Real-time time-dependent density functional theory using higher order finite element methods

Bikash Kanungo[a], Vikram Gavini[a,b]

[a]*Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA*
[b]*Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA*

Abstract

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. To this end, we develop an *a priori* mesh adaption technique, based on the semi-discrete (discrete in space but continuous in time) error estimate on the time-dependent Kohn-Sham orbitals, to construct a close to optimal finite-element discretization. Subsequently, we obtain the full-discrete error estimate to guide our choice of the time-step. We employ spectral finite-elements along with Gauss-Legendre-Lobatto quadrature to render the overlap matrix diagonal, thereby simplifying the inversion of the overlap matrix that features in the evaluation of the discrete time-evolution operator. We use the second-order Magnus operator as the time-evolution operator in all our calculations. Furthermore, the action of the discrete Magnus operator, expressed as exponential of a matrix, on the Kohn-Sham orbitals is obtained efficiently through an adaptive Lanczos iteration. We observe close to optimal rates of convergence of the dipole moment with respect to spatial and temporal discretization, for both pseudopotential and all-electron calculations. We demonstrate a staggering 100-fold reduction in the computational time afforded by higher-order finite-elements over linear finite-elements, for both pseudopotential and all-electron calculations. We present comparative studies, in terms of accuracy and efficiency, of our approach against finite-difference based discretization for pseudopotential calculations, and demonstrate significant computational savings when compared to the finite-difference method. We also demonstrate the competence of higher-order finite-elements for all-electron benchmark systems. Lastly, we observe good parallel scalability of the proposed method on many hundreds of processors.

Keywords: Time-dependent density functional theory; Finite-elements; Higher-order; Mesh adaption; Magnus propagator; Scalability

1. Introduction

Time-dependent density functional theory (TDDFT) 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 [1] 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* [2] 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. Although this entails, analogous to the ground-state case, the use of approximate exchange-correlation functionals, TDDFT offers a great balance of accuracy and computational efficiency to enable the study of a wide-range of time-dependent phenomena—optical [3] and higher-order responses [4, 5], electron transport [6, 7], charge-transfer excitations [8, 9], dynamics of chemical bonds [10], multi-photon ionization [11, 12, 13], to name a few.

Over the past two decades there has been a growing interest in developing faster and more accurate numerical methods for solving the TDKS equations. Broadly, these numerical methods can be classified

Email address: vikramg@umich.edu (Vikram Gavini)

into two categories, characterized by the strength of the light-matter interaction, namely, linear-response time-dependent density functional theory (LR-TDDFT) [14, 15] and real-time time-dependent density functional theory (RT-TDDFT) [16, 17, 18]. 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.

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 [19, 20, 21] and has an initial-state dependence [22]. 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 [3], and, in most cases, is the underlying approximation in existing RT-TDDFT packages. As with most of the numerical implementations in RT-TDDFT, this work is restricted to the adiabatic approximation. The second challenge 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 discretizations.

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 [23, 24]; linear combination of atomic orbitals (LCAO) in Siesta [25, 26] and GPAW [27]; Gaussian basis in NWChem [28, 29]; and finite-difference based approaches in Octopus [30] and GPAW [31, 32]. 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. Furthermore, the nonlocality of the basis substantially limits parallel scalability of their numerical implementations. 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. Additionally, the lack of basis in FD makes an accurate treatment of singular potentials difficult, thereby restricting its applicability to only pseudopotential calculations. On the other hand, the finite-element basis [33, 34], 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.

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 [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46] and all-electron calculations [35, 45, 46, 47, 48, 49, 50, 51, 52, 53]. A similarly comprehensive study on the efficacy of the finite-element basis for RT-TDDFT is, however, lacking. While two recent studies [54, 55] demonstrate the accuracy of finite-elements for RT-TDDFT, they remain restricted to only linear and quadratic

finite-elements. As has been known from prior studies in ground-state DFT [56, 48, 45], 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 [56, 57]. In particular, the works of Motamarri *et. al.* [58, 45] have demonstrated a 100- to 1000-fold improved efficiency of the higher-order finite-elements for ground-state DFT calculations over linear finite-elements. In this work, we extend the use of higher-order finite-elements to RT-TDDFT calculations and demonstrate the resulting advantages.

The keys ideas in this work can be summarized as: (i) developing an *a priori* mesh-adaption, 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 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 draws from the ideas described in [59, 60, 58, 45], wherein the optimal spatial discretization is obtained by minimizing the discretization error in the observable of importance, subject to fixed number of elements in the finite-element mesh. In particular, we proceed by obtaining a semi-discrete error estimate in the dipole moment of the system as a function of the characteristic mesh-size distribution, $h(\mathbf{r})$, and the exact time-dependent electronic fields. Next, we minimize this estimate with respect to $h(\mathbf{r})$, subject to fixed number of elements in the mesh, to obtain the optimal spatial discretization. The resulting expression for the optimal $h(\mathbf{r})$ is defined in terms of the degree of the interpolating polynomials and higher-order derivatives of the exact time-dependent electronic fields. We remark that although the exact electronic fields are *a priori* unknown, the use of asymptotic behavior of the atomic wavefunctions [61] provides a practically useful strategy to coarse-grain the mesh away from the nuclei, thereby efficiently resolving the interstitial and vacuum regions in our numerical studies. We note that the above semi-discrete analysis is also useful in decoupling the spatial discretization from the choice of time-integration scheme. Having obtained the optimal spatial discretization, the suitable temporal discretization is obtained through a full-discrete error analysis, in the context of second-order Magnus time-evolution operator. This work, to the best of our knowledge, is the first work that guides the spatio-temporal discretization for the RT-TDDFT problem using error estimates. In contrast to the aforementioned *a priori* scheme, an *a posteriori* scheme offers an alternative approach by defining the optimal spatio-temporal discretization in terms of the discrete solutions, instead of the exact solutions. However, such an approach involves error analysis based on simultaneous spatial and temporal discretization, as opposed to the semi-discrete approach described above, and thus, has to be treated differently for different time integrators. While schemes for *a posteriori* error control, in the context of nonlinear Schrödinger equation, have been developed [62, 63, 64], they all employ Galerkin schemes for time-integration, which violate the unitary requirement of the time-evolution operator. An *a posteriori* spatio-temporal adaption scheme for RT-TDDFT based on unitary time-integrators (e.g., Crank-Nicolson, Magnus) is, to the best of our knowledge, lacking, and presents a future extension of this work.

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 higher-order finite-element discretization of the RT-TDDFT problem. We remark that although optimal rates of convergence have been demonstrated in the context of nonlinear Schrödinger equation [65, 66, 67, 68, 69], a similar study for RT-TDDFT is missing. In this work, 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 employ the previously mentioned *a priori* mesh-adaption scheme to generate the meshes for all the benchmark systems. We observe numerical rates of convergence in the dipole moment close to the optimal rate of $\mathcal{O}(h^p)$, where p denotes the order of the finite-element. Next, we as-

sess 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, for other material systems. For the pseudopotential case, we use aluminum clusters (Al_2 and Al_{13}), and Buckministerfullerene (C60) molecule as the benchmark systems. The finite-difference based calculations are done using the Octopus package [30]. 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. 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 $\sim 75\%$ efficiency at 768 processors for a benchmark system of a Buckministerfullerene molecule containing 3.5 million degrees of freedom.

The rest of the paper is organized as follows. In Section 2, we briefly discuss the time-dependent Kohn-Sham equations and the form of the exact time-evolution operator. In Section 3, we derive the semi-and full-discrete error estimates, and use them for the spatio-temporal discretization. In Section 4, we describe the various numerical implementation aspects pertaining to spectral finite-elements and the discrete second-order Magnus operator. Section 5 details the convergence, efficiency and parallel scalability of the higher-order finite-elements along with its relative performance against the finite-difference. Finally, we summarize our findings and outline the future scope in Section 6.

2. Time-dependent Kohn-Sham Equations

TDDFT relies on the Runge-Gross theorem [1] and the Kohn-Sham *ansatz* [2] 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. These equations prescribe the evolution of an auxiliary system of non-interacting electrons that yield the same time-dependent electronic charge density, $\rho(\mathbf{r},t)$, as that of the interacting system. The TDKS equations, in atomic units, is given as

$$i\frac{\partial\psi_\alpha(\mathbf{r},t)}{\partial t} = H_{KS}[\rho](\mathbf{r},t;\mathbf{R})\psi_\alpha(\mathbf{r},t) \coloneqq \left[-\frac{1}{2}\nabla^2 + V_{KS}[\rho](\mathbf{r},t;\mathbf{R})\right]\psi_\alpha(\mathbf{r},t)\,, \tag{1}$$

where $H_{KS}[\rho](\mathbf{r},t;\mathbf{R})$, $V_{KS}[\rho](\mathbf{r},t;\mathbf{R})$ 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

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

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 periodic (solids) as well as spin-polarized systems.

The time-dependent Kohn-Sham potential, $V_{KS}[\rho](\mathbf{r},t;\mathbf{R})$ in Eq. 1, is given by

$$V_{KS}[\rho](\mathbf{r},t;\mathbf{R}) = V_{ext}(\mathbf{r},t;\mathbf{R}) + V_H[\rho](\mathbf{r},t) + V_{XC}[\rho](\mathbf{r},t)\,, \tag{3}$$

where $V_{ext}(\mathbf{r},t;\mathbf{R})$ denotes the external potential, $V_H[\rho](\mathbf{r},t)$ denotes the Hartree potential, and $V_{XC}[\rho](\mathbf{r},t)$ represents the exchange-correlation potential. The exchange-correlation potential, $V_{XC}[\rho](\mathbf{r},t)$, in general, is nonlocal in both space and time [19, 20, 21], and has a dependence on the initial many-electron wavefunction [22]. 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) [70], which is local in both space and time. Specifically, we use the Ceperley-Alder form [71].

In Eq. 3, the Hartree potential is given by

$$V_H[\rho](\mathbf{r},t) = \int \frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' \,. \tag{4}$$

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

$$V_N(\mathbf{r};\mathbf{R}) = \begin{cases} V_N^{ae} = -\sum\limits_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}, & \text{for all-electron calculations}\,, \\ \\ V_N^{psp}(\mathbf{R}), & \text{for pseudopotential calculations}\,, \end{cases} \tag{5}$$

where Z_I and $\mathbf{R}_I$ represent the atomic charge and position of the I^{th} nucleus. For a typical pseudopotential calculation, V_N^{psp} comprises of a local part, V_{psp}^{loc}, and a nonlocal part, V_{psp}^{nl}. For the nonlocal part, the action on a function $\phi(\mathbf{r})$, written in the Kleinman-Bylander form [72], is given by

$$V_{psp}^{nl}(\mathbf{R})\phi(\mathbf{r}) = \sum_{I=1}^{N_a}\sum_{l=0}^{L_I}\sum_{m=-l}^{l} \left(\int Y_{lm}(\mathbf{r},\mathbf{R}_I)\delta V_l^I \phi(\mathbf{r})\, d\mathbf{r} \right) Y_{lm}(\mathbf{r},\mathbf{R}_I)\,, \tag{6}$$

where l and m denote the angular and magnetic quantum number, respectively. $Y_{lm}(\mathbf{r},\mathbf{R}_I)$ denotes the spherical harmonics at $\mathbf{r}$ with the origin at $\mathbf{R}_I$, δV_l^I 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, $V_{field}(\mathbf{r},t)$, is typically provided as a monochromatic laser pulse of the form

$$V_{field}(\mathbf{r},t) = -\left(\mathbf{E_0}\cdot\mathbf{r}\right) g(t) cos(\omega t)\,, \tag{7}$$

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

We note that both the electrostatic potentials—Hartree and nuclear (all-electron)—are extended in real space. However, using the fact that the $\frac{1}{|\mathbf{r}|}$ kernel in these extended interactions is the Green's function of the Laplace operator, one can recast their evaluation as the solutions to the following Poisson equations:

$$-\frac{1}{4\pi}\nabla^2 V_H(\mathbf{r},t) = \rho(\mathbf{r},t)\,, \tag{8a}$$

$$-\frac{1}{4\pi}\nabla^2 V_N^{ae}(\mathbf{r};\mathbf{R}) = b(\mathbf{r},\mathbf{R})\,. \tag{8b}$$

In the above equation, $b(\mathbf{r};\mathbf{R}) = -\sum\limits_{I=1}^{N_a} Z_I\delta(\mathbf{r};\mathbf{R}_I)$, 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 $\int \delta(\mathbf{r};\mathbf{R}_I)\, d\mathbf{r} = 1$. We refer to previous works on finite-elements based ground-state DFT calculations [37, 39, 42, 45, 58] for a detailed treatment of the local reformulation of the electrostatic potentials into Poisson equations.

Formally, the solution to Eq. 1 can be written as

$$\psi_\alpha(\mathbf{r},T) = U(T,t_0)\psi_\alpha(\mathbf{r},t_0) = \mathcal{T}exp\left\{-i\int_0^T H_{KS}[\rho](\mathbf{r},\tau)d\tau\right\}\psi_\alpha(\mathbf{r},t_0)\,, \tag{9}$$

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,

$$U(t_2, t_0) = U(t_2, t_1)U(t_1, t_0)\,, \qquad t_0 < t_1 < t_2\,, \tag{10}$$

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

$$U(T, t_0) = \prod_{i=0}^{N-1} U(t_{i+1}, t_i)\,, \tag{11}$$

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 propagation, given by

$$\psi_\alpha(\mathbf{r}, t + \Delta t) = U(t + \Delta t)\psi_\alpha(\mathbf{r}, t) = \mathcal{T}exp\left\{-i \int_t^{t+\Delta t} H_{KS}[\rho](\mathbf{r}, \tau)d\tau\right\}\psi_\alpha(\mathbf{r}, t)\,. \tag{12}$$

In addition to resolving the implicit time-dependence in $H_{KS}[\rho](\mathbf{r}, t)$, the short time propagation provides the numerical advantage of containing the norm of the exponent in Eq. 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, $\mathcal{T}exp\left\{-i \int_t^{t+\Delta t} H_{KS}[\rho](\mathbf{r}, \tau)d\tau\right\}$. We discuss these approximations and their associated time-discretization errors in greater detail in Sec. 3.3.

3. Error estimates and *a priori* mesh adaptation

In this section we provide an *a priori* mesh adaption based on minimizing the discretization error in the Kohn-Sham orbitals. To this end, 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 $\psi_\alpha^h(\mathbf{r}, t)$ and $\psi_\alpha^{h,n}(\mathbf{r}, t)$ represent the semi-discrete (discrete in space but continuous in time) and full-discrete solution to $\psi_\alpha(\mathbf{r}, t)$, respectively, then one can use the triangle inequality to bound the discretization error in ψ_α as follows

$$||\psi_\alpha(\mathbf{r}, t) - \psi_\alpha^{h,n}(\mathbf{r}, t)|| \leq ||\psi_\alpha(\mathbf{r}, t) - \psi_\alpha^{h}(\mathbf{r}, t)|| + ||\psi_\alpha^{h}(\mathbf{r}, t) - \psi_\alpha^{h,n}(\mathbf{r}, t)|| \tag{13}$$

where $||.||$ denotes an appropriately chosen norm.

3.1. Notations, assumptions and preliminaries

Before deriving formal bounds on the two right-hand terms in the above equation, we introduce some notations, key 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 $(f, g)(t) = \int_\Omega f(\mathbf{r}, t)\bar{g}(\mathbf{r}, t)\, d\mathbf{r}$, where $\bar{g}(\mathbf{r}, t)$ 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

$$L^2(\Omega_T) = \left\{ f(\mathbf{r}, t) \,\middle|\, ||f||_{L^2(\Omega)}(t) \leq \infty, \quad \forall t \in [0, T] \right\}\,, \tag{14a}$$

$$H^1(\Omega_T) = \left\{ f(\mathbf{r},t) \,\middle|\, f, \frac{\partial f}{\partial t}, Df \in L^2(\Omega_T) \right\} , \tag{14b}$$

$$H^1_0(\Omega_T) = \left\{ f(\mathbf{r},t) \,\middle|\, f \in H^1(\Omega_T), f(\mathbf{r},t)|_{\partial\Omega} = 0, \quad \forall t \in [0,T] \right\} , \tag{14c}$$

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. 8),

$$H^1_Z(\Omega_T) = \left\{ f(\mathbf{r},t) \,\middle|\, f \in H^1(\Omega_T), f(\mathbf{r},t)|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|} \quad \forall t \in [0,T] \right\} \tag{15a}$$

$$H^1_{-Z}(\Omega) = \left\{ f(\mathbf{r}) \,\middle|\, f \in H^1(\Omega), f(\mathbf{r})|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{-Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|} \right\}. \tag{15b}$$

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 now 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, $\psi_\alpha \in H^1_0(\Omega_T) \cap L^\infty(\Omega_T)$.

