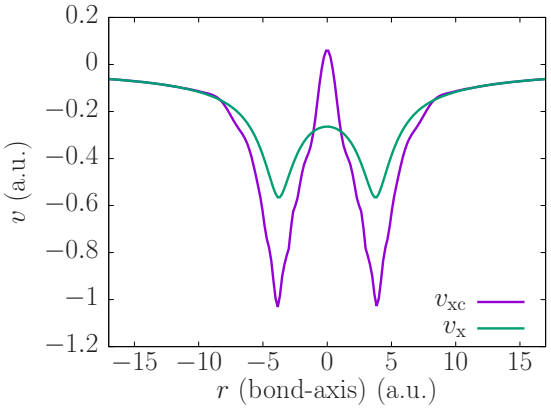
and, in turn, results in a well-posed inverse DFT problem; (ii) use of $\Delta\rho$ correction and appropriate far-field boundary conditions to alleviate the unphysical artifacts associated with Gaussian basis-set densities. We emphasize that the proposed approach can easily drive the error in the target densities, i.e., $||\rho_{\text{data}} - \rho||_{L^2}$, to tight tolerances of $\mathcal{O}(10^{-5} - 10^{-4})$—which represents a stringent accuracy. Notably, as demonstrated through the 1,3-dimethylbenzene and the ortho-benzyne molecules, our approach can competently handle system sizes which have, heretofore, remained challenging for other inverse DFT methods. Furthermore, for all the exact v_{xc}'s obtained from CI densities, we obtain excellent agreement between ϵ_{H} and $-I_{\text{p}}$ (excepting in the case of $H_2(d)$), further validating the accuracy and robustness of the approach. We remark that the larger discrepancy between ϵ_{H} and $-I_{\text{p}}$ in the case of $H_2(d)$ is a consequence of long-range (static) correlations in this system coupled with incorrect far-field asymptotics of Gaussian basis-set densities, and can be remedied with the availability of more accurate densities. The ability to evaluate exact xc potentials from ground-state electron densities, enabled by this method, will provide a powerful tool in the future testing and development of approximate xc functionals. Further, it paves the way for using machine-learning to construct the functional dependence of v_{xc} on ρ, i.e., $v_{\text{xc}}[\rho]$, providing another avenue to develop density functionals [219, 220, 207] that can systematically improve both ground-state densities and energies [221] as well as satisfy the known conditions on the exact functional [222, 223, 224].



(a)

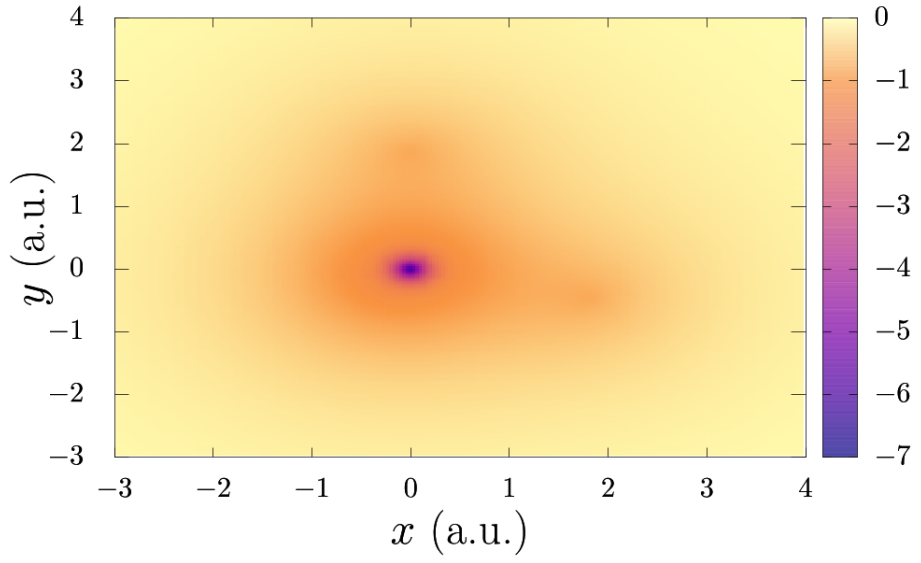


(b)

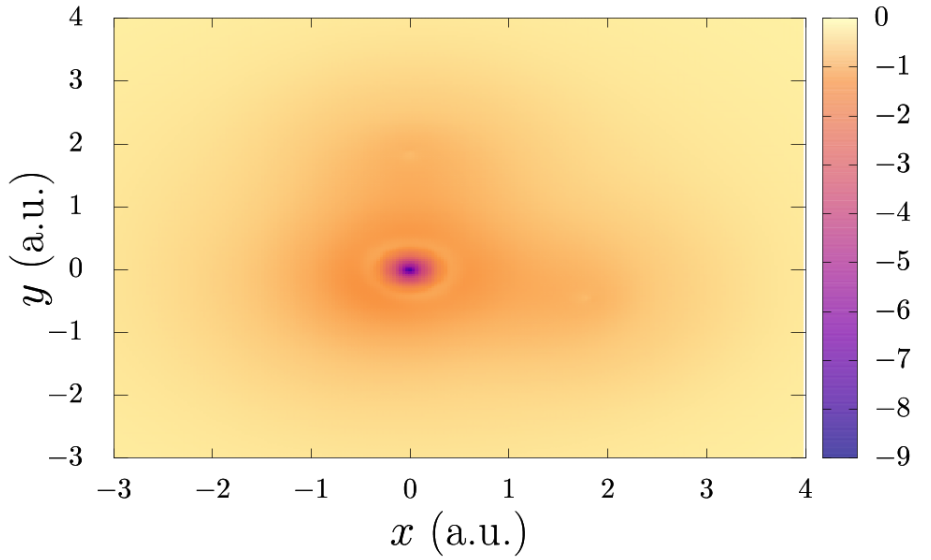


(c)

Figure 5.8: Nature and extent of electronic correlations in H_2 molecules. A comparison of the exact exchange-correlation (v_{xc}) and the exchange-only (v_x) potentials is provided for H_2 molecules at three different bond-lengths: **(a)** equilibrium bond-length ($H_2(eq)$); **(b)** twice the equilibrium bond-length ($H_2(2eq)$); **(c)** at dissociation ($H_2(d)$). The relative difference between v_{xc} and v_x indicates the nature and extent of electronic correlations. The correlations become stronger with bond stretching.



(a)



(b)

Figure 5.9: Comparison of the local density approximation (LDA) based and the exact exchange-correlation potential for H_2O (in a.u.), presented on the plane of the molecule. **(a)** LDA based exchange-correlation potential. **(b)** exact exchange-correlation potential. The O atom is at (0,0), and the two H atoms are at (0,1.89) and (1.83, -0.46) (a.u.), respectively.

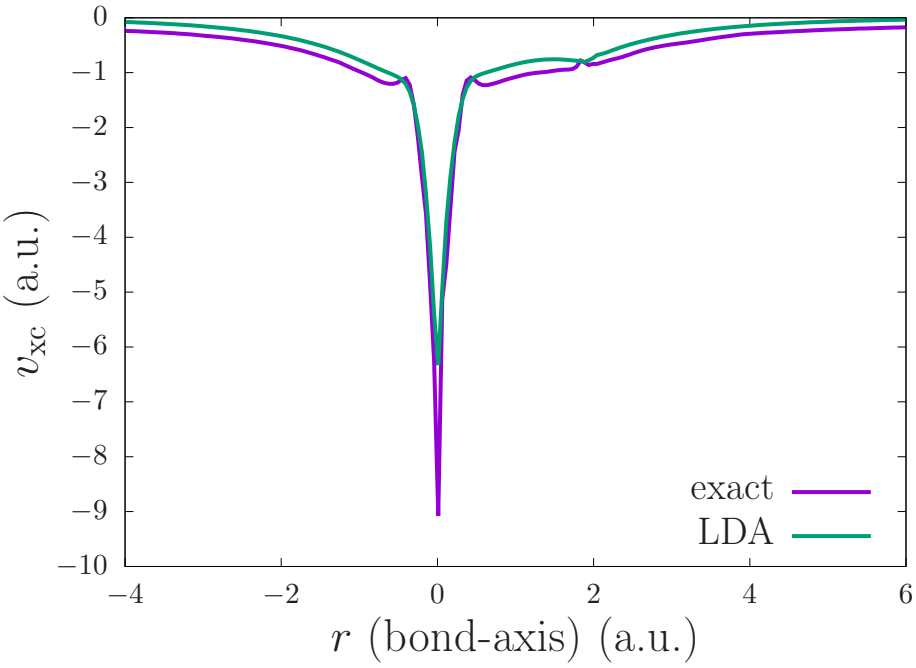
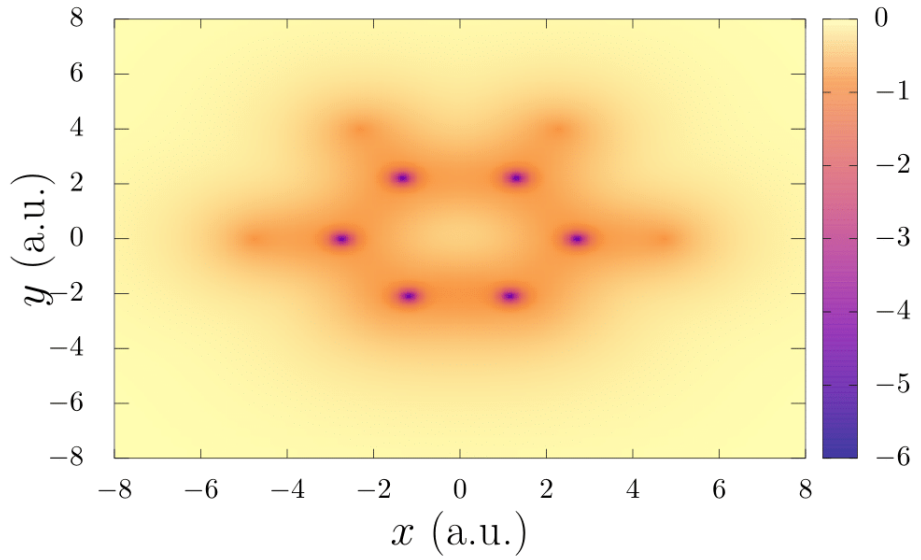
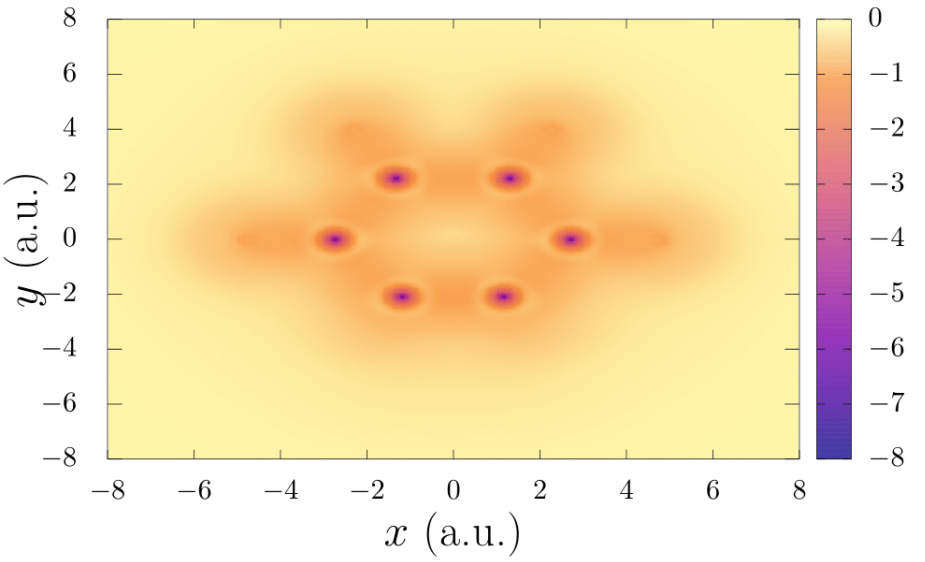


