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

Bikash Kanungo[1] and Vikram Gavini[1,2]
[1]*Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA*
[2]*Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA*

(Received 12 April 2019; revised manuscript received 23 July 2019; published 23 September 2019)

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 semidiscrete (discrete in space but continuous in time) error estimate on the time-dependent Kohn-Sham orbitals, to construct an efficient 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 special reduced order 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, wherein 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. Further, for similar level of accuracy, we obtain significant computational savings by our approach as compared to state-of-the-art finite-difference methods. 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.

DOI: 10.1103/PhysRevB.100.115148

I. INTRODUCTION

Time-dependent density functional theory (TDDFT) extends the key 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. For all practical purposes, it requires the use of approximate exchange-correlation functionals, analogous to the ground-state case. However, TDDFT offers a great balance of accuracy and computational efficiency which have enabled the study of a wide range of time-dependent phenomena: optical [3] and higher order responses [4,5], electron transport [6,7], charge-transfer excitations [8,9], dynamics of chemical bonds [10], multiphoton ionization [11–13], to name a few.

Given the practical significance of TDDFT calculations, there has been a growing interest in developing faster and more accurate numerical methods for solving the TDKS equations over the past two decades. Broadly, these numerical methods can be classified into two categories, characterized by the strength of the light-matter interaction, namely, linear-response time-dependent density functional theory (LR-TDDFT) [14,15] and real time time-dependent density functional theory (RT-TDDFT) [16–18]. 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 nonperturbative regimes (e.g., harmonic generation, electron transport) in a unified manner. This involves propagating the TDKS equations in real time without any restriction to the external field in terms of its frequency, shape, or intensity. This work pertains to the more general RT-TDDFT.

Despite its generality in dealing with various time-dependent processes, there are two major challenges associated with RT-TDDFT. The first stems from the quality of the time-dependent exchange-correlation approximation used in the TDKS equations. The exact exchange-correlation functional is, in general, nonlocal in both space and time [19–21] 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 softwares. 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 nonlinear TDKS equations. Numerical simulations for large length scales and timescales are still computationally challenging, and warrant systematically improvable, accurate, efficient, and scalable spatiotemporal discretization. Addressing these numerical challenges constitutes the main subject of this work.

Significant efforts have been made toward efficient RT-TDDFT numerical schemes as extensions to popular ground-state DFT packages, borrowing from their respective spatial discretization. These include plane-wave 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 plane-wave 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 nonperiodic systems like isolated molecules and nanoclusters. 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 nanoclusters for both pseudopotential as well as all-electron calculations. However, owing to the incompleteness of such basis, systematic convergence for all materials systems remains a concern. The finite-difference discretization (FD) provides systematic convergence, can handle a broad range of boundary conditions, and exhibits improved parallel scalability in comparison to plane-wave and atomic-type orbital basis. However, incorporating adaptive spatial resolution in FD through a nonuniform grid remains nontrivial. The lesser flexibility of FD to vary spatial resolution renders FD less straightforward to employ in the context of singular potentials (as in the case of all-electron calculations). On the other hand, the finite-element basis [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. We add that many of these advantages of finite-element basis are also shared by the wavelets basis [35]. While, at present, the use of wavelets basis has been restricted to LR-TDDFT [36], we expect them to be a competent basis for RT-TDDFT as well.

The efficacy of the finite-element basis in terms of its accuracy, efficiency, scalability, and relative performance with other competing methods (e.g., plane waves, Gaussian basis, FD), have been thoroughly studied in the context of ground-state DFT, for both pseudopotential [37–51] and all-electron calculations [37,48,50–58]. A similarly comprehensive study on the efficacy of the finite-element basis for RT-TDDFT is, however, lacking. While two recent studies [59,60] demonstrate the accuracy of finite elements for RT-TDDFT, they remain restricted to only linear and quadratic finite elements. As known from prior studies in ground-state DFT [48,53,61], lower order (linear and quadratic) finite elements require a large number of basis functions (50 000–500 000 per atom for pseudopotential calculations) to achieve chemical accuracy, and hence, perform poorly in comparison to plane waves 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 [48]. In this work, we extend the use of higher order finite elements to RT-TDDFT calculations and demonstrate the resulting advantages over lower order finite elements as well as finite-difference-based methods.

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

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

clusters (Al_2 and Al_{13}), and the buckminsterfullerene (C60) molecule as our benchmark pseudopotential systems. The finite-difference-based calculations are done using the OCTOPUS package [30]. Depending on the benchmark system, the finite-element discretization shows a 3-fold to 60-fold savings in computational time as compared to the finite-difference approach for pseudopotential calculations. We also demonstrate the efficacy of finite elements for systems subjected to strong perturbation by studying higher harmonic generation in Mg_2. Additionally, we demonstrate the competence of finite elements for all-electron calculations on two benchmark systems: methane and benzene molecule. Lastly, we study the strong scaling of our implementation and observe good parallel scalability with ∼75% efficiency at 768 processors for a benchmark system of a buckminsterfullerene molecule containing 3.5×10^6 degrees of freedom.

The rest of the paper is organized as follows. In Sec. II, we briefly discuss the TDKS equations and the form of the exact time-evolution operator. In Sec. III, we introduce the notion of semidiscrete and full-discrete solutions to the TDKS equation. In Sec. IV, we provide formal spatial- and time-discretization error estimates in the Kohn-Sham orbitals. Section V provides an efficient spatiotemporal-discretization scheme guided by the error estimates. In Sec. VI, we describe the various numerical implementation aspects pertaining to spectral finite elements and the discrete second-order Magnus operator. Section VII details the convergence, accuracy, efficiency, and parallel scalability of the higher order finite elements along with its relative performance against the finite-difference method. Finally, we summarize our findings and outline the future scope in Sec. VIII.