$\mathcal{A}2$ The nuclear potential (in the all-electron case) due to the regularized charge distribution $b(\mathbf{r};\mathbf{R})$ (defined in Eq. 8b) is bounded, i.e., $V_N^{ae} \in L^\infty(\mathbb{R}^3)$.

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

$\mathcal{A}4$ 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., $V_{XC}[\rho], V'_{XC}[\rho] \in L^\infty(\mathbb{R}^3), \forall t \in [0,T]$.

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

$\mathcal{A}7$ 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:
$C_1 \left\|\nabla^2\phi\right\|_{L^2(\Omega)} \leq \|H_{KS}\phi\|_{L^2(\Omega)} \leq C_2 \left\|\nabla^2\phi\right\|_{L^2(\Omega)}$, $\forall \phi \in H^1_0(\Omega)$, $\forall t \in [0,T]$.

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

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 $\|V_{KS}[\rho_{\Psi_1}]\psi_{1,\alpha} - V_{KS}[\rho_{\Psi_2}]\psi_{2,\alpha}\|_{L^2(\Omega)}$ 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

$$\|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^1(\Omega)} \leq C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \,, \tag{16a}$$

$$\|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^2(\Omega)} \leq C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{H^1(\Omega)} \,. \tag{16b}$$

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

$$\left\| |\mathbf{r}|^{-1} * \rho \right\|_{L^\infty(\Omega)} \leq C \left\| |\mathbf{r}|^{-1} \right\|_{L^2(\Omega)} \|\rho\|_{L^2(\Omega)} \,, \tag{17a}$$

$$\left\| |\mathbf{r}|^{-1} * \rho \right\|_{L^2(\Omega)} \leq C \left\| |\mathbf{r}|^{-1} \right\|_{L^2(\Omega)} \|\rho\|_{L^1(\Omega)} \,, \tag{17b}$$

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

We now attempt to bound $\|V_{KS}[\rho_{\Psi_1}]\psi_{1,\alpha} - V_{KS}[\rho_{\Psi_2}]\psi_{2,\alpha}\|_{L^2(\Omega)}$, 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 (see Eq. 4), we have, for $\forall v \in H_0^1(\Omega_T)$,

$$(V_H[\rho_{\Psi_1}]\psi_{1,\alpha} - V_H[\rho_{\Psi_2}]\psi_{2,\alpha}, v) = (V_H[\rho_{\Psi_1}](\psi_{1,\alpha} - \psi_{2,\alpha}), v) + ((V_H[\rho_{\Psi_1}] - V_H[\rho_{\Psi_2}])\psi_{2,\alpha}, v) \,. \tag{18}$$

In order to bound the two terms on the right of the above equation, we make the following observations. Firstly, by definition, $V_H[\rho_{\Psi_1}] - V_H[\rho_{\Psi_2}] = V_H[\rho_{\Psi_1} - \rho_{\Psi_2}]$. Second, using $\mathcal{A}1$, we note that $\psi_{2,\alpha} \in L^\infty(\Omega_T)$. Further, as a consequence of Eq. 17a, $V_H[\rho_{\Psi_1}] \in L^\infty(\Omega_T)$. Thus, using these arguments, we have, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned} |(V_H[\rho_{\Psi_1}]\psi_{1,\alpha} - V_H[\rho_{\Psi_2}]\psi_{2,\alpha}, v)| &\leq \|V_H[\rho_{\Psi_1}]\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\quad + \|\psi_{2,\alpha}\|_{L^\infty(\Omega)} \|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^1(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq C \left(\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|\rho_{\Psi_1} - \rho_{\Psi_2}\|_{L^1(\Omega)} \|v\|_{L^2(\Omega)} \right) \,. \end{aligned} \tag{19}$$

For the first inequality in the above equation, we have used Eq. 17b in conjunction with Hölder inequality.

Next, for the exchange-correlation potential, we begin with the observation that for a given ρ_X defined by the orbitals $\{\chi_1, \chi_2, \ldots, \chi_{N_e}\}$, we have,

$$\frac{\delta(V_{XC}[\rho_X]\chi_\alpha)}{\delta\chi_\alpha} = V_{XC}[\rho_X] + 2\chi_\alpha^2 V'_{XC}[\rho_X] \,. \tag{20}$$

Further, invoking the mean-value theorem, we observe

$$V_{XC}[\rho_{\Psi_1}]\psi_{1,\alpha} - V_{XC}[\rho_{\Psi_2}]\psi_{2,\alpha} = (V_{XC}[\rho_X] + 2\chi_\alpha^2 V'_{XC}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}) \,, \tag{21}$$

for some $\chi_\alpha = \lambda_\alpha \psi_{1,\alpha} + (1-\lambda_\alpha)\psi_{2,\alpha}$, with $\lambda_\alpha \in [0,1]$. Thus, using the above relation, we have, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned} |(V_{XC}[\rho_{\Psi_1}]\psi_{1,\alpha} - V_{XC}[\rho_{\Psi_2}]\psi_{2,\alpha}, v)| &= \left| \left((V_{XC}[\rho_X] + 2\chi_\alpha^2 V'_{XC}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}), v \right) \right| \\ &\leq \left\| (V_{XC}[\rho_X] + 2\chi_\alpha^2 V'_{XC}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}) \right\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq \left\| V_{XC}[\rho_X] + 2\chi_\alpha^2 V'_{XC}[\rho_X] \right\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,, \end{aligned} \tag{22}$$

where we have used the Cauchy-Schwarz inequality in the second line, the Hölder inequality in the third line, and have invoked the boundedness assumption on V_{XC} and V'_{XC} (assumption $\mathcal{A}5$) in the last inequality.

For the nuclear potential in the all-electron case, we have, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned}
|(V_N^{ae}\psi_{1,\alpha} - V_N^{ae}\psi_{2,\alpha}, v)| &\leq \|V_N^{ae}(\psi_{1,\alpha} - \psi_{2,\alpha})\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq \|V_N^{ae}\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,,
\end{aligned} \tag{23}$$

where we have used Cauchy-Schwarz and Hölder inequalities in the first and second line, respectively. In the last line, we have invoked the boundedness assumption on V_N^{ae} (assumption $\mathcal{A}2$).

For the pseudopotential case, we have, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned}
|(V_N^{psp}\psi_{1,\alpha} - V_N^{psp}\psi_{2,\alpha}, v)| &\leq \left|\left(V_{psp}^{loc}\left(\psi_{1,\alpha} - \psi_{2,\alpha}\right), v\right)\right| + \left|\left(V_{psp}^{nl}\psi_{1,\alpha} - V_{psp}^{nl}\psi_{2,\alpha}, v\right)\right| \\
&\leq \left\|V_{psp}^{loc}\right\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\quad + \left|\sum_{I=1}^{N_a}\sum_{l=0}^{L_\alpha}\sum_{m=-l}^{l}\left(Y_{lm}(\mathbf{r},\mathbf{R}_\alpha)\delta V_l^I, \psi_{1,\alpha} - \psi_{2,\alpha}\right)\left(Y_{lm}(\mathbf{r},\mathbf{R}_\alpha), v\right)\right| \\
&\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\quad + \sum_{I=1}^{N_a}\sum_{l=0}^{L_\alpha}\sum_{m=-l}^{l}\left\|Y_{lm}(\mathbf{r},\mathbf{R}_\alpha)\delta V_l^I\right\|_{L^2(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|Y_{lm}(\mathbf{r},\mathbf{R}_\alpha)\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\quad + \sum_{I=1}^{N_a}\sum_{l=0}^{L_\alpha}\sum_{m=-l}^{l}\left\|\delta V_l^I\right\|_{L^\infty(\Omega)} \|Y_{lm}(\mathbf{r},\mathbf{R}_\alpha)\|_{L^2(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|Y_{lm}(\mathbf{r},\mathbf{R}_\alpha)\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq C' \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,,
\end{aligned} \tag{24}$$

where we have, once again, made use of the Cauchy-Schwarz and Hölder inequality along with the boundedness assumptions on V_{psp}^{loc} and δV_l^I (assumptions $\mathcal{A}3$ and $\mathcal{A}4$, respectively).

Lastly, invoking the boundedness assumption on V_{field} (assumption $\mathcal{A}6$), we get, $\forall v \in H_0^1(\Omega_T)$,

$$\begin{aligned}
|(V_{field}\psi_{1,\alpha} - V_{field}\psi_{2,\alpha}, v)| &\leq \|V_{field}\left(\psi_{1,\alpha} - \psi_{2,\alpha}\right)\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq \|V_{field}\|_{L^\infty(\Omega)} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq C \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \,.
\end{aligned} \tag{25}$$

We now define the weak solution of the TDKS equation (Eq. 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

$$i\left(\frac{\partial\psi_\alpha}{\partial t}, v\right) = \frac{1}{2}\left(\nabla\psi_\alpha, \nabla v\right) + \left(V_{KS}[\rho]\psi_\alpha, v\right), \quad \forall v \in H_0^1(\Omega_T)\,, \text{ and } \forall t \in [0,T]\,. \tag{26}$$

Similarly, the weak solutions to the Poisson problems defined in Eq. 8 are defined to be $V_H(\mathbf{r},t) \in H_Z^1(\Omega_T)$, and $V_N^{ae}(\mathbf{r},\mathbf{R}) \in H_{-Z}^1(\Omega)$, satisfying,

$$\left(\nabla V_H, \nabla v\right) = 4\pi(\rho, v), \quad \forall v \in H_0^1(\Omega_T), \text{ and } \forall t \in [0,T] \tag{27a}$$

$$\left(\nabla V_N^{ae}, \nabla v\right) = 4\pi\left(b, v\right), \quad \forall v \in H_0^1(\Omega)\,. \tag{27b}$$

3.2. Semi-discrete error estimate

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. 26 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

$$i\left(\frac{\partial \psi_\alpha^h}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla\psi_\alpha^h, \nabla v^h\right) + \left(V_{KS}^h[\rho^h]\psi_\alpha^h, v^h\right), \quad \forall v^h \in X_0^{h,p} \times [0,T], \text{ and } \forall t \in [0,T], \tag{28}$$

where $\rho^h(\mathbf{r},t) = \sum_{\alpha=1}^{N_e} \left|\psi_\alpha^h(\mathbf{r},t)\right|^2$ and $V_{KS}^h[\rho^h](\mathbf{r},t) = V_H^h[\rho^h](\mathbf{r},t) + V_N^h(\mathbf{r};\mathbf{R}) + V_{XC}[\rho^h](\mathbf{r},t) + V_{field}(\mathbf{r},t)$.

We now elaborate on the different terms appearing in the expression for $V_{KS}^h[\rho^h](\mathbf{r},t)$. To begin with, $V_H^h[\rho^h](\mathbf{r},t) \in X^{h,p} \times [0,T] \cap H_Z^1(\Omega_T)$ represents the discrete counterpart of the weak solution described in Eq. 27a, and is obtained as the solution to

$$(\nabla V_H^h, \nabla v^h) = 4\pi(\rho^h, v^h), \quad \forall v^h \in X_0^{h,p} \times [0,T], \text{and } \forall t \in [0,T]. \tag{29}$$

Similarly, $V_N^h(\mathbf{r};\mathbf{R})$, for the all-electron case, represents the discrete analog of the weak solution defined in Eq. 27b, and is given by $V_N^{ae,h}(\mathbf{r};\mathbf{R}) \in X^{h,p} \cap H_{-Z}^1(\Omega)$ such that

$$(\nabla V_N^{ae,h}, \nabla v^h) = 4\pi(b, v^h), \quad \forall v^h \in X_0^{h,p}. \tag{30}$$

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

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,

$$\left(\nabla(\psi - \mathcal{P}_h\psi), \nabla v^h\right) = 0, \quad \forall \psi \in H_0^1(\Omega_T), \forall v^h \in X_0^{h,p} \times [0,T], \text{ and } \forall t \in [0,T]. \tag{31}$$

This allows us to use some standard finite-element error estimates (see [73]) in order to bound $\|\psi - \mathcal{P}_h\psi\|_{L^2(\Omega)}$. In the following error-estimates we take a triangulation $\mathcal{T}^{h,p}$ of p^{th} order finite-elements covering the domain Ω [73].

Proposition 3.1. *Assuming uniqueness and existence of the solution to Eqs. 26 and 28, we obtain the following bounds on the finite-element semi-discrete approximation error to the Kohn-Sham orbitals:*

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq C_1 e^{C_2 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)$$
$$+ C_1 e^{C_2 t} t \sum_e h_e^{p+1} \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + e^{C_2 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0), \tag{32a}$$

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq C_3 e^{C_2 t} t \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right)$$
$$+ C_3 e^{C_2 t} t \sum_e h_e^p \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right)$$
$$+ C_4 e^{C_2 t} h_{min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0), \tag{32b}$$

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

$$s_{1,\alpha} = \underset{0\leq s\leq t}{\arg\max} \left\|\psi_\alpha - \mathcal{P}_h\psi_\alpha\right\|_{L^2(\Omega)}(s), \tag{33a}$$

$$s_{2,\alpha} = \underset{0\leq s\leq t}{\arg\max} \left\| \frac{\partial\psi_\alpha}{\partial t} - \mathcal{P}_h \frac{\partial\psi_\alpha}{\partial t} \right\|_{L^2(\Omega)}(s), \tag{33b}$$

$$s_3 = \underset{0\leq s\leq t}{\arg\max} \left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}(s). \tag{33c}$$

Proof. Taking $v = v^h \in X_0^{h,p} \times [0,T]$ in Eq. 26 and subtracting it from Eq. 28, we get

$$i\left(\frac{\partial\left(\psi_\alpha - \psi_\alpha^h\right)}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla\left(\psi_\alpha - \psi_\alpha^h\right), \nabla v^h\right) + \left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, v^h\right), \forall v^h \in X_0^{h,p} \times [0,T]. \tag{34}$$

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. 34 as

$$i\left(\frac{\partial w_\alpha}{\partial t}, v^h\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, v^h\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla v^h\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla v^h\right) + \left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, v^h\right). \tag{35}$$

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

$$i\left(\frac{\partial w_\alpha}{\partial t}, w_\alpha\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla w_\alpha\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla w_\alpha\right) + \left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right). \tag{36}$$

Noting that

$$\frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|_{L^2(\Omega)}^2 = \mathrm{Re}\left\{\left(\frac{\partial}{\partial t}w_\alpha, w_\alpha\right)\right\}, \tag{37}$$

and comparing the imaginary parts of Eq. 36, we have

$$\frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|_{L^2(\Omega)}^2 = -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla u_\alpha, \nabla w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla w_\alpha, \nabla w_\alpha\right)\right\}$$
$$+ \mathrm{Im}\left\{\left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right)\right\} \tag{38}$$

In the above equation, we note that $(\nabla u_\alpha, \nabla w_\alpha) = 0$, as a consequence of Eq. 31. Furthermore, $(\nabla w_\alpha, \nabla w_\alpha)$ is real. Thus, Eq. 38 simplifies to,

$$\frac{1}{2}\frac{d}{dt}\left\|w_\alpha\right\|_{L^2(\Omega)}^2 = -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \mathrm{Im}\left\{\left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right)\right\}$$
$$\leq \left|\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right| + \left|\left(V_{KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right)\right|. \tag{39}$$