Figure 5.10: Comparison of the LDA based and the exact exchange-correlation potential for H_2O along an O-H bond. The O and H atoms are located at r=0 and r=1.89 (a.u.), respectively.



(a)



(b)

Figure 5.11: Comparison of exchange-correlation potentials (v_{xc}) for C_6H_4. **(a)** LDA based exchange-correlation potential. **(b)** exact exchange-correlation potential. In both the cases, the v_{xc} (in a.u.) is presented on the plane of the molecule (refer Table A.3 in Appendices for coordinates).

CHAPTER VI

Conclusions

6.1 Summary

This thesis had set out with two challenges questions in DFT (TDDFT)—(i) to dispense with the pseudopotential approximation without incurring heavy computational cost, and (ii) provide key insights into the exchange-correlation potential that can form the basis for future testing and development of exchange-correlation approximation. We have addressed both these challenges, in substantial measure, through multiple numerical schemes.

The first part of this thesis (Chapter III) deals with the first question of dispensing with the pseudopotential approximation and performing efficient all-electron DFT calculations. We have developed a computationally efficient mixed basis, termed as *enriched* finite element basis, which augments the classical finite element basis with numerical atom-centered basis (called *enrichment* functions) [37]. The underlying rationale for use of an enriched finite element basis is to combine the efficiency of the atom-centered basis to capture the sharp variations of the electronic fields closer to the atoms with that of the completeness of the classical finite element basis. In effect, we mitigate the need for a highly refined finite element near the atoms, that is typical for all-electron DFT calculations. The enrichment functions were constructed from single-atom Kohn-Sham orbitals and electrostatic potentials. In particular, the

enrichment functions were inexpensively pre-computed and stored by solving radial Kohn-Sham equations for all atoms in the periodic table. Care was taken to truncate the enrichment functions with smooth cutoff functions so as to ensure locality as well as better conditioning of the enriched finite element basis. Furthermore, we employed a divide and conquer strategy to construct an adaptive quadrature grid that can efficiently evaluate the integrals involving the enrichment functions. We employed a computationally efficient scheme to evaluate the inverse of the overlap matrix in the enriched finite element basis, by exploiting the block-wise matrix inversion, which is instrumental in converting the generalized Kohn-Sham eigenvalue problem to a standard one as well as in performing sparse matrix-vector products. The classical-classical block of the overlap matrix is rendered diagonal with the use of spectral finite elements along with Gauss-Lobatto-Legendre, which, in turn, is crucial to efficient evaluation of the inverse of the full overlap matrix. Finally, we employed the Chebyshev polynomial based filter to compute the occupied eigenstates. Here, we exploited the finite element structure in the Hamiltonian and the inverse overlap matrices to achieve an efficient and scalable implementation of the matrix-vector products involved in the action of the Chebyshev filter on a subspace. In terms of the numerical convergence afforded by the enriched finite element basis, we demonstrated close to optimal rates of convergence for the ground-state energy with respect to the finite element discretization. We demonstrated the accuracy and performance of the proposed enriched finite element basis on: (i) silicon nano-clusters of various sizes, with the largest cluster containing 8694 electrons; and (ii) gold nano-clusters of various sizes, with the largest cluster containing 1817 electrons. We obtained good agreement in the ground-state energies when compared to classical finite element and Gaussian basis. In the larger clusters considered in this study, the enriched finite element basis provides a striking $50-300$ fold speedup compared to the classical finite element basis, owing to a $30-$fold reduction in the degrees of freedom as well as a

20−fold reduction in the Chebyshev polynomial degree. We also attained a significant $3-8$-fold outperformance by the enriched finite element basis relative to Gaussian basis (pc-3 and pc-4). Furthermore, we were able to perform ground-state energy calculations for silicon clusters containing 280 and 621 atoms, for which the Gaussian basis failed to converge owing to linear dependency of the basis. These larger systems underscore the effectiveness of enriched finite element basis to systematically scale to several thousands of electrons, with moderate computational expense. In terms of parallel scalability, we obtained good parallel efficiency with almost linear scaling up to 384 processors for the benchmark system comprising of 280 atoms silicon nano-cluster and containing $\sim$4 million basis functions.

Unlike the case in ground-state DFT, wherein prior works (most notably [38]) lay the ground for an efficient adaptive finite element based spatial discretization, TDDFT lacks such adaptive spatio-temporal schemes. To that end, the second part of this thesis (Chapter IV) built the foundation for an efficient higher-order (classical) finite element based spatio-temporal discretization, for both all-electron and pseudopotential based real-time TDDFT (RT-TDDFT) calculations [56]. Hence, this work forms an important stepping stone towards extending the ideas of the enriched finite element based discretization for all-electron TDDFT calculations. We presented an efficient *a priori* spatio-temporal scheme guided by the discretization errors in the time-dependent Kohn-Sham orbitals, in the context of second-order Magnus propagator. In particular, we employed the knowledge of the ground-state electronic fields to determine an efficient adaptively resolved finite element mesh and to make an economic choice for the time-step. We remark that the adaptive resolution, so afforded by our scheme, is crucial to the use of large simulation domains without significant increase in the number of basis functions. This, in turn, allowed us to circumvent the use of any artificial absorbing boundary conditions. We employed higher-order spectral finite-elements along with special (Gauss-Lobatto-Legendre) quadrature rules to

render the overlap matrix as well as improve the conditioning of the basis. Consequently, it enabled the transformation of a generalized equation to its standard form, and hence, provided for an efficient construction of the Magnus propagator (or any exponential time-integrator) for finite element discretization. Furthermore, we used an adaptive Lanczos subspace projection to efficiently evaluate the action of the Magnus propagator, defined as exponential of a matrix, on the Kohn-Sham orbitals. Numerically, we attained rates of convergence, on both pseudopotential and all-electron systems, that are close to the optimal rates of convergence with respect to both spatial and temporal discretization, as determined by our error estimates. We established the computational efficiency afforded by higher-order finite elements by demonstrating a staggering $10 - 100$ fold speedup for fourth-order finite elements over linear and quadratic finite elements. We assessed the performance of our approach against the finite-difference based method of Octopus software package, for pseudopotential calculations. For all the benchmark systems considered, we obtained good agreement between the absorption spectrum evaluated using the finite elements and finite-difference (Octopus) based calculations. Remarkably, we obtained $3 - 60$ fold speedup over finite-difference, largely attributed to the adaptive spatial resolution afforded by our approach. Further, we demonstrated the efficacy of finite elements, especially its efficient handling of large domains, for nonlinear response by studying the higher harmonic generation under a strong electric field. We also demonstrated the competence of higher-order finite elements for all-electron RT-TDDFT calculations. This highlights the versatility of finite elements in handling both pseudopotential and all-electron calculations on an equal footing, and, thereby, offers a powerful tool for conducting transferability studies on pseudopotentials for RT-TDDFT calculations. Lastly, in terms of parallel scalability, we obtained good parallel efficiency up to 768 processors for a benchmark system comprising of the Buckminsterfullerene molecule containing $\sim$3.5 million basis functions.