II. 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 noninteracting 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, are given as

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

where $H_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})$, $V_{\mathrm{KS}}[\rho](\mathbf{r},t;\mathbf{R})$, and $\psi_\alpha(\mathbf{r},t)$ represent the time-dependent Kohn-Sham Hamiltonian, potential, and orbitals, respectively, with the index α spanning over all the N_e electrons in the system. $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_a}\}$ 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 this work, we restrict ourselves to only nonperiodic (clusters and molecules) as well as spin-unpolarized systems. However, we note that all the ideas discussed subsequently can be generalized to spin-polarized systems as well.

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

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

where $V_{\mathrm{ext}}(\mathbf{r},t;\mathbf{R})$ denotes the external potential, $V_H[\rho](\mathbf{r},t)$ denotes the Hartree potential, and $V_{\mathrm{XC}}[\rho](\mathbf{r},t)$ represents the exchange-correlation potential. The exchange-correlation potential $V_{\mathrm{XC}}[\rho](\mathbf{r},t)$, in general, is nonlocal in both space and time [19–21], and has a dependence on the initial many-electron wave function [22]. However, in absence of the knowledge of its true form, most of the existing approximations use locality in time (adiabatic exchange correlation) and nondependence on the initial many-electron wave function. This allows for direct use of the existing exchange-correlation approximations used in ground-state DFT. In this work, we use the adiabatic local-density approximation (ALDA) [62], which is local in both space and time. Specifically, we use the Ceperley-Alder form [63].

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 $V_{\mathrm{field}}(\mathbf{r},t)$. The nuclear potential is given by

$$V_N(\mathbf{r};\mathbf{R}) = \begin{cases} V_N^{\mathrm{AE}} = -\sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|} & \text{for all-electron,} \\ V_N^{\mathrm{PSP}}(\mathbf{R}) & \text{for pseudopotential,} \end{cases} \tag{5}$$

where Z_I and $\mathbf{R}_I$ represent the atomic charge and position of the Ith nucleus. For a typical pseudopotential calculation, V_N^{PSP} comprises of a local part $V_{\mathrm{PSP}}^{\mathrm{loc}}$ 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 [64], is given by

$$\begin{aligned} &V_{\mathrm{PSP}}^{nl}(\mathbf{R})\phi(\mathbf{r}) \\ &\quad= \sum_{I=1}^{N_a}\sum_{l=0}^{L_I}\sum_{m=-l}^{l}\left(\frac{\int u_{lm}^I(\mathbf{r}')\delta V_l^I(\mathbf{r}')\phi(\mathbf{r}')\,d\mathbf{r}'}{\int u_{lm}^I(\mathbf{r}')\delta V_l^I(\mathbf{r}')u_{lm}^I(\mathbf{r}')\,d\mathbf{r}'}\right) \\ &\qquad\times \delta V_l^I(\mathbf{r})u_{lm}^I(\mathbf{r}), \end{aligned} \tag{6}$$

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

$$V_{\mathrm{field}}(\mathbf{r},t) = -\mathbf{E_0}(t)\cdot\mathbf{r}, \tag{7}$$

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

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:

$$-\frac{1}{4\pi}\nabla^2 V_H(\mathbf{r},t) = \rho(\mathbf{r},t), \quad V_H(\mathbf{r},t)|_{\partial\Omega} = f(\mathbf{r},\mathbf{R}), \tag{8a}$$

$$-\frac{1}{4\pi}\nabla^2 V_N^{\mathrm{AE}}(\mathbf{r};\mathbf{R}) = b(\mathbf{r},\mathbf{R}), \quad V_N^{\mathrm{AE}}(\mathbf{r})|_{\partial\Omega} = -f(\mathbf{r},\mathbf{R}). \tag{8b}$$

In the above equation, $b(\mathbf{r};\mathbf{R}) = -\sum_{I=1}^{N_a} Z_I \delta(\mathbf{r};\mathbf{R}_I)$, where $\delta(\mathbf{r};\mathbf{R}_I)$ is a bounded regularization of the Dirac-delta distribution with compact support in a small ball around $\mathbf{R}_I$ and satisfies $\int \delta(\mathbf{r};\mathbf{R}_I)\,d\mathbf{r} = 1$; $f(\mathbf{r},R) = \sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$; and $\partial\Omega$ denotes the boundary of a sufficiently large bounded domain $\Omega \in \mathbb{R}^3$. We refer to previous works on finite-element-based ground-state DFT calculations [39,41,44,48,65] 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

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

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:

$$U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0), \quad t_0 \leqslant t_1 \leqslant t_2 \tag{10}$$

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

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

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

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

In addition to resolving the implicit time dependence in $H_{\mathrm{KS}}[\rho](\mathbf{r},t)$, the short-time propagation provides the numerical advantage of containing the norm of the exponent in Eq. (9). To elaborate, any efficient numerical scheme to compute the action of the propagator on a wave function 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 $\omega_{\max}$ that one wants to resolve in their calculations, i.e., $\Delta t_{\max} = \frac{1}{\omega_{\max}}$. Typically, $\omega_{\max}$ 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 $\mathcal{T}\exp\{-i\int_t^{t+\Delta t} H_{\mathrm{KS}}[\rho](\mathbf{r},\tau)d\tau\}$. We discuss these approximations and their associated time-discretization errors in greater detail in Secs. III and IV.

III. SEMIDISCRETE AND FULL-DISCRETE SOLUTIONS

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

To begin with, we provide some of the finite-element essentials. In the finite-element method, the spatial domain of interest ($\Omega \in \mathbb{R}^3$) is divided into nonoverlapping subdomains, known as finite elements. Each finite element (e) is characterized by its spatial extent (Ω_e) and size (h_e). Subsequently, the finite-element basis is constructed from piecewise Lagrange interpolating polynomials that have a compact support on the finite elements (i.e., on Ω_e), thus rendering locality to these basis functions. We note that there is an abundance of choice in terms of the form and order of the polynomial functions that can be used in constructing the finite-element basis. We refer to Refs. [34,66] for a comprehensive discourse on the subject.