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

$$\begin{aligned}\left(V_{KS}[\rho]\psi_\alpha - V^h_{KS}[\rho^h]\psi^h_\alpha, w_\alpha\right) &= \left(V_{XC}[\rho]\psi_\alpha - V_{XC}[\rho^h]\psi^h_\alpha, w_\alpha\right) + \left(V_H[\rho]\psi_\alpha - V^h_H[\rho^h]\psi^h_\alpha, w_\alpha\right) \\ &\quad + \left(V_N\psi_\alpha - V^h_N\psi^h_\alpha, w_\alpha\right) + \left(V_{field}\psi_\alpha - V_{field}\psi^h_\alpha, w_\alpha\right) \\ &= \left(V_{XC}[\rho]\psi_\alpha - V_{XC}[\rho^h]\psi^h_\alpha, w_\alpha\right) + \left(V_H[\rho]\psi_\alpha - V_H[\rho^h]\psi^h_\alpha, w_\alpha\right) \\ &\quad + \left(\left(V_H[\rho^h] - V^h_H[\rho^h]\right)\psi^h_\alpha, w_\alpha\right) + \left(V_N\psi_\alpha - V_N\psi^h_\alpha, w_\alpha\right) \\ &\quad + \left(V_N\psi^h_\alpha - V^h_N\psi^h_\alpha, w_\alpha\right) + \left(V_{field}\psi_\alpha - V_{field}\psi^h_\alpha, w_\alpha\right).\end{aligned} \tag{40}$$

We note that the term $\left(V_N\psi^h_\alpha - V^h_N\psi^h_\alpha, w_\alpha\right)$, on the right side of the above equation, is relevant only in the all-electron case (i.e., zero for the pseudopotential case as $V_N = V^h_N$). Combining the results from Eqs. 19, 22, 23, 24, and 25, with $v = w_\alpha$, and using the fact that $\psi^h \in L^\infty(\Omega)$, it is straightforward to show that

$$\begin{aligned}\left|\left(V_{KS}[\rho]\psi_\alpha - V^h_{KS}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right| &\leq C_0 \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)} \\ &\quad + C_1 \left(\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + \left\|\rho - \rho^h\right\|_{L^1(\Omega)}\right) \|w_\alpha\|_{L^2(\Omega)} \\ &\quad + C_2 \left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)} \\ &\quad + C_3 \left\|V_N - V^h_N\right\|_{L^2(\Omega)} \|w_\alpha\|_{L^2(\Omega)}\ .\end{aligned} \tag{41}$$

Using the above result in Eq. 39, we obtain

$$\begin{aligned}\frac{d}{dt}\|w_\alpha\|_{L^2(\Omega)} &\leq \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0 \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + C_1 \left(\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + \left\|\rho - \rho^h\right\|_{L^1(\Omega)}\right) \\ &\quad + C_2 \left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_3 \left\|V_N - V^h_N\right\|_{L^2(\Omega)} \\ &\leq \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0 \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} \\ &\quad + C_2 \left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_3 \left\|V_N - V^h_N\right\|_{L^2(\Omega)} + C_4 \sum_{\beta=1}^{N_e} \left\|\psi_\beta - \psi^h_\beta\right\|_{L^2(\Omega)}\ ,\end{aligned} \tag{42}$$

where we have used Eq. 16a in the second line to simplify the term involving $\left\|\rho - \rho^h\right\|_{L^1(\Omega)}$. Thus, summing the above equation over all index α, we have

$$\begin{aligned}\frac{d}{dt}\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)} &\leq \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5 \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}\right) \\ &\quad + C_6 \left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_7 \left\|V_N - V^h_N\right\|_{L^2(\Omega)} \\ &\leq \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5 \|u_\alpha\|_{L^2(\Omega)} + C_5 \|w_\alpha\|_{L^2(\Omega)}\right) \\ &\quad + C_6 \left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_7 \left\|V_N - V^h_N\right\|_{L^2(\Omega)}\ ,\end{aligned} \tag{43}$$

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

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(0) + C_5 \int_0^t \sum_{\alpha=1}^{N_e} \|w_\alpha\|(s)\, ds \\ + C_5 \int_0^t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)}(s) + \|u_\alpha\|_{L^2(\Omega)}(s) \right) ds \\ + C_8 \int_0^t \left(\left\| V_H[\rho^h] - V_H^h[\rho^h] \right\|_{L^2(\Omega)}(s) + \left\| V_N - V_N^h \right\|_{L^2(\Omega)} \right) ds\,.\end{aligned} \tag{44}$$

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 (see Eq. 33), we can simplify the above equation as

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(0) + C_5 \int_0^t \sum_{\alpha=1}^{N_e} \|w_\alpha\|(s)\, ds \\ + C_5 t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha}) \right) \\ + C_8 t \left(\left\| V_H[\rho^h] - V_H^h[\rho^h] \right\|_{L^2(\Omega)}(s_3) + \left\| V_N - V_N^h \right\|_{L^2(\Omega)} \right).\end{aligned} \tag{45}$$

Invoking the Grönwall's inequality on the above equation yields

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq e^{C_5 t} \left[\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(0) + C_5 t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha}) \right) \right] \\ + C_8 e^{C_5 t} t \left(\left\| V_H[\rho^h] - V_H^h[\rho^h] \right\|_{L^2(\Omega)}(s_3) + \left\| V_N - V_N^h \right\|_{L^2(\Omega)} \right).\end{aligned} \tag{46}$$

Expressing $w_\alpha(0) = \left(\psi_\alpha(0) - \psi_\alpha^h(0)\right) - u_\alpha(0)$, we bound $w_\alpha(0)$ as

$$\|w_\alpha\|_{L^2(\Omega)}(0) \leq \left\| \psi_\alpha - \psi_\alpha^h \right\|_{L^2(\Omega)}(0) + \|u_\alpha\|_{L^2(\Omega)}(0)\,. \tag{47}$$

Plugging the above relation into Eq. 46, we obtain

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq C_5' e^{C_5 t} t \sum_{\alpha=1}^{N_e} \left(\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha}) \right) \\ + C_8 e^{C_5 t} t \left(\left\| V_H[\rho^h] - V_H^h[\rho^h] \right\|_{L^2(\Omega)}(s_3) + \left\| V_N - V_N^h \right\|_{L^2(\Omega)} \right) + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\| \psi_\alpha - \psi_\alpha^h \right\|_{L^2(\Omega)}(0)\,.\end{aligned} \tag{48}$$

We now introduce formal bounds on the terms involving $\|u_\alpha\|_{L^2(\Omega)}$, $\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}$, $\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}$, and $\left\|V_N - V_N^h\right\|_{L^2(\Omega)}$, by using standard finite-element error estimates. To begin with, we use the definition of the Ritz projection along with Ceá's lemma [73] to note that

$$\|u_\alpha\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} |\psi_\alpha|_{p+1,\Omega_e}\,. \tag{49}$$

For the $\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}$ term, we obtain, by taking the time-derivative of Eq. 31,

$$\left(\nabla \left(\frac{\partial \psi_\alpha}{\partial t} - \mathcal{P}_h \frac{\partial \psi_\alpha}{\partial t} \right), \nabla v^h \right) = \left(\nabla \left(\frac{\partial u_\alpha}{\partial t} \right), \nabla v^h \right) = 0, \quad \forall v^h \in X_0^{h,p}\,. \tag{50}$$

We emphasize that, in arriving at the above equation, we choose v^h that is time-independent. The above equation allows us to invoke the Ceá's lemma (see [73]) on $\frac{\partial u_\alpha}{\partial t}$, to obtain

$$\left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left| \frac{\partial \psi_\alpha}{\partial t} \right|_{p+1,\Omega_e} . \tag{51}$$

Furthermore, using Eq. 1, we have,

$$\begin{aligned} \left\| \frac{\partial u_\alpha}{\partial t} \right\|_{L^2(\Omega)} &\leq C \sum_e h_e^{p+1} \left| \frac{\partial \psi_\alpha}{\partial t} \right|_{p+1,\Omega_e} \\ &\leq C \sum_e h_e^{p+1} \left| -\frac{1}{2}\nabla^2 \psi_\alpha + V_{KS}[\rho]\psi_\alpha \right|_{p+1,\Omega_e} \\ &\leq C \sum_e h_e^{p+1} \left(|\psi_\alpha|_{p+3,\Omega_e} + |(V_H + V_N + V_{XC} + V_{field})\psi_\alpha|_{p+1,\Omega_e} \right) \\ &\leq C \sum_e h_e^{p+1} \left(|\psi_\alpha|_{p+3,\Omega_e} + |\psi_\alpha|_{p+1,\Omega_e} \right) , \end{aligned} \tag{52}$$

where we have used the definition of the $|.|_{p+3}$ semi-norm in the third line and the boundedness assumptions on V_N, V_{XC}, and V_{field} (assumptions $\mathcal{A}2$–$\mathcal{A}6$) in the last line. Lastly, for the $\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}$, and $\left\|V_N - V_N^h\right\|_{L^2(\Omega)}$ terms, we make use of standard finite-element error estimates for the Poisson problem to get

$$\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left|V_H[\rho^h]\right|_{p+1,\Omega_e} . \tag{53a}$$

$$\left\|V_N - V_N^h\right\|_{L^2(\Omega)} \leq C \sum_e h_e^{p+1} \left|V_N\right|_{p+1,\Omega_e} . \tag{53b}$$

Thus, using Eqs. 49, 52, and 53 in Eq. 48, we have

$$\begin{aligned} \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t) \leq\ & C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\ & + C_9 e^{C_5 t} t \sum_e h_e^{p+1} \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right) + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) . \end{aligned} \tag{54}$$

Finally, expressing $\psi_\alpha - \psi_\alpha^h = w_\alpha + u_\alpha$ and combining the result of the above equation with Eq. 49, we obtain

$$\begin{aligned} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq\ & C_{10} e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) \\ & + C_{10} e^{C_5 t} t \sum_e h_e^{p+1} \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right) + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) . \end{aligned} \tag{55}$$

This concludes the proof of Eq. 32a.

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 [33] for $w_\alpha = (\mathcal{P}_h\psi_\alpha - \psi_\alpha^h) \in X_0^{h,p}$ to obtain

$$\|w_\alpha\|_{H^1(\Omega_e)}(t) \leq C h_e^{-1} \|w_\alpha\|_{L^2(\Omega_e)}(t) . \tag{56}$$

Squaring the above the inequality and summing over all the elements e, it is straightforward to obtain

$$\|w_\alpha\|^2_{H^1(\Omega)}(t) \coloneqq \sum_e \|w_\alpha\|^2_{H^1(\Omega_e)}(t) \leq Ch^{-2}_{min} \sum_e \|w_\alpha\|^2_{L^2(\Omega_e)}(t) \coloneqq Ch^{-2}_{min} \|w_\alpha\|^2_{L^2(\Omega)}(t). \tag{57}$$

Thus, taking the square-root on the above inequality, we have

$$\|w_\alpha\|_{H^1(\Omega)}(t) \leq Ch^{-1}_{min} \|w_\alpha\|_{L^2(\Omega)}(t). \tag{58}$$

Additionally, we once again invoke the Ceá's lemma [33] on $u_\alpha = (\psi_\alpha - \mathcal{P}_h\psi_\alpha)$ to note that

$$\|u_\alpha\|_{H^1(\Omega)}(t) \leq C \sum_e h^p_e |\psi_\alpha|_{k+1,\Omega_e}(t). \tag{59}$$

Thus, using Eqs. 58 and 59, we get

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi^h_\alpha\right\|_{H^1(\Omega)}(t) &\leq \sum_{\alpha=1}^{N_e} \left(\|u_\alpha\|_{H^1(\Omega)}(t) + \|w_\alpha\|_{H^1(\Omega)}(t)\right) \\ &\leq C_{11} \sum_e \sum_{\alpha=1}^{N_e} h^p_e |\psi_\alpha|_{k+1,\Omega_e}(t) + C_{12} h^{-1}_{min} \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t).\end{aligned} \tag{60}$$

Further, using the inequality obtained in Eq. 54 in the above equation, yields

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi^h_\alpha\right\|_{H^1(\Omega)}(t) \leq\ & C_{11} \sum_e \sum_{\alpha=1}^{N_e} h^p_e |\psi_\alpha|_{k+1,\Omega_e}(t) \\ &+ C_{13} e^{C_5 t} t \sum_e h^{-1}_{min} h^{p+1}_e \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_{13} e^{C_5 t} t \sum_e h^{-1}_{min} h^{p+1}_e \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) \\ &+ C_{12} e^{C_5 t} h^{-1}_{min} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}(0).\end{aligned} \tag{61}$$

Using the fact that $h_e/h_{min} \leq C$ for all the elements in $\mathcal{T}^{h,p}$, we can simplify the above equation as

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi^h_\alpha\right\|_{H^1(\Omega)}(t) \leq\ & C_{14} e^{C_5 t} t \sum_e h^p_e \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_{14} e^{C_5 t} t \sum_e h^p_e \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) \\ &+ C_{12} e^{C_5 t} h^{-1}_{min} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}(0).\end{aligned} \tag{62}$$

This concludes the proof for Eq. 32b. □

Corollary 3.1.1. *If the initial orbitals $\psi_\alpha(\boldsymbol{r}, 0)$ are obtained from a ground-state DFT calculation, wherein [45]*

$$\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}(0) \leq C \sum_e h^{p+1}_e \left(|\psi_\alpha|_{p+1,\Omega_e} + \left|V_H[\rho^h]\right|_{p+1,\Omega_e} + |V_N|_{p+1,\Omega_e}\right), \tag{63}$$

the results of Proposition 3.1 can be simplified, $\forall t \in [0,T]$, to

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leq C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) + C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right) . \tag{64a}$$

$$\sum_{\alpha}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leq C_3' e^{C_2 t}(t+1) \sum_e h_e^{p} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right) + C_3' e^{C_2 t}(t+1) \sum_e h_e^{p} \left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right) \tag{64b}$$

3.3. Full discrete error estimate

We now derive the full-discrete error estimate in the context of the second-order Magnus time-evolution operator. To begin with, we represent the finite-element approximation to the Kohn-Sham orbitals and the Hartree potential as

$$\psi_\alpha^h(\mathbf{r},t) = \sum_{j=1}^{n_h} N_j(\mathbf{r}) \psi_\alpha^j(t) , \tag{65}$$

$$V_H^h(\mathbf{r},t) = \sum_{j=1}^{n_h} N_j(\mathbf{r}) \phi^j(t) , \tag{66}$$

where $N_j : 1 \leq j \leq n_h$ denote a basis for $X^{h,p}$. Subsequently, the weak formulation in Eq. 28 can be written in the form of the following matrix vector equation

$$i\mathbf{M}\dot{\boldsymbol{\psi}}_\alpha(t) = \mathbf{H}\psi_\alpha(t) \coloneqq (\mathbf{U}+\mathbf{V})\,\boldsymbol{\psi}_\alpha(t), \tag{67}$$

where H_{jk} and M_{jk} denote the discrete Hamiltonian and overlap matrix, respectively, and $\boldsymbol{\psi}_\alpha(t)$ denotes the vector containing the coefficients $\psi_\alpha^j(t)$. The Hamiltonian matrix is, in turn, split into two parts, U_{jk} and V_{jk}, given by

$$\mathrm{U}_{jk} = \frac{1}{2}\int_\Omega \nabla N_j(\mathbf{r}) \cdot \nabla N_k(\mathbf{r})\, d\mathbf{r} , \tag{68}$$

$$\mathrm{V}_{jk} = \int_\Omega V_{KS}^h[\rho^h(t)] N_j(\mathbf{r}) N_k(\mathbf{r})\, d\mathbf{r} . \tag{69}$$

We note that the matrix $\mathbf{M}$, owing to its positive definiteness, guarantees the existence of a unique positive definite symmetric square root, denoted by $\mathbf{M}^{1/2}$. This allows us to rewrite Eq. 67 as

$$i\dot{\bar{\boldsymbol{\psi}}}_\alpha(t) = \bar{\mathbf{H}}\bar{\boldsymbol{\psi}}_\alpha(t) \coloneqq \left(\bar{\mathbf{U}}+\bar{\mathbf{V}}\right)\bar{\boldsymbol{\psi}}_\alpha(t) , \tag{70}$$

where $\bar{\boldsymbol{\psi}}_\alpha(t) = \mathbf{M}^{1/2}\boldsymbol{\psi}_\alpha(t)$ and $\{\bar{\mathbf{H}},\bar{\mathbf{U}},\bar{\mathbf{V}}\} = \mathbf{M}^{-1/2}\{\mathbf{H},\mathbf{U},\mathbf{V}\}\mathbf{M}^{-1/2}$. In the following discussion, we drop the superscript h in ρ, for the sake of brevity.