In the third part of the thesis (Chapter V), we turned to the vital question of obtaining the exact exchange-correlation potential corresponding to a ground-state density, otherwise known as the *inverse* DFT problem. Although it constitutes an immensely consequential question for testing and developing approximate exchange-correlation functionals in DFT, the lack of an accurate and systematically convergent approach has left the problem unresolved, heretofore. To this end, we have presented an accurate and robust method to evaluate the exact v_{xc}, solely from the ground-state electron density. We remark that while the continuous form of inverse DFT problem remains well-posed, the discrete one turns ill-posed when discretized using an incomplete basis. Consequently, prior works, which had employed Gaussian basis to discretize the inverse problem, suffered from either spurious oscillations or non-uniqueness in v_{xc}. In this work, we mitigate any such incomplete basis-induced artifact by employing finite element basis, a systematically convergent and complete basis, to discretize both the Kohn-Sham orbitals as well as the v_{xc}. We, first, demonstrated the accuracy of our approach through verification studies against ground-state densities obtained from DFT calculations with known exchange-correlation functionals (e.g., LDA, GGA). Next, we addressed the challenge of obtaining the exact v_{xc} corresponding to ground-state densities obtained from wavefunction based calculations (e.g., configuration interaction (CI)). We remark that, from a practical viewpoint, Gaussian basis is the only viable means, at present, to obtain all-electron correlated *ab initio* densities from any wavefunction based methods (like CI). However, it is known that Gaussian basis-set densities, owing to their lack of cusp at the nuclei as well as incorrect far-field decay, induce highly unphysical features in the v_{xc}'s obtained from inverse calculations. To this end, we remedy these artifacts by—(a) adding a small $\Delta\rho$ correction to the CI densities, so as to correct for the missing cusp at the nuclei, and (b) using approximate far-field boundary conditions for v_{xc} that honors the $-1/r$ decay. Consequently, the combination of these numerical strate-

gies, in conjunction with the use of finite elements, have attained stringent accuracies in the target densities, i.e., $||\rho_{\text{data}} - \rho||_{L^2}$ of $\mathcal{O}(10^{-5} - 10^{-4})$. More importantly, as demonstrated through the 1,3-dimethylbenzene and the ortho-benzyne molecules, our approach can competently handle system sizes which have, heretofore, remained challenging for other inverse DFT methods. Notably, for all the exact v_{xc}'s obtained from CI densities (excepting in the case of $H_2(d)$), we could verify the Koopmans' theorem to remarkable accuracy by demonstrating the close agreement between ϵ_{H} and $-I_{\text{p}}$. This further validates the accuracy and robustness of the approach. In the case of $H_2(d)$, the relatively larger discrepancy between ϵ_{H} and $-I_{\text{p}}$ stems from the presence of long-range (static) correlations in the system, which when coupled with incorrect far-field asymptotics of Gaussian basis-set densities results in relatively higher error in ϵ_{H}. However, this error can be remedied with the availability of more accurate densities. In conclusion, this ability to evaluate exact xc potentials from ground-state electron densities, enabled by our method, offers a powerful tool that can aid in future testing and development of approximate xc functionals. Further, it provides a new avenue of functional development through machine-learning, wherein one can use the $\{\rho(\mathbf{r}), v_{\text{xc}}(\mathbf{r})\}$ data from inverse calculations to construct the functional dependence of v_{xc} on ρ, i.e., $v_{\text{xc}}[\rho]$. Thus, it holds the key to a systematic route to functional development that can simultaneously improve both ground-state densities and energies as well as satisfy the known conditions on the exact functional.

6.2 Future directions

The numerical strategies developed in this thesis, ranging from efficient all-electron DFT calculations to efficient spatio-temporal discretization in TDDFT to accurate evaluation of the $v_{\text{xc}}(\mathbf{r})$ from a given $\rho(\mathbf{r})$, holds a great deal of promise for all-electron DFT (TDDFT) calculations, from a numerical as well as physical standpoint. It paves the way for a number of interesting research directions within the realm of

DFT (TDDFT). We list some of these worthwhile subjects below.

- **Robust othogonalized enriched finite element basis**: As noted in Sec. 3.5.2, one needs to pre-multiply the enrichment functions with smooth cutoff functions to enforce locality as well as avoid ill-conditioning of the resultant enriched finite element basis. Although it provides a practical approach to avoid ill-conditioning in the enriched finite element basis, it demands some degree of care in terms of choosing the cutoff radius and smoothness of the cutoff function. Furthermore, the conditioning of the basis can become sensitive to the choice of these cutoff parameters, especially while working with a refined classical finite element mesh, thereby affecting the accuracy of the ground-state calculation. Thus, in order to alleviate these shortcomings, we have formulated a modified enriched finite element basis, wherein the enrichment functions are orthogonalized to the subspace spanned by the classical finite element basis, without compromising on their locality. In addition to alleviating the the ill-conditioning concerns, the orthogonality of the modified enrichment functions with the classical finite element basis will render the overlap matrix $\mathbf{M}^E$ (see Sec 3.5.1) block-diagonal, and hence, will greatly simplify the evaluation of its inverse.

- **Enriched finite element based all-electron TDDFT**: Having developed an efficient *a priori* spatio-temporal discretization scheme for RT-TDDFT using classical finite elements, it is but natural to extend the key ideas of enriched finite element basis to RT-TDDFT. As demonstrated in the case of all-electron ground-state DFT, we expect similar reduction in the degrees of freedom for an enriched finite element based all-electron TDDFT calculation. Additionally, the ability to use a coarser finite element discretization will also afford the use of larger time-step and/or smaller Krylov subspace to evaluate the action of the Magnus propagator, thereby, allowing for all-electron TDDFT calculations for

larger time-scales.

- **Dynamic subspace based TDDFT calculations**: We remark that the adaptive Krylov subspace projection based approach to evaluate the action of the Magnus propagator (or any exponential propagator) provides an accurate and efficient scheme for the description of the action of the exponential operator on an orbital. However, it still remains sub-optimal from a computational efficiency standpoint due to a different subspace being used for each $\bar{\boldsymbol{\psi}}_\alpha(t)$, and the need to construct these subspaces at each time-step. Consequently, TDDFT calculations on systems containing more than 1000 electrons for over a time-scale of 100 fs remains a serious computational challenge. To this end, one can dispense with these computational costs by: (i) constructing a common subspace that represents all the electrons in the system, (ii) using a multiscale propagation technique, wherein the subspace is updated on a larger time-scale as compared to the time-step of the exponential propagator. The evolution and size of the common subspace can be guided by drawing ideas from time-dependent perturbation theory and the time-dependent variational principle. To elaborate, the time-dependent perturbation theory can allow us to determine an adequate low-lying eigenspace to represent the Kohn-Sham orbitals, and the time-dependent variational principle can govern the optimal evolution of the subspace.

- **Accelerating inverse DFT**: While the challenge of obtaining the exact $v_{\text{xc}}(\mathbf{r})$ from a given $\rho(\mathbf{r})$ have been adequately addressed in this thesis, from an efficiency viewpoint there is much more to be desired. Accelerating the inversion, without compromising on the accuracy, is instrumental to the generation of training data that can be used for a machine-learned exchange-correlation functional. The current numerical implementation, as presented in this work, relies on the use of classical finite element basis to discretize the inverse DFT prob-

lem, which, in turn, remains inefficient for any all-electron DFT calculation. To this end, the idea of enriched finite element basis can play an important role in accelerating the inverse DFT calculations.

- **Machine learned exchange-correlation functional**: As noted earlier, the solution to the inverse DFT problem presented in this thesis provides a powerful tool to constructing accurate exchange-correlation functionals using machine-learning. To elaborate, we can use the inverse DFT scheme to generate $\{\rho^{(i)}, v_{\text{xc}}^{(i)}\}$ pairs, using $\rho^{(i)}$'s from correlated ab-initio calculations. Subsequently, these $\{\rho^{(i)}, v_{\text{xc}}^{(i)}\}$ can be used as training data to model $v_{\text{xc}}[\rho]$ through machine-learning algorithms [207, 208], conforming to the known exact conditions on v_{xc}. Further, the exchange-correlation energy ($E_{\text{xc}}[\rho]$) can be directly evaluated through line-integration on $v_{\text{xc}}[\rho]$. Given the paramountcy of the exchange-correlation to DFT (TDDFT), this work will constitute an exceedingly interesting research direction.