A. Semidiscrete solution

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

$$\psi_\alpha^h(\mathbf{r},t) = \sum_{j=1}^{n_h} N_j(\mathbf{r})\psi_\alpha^j(t), \quad \text{s.t. } \psi_\alpha^h(\mathbf{r},t)|_{\partial\Omega} = 0, \quad \forall\, t \geqslant 0 \tag{13}$$

where $\{N_j(\mathbf{r})\}$ represents the set of finite-element basis functions, each of polynomial order p; and $\psi_\alpha^j(t)$ denote the time-dependent expansion coefficient corresponding to the N_j basis function. We refer to the Appendix for a formal discussion on the appropriate function space for $\psi_\alpha^h(\mathbf{r},t)$. Using the discretization of Eq. (13) in the TDKS equation [Eq. (1)] results in following discrete equation:

$$i\mathbf{M}\dot{\boldsymbol{\psi}}_\alpha(t) = \mathbf{H}\boldsymbol{\psi}_\alpha(t), \tag{14}$$

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

$$\begin{aligned}H_{jk} &= \frac{1}{2}\int_\Omega \nabla N_j(\mathbf{r})\cdot\nabla N_k(\mathbf{r})\,d\mathbf{r}\\ &\quad+ \int_\Omega V_{\mathrm{KS}}^h[\rho^h](\mathbf{r},t)N_j(\mathbf{r})N_k(\mathbf{r})\,d\mathbf{r},\end{aligned} \tag{15}$$

where $V_{\mathrm{KS}}^h[\rho^h](\mathbf{r},t)$ is the discrete Kohn-Sham potential corresponding to the semidiscrete density $\rho^h(\mathbf{r},t)$ [i.e., evaluated

from the solutions of Eq. (14)]. $V^h_{\rm KS}[\rho^h](t)$ is, in turn, given by

$$V^h_{\rm KS}[\rho^h](\mathbf{r},t) = V^h_H[\rho^h](\mathbf{r},t) + V^h_N(\mathbf{r}) + V_{\rm XC}[\rho^h](\mathbf{r},t) + V_{\rm field}(\mathbf{r},t), \tag{16}$$

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

$$V^h_H[\rho^h](\mathbf{r},t) = \sum_{j=1}^{n_h} N_j(\mathbf{r})\phi^j_H(t), \tag{17}$$

$$V^{{\rm AE},h}_N(\mathbf{r}) = \sum_{j=1}^{n_h} N_j(\mathbf{r})\phi^j_N, \tag{18}$$

satisfying the boundary conditions presented in Eq. (8) in a discrete sense. We refer to the Appendix for a formal discussion on the appropriate function spaces for $V^h_H[\rho^h](\mathbf{r},t)$ and $V^h_N(\mathbf{r})$. Subsequently, the expansion coefficients $\phi^j_H(t)$ and ϕ^j_N can be obtained by using the above expressions in Eq. (8), and solving the resulting system of linear equations

$$\boldsymbol{K}\boldsymbol{\phi}_H(t) = 4\pi\mathbf{c}(t), \tag{19}$$

$$\boldsymbol{K}\boldsymbol{\phi}_N = 4\pi\mathbf{d}, \tag{20}$$

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

B. Full-discrete solution

We now discuss the full-discrete solution to the TDKS equations, in the context of second-order Magnus propagator. To begin with, we note that the overlap matrix $\mathbf{M}$, being positive definite, guarantees the existence of a unique positive-definite square root $\mathbf{M}^{1/2}$. This allows us to rewrite Eq. (14) as

$$i\dot{\bar{\boldsymbol{\psi}}}_\alpha(t) = \bar{\mathbf{H}}\bar{\boldsymbol{\psi}}_\alpha(t), \tag{21}$$

where $\bar{\boldsymbol{\psi}}_\alpha(t) = \mathbf{M}^{1/2}\boldsymbol{\psi}_\alpha(t)$ and $\bar{\mathbf{H}} = \mathbf{M}^{-1/2}\mathbf{H}\mathbf{M}^{-1/2}$. To put it differently, $\bar{\boldsymbol{\psi}}_\alpha(t)$ is the representation of $\psi^h_\alpha(\mathbf{r},t)$ in a Löwdin orthonormalized basis [67]. We remark that the practicality of the above reformulation in terms of $\bar{\boldsymbol{\psi}}_\alpha$ is contingent upon the efficient evaluation of $\mathbf{M}^{-1/2}$. To that end, we present an efficient scheme for computing $\mathbf{M}^{-1/2}$ in Sec. VI.

We invoke the Magnus ansatz to write the solution of Eq. (21) as

$$\bar{\boldsymbol{\psi}}_\alpha(t) = \exp(\mathbf{A}(t))\bar{\boldsymbol{\psi}}_\alpha(0), \quad \forall\, t \geqslant 0. \tag{22}$$

The operator $\exp(\mathbf{A}(t))$ is termed as the *Magnus propagator*, where $\mathbf{A}(t)$ is given explicitly as [68,69]

$$\mathbf{A}(t) = \int_0^t -i\bar{\mathbf{H}}(\tau)d\tau - \frac{1}{2}\int_0^t\left[\int_0^\tau -i\bar{\mathbf{H}}(\sigma)d\sigma, -i\bar{\mathbf{H}}(\tau)\right]d\tau + \cdots, \tag{23}$$

where $[\mathbf{X},\mathbf{Y}] = \mathbf{XY} - \mathbf{YX}$ denotes the commutator. Although known explicitly, the above equation is not practically useful, given the difficulty in resolving the implicit dependence of $\bar{\mathbf{H}}(t)$ on $\rho^h(\mathbf{r},t)$. As mentioned in Sec. II, we resolve the implicit dependence by using the composition property of a propagator [cf. Eq. (11)]. This allows us to rewrite the Magnus propagator as

$$\exp(\mathbf{A}(t)) = \prod_{n=1}^{N}\exp(\mathbf{A}_n), \tag{24}$$

where $\mathbf{A}_n$ is given by Eq. (23), albeit with the limits of integration modified to $[t_{n-1},t_n]$.