We now invoke the Magnus *ansatz* to write the solution of Eq. 70 as

$$\bar{\boldsymbol{\psi}}_\alpha(t) = \exp(\mathbf{A}(t))\bar{\boldsymbol{\psi}}_\alpha(0) , \qquad \forall t \geq 0 . \tag{71}$$

The operator $\exp(\mathbf{A}(t))$ is termed as the *Magnus propagator*. Differentiating the above equation and com-

paring with Eq. 70 gives [74]

$$\dot{\bar{\boldsymbol{\psi}}}_\alpha(t) = -i\bar{\mathbf{H}}(t)\bar{\boldsymbol{\psi}}_\alpha(t) = \mathrm{dexp}_{\mathbf{A}(t)}\left(\dot{\mathbf{A}}(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\,, \tag{72}$$

where $\mathrm{dexp}_{\mathbf{X}}(\mathbf{Y}) = \sum_0^\infty \frac{1}{(k+1)!}\mathrm{ad}^k_{\mathbf{X}}(\mathbf{Y})$. The operator $\mathrm{ad}^k_{\mathbf{X}}(\mathbf{Y})$ is defined recursively as

$$\mathrm{ad}^k_{\mathbf{X}}(\mathbf{Y}) = \mathrm{ad}_{\mathbf{X}}\left(\mathrm{ad}^{k-1}_{\mathbf{X}}(\mathbf{Y})\right)\,, \tag{73}$$

with $\mathrm{ad}^1_{\mathbf{X}}(\mathbf{Y}) = \mathbf{XY} - \mathbf{YX}$, and $\mathrm{ad}^0_{\mathbf{X}}(\mathbf{Y}) = \mathbf{Y}$. Now, inverting the operator on the right side of the above equation provides the following differential equation for $\mathbf{A}(t)$ [74]:

$$\dot{\mathbf{A}}(t) = \sum_{k\geq 0}\frac{\beta_k}{k!}\mathrm{ad}^k_{\mathbf{A}(t)}(-i\bar{\mathbf{H}}(t))\,, \tag{74}$$

where β_k is given by the series $\frac{e^z-1}{z} = \sum_0^\infty(\beta_k/k!)z^k$. Finally, invoking the Picard iteration on the above equation yields the following explicit form of the Magnus expansion [75],

$$\mathbf{A}(t) = \int_0^t -i\bar{\mathbf{H}}(\tau)d\tau - \frac{1}{2}\int_0^t\left[\int_0^\tau -i\bar{\mathbf{H}}(\sigma)d\sigma, -i\bar{\mathbf{H}}(\tau)\right]d\tau + \ldots \quad . \tag{75}$$

From the above expression, we note that $\mathbf{A}$ is skew-Hermitian, owing to the Hermiticity of $\bar{\mathbf{H}}$. As mentioned in Section 2, we resolve the implicit dependence of $\bar{\mathbf{H}}(t)$ on $\rho(\mathbf{r},t)$ by using the composition property of a propagator (see Eq. 11). Thus, the Magnus propagator can be rewritten as

$$\exp(\mathbf{A}(t)) = \prod_{n=1}^{N}\exp(\mathbf{A}_n)\,, \tag{76}$$

where $\mathbf{A}_n$ is given by Eq. 75 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, and second, an approximation for the time integrals in the truncated Magnus expansion. Truncating the Magnus expansion (defined in Eq. 75) after the first p terms results in a time-integration scheme of order $2p$. In this work, for simplicity, we restrict ourselves to the second-order Magnus propagator, i.e., obtained by truncating the Magnus expansion after the first term.

We now express the time-discrete approximation to $\bar{\boldsymbol{\psi}}_\alpha(t_{n+1})$, denoted by $\bar{\boldsymbol{\psi}}^{n+1}_\alpha$, as

$$\bar{\boldsymbol{\psi}}^{n+1}_\alpha = \exp(\tilde{\mathbf{A}}_n)\bar{\boldsymbol{\psi}}^{n}_\alpha. \tag{77}$$

In particular, the action of the second-order Magnus propagator, with a mid-point integration rule, on a set of Kohn-Sham orbitals $\boldsymbol{\Phi} = \{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_{N_e}\}$ which defines the density $\rho_{\boldsymbol{\Phi}}$, is given by

$$e^{\tilde{\mathbf{A}}_n}\boldsymbol{\phi}_\alpha(t) = e^{-i\bar{\mathbf{H}}\left[\rho\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]\Delta t}\boldsymbol{\phi}_\alpha(t)\,, \tag{78}$$

where $\Delta t = t_n - t_{n-1}$ and $\bar{\mathbf{H}}\left[\rho\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]$ is the time-continuous Kohn-Sham Hamiltonian described by the Kohn-Sham orbitals $\boldsymbol{\Phi}$ at the future time instance $t_{n-1}+\Delta t/2$. We remark that the operator $e^{-i\bar{\mathbf{H}}\left[\rho\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]\Delta t}$ results from two successive approximations: a) truncation of the Magnus expansion after the first term; and b) from the use of the mid-point integration to evaluate the time integral. Furthermore, $\bar{\mathbf{H}}\left[\rho\left(t_{n-1}+\frac{\Delta t}{2}\right)\right]$ 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.

In the following error-estimates, we assume each time interval $[t_{n-1}, t_n]$ to be of length Δt. Moreover, for simple terminology, we term $e^{\tilde{\mathbf{A}}_n}$ as the second-order Magnus propagator without explicitly spelling out the mid-point integration rule invoked in it.

Proposition 3.2. *For a second-order Magnus propagator with a mid-point integration rule, we obtain the following bound for the time-discretization error to the vectors* $\bar{\boldsymbol{\psi}}_\alpha$

$$\left\|\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n\right\| \leq C(\Delta t)^2 t_n \max_{0\leq t\leq t_n} \left\|\bar{\boldsymbol{\psi}}_\alpha(t)\right\|_{\bar{\boldsymbol{U}}}, \tag{79}$$

where $\|\mathbf{x}\|_{\bar{\boldsymbol{U}}} = \sqrt{\mathbf{x}^\dagger \bar{\boldsymbol{U}}\mathbf{x}}$.

Proof. To begin with, we introduce the following operators

$$\mathbf{S}_0^k = e^{\mathbf{A}_k} e^{\mathbf{A}_{k-1}} \dots e^{\mathbf{A}_1} = \prod_{l=0}^{k-1} e^{\mathbf{A}_{k-l}} \quad \text{for } 0 < k \leq n\,, \quad \mathbf{S}_0^0 = I \tag{80a}$$

$$\mathbf{R}_k^n = e^{\tilde{\mathbf{A}}_n} e^{\tilde{\mathbf{A}}_{n-1}} \dots e^{\tilde{\mathbf{A}}_{k+1}} = \prod_{l=0}^{n-k-1} e^{\mathbf{A}_{n-l}} \quad \text{for } 0 \leq k < n\,, \quad \mathbf{R}_n^n = I\,. \tag{80b}$$

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. We now rewrite the time-discretization error in $\bar{\boldsymbol{\psi}}_\alpha$ at t_n using the following telescopic series,

$$\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n = \left(\mathbf{R}_n^n \mathbf{S}_0^n - \mathbf{R}_0^n \mathbf{S}_0^0\right) \bar{\boldsymbol{\psi}}_\alpha(0) = \sum_{k=1}^{n} \left(\mathbf{R}_k^n \mathbf{S}_0^k - \mathbf{R}_{k-1}^n \mathbf{S}_0^{k-1}\right) \bar{\boldsymbol{\psi}}_\alpha(0)\,. \tag{81}$$

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

$$\begin{aligned}
\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n &= \sum_{k=1}^{n} \left(\mathbf{R}_k^n \mathbf{S}_0^k - \mathbf{R}_{k-1}^n \mathbf{S}_0^{k-1}\right) \bar{\boldsymbol{\psi}}_\alpha(0) = \sum_{k=1}^{n} \left(\mathbf{R}_k^n e^{\mathbf{A}_k} \mathbf{S}_0^{k-1} - \mathbf{R}_k^n e^{\tilde{\mathbf{A}}_k} \mathbf{S}_0^{k-1}\right) \bar{\psi}_\alpha(0) \\
&= \sum_{k=1}^{n} \mathbf{R}_k^n \left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right) \mathbf{S}_0^{k-1} \bar{\boldsymbol{\psi}}_\alpha(0) \\
&= \sum_{k=1}^{n} \mathbf{R}_k^n \left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\,.
\end{aligned} \tag{82}$$

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 [75] to the non-linear case of the TDKS equations. We let $\boldsymbol{a}_k^\alpha = \left(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ and split it as

$$\boldsymbol{a}_k^\alpha = \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) + \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\,, \tag{83}$$

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. 83, we introduce the following auxiliary function

$$\boldsymbol{\xi}_\alpha^k(t) = e^{\mathbf{B}_k(t)} \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}), \qquad \forall t \in [t_{k-1}, t_k]\,, \tag{84}$$

where $\mathbf{B}_k(t) = \int_{t_{k-1}}^{t} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau$. We remark that $\boldsymbol{\xi}_\alpha^k(t)$ encapsulates the effect of truncation of the Magnus expansion in the time interval $[t_{k-1}, t_k]$. Differentiating the above equation gives

$$\dot{\boldsymbol{\xi}}_\alpha^k(t) = -i\bar{\mathbf{G}}_k(t) \boldsymbol{\xi}_\alpha^k(t)\,, \qquad \forall t \in [t_{k-1}, t_k], \tag{85}$$

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 $\mathrm{ad}_{\mathbf{X}}(\mathbf{Y})$ 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 $\mathrm{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))$ and using the above two arguments, it can be shown that $\bar{\mathbf{G}}_k$ is Hermitian. Letting $\boldsymbol{\gamma}_\alpha^k(t) = \bar{\boldsymbol{\psi}}_\alpha(t) - \boldsymbol{\xi}_\alpha^k(t)$, we use the Eqs. 72 and 85 to write the time-derivative of $\boldsymbol{\gamma}_\alpha^k$ as

$$\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t))\bar{\boldsymbol{\psi}}_\alpha(t), \qquad \forall t \in [t_{k-1}, t_k]\,. \tag{86}$$

Now, taking the dot product with $\boldsymbol{\gamma}_\alpha^k(t)^\dagger$ on both sides yields

$$\boldsymbol{\gamma}_\alpha^k(t)^\dagger\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\boldsymbol{\gamma}_\alpha^k(t)^\dagger\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i\boldsymbol{\gamma}_\alpha^k(t)^\dagger\left(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\,. \tag{87}$$

We note that $2\mathrm{Re}\{\boldsymbol{\gamma}_\alpha^k(t)^\dagger\dot{\boldsymbol{\gamma}}_\alpha^k(t)\} = \frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\|^2$, 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

$$\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\|^2 = \mathrm{Im}\{\boldsymbol{\gamma}_\alpha^k(t)^\dagger\left(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\}\,. \tag{88}$$

Consequently,

$$\frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\| \leq \left\|\left(\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)\right)\bar{\boldsymbol{\psi}}_\alpha(t)\right\|\,. \tag{89}$$

Time integrating the above equation yields

$$\left\|\boldsymbol{\gamma}_\alpha^k\right\|(t) \leq \int_{t_{k-1}}^{t}\left\|\left(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)\right)\bar{\boldsymbol{\psi}}_\alpha(\tau)\right\|d\tau\,, \tag{90}$$

where we have used 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 $\boldsymbol{\xi}_\alpha^k(t)$ (see Eq. 84). Using $\bar{\mathbf{G}}_k(\tau) = i\mathrm{dexp}_{\mathbf{B}_k(\tau)}(\dot{\mathbf{B}}_k(\tau))$ and the definition of the operator $\mathrm{dexp}_{\mathbf{X}}(\mathbf{Y})$, we get

$$\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau) = -\frac{i}{2}[\mathbf{B}_k(\tau), \dot{\mathbf{B}}_k(\tau)] + h.o.t = -\frac{i}{2}\int_{t_{k-1}}^{\tau}[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)]d\sigma + h.o.t.\,. \tag{91}$$

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

$$[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)] = [\bar{\mathbf{H}}(\tau), \bar{\mathbf{V}}'(\tau)](\sigma - \tau) + \mathcal{O}((\sigma - \tau)^2)\,, \tag{92}$$

where $\bar{\mathbf{V}}'(\tau) = \frac{d}{dt}(\bar{\mathbf{V}}(\sigma))|_{\sigma=\tau}$. Thus using the above relation in Eq. 91 we get

$$(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau))\bar{\boldsymbol{\psi}}_\alpha(\tau) = \frac{i}{4}\left([\bar{\mathbf{H}}, \bar{\mathbf{V}}'(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau)\right)(\tau - t_{k-1})^2 + \mathcal{O}((\tau - t_{k-1})^3)\,. \tag{93}$$

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

$$\left\|(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau))\bar{\boldsymbol{\psi}}_\alpha(\tau)\right\| \leq C(\tau - t_{k-1})^2\left\|\bar{\boldsymbol{\psi}}_\alpha(\tau)\right\|_{\bar{\mathbf{U}}} + \mathcal{O}((\tau - t_{k-1})^3). \tag{94}$$

We now turn to the definition of $\boldsymbol{\xi}_\alpha^k$ (defined in Eq. 84) to note that

$$\boldsymbol{\xi}_\alpha^k(t_k) = e^{\mathbf{B}_k(t_k)}\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) = e^{\bar{\mathbf{A}}_k}\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \tag{95}$$

where the second equality follows from 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, we write

$\boldsymbol{\gamma}_\alpha^k(t_k) = \bar{\boldsymbol{\psi}}_\alpha(t_k) - \boldsymbol{\xi}_\alpha^k(t_k) = \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$. We use this relation in Eq. 90 to get

$$\left\| \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| = \left\| \boldsymbol{\gamma}_\alpha^k \right\| (t_k) \leq \int_{t_{k-1}}^{t_k} \left\| \left(\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)\right) \bar{\boldsymbol{\psi}}_\alpha(\tau) \right\| d\tau \,. \tag{96}$$

Plugging in the result of Eq. 94 into the above equation provides the following bound

$$\left\| \left(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| \leq C(\Delta t)^3 \max_{t_{k-1} \leq t \leq t_k} \left\| \bar{\boldsymbol{\psi}}_\alpha(t) \right\|_{\bar{\mathbf{U}}} \,. \tag{97}$$

We note that the above equation provides a bound for the first term (truncation error) on the right side of Eq. 83. In order to bound the second term on the right side of Eq. 83, i.e., the error due to mid-point quadrature rule, we begin with the following identity

$$e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k} = \int_0^1 \frac{d}{dx} \left(e^{(1-x)\tilde{\mathbf{A}}_k} e^{x\bar{\mathbf{A}}_k} \right) \, dx = \int_0^1 e^{(1-x)\tilde{\mathbf{A}}_k} (\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k) e^{x\bar{\mathbf{A}}_k} \, dx. \tag{98}$$

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 $\left\| \bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k \right\| \leq C(\Delta t)^3 \left\| \frac{d^3}{dt^3}(\bar{\mathbf{A}})|_{t'} \right\|$ 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

$$\left\| \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| \leq \left\| (\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k) \right\| \leq C(\Delta t)^3 \left\| \frac{d^3}{dt^3}(\bar{\mathbf{A}})|_{t'} \right\|, \text{ for some } t' \in [t_{k-1}, t_k] \,. \tag{99}$$

Noting that $\frac{d^3}{dt^3}\bar{\mathbf{A}} = -i\frac{d^2}{dt^2}\bar{\mathbf{H}} = -i\frac{d^2}{dt^2}\bar{\mathbf{V}}$, $\forall t \in [t_{k-1}, t_k]$, and invoking the boundedness assumption on $\frac{d^2}{dt^2}\bar{\mathbf{V}}$ (assumption $\mathcal{A}8$), we get

$$\left\| \left(e^{\bar{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k}\right) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) \right\| \leq C(\Delta t)^3 \,. \tag{100}$$

Finally, using the results of Eqs. 97 and 100 in Eq. 82 along with the unitarity of the operators $\mathbf{R}_k^n$, yields

$$\left\| \bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n \right\| \leq C(\Delta t)^2 t_n \max_{t_0 \leq t \leq t_n} \left\| \bar{\boldsymbol{\psi}}_\alpha(t) \right\|_{\bar{\mathbf{U}}} \,. \tag{101}$$

This concludes the proof of Eq. 79.