APPENDIX

APPENDIX A

Coordinates

Table A.1: Coordinates (in a.u.) of systems used in Chapter III

Au_6			
Au	2.9185	−5.0550	0.0000
Au	2.9185	0.0000	0.0000
Au	2.9185	5.0550	0.0000
Au	−1.4593	2.5275	0.0000
Au	−5.8370	0.0000	0.0000
Au	−1.4593	−2.5275	0.0000

Table A.2: Coordinates (in a.u.) of systems used in Chapter IV

Al_{13}			
Al	3.2046	1.8002	3.5079
Al	−3.2046	−1.8002	−3.5079
Al	0.0000	0.0000	0.0000
Al	−3.8897	−2.7943	1.6963
Al	4.5987	−2.1518	0.1944
Al	0.9330	−3.0116	3.9844
Al	0.2141	−4.9914	−0.9252
Al	−0.9330	3.0116	−3.9844
Al	−4.5987	2.1518	−0.1944
Al	3.8897	2.7943	−1.6963
Al	−0.2141	4.9914	0.9252
Al	−2.0415	1.4031	4.4361
Al	2.0415	−1.4031	−4.4361
C_6H_6 (benzene)			
C	0.0000	2.6381	0.0000
C	2.2847	1.3190	0.0000
C	2.2847	−1.3190	0.0000

Continued on next column

Continued from previous column

C	0.0000	−2.6381	0.0000
C	−2.2847	−1.3190	0.0000
C	−2.2847	1.3190	0.0000
H	0.0000	4.6846	0.0000
H	4.0573	2.3433	0.0000
H	4.0573	−2.3433	0.0000
H	0.0000	−4.6846	0.0000
H	−4.0573	−2.3433	0.0000
H	−4.0573	2.3433	0.0000

C_{60} (Buckminsterfullerene)

C	−2.6748	5.6976	−2.2071
C	4.8825	4.3314	−1.3628
C	1.3628	−4.8825	4.3314
C	5.6976	−2.2071	2.6748
C	−5.6976	−2.2071	2.6748
C	−1.3627	−4.8825	4.3314
C	−4.8825	4.3314	−1.3628
C	−6.5411	0.0000	1.3121
C	2.6748	5.6976	−2.2071

Continued on next column

Continued from previous column

C	6.5411	0.0000	1.3121
C	−2.6748	−5.6976	−2.2071
C	−4.3314	−1.3628	−4.8825
C	0.0000	1.3121	−6.5411
C	4.3314	−1.3628	−4.8825
C	2.6748	−5.6976	−2.2071
C	−4.3314	1.3628	−4.8825
C	2.2071	2.6748	−5.6976
C	5.6976	−2.2071	−2.6748
C	1.3121	−6.5411	0.0000
C	−4.8825	−4.3314	−1.3628
C	1.3628	4.8825	4.3314
C	2.2071	2.6748	5.6976
C	0.0000	1.3121	6.5411
C	−2.2071	2.6748	5.6976
C	−1.3628	4.8825	4.3314
C	−1.3121	6.5411	0.0000
C	4.8825	4.3314	1.3628
C	4.3314	−1.3628	4.8825

Continued on next column

Continued from previous column

C	−2.2071	−2.6748	5.6976
C	−5.6976	2.2071	2.6748
C	2.2071	−2.6748	5.6976
C	−4.3314	−1.3628	4.8825
C	−4.8825	4.3314	1.3628
C	1.3121	6.5411	0.0000
C	5.6976	2.2071	2.6748
C	−1.3121	−6.5411	0.0000
C	−5.6976	−2.2071	−2.6748
C	−2.2071	2.6748	−5.6976
C	4.3314	1.3628	−4.8825
C	4.8825	−4.3314	−1.3628
C	2.2071	−2.6748	−5.6976
C	1.3628	−4.8825	−4.3314
C	−1.3628	−4.8825	−4.3314
C	−2.2071	−2.6748	−5.6976
C	0.0000	−1.3121	−6.5411
C	1.3627	4.8825	−4.3314
C	6.5411	0.0000	−1.3121

Continued on next column

Continued from previous column

C	2.6748	−5.6976	2.2071
C	−4.8825	−4.3314	1.3628
C	−5.6976	2.2071	−2.6748
C	−4.3314	1.3628	4.8825
C	−2.6748	5.6976	2.2071
C	2.6748	5.6976	2.2071
C	4.3314	1.3628	4.8825
C	0.0000	−1.3121	6.5411
C	−2.6748	−5.6976	2.2071
C	−6.5411	0.0000	−1.3121
C	−1.3628	4.8825	−4.3314
C	5.6976	2.2071	−2.6748
C	4.8825	−4.3314	1.3628

Table A.3: Coordinates (in a.u.) of systems used in Chapter V

$H_2(eq)$			
H	−0.6990	0.0000	0.0000
H	0.6990	0.0000	0.0000
$H_2(2eq)$			
H	−1.4150	0.0000	0.0000
H	1.4150	0.0000	0.0000
$H_2(d)$			
H	−3.7795	0.0000	0.0000
H	3.7795	0.0000	0.0000
H_2O			
O	0.0000	0.0000	0.0000
H	0.0000	1.8897	0.0
H	1.8327	−0.4607	0.000
C_6H_4 (ortho-benzyne)			
C	1.17276	−2.085	0.0000
C	−1.17276	−2.0850	0.0000
C	−2.7274	−0.0054	0.0000
C	−1.3202	2.2231	0.0000

Continued on next column

Continued from previous column

C	1.3202	2.2231	0.0000
C	2.7274	−0.0054	0.0000
H	−4.7590	−0.00896	0.0000
H	−2.303	4.0071	0.0000
H	2.303	4.0071	0.0000
H	4.7590	−0.00896	0.0000
C_8H_{10} (1,3-dimethylbenzene)			
C	2.2976	0.2359	0.0000
C	−0.00276	1.50778	0.0000
C	−2.3137	0.2284	0.0000
C	−2.284	−2.3996	0.0000
C	−0.00226	−3.7077	0.0000
C	2.2689	−2.4046	0.0000
C	4.7677	1.6546	0.0000
C	−4.7636	1.6818	0.0000
H	−0.00426	3.5568	0.0000
H	−4.0496	−3.4334	0.0000
H	−0.00443	−5.7529	0.0000
H	4.0371	−3.4356	0.0000

Continued on next column

Continued from previous column

H	5.8985	1.1838	1.6606
H	4.464	3.6909	0.0000
H	5.89852	1.1838	−1.6606
H	−4.9107	2.8976	1.6608
H	−6.3865	0.4148	0.0000
H	−4.9107	2.8976	−1.6608

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Kieron Burke. Perspective on density functional theory. *J. Chem. Phys.*, 136(15):150901, 2012.

[2] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, 87:897–923, 2015.

[3] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[4] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:1133–1138, 1965.

[5] Erich Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52:997–1000, 1984.

[6] Artem R. Oganov and Peter I. Dorogokupets. All-electron and pseudopotential study of MgO: Equation of state, anharmonicity, and stability. *Phys. Rev. B*, 67:224110, 2003.

[7] M. Abu-Jafar, A. I. Al-Sharif, and A. Qteish. FP-LAPW and pseudopotential calculations of the structural phase transformations of GaN under high-pressure. *Solid State Commun.*, 116(7):389–393, 2000.

[8] H. Y. Xiao, X. D. Jiang, G. Duan, Fei Gao, X. T. Zu, and W. J. Weber. First-principles calculations of pressure-induced phase transformation in AlN and GaN. *Comput. Mater. Sci.*, 48(4):768–772, 2010.

[9] P R Levashov, G V Sin'ko, N A Smirnov, D V Minakov, O P Shemyakin, and K V Khishchenko. Pseudopotential and full-electron DFT calculations of thermodynamic properties of electrons in metals and semiempirical equations of state. *J. Phys. Condens. Matter*, 22(50):505501, 2010.

[10] Wenjian Liu, Wolfgang Küchle, and Michael Dolg. *Ab initio* pseudopotential and density-functional all-electron study of ionization and excitation energies of actinide atoms. *Phys. Rev. A*, 58:1103–1110, 1998.

[11] Peter Schwerdtfeger, Behnam Assadollahzadeh, Urban Rohrmann, Rolf Schäfer, and James R. Cheeseman. Breakdown of the pseudopotential approximation for magnetizabilities and electric multipole moments: Test calculations for Au, AuF, and Sn_n cluster ($n \leq 20$). *J. Chem. Phys.*, 134(20), 2011.

[12] Peter Schwerdtfeger, Thomas Fischer, Michael Dolg, Gudrun Igel-Mann, Andreas Nicklass, Hermann Stoll, and Arne Haaland. The accuracy of the pseudopotential approximation. I. An analysis of the spectroscopic constants for the electronic ground states of InCl and InCl3 using various three valence electron pseudopotentials for indium. *J. Chem. Phys.*, 102(1996):2050, 1995.

[13] Peter Schwerdtfeger, J Reuben Brown, Jon K Laerdahl, and Hermann Stoll. between pseudopotential and all-electron methods for Au and AuH. *J. Chem. Phys.*, 113(2000):7110–7118, 2000.

[14] Ricardo Gómez-Abal, Xinzheng Li, Matthias Scheffler, and Claudia Ambrosch-Draxl. Influence of the core-valence interaction and of the pseudopotential approximation on the electron self-energy in semiconductors. *Phys. Rev. Lett.*, 101:106404, 2008.

[15] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–794, 2008.

[16] Aron J. Cohen, Paula Mori-Snchez, and Weitao Yang. Challenges for density functional theory. *Chem. Rev.*, 112(1):289–320, 2012. PMID: 22191548.

[17] David J. Tozer and Nicholas C. Handy. On the determination of excitation energies using density functional theory. *Phys. Chem. Chem. Phys.*, 2:2117–2121, 2000.

[18] Andreas Dreuw and Martin Head-Gordon. Failure of time-dependent density functional theory for long-range charge-transfer excited states: the zincbacteriochlorin-zacteriochlorin and bacteriochlorophyll-zpheroidene complexes. *J. Am. Chem. Soc.*, 126(12):4007–4016, 2004.

[19] Kieron Burke, Jan Werschnik, and E. K. U. Gross. Time-dependent density functional theory: Past, present, and future. *J. Chem. Phys.*, 123(6):062206, 2005.

[20] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. I. use of Gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.*, 51(6):2657–2664, 1969.

[21] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli,

J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian09 Revision A.1. Gaussian Inc. Wallingford CT 2009.