In practice, one replaces the exact $\mathbf{A}_n$ with an approximate operator $\tilde{\mathbf{A}}_n$, which involves, first, a truncation of the Magnus expansion [defined in Eq. (23)], and second, an approximation for the time integrals in the truncated Magnus expansion. Truncating the Magnus expansion after the first p terms results in a time-integration scheme of order $2p$. In this work, we restrict ourselves to the second-order Magnus propagator, i.e., obtained by truncating the Magnus expansion after the first term. Furthermore, we use a midpoint integration rule to evaluate $\int_{t_{n-1}}^{t_n} -i\bar{\mathbf{H}}(\tau)d\tau$. In particular, the action of the second-order Magnus propagator, with a midpoint integration rule, on the set of Kohn-Sham orbitals $\{\bar{\boldsymbol{\psi}}_1,\bar{\boldsymbol{\psi}}_2,\ldots,\bar{\boldsymbol{\psi}}_{N_e}\}$ which defines the density $\rho^h(\mathbf{r},t)$, is given by

$$e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}}_\alpha(t) = e^{-i\bar{\mathbf{H}}[\rho^h(t_{n-1}+\frac{\Delta t}{2})]\Delta t}\bar{\boldsymbol{\psi}}_\alpha(t), \tag{25}$$

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

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

$$\bar{\boldsymbol{\psi}}^n_\alpha = \exp(\tilde{\mathbf{A}}_n)\bar{\boldsymbol{\psi}}^{n-1}_\alpha. \tag{26}$$

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

IV. DISCRETIZATION ERRORS

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

$$\psi_\alpha(\mathbf{r},t_n) - \psi^{h,n}_\alpha(\mathbf{r}) = \big(\psi_\alpha(\mathbf{r},t_n) - \psi^h_\alpha(\mathbf{r},t_n)\big) + \big(\psi^h_\alpha(\mathbf{r},t_n) - \psi^{h,n}_\alpha(\mathbf{r})\big). \tag{27}$$

In the following two subsections, we present error estimates on the two right-hand terms in the above equation, respectively.

A. Spatial-discretization error

If $\psi_\alpha^h(\mathbf{r},t)$ denotes the semidiscrete solution to $\psi_\alpha(\mathbf{r},t)$, then the following bound holds true (see Appendix for the derivation):

$$\begin{aligned}\sum_\alpha^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) &\leqslant C(t)\sum_e h_e^p \sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)\\ &\quad + C(t)\sum_e h_e^p\left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right)\\ &\quad \text{for some } \{s_{1,\alpha}\}, \{s_{2,\alpha}\}, \text{and } s_3 \in [0,t]. \end{aligned} \tag{28}$$

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

B. Time-discretization error

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

$$\left\|\psi_\alpha^h(t_n) - \psi_\alpha^{h,n}\right\|_{L^2(\Omega)} \leqslant C(\Delta t)^2 t_n \max_{0\leqslant t\leqslant t_n}\left\|\psi_\alpha^h(t)\right\|_{H^1(\Omega)}. \tag{29}$$

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

V. EFFICIENT SPATIOTEMPORAL DISCRETIZATION

We now utilize our spatial- and temporal-discretization error estimates [Eqs. (28) and (29)] to obtain an efficient spatiotemporal discretization. We follow along the lines of [48,65,70] to obtain an efficient spatial discretization by minimizing the semidiscrete error in the dipole moment at a given time, subject to a fixed number of finite elements. We remark that the choice of dipole moment as an observable for this exercise is solely a matter of convenience, and any observable which can be inexpensively evaluated in terms of the density or the Kohn-Sham orbitals can be used instead. Representing the x component of the continuous and the semidiscrete dipole moments at time t as $\mu_x(t)$ and $\mu_x^h(t)$, respectively, we have

$$\begin{aligned}\left|\mu_x(t) - \mu_x^h(t)\right| &\leqslant \|x\|_{L^2(\Omega)}\|\rho(t) - \rho^h(t)\|_{L^2(\Omega)}\\ &\leqslant C\|\rho(t) - \rho^h(t)\|_{L^2(\Omega)}\\ &\leqslant C\sum_{\alpha=1}^{N_e}\left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t).\end{aligned} \tag{30}$$

Now, using Eq. (28) in the above equation results in

$$\begin{aligned}\left|\mu_x(t) - \mu_x^h(t)\right| &\leqslant C_1(t)\sum_e h_e^p\sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)\\ &\quad + C_1(t)\sum_e h_e^p\left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right)\end{aligned} \tag{31}$$

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

$$\begin{aligned}\left|\mu_x(t) - \mu_x^h(t)\right| &\leqslant C_1(t)\int_\Omega h^p(\mathbf{r})\left[\sum_{\alpha=1}^{N_e}(\bar{D}^{p+1}\psi_\alpha(s_{1,\alpha}) + \bar{D}^{p+1}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+3}\psi_\alpha(s_{2,\alpha}))\right]d\mathbf{r}\\ &\quad + C_1(t)\int_\Omega h^p(\mathbf{r})(\bar{D}^{p+1}V_H[\rho^h] + \bar{D}^{p+1}V_N)\,d\mathbf{r},\end{aligned} \tag{32}$$

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

$$\begin{aligned}\min_{h(\mathbf{r})}\int_\Omega h^p(\mathbf{r})\left[\sum_{\alpha=1}^{N_e}(\bar{D}^{p+1}\psi_\alpha(s_{1,\alpha}) + \bar{D}^{p+1}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+3}\psi_\alpha(s_{2,\alpha})) + \bar{D}^{p+1}V_H[\rho^h](s_3) + \bar{D}^{p+1}V_N\right]d\mathbf{r}\\ \text{subject to}\quad \int_\Omega \frac{d\mathbf{r}}{h^3(\mathbf{r})} = N_{\text{elem}}.\end{aligned} \tag{33}$$