□

3.4. Optimal spatio-temporal discretization

We follow on the lines of [60, 58, 45] to obtain the optimal spatial discretization by minimizing the semi-discrete error in the dipole moment at a given time, subject to a fixed number of finite-elements. The dipole moment is a vector-valued function defined at time t by $\boldsymbol{\mu}(t) = \int_\Omega \mathbf{x}\ \rho(t)\, d\mathbf{x}$. 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 use the Cauchy-Schwarz inequality to obtain

$$|\mu_x(t) - \mu_x^h(t)| \leq \|x\|_{L^2(\Omega)} \left\| \rho(t) - \rho^h(t) \right\|_{L^2(\Omega)} \leq C \left\| \rho(t) - \rho^h(t) \right\|_{L^2(\Omega)} , \tag{102}$$

where we have used the fact that $x \in L^2(\Omega)$, for a finite domain Ω. Plugging Eq. 16b into the above equation, we have

$$|\mu_x(t) - \mu_x^h(t)| \leq C \sum_{i=1}^{N_e} \left\| \psi_\alpha - \psi_\alpha^h \right\|_{H^1(\Omega)} (t) \,. \tag{103}$$

Now using Eq. 64b in the above equation, we have

$$\begin{aligned}|\mu_x(t)-\mu_x^h(t)| \leq C_4' e^{C_2 t}(t+1)\sum_e h_e^p \sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha})+|\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha})+|\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)\\ + C_4' e^{C_2 t}(t+1)\sum_e h_e^p\left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3)+|V_N|_{p+1,\Omega_e}\right),\end{aligned} \tag{104}$$

where $s_{1,\alpha}, s_{2,\alpha},$ and s_3 are defined in the Proposition 3.1. Using the definition of semi-norms, we rewrite the above equation as

$$\begin{aligned}|\mu_x(t)-\mu_x^h(t)| \leq C_4' e^{C_2 t}(t+1)\sum_{e=1}^{N_{elem}} h_e^p \int_{\Omega_e}\left[\sum_{\alpha=1}^{N_e}\left(\left|D^{p+1}\psi_\alpha\right|(s_{1,\alpha})+\left|D^{p+1}\psi_\alpha\right|(s_{2,\alpha})+\left|D^{p+3}\psi_\alpha\right|(s_{2,\alpha})\right)\right]d\mathbf{r}\\ + C_4' e^{C_2 t}(t+1)\sum_{e=1}^{N_{elem}} h_e^p \int_{\Omega_e}\left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3)+|V_N|_{p+1,\Omega_e}\right)d\mathbf{r}\,,\end{aligned} \tag{105}$$

where N_{elem} denotes the number of elements in the finite-element mesh (triangulation), and D^{p+1} and D^{p+3} denote the $(p+1)th$ and $(p+3)th$ spatial derivative of a function, respectively. Taking an element size distribution function, $h(\mathbf{r})$, the above inequality can be further rewritten as,

$$\begin{aligned}|\mu_x(t)-\mu_x^h(t)| \leq C_4' e^{C_2 t}(t+1)\int_\Omega h^p(\mathbf{r})\left[\sum_{\alpha=1}^{N_e}\left(\left|D^{p+1}\psi_\alpha\right|(s_{1,\alpha})+\left|D^{p+1}\psi_\alpha\right|(s_{2,\alpha})+\left|D^{p+3}\psi_\alpha\right|(s_{2,\alpha})\right)\right]d\mathbf{r}\\ + C_4' e^{C_2 t}(t+1)\int_\Omega h^p(\mathbf{r})\left(\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3)+|V_N|_{p+1,\Omega_e}\right)d\mathbf{r}\,.\end{aligned} \tag{106}$$

Obtaining the optimal element size distribution, for a fixed N_{elem}, now reduces to the following minimization problem,

$$\begin{aligned}\min_{h(\mathbf{r})}\int_\Omega h^p(\mathbf{r})\left[\sum_{\alpha=1}^{N_e}\left(\left|D^{p+1}\psi_\alpha\right|(s_{1,\alpha})+\left|D^{p+1}\psi_\alpha\right|(s_{2,\alpha})+\left|D^{p+3}\psi_\alpha\right|(s_{2,\alpha})\right)+\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3)+|V_N|_{p+1,\Omega_e}\right]d\mathbf{r}\\ \text{subject to}: \int_\Omega \frac{d\mathbf{r}}{h^3(\mathbf{r})} = N_{elem}\,.\end{aligned} \tag{107}$$

Solving the Euler-Lagrange equation for the above objective function with respect to $h(\mathbf{r})$ yields the following optimal element size distribution,

$$h(\mathbf{r}) = E\left[\sum_{\alpha=1}^{N_e}\left(\left|D^{p+1}\psi_\alpha\right|(s_{1,\alpha})+\left|D^{p+1}\psi_\alpha\right|(s_{2,\alpha})+\left|D^{p+3}\psi_\alpha\right|(s_{2,\alpha})\right)+\left|V_H[\rho^h]\right|_{p+1,\Omega_e}(s_3)+|V_N|_{p+1,\Omega_e}\right]^{-1/p+3}, \tag{108}$$

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 $V_H[\rho^h](s_3)$, we use the knowledge of their asymptotics from ground-state atomic calculations to coarse-grain the 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.

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 $\mu_{n,x}^h$ denotes the x-component of the full-discrete dipole moment then using the result in Eq. 101

it is straightforward to show,

$$|\mu_x^h(t_n) - \mu_{n,x}^h| \leq C t_n (\Delta t)^2 \left\| \bar{\boldsymbol{\psi}}_\alpha \right\|_{\bar{\mathbf{U}}} . \tag{109}$$

As is evident from the above relation, our choice of Δt is intrinsically tied to the spectral radius of the Kohn-Sham Hamiltonian, through $\left\| \bar{\boldsymbol{\psi}}_\alpha \right\|_{\bar{\mathbf{U}}}$. 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 $\|\boldsymbol{\psi}_\alpha(t)\|_{\bar{\mathbf{U}}}$ and the value of the constant, C, featuring in it. The value of $\left\| \bar{\boldsymbol{\psi}}_\alpha(t) \right\|_{\bar{\mathbf{U}}}$ can be reliably approximated from its ground-state value, i.e., $\left\| \bar{\boldsymbol{\psi}}_\alpha(0) \right\|_{\bar{\mathbf{U}}}$. The characteristic value of the constant is determined from atomic RT-TDDFT calculations at different Δt. To elaborate, the constant can be evaluated from a linear fit to the log-log plot of the error $|\mu_x^h(t_n) - \mu_{n,x}^h|$ with respect to Δt. For a multi-atom system, we use the least Δt obtained for each of its constituent atomic species.

4. Numerical Implementation

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

4.1. Higher-order spectral finite-elements

Finite-elements, with their varied choices of forms and orders [76, 34], 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 [56, 48]. However, this shortcoming has been demonstrably mitigated by the use of higher-order finite-elements [58, 45]. 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. To elaborate, while the conventional finite-elements are constructed from a tensor product of Lagrange polynomials interpolated through equidistant nodal points in an element, spectral finite elements employ a distribution of nodes obtained from the roots of the derivative of Legendre polynomials or the Chebyshev polynomials [77]. In our work, we use the roots of the derivative of Legendre polynomials along with boundary nodes, so as to maintain C^0 continuity. The resulting distribution is known as the Gauss-Legendre-Lobatto node distribution. The spectral finite-elements are known to afford better conditioning with increasing polynomial degree [77], and have been used to gain significant computational efficiency in ground-state DFT calculations [58]. Moreover, 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 $\mathbf{M}^{-1/2}$ that features in the discrete TDKS equations (see Eq. 70). 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. [45] 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.

4.2. Approximating the second-order Magnus operator

The form of the Magnus operator, as shown in Eq. 77, calls for efficient means of evaluating $\exp(\tilde{\mathbf{A}}_n)\boldsymbol{\psi}_\alpha^n$. Direct means of evaluating the matrix $\exp(\tilde{\mathbf{A}}_n)$ remain computationally prohibitive beyond small sizes of $\tilde{\mathbf{A}}_n$. Alternatively, one can attempt to evaluate the action of $\exp(\tilde{\mathbf{A}}_n)$ on $\boldsymbol{\psi}_\alpha^n$ 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. [78]

and references therewithin 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 *a priori* estimate, based on the spectral radius of $\tilde{\mathbf{A}}_n$, the number of terms required in the polynomial expansion for a desired accuracy, they 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 the kinetic and the potential part of the Kohn-Sham Hamiltonian in succession. 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. 5.4).

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 $\bar{\boldsymbol{\psi}}$ is given by

$$\mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}}) = span\{\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n^2\bar{\boldsymbol{\psi}}, \ldots, \tilde{\mathbf{A}}_n^{k-1}\bar{\boldsymbol{\psi}}\}. \tag{110}$$

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 $\mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}})$, and has the following property [79]:

$$\tilde{\mathbf{A}}_n\mathbf{Q}_k = \mathbf{Q}_k\mathbf{T}_k + \beta_{k+1,k}q_{k+1}e_k^T\,, \tag{111}$$

where $\mathbf{T}_k$ is a tridiagonal matrix (upper Hessenberg matrix if $\tilde{\mathbf{A}}_n$ is non-Hermitian) and represents the projection of $\tilde{\mathbf{A}}_n$ into $\mathbf{Q}_k$ (i.e., $\mathbf{T}_k = \mathbf{Q}_k^T\tilde{\mathbf{A}}_n\mathbf{Q}_k$), e_i is the i-th unit vector in $\mathbb{C}^k$, and $\beta_{k+1,k}$ is the $(k+1,k)$ entry of $\mathbf{T}_{k+1} = \mathbf{Q}_{k+1}^T\tilde{\mathbf{A}}_n\mathbf{Q}_{k+1}$. By construction, we have $\bar{\boldsymbol{\psi}} = \left\|\bar{\boldsymbol{\psi}}\right\|\mathbf{Q}_k e_1$. Furthermore, the operator $\mathbf{Q}_k\mathbf{Q}_k^T$ denotes the orthogonal projector onto the subspace $\mathcal{K}_k(\tilde{\mathbf{A}}_n, \boldsymbol{\psi})$. Thus, one can obtain the best approximation to $e^{\tilde{\mathbf{A}}_n}\boldsymbol{\psi}$ in the subspace $\mathcal{K}_k(\tilde{\mathbf{A}}_n, \boldsymbol{\psi})$, denoted by $\boldsymbol{y}_k$, through a Galerkin projection, and is given by

$$\boldsymbol{y}_k = \|\boldsymbol{\psi}\|\,\mathbf{Q}_k\mathbf{Q}_k^T e^{\tilde{\mathbf{A}}_n}\mathbf{Q}_k e_1. \tag{112}$$

However, as evident, the above expression requires the evaluation of the action of $e^{\tilde{\mathbf{A}}_n}$ on $\mathbf{Q}_k$ which remains computationally impractical beyond small system sizes, owing to the large size of the matrix $\tilde{\mathbf{A}}_n$. To this end, one can use a different approximation to $e^{\tilde{\mathbf{A}}_n}\boldsymbol{\psi}$, denoted by $\boldsymbol{z}_k \in \mathcal{K}_k(\tilde{\mathbf{A}}_n, \boldsymbol{\psi})$, given by

$$\boldsymbol{z}_k = \|\boldsymbol{\psi}\|\,\mathbf{Q}_k e^{\mathbf{Q}_k^T\tilde{\mathbf{A}}_n\mathbf{Q}_k}e_1 = \|\boldsymbol{\psi}\|\,\mathbf{Q}_k e^{\mathbf{T}_k}e_1\,. \tag{113}$$

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 [80]

$$\epsilon_k = \left\|e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}} - \left\|\bar{\boldsymbol{\psi}}\right\|\mathbf{Q}_k e^{\mathbf{T}_k}e_1\right\| = \left\|\bar{\boldsymbol{\psi}}\right\|\beta_{k+1,k}\left|\left[e^{\mathbf{T}_k}\right]_{k,1}\right|. \tag{114}$$

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. 3.3, 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 (Eq. 111) 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 [81] 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 estimates obtained in our full-discrete error analysis.

5. Results

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. 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 [30] software package. In all our finite-difference based calculations, we have used a stencil of order 4 in each direction, which is the default stencil order in Octopus. All the pseudopotential calculations are done using the norm-conserving Troullier-Martins pseudopotentials [82]. For all calculations, the ground-state Kohn-Sham orbitals are used as the initial states. We use the Chebyshev filter acceleration technique (see [45]) 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.

5.1. 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 (see Eq. 114). In effect, we mimic a semi-discrete (discrete in space but continuous in time) error analysis. Furthermore, to determine the coarsening rate of the mesh, as per Eq. 108, we use the *a priori* knowledge of the asymptotics of the ground-state atomic orbitals and electrostatic potentials. We remark that although we propose for using the far-field asymptotics of the ground-state electronic fields to determine the mesh coarsening rate, it, nevertheless, forms a reliable and cost effective way for discretizing the mesh away from the nuclei as opposed to any ad-hoc coarsening or using uniform discretization.

We now briefly touch upon on how to inexpensively obtain the single-atom ground-state electronic fields in the all-electron case. We note that for the single-atom case, under the assumption of equal fractional occupancy for degenerate orbitals, the electron density, and hence, the Kohn-Sham potential is spherically symmetric. Thus, one can rewrite the Kohn-Sham eigenvalue problem in spherical coordinates and invoke separation of variables to express a Kohn-Sham orbital as

$$\psi(\mathbf{r}) = R_{nl}(r)Y_{ml}(\theta,\phi)\,. \tag{115}$$

In the above equation, $R_{nl}(r)$ represents the radial part of the Kohn-Sham orbital corresponding to the principal quantum number n and azimuthal quantum number l, $Y_{lm}(\theta,\phi)$ denotes the real form of spherical harmonics for the pair of azimuthal quantum number l and magnetic quantum number m, and θ and ϕ

represent the polar and azimuthal angles, respectively. The radial part, $R_{nl}(r)$, can be solved using the following 1D radial eigenvalue equation

$$\Big[-\frac{1}{2}\frac{1}{r^2}\frac{d}{dr}(r^2\frac{d}{dr}) + \frac{l(l+1)}{r^2} + V_{KS}(r)\Big] R_{nl}(r) = \epsilon_{nl} R_{nl}(r)\,, \tag{116}$$

where ϵ_{nl} is the eigenvalue corresponding to $R_{nl}(r)$. Similarly, the spherically symmetric electrostatic potentials, $V_H(r)$ and $V_N^{ae}(r)$, can be obtained from the following 1D radial equations

$$-\frac{1}{4\pi}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}V_H(r)\right) = \rho(r)\,, \tag{117a}$$

$$-\frac{1}{4\pi}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}V_N^{ae}(r)\right) = -Z\delta(r)\,, \tag{117b}$$

where Z denotes the atomic number of the atom under consideration and $\delta(r)$ is the Dirac-delta distribution. We note that all the above radial equations can be pre-computed inexpensively using 1D finite-elements, and stored for various atomic species for future use in obtaining optimal mesh coarsening rates. Subsequently, for a multi-atom case, we use the superposition of atomic orbitals and electrostatic potentials, in Eq. 108, to determine the coarsening rate. The resultant mesh forms the coarsest mesh for our convergence studies with respect to mesh-size. We then refine the coarsest mesh near the nuclei to obtain a sequence of increasingly refined meshes. We employ finite-elements of three different orders—HEX8, HEX27, and HEX125SPEC—in all our convergence studies. We remark that, although in the above procedure the single electronic fields are obtained for all-electron case, the same single-atom fields can be used for obtaining the coarsening rates for pseudopotential calculations as well. This is owing to the fact that the pseudo-wavefunctions and the pseudo-electrostatic potentials have the same far-field decay rates as its all-electron counterparts. Alternatively, one can directly use the single-atom pseudopotentials and pseudo-wavefunctions, obtained from a pseudopotential generation program, like FHI98PP [83].