[22] Reinhart Ahlrichs, Michael Bär, Marco Häser, Hans Horn, and Christoph Kölmel. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.*, 162(3):165 – 169, 1989.

[23] Michael W. Schmidt, Kim K. Baldridge, Jerry A. Boatz, Steven T. Elbert, Mark S. Gordon, Jan H. Jensen, Shiro Koseki, Nikita Matsunaga, Kiet A. Nguyen, Shujun Su, Theresa L. Windus, Michel Dupuis, and John A. Montgomery. General atomic and molecular electronic structure system. *J. Comput. Chem.*, 14(11):1347–1363, 1993.

[24] Yihan Shao, Zhengting Gan, Evgeny Epifanovsky, Andrew TB Gilbert, Michael Wormit, Joerg Kussmann, Adrian W Lange, Andrew Behn, J ia Deng, Xintian Feng, et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.*, 113(2):184–215, 2015.

[25] Volker Blum, Ralf Gehrke, Felix Hanke, Paula Havu, Ville Havu, Xinguo Ren, Karsten Reuter, and Matthias Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.*, 180(11):2175 – 2196, 2009.

[26] 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.A. de Jong. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.*, 181:1477–1489, 2010.

[27] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz. Molpro: a generalpurpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2:242–253, 2012.

[28] Kestutis Aidas, Celestino Angeli, Keld L. Bak, Vebjørn Bakken, Radovan Bast, Linus Boman, Ove Christiansen, Renzo Cimiraglia, Sonia Coriani, Pål Dahle, Erik K. Dalskov, Ulf Ekström, Thomas Enevoldsen, Janus J. Eriksen, Patrick Ettenhuber, Berta Fernández, Lara Ferrighi, Heike Fliegl, Luca Frediani, Kasper Hald, Asger Halkier, Christof Hättig, Hanne Heiberg, Trygve Helgaker, Alf Christian Hennum, Hinne Hettema, Eirik Hjertenæs, Stinne Høst, Ida-Marie Høyvik, Maria Francesca Iozzi, Branislav Jansík, Hans Jørgen Aa. Jensen, Dan Jonsson, Poul Jørgensen, Joanna Kauczor, Sheela Kirpekar, Thomas Kjærgaard, Wim Klopper, Stefan Knecht, Rika Kobayashi, Henrik Koch, Jacob Kongsted, Andreas Krapp, Kasper Kristensen, Andrea Ligabue, Ola B. Lutnæs, Juan I. Melo, Kurt V. Mikkelsen, Rolf H. Myhre, Christian Neiss, Christian B. Nielsen, Patrick Norman, Jeppe Olsen, Jógvan Magnus H. Olsen, Anders Osted, Martin J. Packer, Filip Pawlowski, Thomas B. Pedersen, Patricio F. Provasi, Simen Reine, Zilvinas Rinkevicius, Torgeir A. Ruden, Kenneth

Ruud, Vladimir V. Rybkin, Pawel Sałek, Claire C. M. Samson, Alfredo Sánchez de Merás, Trond Saue, Stephan P. A. Sauer, Bernd Schimmelpfennig, Kristian Sneskov, Arnfinn H. Steindal, Kristian O. Sylvester-Hvid, Peter R. Taylor, Andrew M. Teale, Erik I. Tellgren, David P. Tew, Andreas J. Thorvaldsen, Lea Thøgersen, Olav Vahtras, Mark A. Watson, David J. D. Wilson, Marcin Ziolkowski, and Hans Ågren. The dalton quantum chemistry program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 4(3):269–284, 2014.

[29] T. L. Loucks. *Augmented Plane Wave Method: A Guide to Performing Electronic Structure Calculations*. Frontiers in Physics: Lecture note and reprint series, A. Benjamin, 1967.

[30] D D Koelling and G O Arbman. Use of energy derivative of the radial solution in an augmented plane wave method: application to copper. *J. Phys. F: Met. Phys.*, 5(11):2041, 1975.

[31] O. Krogh Andersen. Linear methods in band theory. *Phys. Rev. B*, 12:3060–3083, 1975.

[32] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman. Full-potential self-consistent linearized-augmented-plane-wave method for calculating the electronic structure of molecules and surfaces: O_2 molecule. *Phys. Rev. B*, 24:864–875, 1981.

[33] M. Weinert, E. Wimmer, and A. J. Freeman. Total-energy all-electron density functional method for bulk solids and surfaces. *Phys. Rev. B*, 26:4571–4578, 1982.

[34] Elisabeth Sjöstedt, Lars Nordström, and D. J. Singh. An alternative way of linearizing the augmented plane-wave method. *Solid State Commun.*, 114(1):15–20, 2000.

[35] Georg K. H. Madsen, Peter Blaha, Karlheinz Schwarz, Elisabeth Sjöstedt, and Lars Nordström. Efficient linearization of the augmented plane-wave method. *Phys. Rev. B*, 64:195134, 2001.

[36] Andris Gulans, Stefan Kontur, Christian Meisenbichler, Dmitrii Nabok, Pasquale Pavone, Santiago Rigamonti, Stephan Sagmeister, Ute Werner, and Claudia Draxl. exciting: a full-potential all-electron package implementing density-functional theory and many-body perturbation theory. *J. Phys. Condens. Matter*, 26(36):363202, 2014.

[37] Bikash Kanungo and Vikram Gavini. Large-scale all-electron density functional theory calculations using an enriched finite-element basis. *Phys. Rev. B*, 95:035112, 2017.

[38] P. Motamarri, M. R. Nowak, K. Leiter, J. Knap, and V. Gavini. Higher-order adaptive finite-element methods for Kohn–Sham density functional theory. *J. Comput. Phys.*, 253:308–343, 2013.

[39] Steven R. White, John W. Wilkins, and Michael P. Teter. Finite-element method for electronic structure. *Phys. Rev. B*, 39:5819–5833, 1989.

[40] Eiji Tsuchida and Masaru Tsukada. Large-scale electronic-structure calculations based on the adaptive finite-element method. *J. Phys. Soc. Jpn.*, 67(11):3844–3858, 1998.

[41] J. E. Pask, B. M. Klein, C. Y. Fong, and P. A. Sterne. Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach. *Phys. Rev. B*, 59:12352–12358, 1999.

[42] J.E. Pask, B.M. Klein, P.A. Sterne, and C.Y. Fong. Finite-element methods in electronic-structure theory. *Comput. Phys. Commun.*, 135:1–34, 2001.

[43] Shunsuke Yamakawa and Shi-aki Hyodo. Gaussian finite-element mixed-basis method for electronic structure calculations. *Phys. Rev. B*, 71:035113, 2005.

[44] J E Pask and P A Sterne. Finite element methods in ab initio electronic structure calculations. *Model. Simul. Mater. Sci. Eng.*, 13(3):R71, 2005.

[45] J-L Fattebert, Richard D Hornung, and Andrew M Wissink. Finite element approach for density functional theory calculations on locally-refined meshes. *J. Comput. Phys.*, 223(2):759–773, 2007.

[46] Dier Zhang, Lihua Shen, Aihui Zhou, and Xin-Gao Gong. Finite element method for solving Kohn–Sham equations based on self-adaptive tetrahedral mesh. *Phys. Lett. A*, 372:5071–5076, 2008.

[47] Eric J. Bylaska, Michael Holst, and John H. Weare. Adaptive finite element method for solving the exact Kohn–Sham equation of density functional theory. *J. Chem. Theory Comput.*, 5(4):937–948, 2009.

[48] Lauri Lehtovaara, Ville Havu, and Martti Puska. All-electron density functional theory and time-dependent density functional theory with high-order finite elements. *J. Chem. Phys.*, 131:054103, 2009.

[49] Phanish Suryanarayana, Vikram Gavini, Thomas Blesgen, Kaushik Bhattacharya, and Michael Ortiz. Non-periodic finite-element formulation of Kohn–Sham density functional theory. *J. Mech. Phys. Solids*, 58:256–280, 2010.

[50] Jun Fang, Xingyu Gao, and Aihui Zhou. A Kohn–Sham equation solver based on hexahedral finite elements. *J. Comput. Phys.*, 231:3166–3180, 2012.

[51] Gang Bao, Guanghui Hu, and Di Liu. An h-adaptive finite element solver for the calculations of the electronic structures. *J. Comput. Phys.*, 231:4967–4979, 2012.

[52] Volker Schauer and Christian Linder. All-electron Kohn–Sham density functional theory on hierarchic finite element spaces. *J. Comput. Phys.*, 250:644–664, 2013.

[53] Phani Motamarri and Vikram Gavini. Subquadratic-scaling subspace projection method for large-scale Kohn-Sham density functional theory calculations using spectral finite-element discretization. *Phys. Rev. B*, 90:115127, 2014.

[54] Phani Motamarri, Vikram Gavini, Kaushik Bhattacharya, and Michael Ortiz. Spectrum-splitting approach for fermi-operator expansion in all-electron Kohn-Sham DFT calculations. *Phys. Rev. B*, 95:035111, 2017.

[55] Phani Motamarri and Vikram Gavini. Configurational forces in electronic structure calculations using Kohn-Sham density functional theory. *Phys. Rev. B*, 97:165132, 2018.

[56] Bikash Kanungo and Vikram Gavini. Real-time time-dependent density functional theory using higher order finite element methods. *Phys. Rev. B*, 2019. (in press).

[57] Yue Wang and Robert G. Parr. Construction of exact Kohn-Sham orbitals from a given electron density. *Phys. Rev. A*, 47:R1591–R1593, 1993.