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

$$h(\mathbf{r}) = E\left[\sum_{\alpha=1}^{N_e}(\bar{D}^{p+1}\psi_\alpha(s_{1,\alpha}) + \bar{D}^{p+1}\psi_\alpha(s_{2,\alpha}) + \bar{D}^{p+3}\psi_\alpha(s_{2,\alpha})) + \bar{D}^{p+1}V_H[\rho^h](s_3) + \bar{D}^{p+1}V_N\right]^{-1/p+3}, \quad (34)$$

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 ground-state atomic solutions to the Kohn-Sham orbitals and the electrostatic potentials to construct the adaptive finite-element mesh for all calculations. Although this does not inform us about the optimal discretization required near the nuclei, this affords an efficient strategy to accurately handle the regions away from the nuclei, wherein the time-dependent Kohn-Sham orbitals, typically, have similar decay properties as their ground-state counterparts. We note that, in practice, the finite-element mesh obtained can deviate from $h(\mathbf{r})$ due to conformity and quality requirements, especially in the context of hexahedral elements that are employed in this work. However, the resulting finite-element mesh broadly captures the optimal coarse-graining rate, and has the general adaptive characteristics that significantly enhance the computational efficiency, as demonstrated in Sec. VII. Figure 1 shows an adaptive mesh for all-electron Al_2 generated using this approach.

Having determined the required spatial discretization, a suitable temporal discretization for the given finite-element mesh can be estimated by using the time-discrete error bound for the dipole moment. To elaborate, if $\mu_x^{h,n}$ denotes the x component of the full-discrete dipole moment, then using the result in Eq. (29) it is straightforward to show

$$\left|\mu_x^h(t_n) - \mu_x^{h,n}\right| \leqslant Ct_n(\Delta t)^2 \sum_{\alpha=1}^{N_e} \max_{0\leqslant t\leqslant t_n} \left\|\psi_\alpha^h(t)\right\|_{H^1(\Omega)}. \quad (35)$$

As is evident from the above relation, our choice of Δt is intrinsically tied to the spatial discretization through $\|\psi_\alpha^h(t)\|_{H^1(\Omega)}$. Furthermore, we remark that although the presence of t_n in the above inequality indicates increasing time-discretization error with time, it does not pose a limitation in containing the errors due to the fact that $t_n \leqslant T$, 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 $\|\psi_\alpha^h(t)\|_{H^1(\Omega)}$ and the value of the constant C featuring in it. The value of $\|\psi_\alpha^h(t)\|_{H^1(\Omega)}$ can be reliably approximated from its ground-state value, i.e., $\|\psi_\alpha^h(0)\|_{H^1(\Omega)}$. The characteristic value of the constant C 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 $|\mu_x^h(t_n) - \mu_x^{h,n}|$ with respect to Δt. For a multiatom system, we use the least Δt obtained for each of its constituent atomic species.

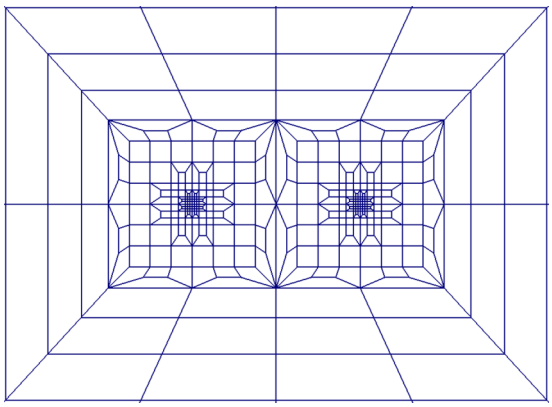

FIG. 1. Adaptive finite-element mesh for all-electron Al_2 (slice shown on the plane of the molecule).

VI. NUMERICAL IMPLEMENTATION

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

A. Higher order spectral finite elements

Finite elements, with their varied choices of forms and orders [34,66], 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 (∼100 000) to achieve chemical accuracy in ground-state DFT calculations [53,61]. However, this shortcoming has been demonstrably mitigated by the use of higher order finite elements [48,65]. 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 [71]. 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 [71], and have been used to gain significant computational efficiency in ground-state DFT calculations [65]. 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 [cf. Eq. (21)]. We note that while an n-point rule in the conventional Gauss quadrature rule integrates polynomials exactly up to degree $2n-1$, an n-point GLL quadrature rule integrates polynomials exactly only up to degree $2n-3$. Therefore, we employ the GLL quadrature rule only in the construction of $\mathbf{M}$, while the more accurate Gauss quadrature rule is used for all other integrals featuring in the discrete TDKS equations. We refer to Motamarri *et al.* [48] 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.

B. Approximating the second-order Magnus operator

The form of the Magnus operator, as shown in Eq. (26), calls for efficient means of evaluating $\exp(\tilde{\mathbf{A}}_n)\bar{\boldsymbol{\psi}}_\alpha^{n-1}$. Direct means of evaluating the matrix $\exp(\tilde{\mathbf{A}}_n)$ remain computationally prohibitive beyond small sizes of $\tilde{\mathbf{A}}_n$. Alternatively, one can attempt to evaluate the action of $\exp(\tilde{\mathbf{A}}_n)$ on $\bar{\boldsymbol{\psi}}_\alpha^{n-1}$ in an iterative fashion. Several such schemes are in use in RT-TDDFT calculations, namely, Taylor series expansion, Chebyshev polynomial expansion, split-operator techniques, and Krylov subspace projection method. We refer to Castro *et al.* [72] 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 use *a priori* estimates, based on the spectral radius of $\tilde{\mathbf{A}}_n$, the number of terms required in the polynomial expansion, for a desired accuracy, remain highly overestimated. 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 parts 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. VII D).