We note that in order to perform the convergence study with respect to mesh-size, and compare to the semi-discrete error estimate obtained in Eq. 105, 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

$$\frac{\left|\mu_{x,0}(t) - \mu_x^h(t)\right|}{|\mu_{x,0}(t)|} = C h_{min}^q\,. \tag{118}$$

to determine $\mu_{x,0}(t)$, C, and q. In the above equation h_{min} represents the minimum element size, h_e, present in the mesh. The obtained $\mu_{x,0}(t)$ 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 $\frac{\left|\mu_{x,0}(t) - \mu_x^h(t)\right|}{|\mu_{x,0}(t)|}$ 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

$$\frac{\left|\mu_x^h(t_n) - \mu_{n,x}^h\right|}{|\mu_x^h(t_n)|} = C(\Delta t)^q\,. \tag{119}$$

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.

5.1.1. 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 mesh-

coarsening strategy based on the knowledge of ground-state atomic electronic fields to adaptively coarsen the mesh away from the nuclei. We remark that this adaptive coarsening of the mesh away from the nuclei 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, which are otherwise essential to tackle wave reflection effects that are observed while dealing with small computational domains. 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 1 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 1, much higher accuracies are obtained with HEX125SPEC when compared to HEX8 and HEX27 of the same mesh size. In particular, $200,000$ HEX8 elements are required to achieve relative errors of 10^{-2}, whereas we achieve relative error of 10^{-3} with just $3,000$ HEX125SPEC elements. 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 2 depicts the rate of convergence of the dipole moment with respect to Δt at $t_n = 1.0$ a.u.. We obtain the rate of convergence, $q = 1.96$ (defined in Eq. 119), which agrees remarkably well with the theoretical rate of convergence for second-order Magnus propagator (see Proposition 3.2).

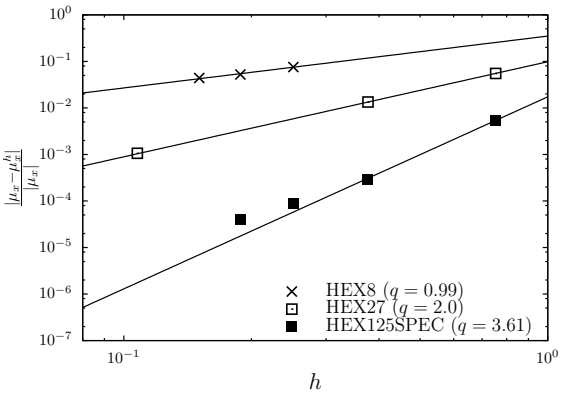

Figure 1: Rates of convergence with respect to spatial discretization for LiH

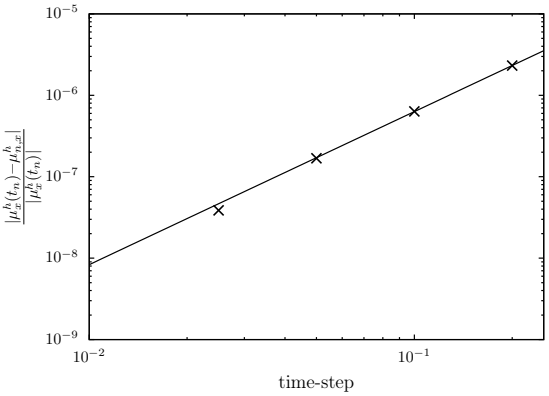

Figure 2: Rate of convergence with respect to temporal discretization for LiH

5.1.2. 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 C-H-C tetrahedral angle of 109.4712° as our benchmark system. As was the case with lithium hydride, we use ground-state single-atom electronic fields to coarsen the mesh away from the nuclei. 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 3 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. To this end, we use a sufficiently refined HEX125SPEC mesh and employ the second-order Magnus propagator with different Δt to propagate the ground-state Kohn-Sham orbitals under the influence of the Gaussian electric field. Figure 4 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 that is remarkably close to optimal rate of

convergence, with $q = 1.98$ (defined in Eq. 119).

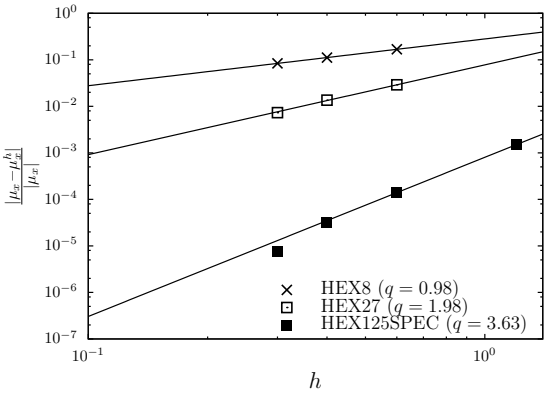

Figure 3: Rates of convergence with respect to spatial discretization for CH_4

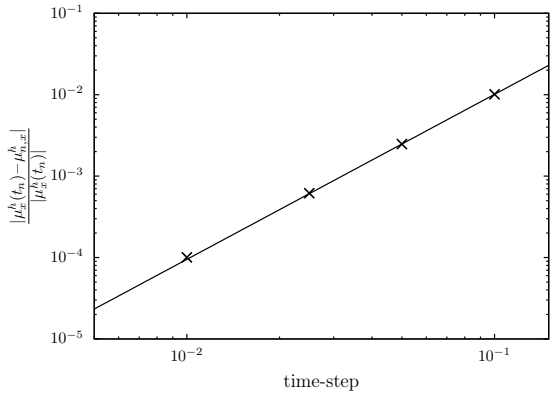

Figure 4: Rate of convergence with respect to temporal discretization for CH_4

5.2. 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 strategy of mesh coarsening as detailed in Sec. 5.1. 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 (see Eq. 114) 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. 5.1.1), and the methane molecule with a Gaussian electric field (see Sec. 5.1.2). Figures 5 and 6 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.

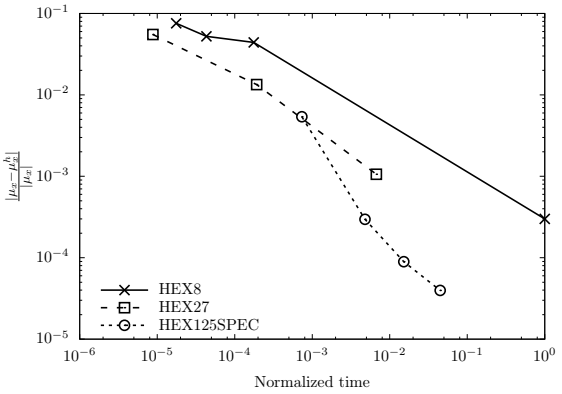

Figure 5: Computational efficiency of various orders finite-elements for LiH

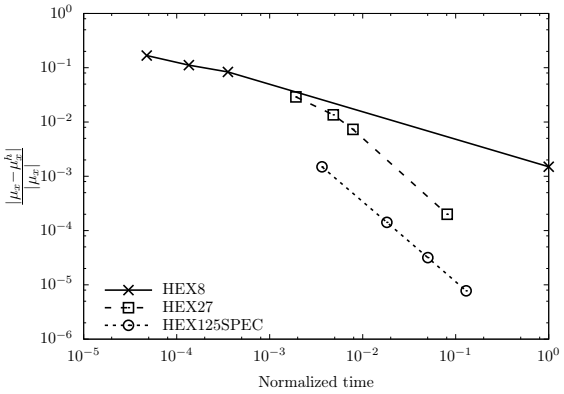

Figure 6: Computational efficiency of various orders finite-elements for CH_4

5.3. 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 also provide comparison, wherever possible, against pseudopotential calculations based on a finite-difference discretization. We use Al_2, Al_{13}, and Mg_2 as the benchmark metallic systems for pseudopotential calculations. Furthermore, we use Buckministerfullerene (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. We remark that 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.

5.3.1. 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 HEX64SPEC finite-elements with local refinement around the nuclei and coarsening away from the nuclei. The coarsening, as in the case of previous examples, was done by using the superposition of ground-state atomic electronic fields in Eq. 108. The local refinement around the atom is done such that it achieves 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 system from its ground-state by using a Gaussian electric 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.. We use the second-order Magnus propagator with a time-step of $\Delta t = 0.05$ a.u. $= 0.001209$ fs to propagate the Kohn-Sham orbitals and allow the system to evolve for 10 fs. We use a tolerance of 10^{-8} for the Krylov subspace projection error (see Eq. 114). Next, we take 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 cross-section tensor $\boldsymbol{\sigma}(\omega) = \frac{4\pi\omega}{c}\mathrm{Im}[\boldsymbol{\alpha}(\omega)]$, where c is the speed of light. Finally, the dipole strength function (absorption spectrum), $S(\omega)$, is given by $S(\omega) = \frac{1}{3}\mathrm{Tr}[\boldsymbol{\sigma}(\omega)]$. The peaks in the absorption spectrum correspond to the excitation energies. In this study, we also assess the performance of higher-order finite-elements by comparing against finite-difference based method. To this end, we use the Octopus [30] software package to perform the finite-difference calculation. 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. Typically, one needs a larger domain for RT-TDDFT calculations than ground-state calculations, so as to avoid reflection at the domain boundaries. However, unlike the case in finite-elements, Octopus uses a uniform mesh, which makes the use of large domain sizes prohibitively expensive. A typical workaround is to use a smaller domain with absorbing boundary conditions. In order to better assess the effects of domain size for the finite-difference mesh, we use two cubical domains of sizes 38 a.u and 46 a.u, both 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 atomic units). For clarity, we refer to the three finite-difference calculations, namely, 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-46.0, FD-38.0, and FD-38.0-ABS, respectively. All the other simulation details, namely, time-step, duration of propagation, choice of propagator, and tolerance for Krylov subspace, are same as those used for the finite-element case. Fig. 7 compares absorption spectrum obtained from finite-elements against finite-difference. We have a 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, which is a standard practice in the computation of the absorption cross-section. As is evident from the figure, we get good agreement with the finite-difference based results for the domain size of 46 a.u.. 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 1 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 $\sim$ 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 coarsening 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.

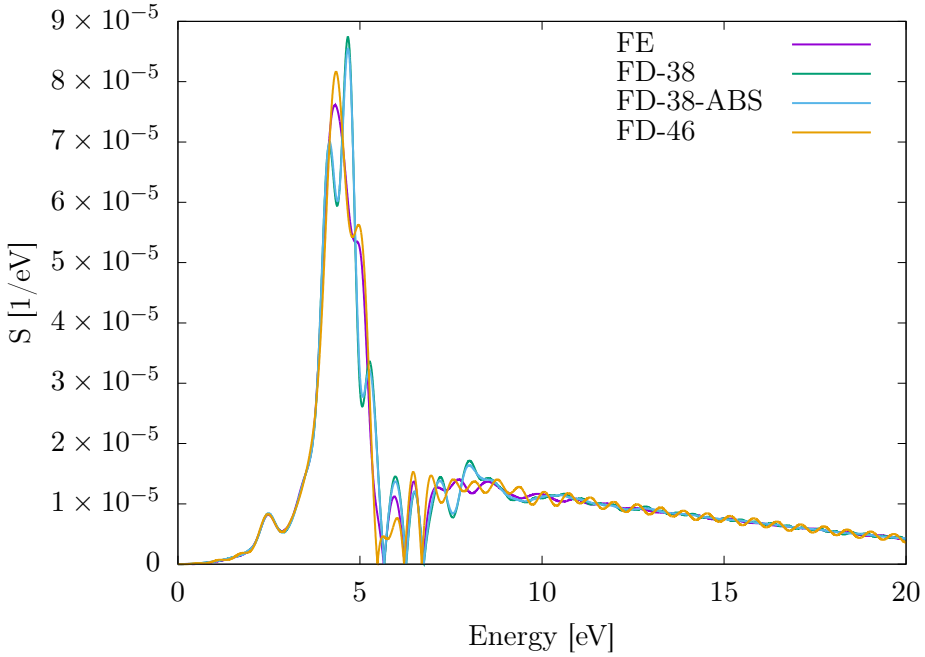

Figure 7: Absorption spectra for Al_2

Table 1: 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

5.3.2. Pseudopotential calculations: Al_{13}

We now consider a 13 atom aluminum cluster with an icosahedral symmetry. We use the same characteristic finite-element mesh as that of Al_2 but with a cubical domain of length 70 a.u., to avoid reflection effects. We repeat the same study as in the case of Al_2. To elaborate, we use the same electric field, time-step, duration of propagation, and tolerance for the Krylov subspace projection as in the case of Al_2.

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. 8 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 2 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 a $\sim$ 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 coarsening of the mesh. In particular, the finite-elements afford $\sim$ 30-fold fewer degrees of freedom as compared to finite-difference.

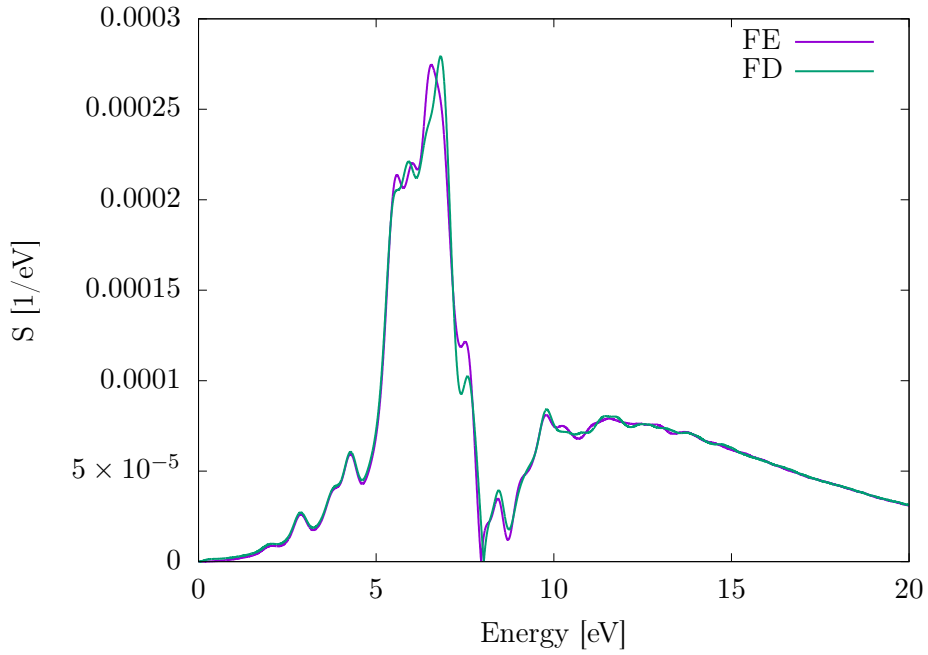

Figure 8: Absorption spectra for Al_{13}.

Table 2: 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

5.3.3. Pseudopotential calculations: Buckministerfullerene

In this example, we consider the Buckministerfullerene molecule, as shown in Figure 9. It comprises of 60 carbon atoms (240 electrons) packed into the shape of a buckyball. In order to generate a suitable finite-element mesh, we use HEX64SPEC finite-elements with the local refinement around the nuclei and coarsening away from the nuclei, governed by Eq. 108. We use a cubical domain of length 50 a.u. to eliminate any reflection effects from the boundaries. The ground-state is excited with a Gaussian electric field of the same form as used for Al_2. We use a time-step of 0.05 a.u. and propagate the system for 10 fs. 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 (see Sec. 5.3.1), 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 10 shows the absorption spectrum obtained from finite-element and the three different finite-difference based calculations. We have used a 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 is not improving the absorption spectrum. This, once again, indicates that one cannot always dispense using a larger domain by, solely, using absorbing boundary conditions. Table 3 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. In terms of computational efficiency, finite-elements attain a $\sim$ 3-fold speedup over FD-36. This higher efficiency of the finite-elements, is once again, attributed to a $\sim$ 9-fold fewer degrees of freedom required by the finite-elements against that of finite-difference.