[58] R. van Leeuwen and E. J. Baerends. Exchange-correlation potential with correct asymptotic behavior. *Phys. Rev. A*, 49:2421–2431, 1994.

[59] Oleg V. Gritsenko, Robert van Leeuwen, and Evert Jan Baerends. Molecular Kohn-Sham exchange-correlation potential from the correlated ab initio electron density. *Phys. Rev. A*, 52:1870–1874, 1995.

[60] Ilya G. Ryabinkin and Viktor N. Staroverov. Determination of Kohn-Sham effective potentials from electron densities using the differential virial theorem. *J. Chem. Phys.*, 137(16):164113, 2012.

[61] Qingsheng Zhao, Robert C. Morrison, and Robert G. Parr. From electron densities to Kohn-Sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies. *Phys. Rev. A*, 50:2138–2142, 1994.

[62] David J. Tozer, Victoria E. Ingamells, and Nicholas C. Handy. Exchange-correlation potentials. *J. Chem. Phys.*, 105(20):9200–9213, 1996.

[63] Qin Wu and Weitao Yang. A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities. *J. Chem. Phys.*, 118(6):2498–2509, 2003.

[64] P. R. T. Schipper, O. V. Gritsenko, and E. J. Baerends. Kohn-Sham potentials corresponding to Slater and Gaussian basis set densities. *Theor. Chem. Acc.*, 98(1):16–24, 1997.

[65] Michael E. Mura, Peter J. Knowles, and Christopher A. Reynolds. Accurate numerical determination of Kohn-Sham potentials from electronic densities: I. Two-electron systems. *J. Chem. Phys.*, 106(23):9659–9667, 1997.

[66] Tim Heaton-Burgess, Felipe A. Bulat, and Weitao Yang. Optimized effective potentials in finite basis sets. *Phys. Rev. Lett.*, 98:256401, 2007.

[67] Felipe A. Bulat, Tim Heaton-Burgess, Aron J. Cohen, and Weitao Yang. Optimized effective potentials from electron densities in finite basis sets. *J. Chem. Phys.*, 127(17):174101, 2007.

[68] Alex P. Gaiduk, Ilya G. Ryabinkin, and Viktor N. Staroverov. Removal of basis-set artifacts in KohnSham potentials recovered from electron densities. *J. Chem. Theory Comput.*, 9(9):3959–3964, 2013. PMID: 26592391.

[69] Alexei A. Kananenka, Sviataslau V. Kohut, Alex P. Gaiduk, Ilya G. Ryabinkin, and Viktor N. Staroverov. Efficient construction of exchange and correlation potentials by inverting the KohnSham equations. *J. Chem. Phys.*, 139(7):074112, 2013.

[70] Ilya G. Ryabinkin, Sviataslau V. Kohut, and Viktor N. Staroverov. Reduction of electronic wave functions to kohn-sham effective potentials. *Phys. Rev. Lett.*, 115:083001, 2015.

[71] Egor Ospadov, Ilya G. Ryabinkin, and Viktor N. Staroverov. Improved method for generating exchange-correlation potentials from electronic wave functions. *J. Chem. Phys.*, 146(8):084103, 2017.

[72] Bikash Kanungo, Paul Zimmerman, and Vikram Gavini. Exact exchange-correlation potentials from ground-state electron densities. *Nat. Commun.*, 2019. (in press).

[73] John P. Perdew and Mel Levy. Comment on "Significance of the highest occupied Kohn-Sham eigenvalue". *Phys. Rev. B*, 56:16021–16028, 1997.

[74] Mark E. Casida. Correlated optimized effective-potential treatment of the derivative discontinuity and of the highest occupied Kohn-Sham eigenvalue: A Janak-type theorem for the optimized effective-potential model. *Phys. Rev. B*, 59:4694–4698, 1999.

[75] Robert van Leeuwen. Density functional approach to the many-body problem: key concepts and exact functionals. *Adv. Quantum Chem.*, 43:25–94, 2003.

[76] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U.S.A*, 76(12):6062–6065, 1979.

[77] T. L. Gilbert. Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B*, 12:2111–2120, 1975.

[78] J. C. Slater and K. H. Johnson. Self-consistent-field $X\alpha$ cluster method for polyatomic molecules and solids. *Phys. Rev. B*, 5:844–853, 1972.

[79] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45:13244–13249, 1992.

[80] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[81] Axel D. Becke. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.*, 98(2):1372–1377, 1993.

[82] Axel D. Becke. Density-functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.

[83] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.

[84] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980.

[85] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98(45):11623–11627, 1994.

[86] John P. Perdew, Matthias Ernzerhof, and Kieron Burke. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.*, 105(22):9982–9985, 1996.

[87] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened coulomb potential. *J. Chem. Phys.*, 118(18):8207–8215, 2003.

[88] W. Kohn, A. D. Becke, and R. G. Parr. Density functional theory of electronic structure. *J. Phys. Chem.*, 100(31):12974–12980, 1996.

[89] Axel D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.*, 140(18):18A301, 2014.

[90] Narbe Mardirossian and Martin Head-Gordon. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.*, 115(19):2315–2372, 2017.

[91] Miguel A L Marques, Carsten A Ullrich, Fernando Nogueira, Angel Rubio, Kieron Burke, and Eberhard K U Gross. *Time-dependent density functional theory*, volume 706. Springer, Berlin Heidelberg, 2006.

[92] Carsten A Ullrich. *Time-dependent density-functional theory: concepts and applications*. OUP Oxford, 2011.

[93] Giovanni Vignale. Mapping from current densities to vector potentials in time-dependent current density functional theory. *Phys. Rev. B*, 70:201102, 2004.

[94] Robert van Leeuwen. Mapping from densities to potentials in time-dependent density-functional theory. *Phys. Rev. Lett.*, 82:3863–3866, 1999.

[95] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64:1045–1097, 1992.

[96] G. B. Bachelet, D. R. Hamann, and M. Schlüter. Pseudopotentials that work: From H to Pu. *Phys. Rev. B*, 26:4199–4228, 1982.

[97] James R Chelikowsky. The pseudopotential-density functional method applied to nanostructures. *J. Phys. D: Appl. Phys.*, 33(8):R33, 2000.

[98] Peter Schwerdtfeger. The pseudopotential approximation in electronic structure theory. *ChemPhysChem*, 12(17):3143–3155, 2011.

[99] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43:1494–1497, 1979.

[100] David Vanderbilt. Optimally smooth norm-conserving pseudopotentials. *Phys. Rev. B*, 32:8412–8415, 1985.

[101] D. R. Hamann. Generalized norm-conserving pseudopotentials. *Phys. Rev. B*, 40:2980–2987, 1989.

[102] N. Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43:1993–2006, 1991.

[103] David Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41:7892–7895, 1990.

[104] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, 1994.

[105] M. Marsman and G. Kresse. Relaxed core projector-augmented-wave method. *J. Chem. Phys.*, 125(10):104101, 2006.

[106] James R. Chelikowsky, N. Troullier, and Y. Saad. Finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Phys. Rev. Lett.*, 72:1240–1243, 1994.

[107] Leeor Kronik, Adi Makmal, Murilo L. Tiago, M. M. G. Alemany, Manish Jain, Xiangyang Huang, Yousef Saad, and James R. Chelikowsky. PARSEC—the pseudopotential algorithm for real-space electronic structure calculations: recent advances and novel applications to nano-structures. *Phys. Status Solidi B*, 243(5):1063–1079, 2006.

[108] Susanne Brenner and Ridgway Scott. *The mathematical theory of finite element methods*, volume 15. Springer Science & Business Media, 2007.

[109] Thomas J R Hughes. *The finite element method: linear static and dynamic finite element analysis*. Courier Corporation, 2012.

[110] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[111] Richard M. Martin. *Electronic structure : basic theory and practical methods*. Cambridge University Press, Cambridge, New York, 2004.

[112] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, 1980.

[113] Phani Motamarri, Mrinal Iyer, Jaroslaw Knap, and Vikram Gavini. Higher-order adaptive finite-element methods for orbital-free density functional theory. *J. Comput. Phys.*, 231:6596–6621, 2012.

[114] G. Kresse and J. Furthmüller. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, 1996.

[115] Stefan Goedecker. Linear scaling electronic structure methods. *Rev. Mod. Phys.*, 71:1085–1123, 1999.

[116] K. J. Bathe. *Finite Element Procedures*. Prentice Hall International editions. Prentice Hall International, 1996.

[117] J. P. Boyd. *Chebyshev and Fourier Spectral Methods: Second Revised Edition*. Dover Books on Mathematics. Dover Publications, 2013.

[118] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics*. Scientific Computation. Springer Berlin Heidelberg, 2007.

[119] Otto F. Sankey and David J. Niklewski. Ab initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems. *Phys. Rev. B*, 40:3979–3995, Aug 1989.

[120] Andrew P. Horsfield. Efficient ab initio tight binding. *Phys. Rev. B*, 56:6594–6602, Sep 1997.

[121] D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert, and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B*, 51:12947–12957, May 1995.

[122] Javier Junquera, Óscar Paz, Daniel Sánchez-Portal, and Emilio Artacho. Numerical atomic orbitals for linear-scaling calculations. *Phys. Rev. B*, 64:235111, 2001.

[123] Eduardo Anglada, José M. Soler, Javier Junquera, and Emilio Artacho. Systematic generation of finite-range atomic basis sets for linear-scaling calculations. *Phys. Rev. B*, 66:205101, 2002.

[124] S. E. Mousavi, J. E. Pask, and N. Sukumar. Efficient adaptive integration of functions with sharp gradients and cusps in n-dimensional parallelepipeds. *Int. J. Numer. Methods Eng.*, 91(4):343–357, 2012.