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

$$\mathcal{K}_k(\tilde{\mathbf{A}}_n, \bar{\boldsymbol{\psi}}) = \text{span}\{\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n\bar{\boldsymbol{\psi}}, \tilde{\mathbf{A}}_n^2\bar{\boldsymbol{\psi}}, \ldots, \tilde{\mathbf{A}}_n^{k-1}\bar{\boldsymbol{\psi}}\}. \tag{36}$$

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

$$z_k = \|\bar{\boldsymbol{\psi}}\|\mathbf{Q}_k e^{\mathbf{Q}_k^T\tilde{\mathbf{A}}_n\mathbf{Q}_k}e_1 = \|\bar{\boldsymbol{\psi}}\|\mathbf{Q}_k e^{\mathbf{T}_k}e_1, \tag{37}$$

where $\mathbf{T}_k = \mathbf{Q}_k^T\tilde{\mathbf{A}}_n\mathbf{Q}_k$ is a tridiagonal matrix, and e_1 is the 1st unit vector in $\mathbb{C}^k$. As is evident from the above form, the problem is now reduced to the evaluation of $\exp(\mathbf{T}_k)$, wherein $\mathbf{T}_k$ is a small matrix of size $k \times k$ and, hence, $\exp(\mathbf{T}_k)$ can be evaluated inexpensively either through Taylor series expansion or exact eigenvalue decomposition of $\mathbf{T}_k$. The error ϵ_k incurred in the above approximation is given by [73]

$$\epsilon_k = \|e^{\tilde{\mathbf{A}}_n}\bar{\boldsymbol{\psi}} - \|\bar{\boldsymbol{\psi}}\|\mathbf{Q}_k e^{\mathbf{T}_k}e_1\| \approx \beta_{k+1,k}\|\bar{\boldsymbol{\psi}}\||[e^{\mathbf{T}_k}]_{k,1}|, \tag{38}$$

where $\beta_{k+1,k}$ is the $(k+1,k)$ entry of $\mathbf{T}_{k+1} = \mathbf{Q}_{k+1}^T\tilde{\mathbf{A}}_n\mathbf{Q}_{k+1}$. Thus, the above relation provides a robust and inexpensive scheme to adaptively determine the dimension of the Krylov subspace by checking if ϵ_k is below a set tolerance. An economic choice for the tolerance for ϵ_k is determined from atomic RT-TDDFT calculations, such that it achieves <10 meV accuracy in the excitation energies. For a multiatom 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. III B, the use of second-order Magnus propagator with midpoint integration rule, i.e., $e^{\tilde{\mathbf{A}}_n}$, requires the knowledge of $\tilde{\mathbf{H}}$ at a future time instant, i.e., $\tilde{\mathbf{H}}[t_{n-1} + \Delta t/2]$, which is *a priori* unknown. A fully consistent approach involves, for a given $\bar{\boldsymbol{\psi}}^{n-1}$, the following steps: (i) approximate $\tilde{\mathbf{H}}[t_{n-1} + \Delta t/2]$ through extrapolation over previous instants of $\tilde{\mathbf{H}}$; (ii) use it to obtain $\bar{\boldsymbol{\psi}}^n$, and then evaluate $\tilde{\mathbf{H}}[t_n]$; (iii) reevaluate $\tilde{\mathbf{H}}[t_{n-1} + \Delta t/2]$ by interpolating between $\tilde{\mathbf{H}}[t_{n-1}]$ and $\tilde{\mathbf{H}}[t_n]$; and (iv) repeat steps (ii) and (iii) until convergence. Although robust and accurate, this approach comes at a huge computational cost arising out of the Lanczos procedure at each iterate of the self-consistent iteration. An efficient and sufficiently accurate approach is to use a predictor-corrector method to, first, predict $\tilde{\mathbf{H}}[t_{n-1} + \Delta t/4]$ through an extrapolation (linear or higher order) from previous instants of $\tilde{\mathbf{H}}$, use it to propagate $\bar{\boldsymbol{\psi}}^{n-1}$ to $\bar{\boldsymbol{\psi}}^{n-1/2}$, which is then used to evaluate $\tilde{\mathbf{H}}[t_{n-1} + \Delta t/2]$. We refer to [74] for the details of the predictor-corrector scheme. We remark that this predictor-corrector scheme is accurate to $\mathcal{O}(\Delta t^2)$ and, hence, does not affect the results of our time-discretization error estimates.

VII. 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, respectively. For the pseudopotential calculations, we provide comparison, in terms of accuracy and performance, of the higher order finite elements against the finite-difference method. The

finite-difference-based calculations are performed using the OCTOPUS [30] software package. In all our finite-difference-based calculations, we have used a stencil of order 4 in each direction (default stencil order in OCTOPUS). All the pseudopotential calculations are done using the norm-conserving Troullier-Martins pseudopotentials [75]. For all calculations, the ground-state Kohn-Sham orbitals are used as the initial states. We use the Chebyshev filter acceleration technique (refer [48,76,77]) 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.

A. Rates of convergence

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

We note that in order to study the convergence with respect to mesh size, the dominant error must arise from spatial discretization. To this end, we contain other sources of error, namely, time-discretization error and Krylov subspace projection error, by choosing a very small time step of $\Delta t = 10^{-4}$, and using a small tolerance of 10^{-12} for the Krylov subspace error [cf. Eq. (38)]. In effect, we mimic a semidiscrete (discrete in space but continuous in time) error analysis. We employ finite elements of three different orders (p), HEX8, HEX27, and HEX125SPEC, in all our convergence studies. For each p, we start with a given N_{elem} and use the superposition of ground-state atomic orbitals and electrostatic potentials, to determine the coarsening rate of the mesh, as per Eq. (34). The resultant mesh forms the coarsest mesh. Subsequently, we increase the value of N_{elem} to obtain a sequence of increasingly refined meshes. We remark that although we propose for using the ground-state electronic fields to determine the mesh-coarsening rate, it nevertheless forms a reliable and cost effective way for discretizing the mesh as opposed to any *ad hoc* coarsening or using uniform discretization. Furthermore, it allows us to use large computational domain sizes without significantly increasing the number of elements. This, in turn, allows us to circumvent the need for artificial absorbing boundary conditions, otherwise essential to tackle wave reflection effects that are observed while dealing with small computational domains.