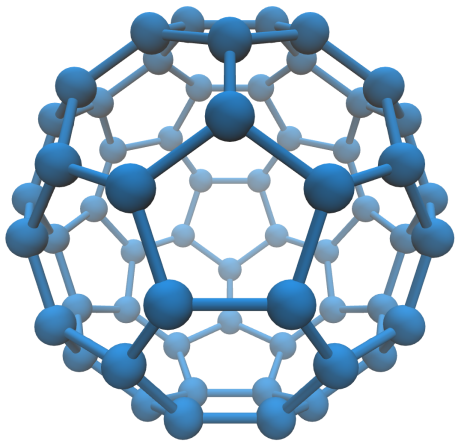

Figure 9: Buckministerfullerene molecule

Table 3: 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	$13,997,521$	$15,361$

5.3.4. 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. In particular, we use an electric field of the form

$$\mathbf{E}(t) = E_0 sin^2(\pi/T) sin(\omega_l t)\hat{x}\,, \tag{120}$$

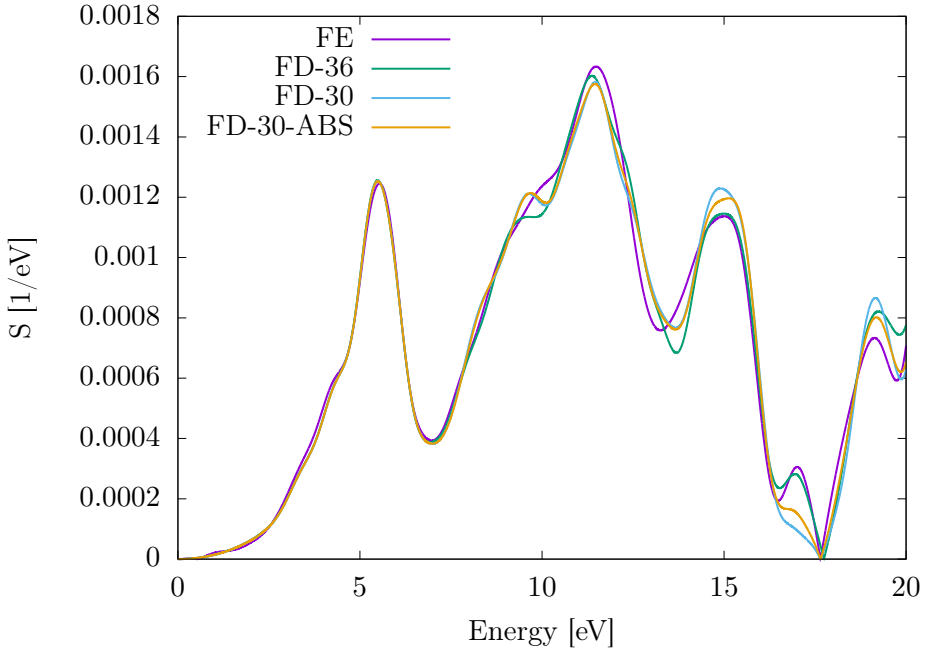

Figure 10: Absorption spectra of Buckministerfullerene

where E_0, ω_l, and T denote the amplitude, frequency, and duration of the laser pulse, respectively. In our example, we use $E_0 = 0.01$, $\omega_l = 0.03$, and $T = 5\left(\frac{2\pi}{\omega_l}\right)$ (all in atomic units). We use HEX125SPEC finite-elements with the local refinement around the nuclei and coarsening away from the nuclei, governed by Eq. 108. Furthermore, we use a cubical domain of length 100 a.u. to eliminate any reflection effects from the boundaries. For time-propagation, we use the second-order Magnus propagator with $\Delta t = 0.025$ and propagate it for 25.33 fs (i.e., T fs). We use a tolerance of 10^{-8} for the Krylov subspace projection error (see Eq. 114). Next, 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 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 [84]. 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.

5.3.5. 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 as described in Sec. 5.1.2. 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 coarsening 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 same Gaussian electric field as used for Al_2 (see Sec. 5.3.1) and propagate them for 10 fs. We use time-steps of 0.05 and 0.025 (both in a.u.) for the pseudopotential and all-electron calculations, respectively. Furthermore, a tolerance of 10^{-8} for the Krylov subspace projection error (see Eq. 114) is used for both the calculations. The absorption spectra for both the calculations are shown in Figure 12. We used the same Gaussian window as in the case of

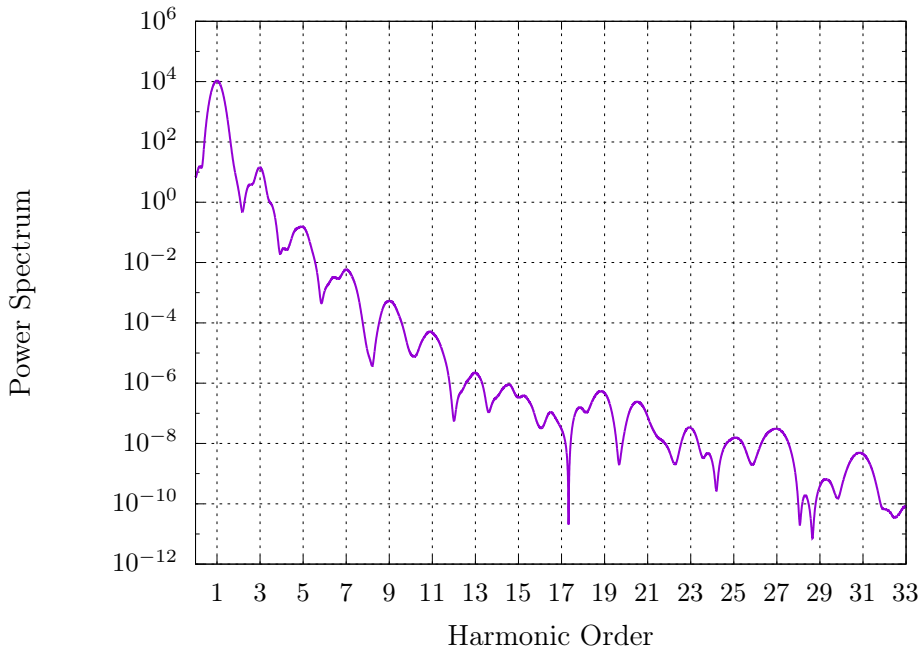

Figure 11: Dipole power spectrum of Mg_2

Buckministerfullerene (see Sec. 5.3.3), 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 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: a) an increase in the degrees of freedom; and b) widening of the spectral radius of the Kohn-Sham Hamiltonian, which, in turn, warrants a smaller time-step (see Eq. 109) as well as a larger Krylov subspace to achieve the prescribed accuracy (see Eq. 114). In particular, for the case of 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. 5.1.1). Moreover, one can mitigate the need of a refined mesh for the all-electron calculation by using an enriched finite-element basis, wherein the standard (classical) finite-element basis are augmented with numerical atom-centered basis. This idea has successfully attained $100-300$x speedup over the standard (classical) finite-elements for ground-state DFT calculations [52], and can be extended to RT-TDDFT to further the capabilities of finite-elements.

Table 4: 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

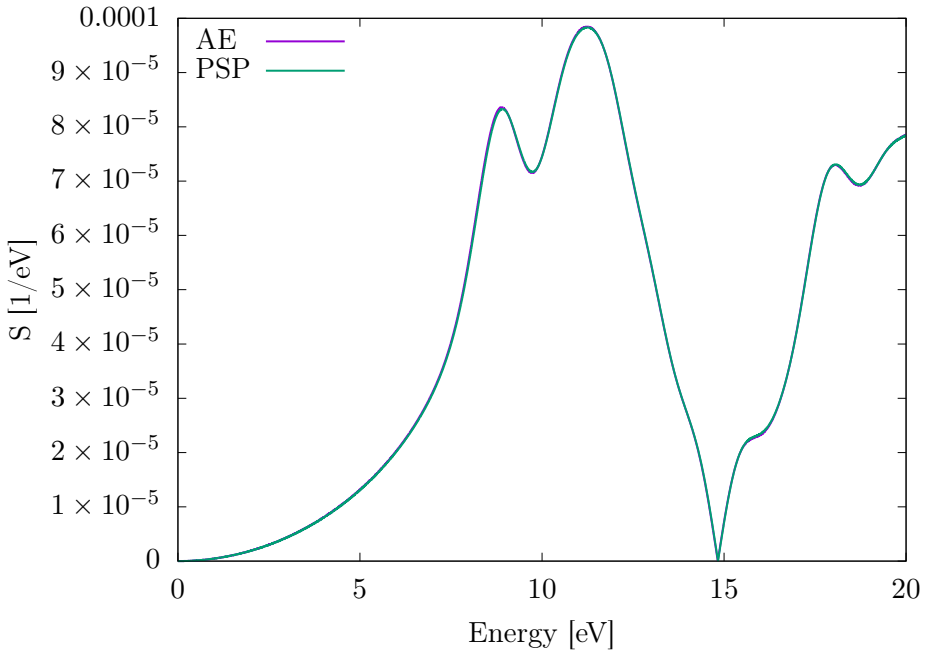

Figure 12: Absorption spectra of methane

5.3.6. All-electron calculations: Benzene

In this example, we perform similar comparative studies between all-electron and pseudopotential calculations for benzene molecule. As with the methane molecule, we use HEX64SPEC and HEX125SPEC 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 (see Sec. 5.3.5). Rest of the simulation details, namely, time-steps, duration of propagation, and the tolerance for the Krylov subspace projection error, remain the same as that used for methane. Figure 13 compares the absorption spectra from the all-electron and pseudopotential calculations. While there is qualitative agreement between both calculations, quantitatively the predictions from all-electron and pseudopotential calculations differ. In particular, the first two excitation peaks (see Table 5) 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.

Table 5: 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	$257,473$	1574

5.4. Scalability

We demonstrate the parallel scalability (strong scaling) of the proposed enriched finite element basis in Figure 14. We choose the Buckministerfullerene molecule containing ~ 3.5 million degrees of freedom

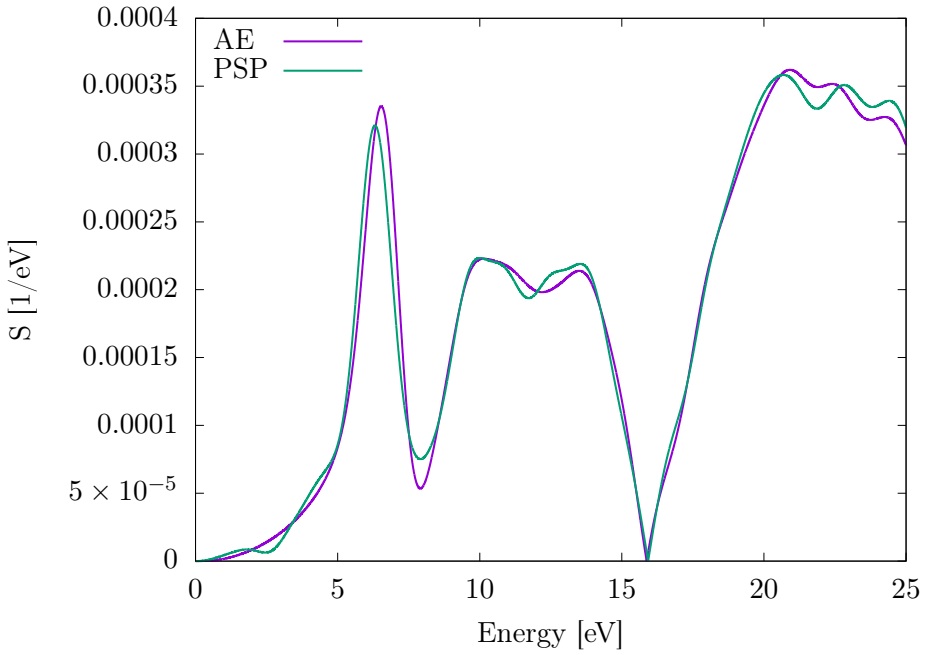

Figure 13: Absorption spectra of benzene

(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.

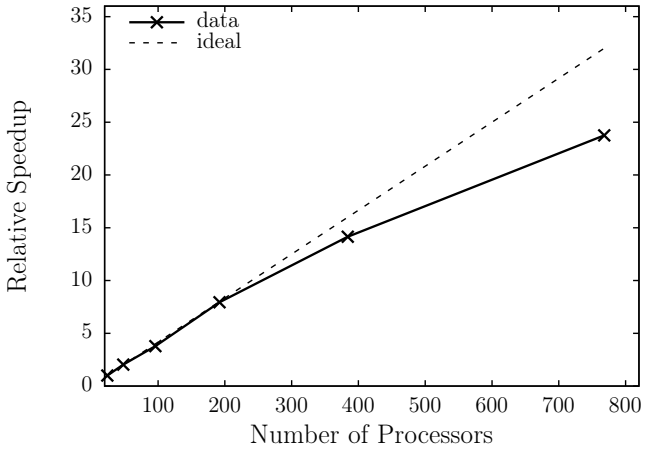

Figure 14: Parallel scalability of the higher-order finite-element implementation.

6. Summary

In summary, in this work, we have investigated the accuracy, computational efficiency and scalability of higher-order finite-elements for the RT-TDDFT problem, for both pseudopotential and all-electron calculations. We begin by decomposing the discretization error in the time-dependent Kohn-Sham orbitals into two parts, one arising due to spatial discretization and the other due to temporal discretization. We, then, developed semi-discrete error estimates (discrete in space, continuous in time) in the Kohn-Sham orbitals in terms of the electronic fields, and the order and size of the finite-elements used. Subsequently, we developed full-discrete error-estimates in the Kohn-Sham orbitals, in the context of second-order Magnus propagator with a midpoint integration rule, in terms of the time-step size and the spatial derivative of the Kohn-Sham orbitals. We note that finite-elements form a non-orthogonal basis, and hence, require efficient strategies to numerically evaluate the inverse of the overlap matrix in order to construct the Magnus propagator (or any exponential time-integrator). To this end, we employed spectral finite-elements in conjunction with Gauss-Legendre-Lobatto quadrature rules that results in a diagonal overlap matrix, thereby making the evaluation of its inverse trivial. We note that the second-order Magnus propagator with a midpoint integration rule warrants the knowledge of the time-dependent Kohn-Sham Hamiltonian at half time-step in future, which is unknown *a priori*. To this end, we used a second-order predictor-corrector scheme to accurately evaluate the Hamiltonian at a future time instance. 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.

In our numerical studies, we used the semi-discrete error estimates along with the *a priori* knowledge of the asymptotics of the electronic fields, using superposition of ground-state atomic solutions, to adaptively coarsen the mesh away from the nuclei and obtain close to optimal finite-element meshes, for both pseudopotential and all-electron systems. We used methane and lithium hydride molecule as our benchmark systems for pseudopotential and all-electron calculations, respectively. To begin with, we investigated the rates of convergence of the dipole moment with respect to mesh size for linear, quadratic and fourth-order spectral finite-elements. In both the benchmark systems, we observed close to optimal rates of convergence for all the three different orders of finite-elements employed. Furthermore, for a fixed order of finite-element, we investigated the rate of convergence of the dipole moment with respect to time-step as afforded by the second-order Magnus propagator. We attained close to optimal rates of convergence with respect to time discretization for both pseudopotential and all-electron calculations. Next, we assessed the computational efficiency afforded by higher-order finite-elements over linear finite-elements, using, once again, methane and lithium hydride as our benchmark systems for pseudopotential and all-electron calculations, respectively. To this end, we used the developed mesh coarsening scheme and studied the computational time required to solve the benchmark systems. We find a staggering $100-$fold and $10-$fold speedup afforded by the fourth-order finite-element over linear and quadratic finite-elements, respectively. Furthermore, we assessed the accuracy and efficiency afforded by higher-order finite-elements for other materials systems. We used Al_2, Al_{13}, and Buckministerfullerene (C_{60}) as the benchmark systems for pseudopotential calculations and compared the resulting absorption spectrum to that obtained from finite-difference based calculations done using the Octopus package. We observed good agreement in the excitation energies obtained from both the methods. In terms of computational efficiency, we obtained 40x, 8x, 3x speedup over finite-difference for Al_2, Al_{13}, and Buckministerfullerene, respectively. Thus, we underscore that higher-order finite-elements can compete with finite-difference, especially for non-periodic systems. Furthermore, we remark that, with the adaptive mesh coarsening afforded by finite-elements, one can use large simulation domains without significant increase in the number of basis functions, and hence, circumvent the use of any artificial absorbing boundary conditions. We also demonstrated the efficacy of higher-order finite-elements, especially its efficient handling of large domains, for nonlinear response by studying the higher harmonic generation in Mg_2 under a strong electric field. Next, we conducted similar studies for the all-electron case, using methane and benzene as our benchmark systems. For both the systems, we obtained good agreement with corresponding pseudopotential based results, thereby, establishing the capabilities of finite-elements for all-electron calculations. Additionally, the all-electron *vis-á-vis* pseudopotential results highlight the versatility of the 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 with almost linear scaling up to 384 processors for a benchmark system

comprising of the Buckministerfullerene molecule containing ~ 3.5 million basis functions, and good scaling up to 768 processors.