[125] Wilhelm M. Pieper. Recursive Gauss integration. *Commun. Numer. Methods Eng.*, 15(2):77–90, 1999.

[126] Jarle Berntsen, Terje O. Espelid, and Alan Genz. An adaptive algorithm for the approximate calculation of multiple integrals. *ACM Trans. Math. Softw.*, 17(4):437–451, 1991.

[127] Fuzhen Zhang. *The Schur complement and its applications*, volume 4. Springer Science & Business Media, 2006.

[128] Branislav Jansík, Stinne Høst, Poul Jørgensen, Jeppe Olsen, and Trygve Helgaker. Linear-scaling symmetric square-root decomposition of the overlap matrix. *J. Chem. Phys.*, 126(12), 2007.

[129] Anders M. N. Niklasson. Iterative refinement method for the approximate factorization of a matrix inverse. *Phys. Rev. B*, 70:193102, 2004.

[130] Nicholas J. Higham. Stable iterations for the matrix square root. *Numer. Algorithms*, 15(2):227–242, 1997.

[131] Francesco Mauri, Giulia Galli, and Roberto Car. Orbital formulation for electronic-structure calculations with linear system-size scaling. *Phys. Rev. B*, 47:9973–9976, 1993.

[132] J. M. Melenk and I. Babuška. The partition of unity finite element method: Basic theory and applications. *Comput. Methods Appl. Mech. Eng.*, 139(14):289 – 314, 1996.

[133] I. Babuška and J. M. Melenk. The partition of unity method. *Int. J. Numer. Methods Eng.*, 40(4):727–758, 1997.

[134] N. Sukumar and J. E. Pask. Classical and enriched finite element formulations for Bloch-periodic boundary conditions. *Int. J. Numer. Methods Eng.*, 77(8):1121–1138, 2009.

[135] J.E. Pask and N. Sukumar. Partition of unity finite element method for quantum mechanical materials calculations. *Extreme Mech. Lett.*, 11:8–17, 2016.

[136] C K Gan, P D Haynes, and M C Payne. Preconditioned conjugate gradient method for the sparse generalized eigenvalue problem in electronic structure calculations. *Comput. Phys. Commun.*, 134(1):33–40, 2001.

[137] Donald G. Anderson. Iterative procedures for nonlinear integral equations. *J. ACM*, 12(4):547–560, 1965.

[138] C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Math. Comput.*, 19(92):577–593, 1965.

[139] V. Eyert. A comparative study on methods for convergence acceleration of iterative vector sequences. *J. Comput. Phys.*, 124(2):271 – 285, 1996.

[140] Konstantin N. Kudin, Gustavo E. Scuseria, and Eric Cancés. A black-box self-consistent field convergence algorithm: One step closer. *J. Chem. Phys.*, 116(19):8255–8261, 2002.

[141] G. L. G. Sleijpen and H. A. Van Der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 17(2):401–425, 1996. cited By 478.

[142] G. W. Stewart. A Krylov–Schur algorithm for large eigenproblems. *SIAM J. Matrix Anal. Appl.*, 23(3):601–614, 2002.

[143] Yunkai Zhou, Yousef Saad, Murilo L. Tiago, and James R. Chelikowsky. Self-consistent-field calculations using Chebyshev-filtered subspace iteration. *J. Comput. Phys.*, 219(1):172 – 184, 2006.

[144] T. J. Rivlin. *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1990.

[145] Frank Jensen. Polarization consistent basis sets. ii. estimating the kohn-sham basis set limit. *J. Chem. Phys.*, 116(17):7372–7379, 2002.

[146] Peter Pulay. Improved SCF convergence acceleration. *J. Comput. Chem.*, 3(4):556–560, 1982.

[147] George B Bacskay. A quadratically convergent HartreeFock (QC-SCF) method. application to closed shell systems. *Chem. Phys.*, 61(3):385–404, 1981.

[148] Jinlan Wang, Guanghou Wang, and Jijun Zhao. Density-functional study of $au_n(n = 2 - 20)$ clusters: Lowest-energy structures and electronic properties. *Phys. Rev. B*, 66:035418, 2002.

[149] Phani Motamarri, Vikram Gavini, and Thomas Blesgen. Tucker-tensor algorithm for large-scale Kohn-Sham density functional theory calculations. *Phys. Rev. B*, 93:125104, 2016.

[150] H. Appel, E. K. U. Gross, and K. Burke. Excitations in time-dependent density-functional theory. *Phys. Rev. Lett.*, 90:043005, 2003.

[151] X. Gonze and J.-P. Vigneron. Density-functional approach to nonlinear-response coefficients of solids. *Phys. Rev. B*, 39:13120–13128, 1989.

[152] S. J. A. van Gisbergen, J. G. Snijders, and E. J. Baerends. Time-dependent density functional results for the dynamic hyperpolarizability of C_{60}. *Phys. Rev. Lett.*, 78:3097–3100, 1997.

[153] G. Stefanucci and C.-O. Almbladh. Time-dependent quantum transport: An exact formulation based on TDDFT. *EPL*, 67(1):14, 2004.

[154] S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio, and E. K. U. Gross. Time-dependent quantum transport: A practical scheme using density functional theory. *Phys. Rev. B*, 72:035308, 2005.

[155] Christine Jamorski Jödicke and Hans Peter Lthi. Time-dependent density functional theory (TDDFT) study of the excited charge-transfer state formation of a series of aromatic donor–acceptor systems. *J. Am. Chem. Soc.*, 125:252–264, 2003.

[156] Tamar Stein, Leeor Kronik, and Roi Baer. Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory. *J. Am. Chem. Soc.*, 131(8):2818–2820, 2009.

[157] T. Burnus, M. A. L. Marques, and E. K. U. Gross. Time-dependent electron localization function. *Phys. Rev. A*, 71:010501, 2005.

[158] Xiao-Min Tong and Shih-I Chu. Time-dependent density-functional theory for strong-field multiphoton processes: Application to the study of the role of dynamical electron correlation in multiple high-order harmonic generation. *Phys. Rev. A*, 57:452–461, 1998.

[159] Xiao-Min Tong and Shih-I Chu. Multiphoton ionization and high-order harmonic generation of He, Ne, and Ar atoms in intense pulsed laser fields: Self-interaction-free time-dependent density-functional theoretical approach. *Phys. Rev. A*, 64:013417, 2001.

[160] Dmitry A. Telnov and Shih-I Chu. Effects of electron structure and multielectron dynamical response on strong-field multiphoton ionization of diatomic molecules with arbitrary orientation: An all-electron time-dependent density-functional-theory approach. *Phys. Rev. A*, 79:041401, 2009.

[161] Mark E Casida. Time-dependent density functional response theory for molecules. In *Recent Advances In Density Functional Methods: (Part I)*, pages 155–192. World Scientific, 1995.

[162] M. Petersilka, U. J. Gossmann, and E. K. U. Gross. Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.*, 76:1212–1215, 1996.

[163] Joachim Theilhaber. Ab initio simulations of sodium using time-dependent density-functional theory. *Phys. Rev. B*, 46:12990–13003, 1992.

[164] K. Yabana and G. F. Bertsch. Time-dependent local-density approximation in real time. *Phys. Rev. B*, 54:4484–4487, 1996.

[165] Roi Baer and Recca Gould. A method for ab initio nonlinear electron-density evolution. *J. Chem. Phys.*, 114:3385–3392, 2001.

[166] G. Vignale. Center of mass and relative motion in time dependent density functional theory. *Phys. Rev. Lett.*, 74:3233–3236, 1995.

[167] Neepa T. Maitra, Kieron Burke, and Chris Woodward. Memory in time-dependent density functional theory. *Phys. Rev. Lett.*, 89:023002, 2002.

[168] Neepa T Maitra and Kieron Burke. Demonstration of initial-state dependence in time-dependent density-functional theory. *Phys. Rev. A*, 63(4):042501, 2001.

[169] Francois Gygi. Architecture of Qbox: A scalable first-principles molecular dynamics code. *IBM J. Res. Dev.*, 52(1.2):137–144, 2008.

[170] Andr Schleife, Erik W. Draeger, Yosuke Kanai, and Alfredo A. Correa. Plane-wave pseudopotential implementation of explicit integrators for time-dependent Kohn-Sham equations in large-scale simulations. *J. Chem. Phys.*, 137:22A546, 2012.

[171] José M Soler, Emilio Artacho, Julian D Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter*, 14(11):2745, 2002.

[172] Y. Takimoto, F. D. Vila, and J. J. Rehr. Real-time time-dependent density functional theory approach for frequency-dependent nonlinear optical response in photonic molecules. *J. Chem. Phys.*, 127:154114, 2007.

[173] M. Kuisma, A. Sakko, T. P. Rossi, A. H. Larsen, J. Enkovaara, L. Lehtovaara, and T. T. Rantala. Localized surface plasmon resonance in silver nanoparticles: Atomistic first-principles time-dependent density-functional theory calculations. *Phys. Rev. B*, 91:115431, 2015.

[174] Kenneth Lopata and Niranjan Govind. Modeling fast electron dynamics with real-time time-dependent density functional theory: application to small molecules and chromophores. *J. Chem. Theory Comput.*, 7(5):1344–1355, 2011.

[175] Alberto Castro, Heiko Appel, Micael Oliveira, Carlo A. Rozzi, Xavier Andrade, Florian Lorenzen, M. A. L. Marques, E. K. U. Gross, and Angel Rubio. octopus: a tool for the application of time-dependent density functional theory. *Phys. Status Solidi B*, 243(11):2465–2488, 2006.