In order to perform the convergence study with respect to mesh size, and compare to the semidiscrete error estimate obtained in Eq. (31), we require the knowledge of the continuous value of the dipole moment $\mu_x(t)$. To this end, we use the discrete dipole moment $\mu_x^h(t)$, obtained from a sequence of increasingly refined HEX125SPEC finite-element meshes, to obtain a least-square fit of the form

$$\frac{\left|\mu_x(t) - \mu_x^h(t)\right|}{\left|\mu_x(t)\right|} = C h_{\text{min}}^q \tag{39}$$

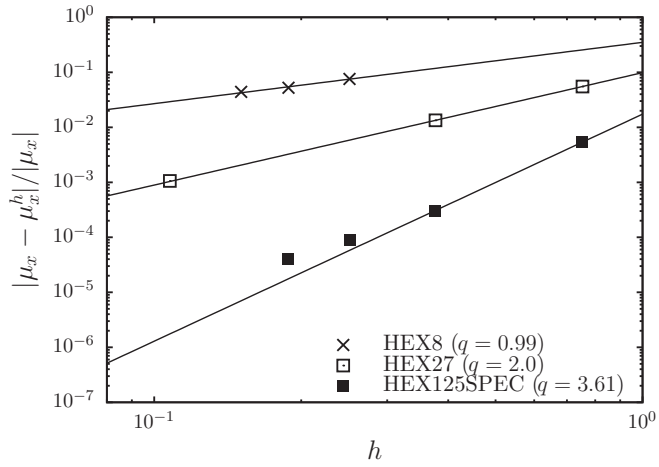

FIG. 2. Rates of convergence with respect to spatial discretization for LiH.

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

Next, we consider the convergence with respect to temporal discretization, i.e., Δt. To this end, we use a sufficiently refined HEX125SPEC finite-element mesh and use increasingly refined Δt to obtain a least-square fit of the form

$$\frac{\left|\mu_x^h(t_n) - \mu_x^{h,n}\right|}{\left|\mu_x^h(t_n)\right|} = C(\Delta t)^q \tag{40}$$

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.

1. All-electron calculations: Lithium hydride

In this example, we conduct all-electron RT-TDDFT study on a lithium hydride molecule (LiH) with Li-H bond length of 3.014 a.u. A large cubical domain of length of 50 a.u. is chosen to ensure that the electron density decays to zero on the domain boundary, thereby allowing us to impose Dirichlet boundary condition on the time-dependent Kohn-Sham orbitals and the Hartree potential. We use the aforementioned adaptive mesh generation strategy to construct a sequence of HEX8, HEX27, and HEX125SPEC meshes. For all the meshes under consideration, we first obtain the ground state and employ a weak delta kick to excite the system. To elaborate, we use an electric field of the form $\mathbf{E}_0(t) = \kappa\delta(t)\hat{x}$, with $\kappa = 10^{-3}$ a.u., where $\delta(t)$ is the Dirac-delta distribution and $\hat{x}$ is the unit vector along the 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 $\psi_\alpha(0) = e^{-i\kappa x}\psi_\alpha^{\text{GS}}$. Figure 2 depicts the rates of convergence for the dipole moment at $t = 1.0$ a.u. for different orders of finite elements. For all the three types of finite elements under consideration, we observe close to optimal rates of convergence, $\mathcal{O}(h^p)$, where

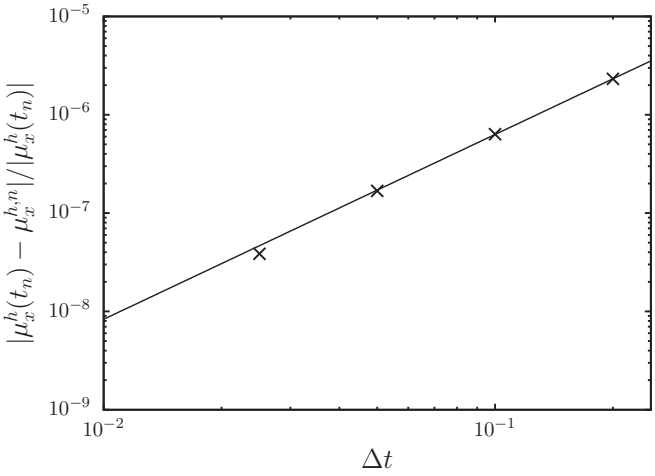

FIG. 3. Rate of convergence with respect to temporal discretization for LiH.

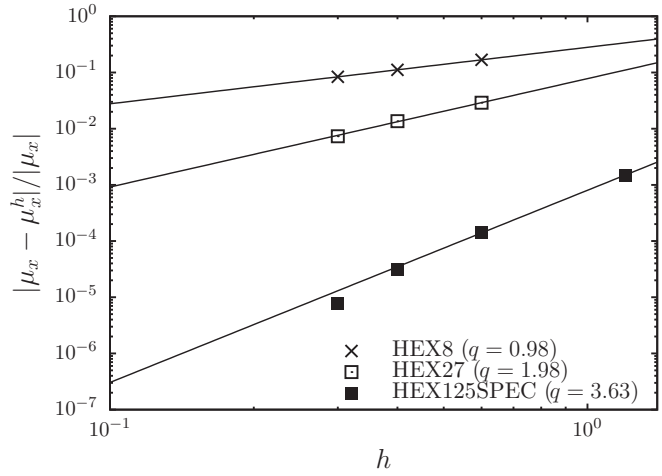

FIG. 4. Rates of convergence with respect to spatial discretization for CH_4.

p is the degree of the finite-element interpolating polynomial. As is evident from Fig. 2, much higher accuracies are obtained with HEX125SPEC when compared to HEX8 and HEX27 of the same mesh size. In particular, 200 000 HEX8 elements (210 644 degrees of freedom) are required to achieve relative errors of 10^{-2}, whereas we achieve relative error of 10^{-3} with just 3000 HEX125SPEC elements (83 156 degrees of freedom). Next, we study the rate of convergence of the dipole moment with respect to temporal discretization. To this end, we used a sufficiently refined HEX125SPEC mesh which affords 10^{-4} relative error with respect to spatial discretization. We then propagate the initial states using second-order Magnus propagator with different Δt. Figure 3 depicts the rate of convergence of the dipole moment with respect to Δt at $t_n = 1.0$ a.u. We obtain a rate of convergence of $q = 1.96$ [defined in Eq. (40)], which agrees remarkably well with the quadratic rate of convergence for second-order Magnus propagator [cf. Eq. (35)].