Thus, higher-order finite-elements offer a computationally efficient, systematically improvable, and scalable basis for RT-TDDFT calculations, applicable to both pseudopotential and all-electron cases. We note that, while for all-electron calculations one needs a large number of basis functions in order to capture the sharp electronic fields near the nuclei, this shortcoming can be alleviated by augmenting the finite-element basis with numerical atom-centered basis. This idea has been successfully used for ground-state DFT [52], and its extension to RT-TDDFT is currently being investigated. Furthermore, the key ideas behind the *a priori* mesh adaptation developed in this work can be used to formulate an *a posteriori* mesh refinement scheme, thereby, allowing for more robust error control. Lastly, the prospect of using higher-order finite-elements to develop linear scaling RT-TDDFT algorithms is also of significance, and is a subject of our future investigation.

7. Acknowledgements

We are grateful for the support of Army Research Office through Grant number W911NF-15-1-0158, under the auspices of which the mathematical formulation, error analysis and numerical implementation was developed, and the pseudopotential studies were conducted. We also gratefully acknowledge the support from the Department of Energy, Office of Basic Energy Sciences, under grant number DE-SC0017380, for supporting the all-electron studies. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant number ACI-1053575. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We also acknowledge the Advanced Research Computing at University of Michigan for providing additional computing resources through the Flux computing platform, part of which were performed using the computing cluster constructed from the DURIP Grant number W911NF-18-1-0242.

References

[1] E. Runge, E. K. U. Gross, Density-functional theory for time-dependent systems, Phys. Rev. Lett. 52 (1984) 997–1000.

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

[3] H. Appel, E. K. U. Gross, K. Burke, Excitations in time-dependent density-functional theory, Phys. Rev. Lett. 90 (2003) 043005.

[4] X. Gonze, J.-P. Vigneron, Density-functional approach to nonlinear-response coefficients of solids, Phys. Rev. B 39 (1989) 13120–13128.

[5] S. J. A. van Gisbergen, J. G. Snijders, E. J. Baerends, Time-dependent density functional results for the dynamic hyperpolarizability of C_{60}, Phys. Rev. Lett. 78 (1997) 3097–3100.

[6] G. Stefanucci, C.-O. Almbladh, Time-dependent quantum transport: An exact formulation based on TDDFT, EPL 67 (2004) 14.

[7] S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio, E. K. U. Gross, Time-dependent quantum transport: A practical scheme using density functional theory, Phys. Rev. B 72 (2005) 035308.

[8] C. Jamorski Jödicke, H. P. 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 (2003) 252–264.

[9] T. Stein, L. Kronik, R. Baer, Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory, J. Am. Chem. Soc. 131 (2009) 2818–2820.

[10] T. Burnus, M. A. L. Marques, E. K. U. Gross, Time-dependent electron localization function, Phys. Rev. A 71 (2005) 010501.

[11] X.-M. Tong, S.-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 (1998) 452–461.

[12] X.-M. Tong, S.-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 (2001) 013417.

[13] D. A. Telnov, S.-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 (2009) 041401.

[14] M. E. Casida, Time-dependent density functional response theory for molecules, in: Recent Advances In Density Functional Methods: (Part I), World Scientific, 1995.

[15] M. Petersilka, U. J. Gossmann, E. K. U. Gross, Excitation energies from time-dependent density-functional theory, Phys. Rev. Lett. 76 1212–1215.

[16] J. Theilhaber, Ab initio simulations of sodium using time-dependent density-functional theory, Phys. Rev. B 46 (1992) 12990–13003.

[17] K. Yabana, G. F. Bertsch, Time-dependent local-density approximation in real time, Phys. Rev. B 54 (1996) 4484–4487.

[18] R. Baer, R. Gould, A method for ab initio nonlinear electron-density evolution, J. Chem. Phys. 114 (2001) 3385–3392.

[19] G. Vignale, Center of mass and relative motion in time dependent density functional theory, Phys. Rev. Lett. 74 (1995) 3233–3236.

[20] N. T. Maitra, K. Burke, C. Woodward, Memory in time-dependent density functional theory, Phys. Rev. Lett. 89 (2002) 023002.

[21] M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, E. K. U. Gross, Time-dependent density functional theory, Springer, 2006.

[22] N. T. Maitra, K. Burke, Demonstration of initial-state dependence in time-dependent density-functional theory, Phys. Rev. A 63 (2001) 042501.

[23] F. Gygi, Architecture of qbox: A scalable first-principles molecular dynamics code, IBM J. Res. Dev. 52 (2008) 137–144.

[24] A. Schleife, E. W. Draeger, Y. Kanai, A. A. Correa, Plane-wave pseudopotential implementation of explicit integrators for time-dependent kohn-sham equations in large-scale simulations, J. Chem. Phys. 137 (2012) 22A546.

[25] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, The siesta method for ab initio order-n materials simulation, J. Phys.: Condens. Matter 14 (2002) 2745.

[26] Y. Takimoto, F. D. Vila, J. J. Rehr, Real-time time-dependent density functional theory approach for frequency-dependent nonlinear optical response in photonic molecules, J. Chem. Phys. 127 (2007) 154114.

[27] M. Kuisma, A. Sakko, T. P. Rossi, A. H. Larsen, J. Enkovaara, L. Lehtovaara, T. T. Rantala, Localized surface plasmon resonance in silver nanoparticles: Atomistic first-principles time-dependent density-functional theory calculations, Phys. Rev. B 91 (2015) 115431.

[28] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, W. de Jong, NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations, Comput. Phys. Commun. 181 (2010) 1477–1489.

[29] K. Lopata, N. Govind, Modeling fast electron dynamics with real-time time-dependent density functional theory: application to small molecules and chromophores, J. Chem. Theory Comput. 7 (2011) 1344–1355.

[30] A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. A. L. Marques, E. K. U. Gross, A. Rubio, octopus: a tool for the application of time-dependent density functional theory, Phys. Status Solidi B 243 (11) 2465–2488.

[31] J. J. Mortensen, L. B. Hansen, K. W. Jacobsen, Real-space grid implementation of the projector augmented wave method, Phys. Rev. B 71 (2005) 035109.

[32] M. Walter, H. Häkkinen, L. Lehtovaara, M. Puska, J. Enkovaara, C. Rostgaard, J. J. Mortensen, Time-dependent density-functional theory in the projector augmented-wave method, J. Chem. Phys. 128 (2008) 244101.

[33] S. Brenner, R. Scott, The mathematical theory of finite element methods, Springer, 2007.

[34] T. J. R. Hughes, The finite element method: linear static and dynamic finite element analysis, Courier Corporation, 2012.

[35] S. R. White, J. W. Wilkins, M. P. Teter, Finite-element method for electronic structure, Phys. Rev. B 39 (1989) 5819–5833.

[36] E. Tsuchida, M. Tsukada, Large-scale electronic-structure calculations based on the adaptive finite-element method, J. Phys. Soc. Jpn. 67 (11) (1998) 3844–3858.

[37] J. E. Pask, B. M. Klein, C. Y. Fong, P. A. Sterne, Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach, Phys. Rev. B 59 (1999) 12352–12358.

[38] J. Pask, B. Klein, P. Sterne, C. Fong, Finite-element methods in electronic-structure theory, Comput Phys Commun 135 (2001) 1–34.

[39] J. E. Pask, P. A. Sterne, Finite element methods in ab initio electronic structure calculations, Model. Simul. Mater. Sci. Eng. 13 (2005) R71.

[40] J.-L. Fattebert, R. D. Hornung, A. M. Wissink, Finite element approach for density functional theory calculations on locally-refined meshes, J. Comput. Phys. 223 (2007) 759–773.

[41] D. Zhang, L. Shen, A. Zhou, X.-G. Gong, Finite element method for solving Kohn–Sham equations based on self-adaptive tetrahedral mesh, Phys. Lett. A 372 (2008) 5071–5076.

[42] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, M. Ortiz, Non-periodic finite-element formulation of Kohn–Sham density functional theory, J. Mech. Phys. Solids 58 (2010) 256–280.

[43] J. Fang, X. Gao, A. Zhou, A Kohn–Sham equation solver based on hexahedral finite elements, J. Comput. Phys. 231 (2012) 3166–3180.

[44] G. Bao, G. Hu, D. Liu, An h-adaptive finite element solver for the calculations of the electronic structures, J. Comput. Phys. 231 (2012) 4967–4979.

[45] P. Motamarri, M. R. Nowak, K. Leiter, J. Knap, V. Gavini, Higher-order adaptive finite-element methods for Kohn–Sham density functional theory, J. Comput. Phys. 253 (2013) 308–343.

[46] P. Motamarri, V. Gavini, Configurational forces in electronic structure calculations using Kohn-Sham density functional theory, Phys. Rev. B 97 (2018) 165132.

[47] S. Yamakawa, S.-a. Hyodo, Gaussian finite-element mixed-basis method for electronic structure calculations, Phys. Rev. B 71 (2005) 035113.

[48] E. J. Bylaska, M. Holst, J. H. Weare, Adaptive finite element method for solving the exact Kohn–Sham equation of density functional theory, J. Chem. Theory Comput. 5 (2009) 937–948.

[49] L. Lehtovaara, V. Havu, M. Puska, All-electron density functional theory and time-dependent density functional theory with high-order finite elements, J. Chem. Phys. 131 (2009) 054103.

[50] V. Schauer, C. Linder, All-electron Kohn–Sham density functional theory on hierarchic finite element spaces, J. Comput. Phys. 250 (2013) 644–664.

[51] P. Motamarri, V. Gavini, Subquadratic-scaling subspace projection method for large-scale Kohn-Sham density functional theory calculations using spectral finite-element discretization, Phys. Rev. B 90 (2014) 115127.

[52] B. Kanungo, V. Gavini, Large-scale all-electron density functional theory calculations using an enriched finite-element basis, Phys. Rev. B 95 (2017) 035112.

[53] P. Motamarri, V. Gavini, K. Bhattacharya, M. Ortiz, Spectrum-splitting approach for fermi-operator expansion in all-electron Kohn-Sham DFT calculations, Phys. Rev. B 95 (2017) 035111.

[54] L. Lehtovaara, V. Havu, M. Puska, All-electron time-dependent density functional theory with finite elements: Time-propagation approach, J. Chem. Phys. 135 (2011) 154104.

[55] G. Bao, G. Hu, D. Liu, Real-time adaptive finite element solution of time-dependent Kohn–Sham equation, J. Comput. Phys. 281 (2015) 743–758.

[56] B. Hermansson, D. Yevick, Finite-element approach to band-structure analysis, Phys. Rev. B 33 (1986) 7241.

[57] P. F. Batcho, Computational method for general multicenter electronic structure calculations, Phys. Rev. E 61 (2000) 7169–7183.

[58] P. Motamarri, M. Iyer, J. Knap, V. Gavini, Higher-order adaptive finite-element methods for orbital-free density functional theory, J. Comput. Phys. 231 (2012) 6596–6621.

[59] Z. H. Levine, J. W. Wilkins, An energy-minimizing mesh for the schrödinger equation, J. Comput. Phys. 83 (1989) 361–372.

[60] R. Radovitzky, M. Ortiz, Error estimation and adaptive meshing in strongly nonlinear dynamic problems, Comput. Methods Appl. Mech. Eng. 172 (1999) 203–240.

[61] T. Ishida, K. Ohno, On the asymptotic behavior of hartree-fock orbitals, Theor. Chim. Acta 81 (1992) 355–364.

[62] O. Karakashian, C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method, Math. Comput. 67 (1998) 479–499.

[63] O. Karakashian, C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method, SIAM J. Numer. Anal. 36 (1999) 1779–1807.

[64] A. Ávila, A. Meister, M. Steigemann, An adaptive Galerkin method for the time-dependent complex Schrödinger equation, Appl. Numer. Math. 121 (2017) 149–169.

[65] J. Wang, A new error analysis of Crank–Nicolson Galerkin FEMs for a generalized nonlinear Schrödinger equation, J. Sci. Comput. 60 (2014) 390–407.

[66] D. Shi, X. Liao, L. Wang, Superconvergence analysis of conforming finite element method for nonlinear Schrödinger equation, Appl. Math. Comput. 289 (2016) 298–310.

[67] D. Shi, J. Wang, Unconditional superconvergence analysis of a Crank–Nicolson Galerkin FEM for nonlinear Schrödinger equation, J. Sci. Comput. 72 (2017) 1093–1118.

[68] W. Auzinger, T. Kassebacher, O. Koch, M. Thalhammer, Convergence of a Strang splitting finite element discretization for the Schrödinger–Poisson equation, ESAIM Math. Model Numer. Anal. 51 (2017) 1245–1278.

[69] H. Zhang, J. Wang, Superconvergence analysis of Crank–Nicolson Galerkin FEMs for a generalized nonlinear Schrödinger equation, Numer. Methods Part. D. E. 34 (2018) 799–820.

[70] E. K. U. Gross, W. Kohn, Local density-functional theory of frequency-dependent linear response, Phys. Rev. Lett. 55 (1985) 2850–2852.

[71] D. M. Ceperley, B. J. Alder, Ground state of the electron gas by a stochastic method, Phys. Rev. Lett. 45 (1980) 566–569.

[72] L. Kleinman, D. M. Bylander, Efficacious form for model pseudopotentials, Phys. Rev. Lett. 48 (1982) 1425–1428.

[73] P. Ciarlet, The Finite Element Method for Elliptic Problems, SIAM, Philadelphia, 2002.

[74] S. Blanes, F. Casas, J. Oteo, J. Ros, The magnus expansion and some of its applications, Phys. Rep. 470 (2009) 151–238.

[75] M. Hochbruck, C. Lubich, On magnus integrators for time-dependent schrdinger equations, SIAM J. Numer. Anal. 41 (2003) 945–963.

[76] K. Bathe, Finite Element Procedures, Prentice Hall International, 1996.

[77] J. P. Boyd, Chebyshev and Fourier spectral methods, Courier Corporation, 2001.

[78] A. Castro, M. A. Marques, A. Rubio, Propagators for the time-dependent Kohn–Sham equations, J. Chem. Phys. 121 (2004) 3425–3433.

[79] L. N. Trefethen, D. Bau III, Numerical linear algebra, Philadelphia, 1997.

[80] M. Hochbruck, C. Lubich, H. Selhofer, Exponential integrators for large systems of differential equations, SIAM J. Sci. Comput. 19 (1998) 1552–1574.

[81] C.-L. Cheng, J. S. Evans, T. Van Voorhis, Simulating molecular conductance using real-time density functional theory, Phys. Rev. B 74 (2006) 155112.

[82] N. Troullier, J. L. Martins, Efficient pseudopotentials for plane-wave calculations, Phys. Rev. B 43 (1991) 1993–2006.

[83] M. Fuchs, M. Scheffler, Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory, Comput. Phys. Commun. 119 (1999) 67–98.

[84] T. Brabec, F. Krausz, Intense few-cycle laser fields: Frontiers of nonlinear optics, Rev. Mod. Phys. 72 (2000) 545.