[176] Jens Jørgen Mortensen, Lars Bruno Hansen, and Karsten Wedel Jacobsen. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B*, 71(3):035109, 2005.

[177] Michael Walter, Hannu Häkkinen, Lauri Lehtovaara, Martti Puska, Jussi Enkovaara, Carsten Rostgaard, and Jens Jørgen Mortensen. Time-dependent density-functional theory in the projector augmented-wave method. *J. Chem. Phys.*, 128(24):244101, 2008.

[178] Susanne Brenner and Ridgway Scott. *The mathematical theory of finite element methods*, volume 15. Springer, 2007.

[179] Luigi Genovese, Alexey Neelov, Stefan Goedecker, Thierry Deutsch, Seyed Alireza Ghasemi, Alexander Willand, Damien Caliste, Oded Zilberberg, Mark Rayson, Anders Bergman, et al. Daubechies wavelets as a basis set for density functional pseudopotential calculations. *J. Chem. Phys.*, 129(1):014109, 2008.

[180] Bhaarathi Natarajan, Luigi Genovese, Mark E Casida, Thierry Deutsch, Olga N Burchak, Christian Philouze, and Maxim Y Balakirev. Wavelet-based linear-response time-dependent density-functional theory. *Chem. Phys.*, 402:29–40, 2012.

[181] Huajie Chen, Xingao Gong, Lianhua He, Zhang Yang, and Aihui Zhou. Numerical analysis of finite dimensional approximations of Kohn–Sham models. *Adv. Comput. Math.*, 38(2):225–256, 2013.

[182] Huajie Chen, Xiaoying Dai, Xingao Gong, Lianhua He, and Aihui Zhou. Adaptive finite element approximations for Kohn–Sham models. *Multiscale Model. Simul.*, 12(4):1828–1869, 2014.

[183] Phani Motamarri, Sambit Das, Shiva Rudraraju, Krishnendu Ghosh, Denis Davydov, and Vikram Gavini. DFT-FE – A massively parallel adaptive finite-element code for large-scale density functional theory calculations. *Comput. Phys. Commun.*, 2019. (in press).

[184] Lauri Lehtovaara, Ville Havu, and Martti Puska. All-electron time-dependent density functional theory with finite elements: Time-propagation approach. *J. Chem. Phys.*, 135(15):154104, 2011.

[185] Gang Bao, Guanghui Hu, and Di Liu. Real-time adaptive finite element solution of time-dependent Kohn–Sham equation. *J. Comput. Phys.*, 281:743–758, 2015.

[186] Björn Hermansson and David Yevick. Finite-element approach to band-structure analysis. *Phys. Rev. B*, 33(10):7241, 1986.

[187] E. K. U. Gross and Walter Kohn. Local density-functional theory of frequency-dependent linear response. *Phys. Rev. Lett.*, 55:2850–2852, 1985.

[188] Leonard Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.*, 48:1425–1428, 1982.

[189] K.J. Bathe. *Finite Element Procedures*. Prentice Hall International editions. Prentice Hall International, 1996.

[190] Per-Olov Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.*, 18(3):365–375, 1950.

[191] Sergio Blanes, Fernando Casas, JA Oteo, and José Ros. The magnus expansion and some of its applications. *Phys. Rep.*, 470(5-6):151–238, 2009.

[192] Marlis Hochbruck and Christian Lubich. On magnus integrators for time-dependent schrdinger equations. *SIAM J. Numer. Anal.*, 41:945–963, 2003.

[193] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM, Philadelphia, 2002.

[194] Raul Radovitzky and Michael Ortiz. Error estimation and adaptive meshing in strongly nonlinear dynamic problems. *Comput. Methods Appl. Mech. Eng.*, 172(1-4):203–240, 1999.

[195] Alberto Castro, Miguel AL Marques, and Angel Rubio. Propagators for the time-dependent Kohn–Sham equations. *J. Chem. Phys.*, 121(8):3425–3433, 2004.

[196] Marlis Hochbruck, Christian Lubich, and Hubert Selhofer. Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.*, 19(5):1552–1574, 1998.

[197] Chiao-Lun Cheng, Jeremy S Evans, and Troy Van Voorhis. Simulating molecular conductance using real-time density functional theory. *Phys. Rev. B*, 74(15):155112, 2006.

[198] Yunkai Zhou, Yousef Saad, Murilo L. Tiago, and James R. Chelikowsky. Parallel self-consistent-field calculations via Chebyshev-filtered subspace acceleration. *Phys. Rev. E*, 74:066704, 2006.

[199] Y Kawashita, K Yabana, M Noda, K Nobusada, and T Nakatsukasa. Oscillator strength distribution of C60 in the time-dependent density functional theory. *J. Mol. Struct. THEOCHEM*, 914(1-3):130–135, 2009.

[200] Thomas Brabec and Ferenc Krausz. Intense few-cycle laser fields: Frontiers of nonlinear optics. *Rev. Mod. Phys.*, 72(2):545, 2000.

[201] JE Pask, N Sukumar, M Guney, and W Hu. Partition-of-unity finite-element method for large scale quantum molecular dynamics on massively parallel computational platforms. Technical report, Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 2011.

[202] JE Pask, N Sukumar, and SE Mousavi. Linear scaling solution of the all-electron coulomb problem in solids. *Int. J. Multiscale Comput. Eng.*, 10(1), 2012.

[203] Karel Peirs, Dimitri Van Neck, and Michel Waroquier. Algorithm to derive exact exchange-correlation potentials from correlated densities in atoms. *Phys. Rev. A*, 67(1):012505, 2003.

[204] Daniel S Jensen and Adam Wasserman. Numerical methods for the inverse problem of density functional theory. *Int. J. Quantum Chem.*, 118(1):e25425, 2018.

[205] Christoph R Jacob. Unambiguous optimization of effective potentials in finite basis sets. *J. Chem. Phys.*, 135(24):244102, 2011.

[206] Rogelio Cuevas-Saavedra, Paul W Ayers, and Viktor N Staroverov. Kohn–Sham exchange-correlation potentials from second-order reduced density matrices. *J. Chem. Phys.*, 143(24):244116, 2015.

[207] Felix Brockherde, Leslie Vogt, Li Li, Mark E Tuckerman, Kieron Burke, and Klaus-Robert Müller. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.*, 8(1):872, 2017.

[208] Ryo Nagai, Ryosuke Akashi, Shu Sasaki, and Shinji Tsuneyuki. Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability. *J. Chem. Phys.*, 148(24):241737, 2018.

[209] Robert van Leeuwen and Evert Jan Baerends. Energy expressions in density-functional theory using line integrals. *Phys. Rev. A*, 51(1):170, 1995.

[210] Jorge Nocedal. Updating quasi-newton matrices with limited storage. *Math. Comput.*, 35(151):773–782, 1980.

[211] Paul M Zimmerman. Incremental full configuration interaction. *J. Chem. Phys.*, 146(10):104102, 2017.

[212] E. V. R. de Castro and F. E. Jorge. Accurate universal gaussian basis set for all atoms of the periodic table. *The Journal of Chemical Physics*, 108(13):5225–5229, 1998.

[213] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90(2):1007–1023, 1989.

[214] Paul W. Ayers, Robert C. Morrison, and Robert G. Parr. Fermi-Amaldi model for exchange-correlation: atomic excitation energies from orbital energy differences. *Mol. Phys.*, 103(15-16):2061–2072, 2005.

[215] Oleg Gritsenko, Robert van Leeuwen, and Evert Jan Baerends. Analysis of electron interaction and atomic shell structure in terms of local potentials. *J. Chem. Phys.*, 101(10):8955–8963, 1994.

[216] Oleg V Gritsenko, Robert van Leeuwen, and Evert Jan Baerends. Molecular exchange-correlation Kohn–Sham potential and energy density from ab initio first-and second-order density matrices: Examples for XH (X= Li, B, F). *J. Chem. Phys.*, 104(21):8535–8545, 1996.

[217] Paul M Zimmerman. Singlet–Triplet gaps through incremental full configuration interaction. *J. Phys. Chem. A*, 121(24):712–4720, 2017.

[218] Eugene S. Kadantsev and M. J. Stott. Variational method for inverting the Kohn-Sham procedure. *Phys. Rev. A*, 69:012502, 2004.

[219] John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, and Kieron Burke. Finding density functionals with machine learning. *Phys. Rev. Lett.*, 108:253002, Jun 2012.

[220] Li Li, Thomas E. Baker, Steven R. White, and Kieron Burke. Pure density functional for strong correlation and the thermodynamic limit from machine learning. *Phys. Rev. B*, 94:245129, Dec 2016.

[221] Michael G. Medvedev, Ivan S. Bushmarinov, Jianwei Sun, John P. Perdew, and Konstantin A. Lyssenko. Density functional theory is straying from the path toward the exact functional. *Science*, 355(6320):49–52, 2017.

[222] Jianwei Sun, Adrienn Ruzsinszky, and John P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.*, 115:036402, Jul 2015.

[223] Chen Li, Xiao Zheng, Neil Qiang Su, and Weitao Yang. Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations. *Natl. Sci. Rev.*, 5(2):203–215, 09 2017.

[224] Neil Qiang Su, Chen Li, and Weitao Yang. Describing strong correlation with fractional-spin correction in density functional theory. *Proc. Natl. Acad. Sci. U.S.A*, 115(39):9678–9683, 2018.