2. Pseudopotential calculations: Methane (CH_4)

We now turn to examining rates of convergence for the pseudopotential case. We use a methane molecule with C-H bond length of 2.07846 a.u. and a H-C-H tetrahedral angle of 109.4712° as our benchmark system. Similar to lithium hydride, we use the ground-state single-atom electronic fields to obtain a sequence of adaptively refined HEX8, HEX27, and HEX125SPEC meshes. We, once again, make use of a large cubical domain of length 50 a.u. to mimic simulations in $\mathbb{R}^3$. For all the meshes, we first obtain the ground state and then excite the system using a Gaussian electric field of the form $\mathbf{E}_0(t) = \kappa e^{(t-t_0)^2/w^2}\hat{x}$, with $\kappa = 2 \times 10^{-5}$ a.u., $t_0 = 3.0$ a.u., and $w = 0.2$ a.u. Figure 4 illustrates the rates of convergence of the dipole moment at $t = 5.0$ a.u. for different orders of finite elements. As in the case of lithium hydride, we obtain close to optimal rates of convergence, and observe significantly higher accuracies for HEX125SPEC over HEX8 and HEX27. Next, we study the rate of convergence afforded by the second-order Magnus propagator with respect to the time step using a sufficiently refined HEX125SPEC mesh. We propagate the ground-state Kohn-Sham orbitals under the influence of the same Gaussian electric field using different Δt. Figure 5 shows the rate of convergence of the dipole moment with respect to Δt at $t_n = 5.0$ a.u. As was the case with lithium hydride, we obtain a convergence rate of $q = 1.98$, which is remarkably close to the optimal (i.e., quadratic) rate of convergence [cf. Eq. (35)].

B. Computational cost

In this section, we investigate the relative computational efficiency afforded by higher order finite elements over linear finite element. We consider the previous two systems, lithium hydride and methane, for all-electron and pseudopotential calculations, respectively. We use the same mesh adaption strategy as detailed in Sec. VII A. Since the objective of this study is to compare the relative performance of various orders of finite elements, we eliminate any time-discretization effect by setting $\Delta t = 10^{-4}$. Furthermore, we use a tolerance of 10^{-12} for the adaptive Lanczos [cf. Eq. (38)] in order to

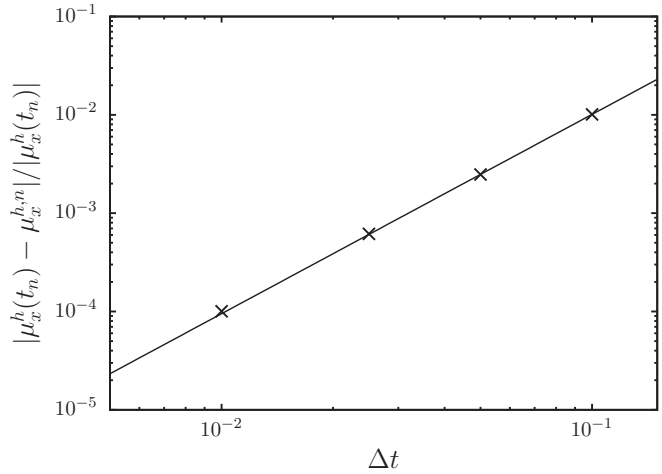

FIG. 5. Rate of convergence with respect to temporal discretization for CH_4.

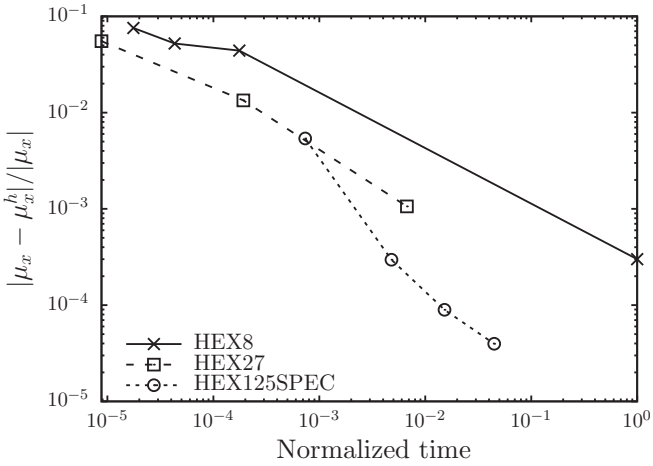

FIG. 6. Computational efficiency of various orders of finite elements for LiH.

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. VII A 1), and the methane molecule with a Gaussian electric field (see Sec. VII A 2). Figures 6 and 7 show the relative error in the dipole moment against the normalized computational time, for three different orders of finite elements. The normalization of the computational time is done with respect to the longest time among the various meshes under consideration. As is evident, the relative computational efficiency afforded by higher order finite elements improves as the desired accuracy is increased. In particular, for a relative accuracy of 10^{-3}, HEX125SPEC outperforms HEX8 and HEX27 by factor 150–200 and 10–18, respectively. This underscores the efficacy of higher order finite elements for RT-TDDFT calculations, an aspect which had, heretofore, remained unexplored for RT-TDDFT.

C. Other materials systems

In this section, we investigate the accuracy and computational efficiency afforded by higher order finite elements

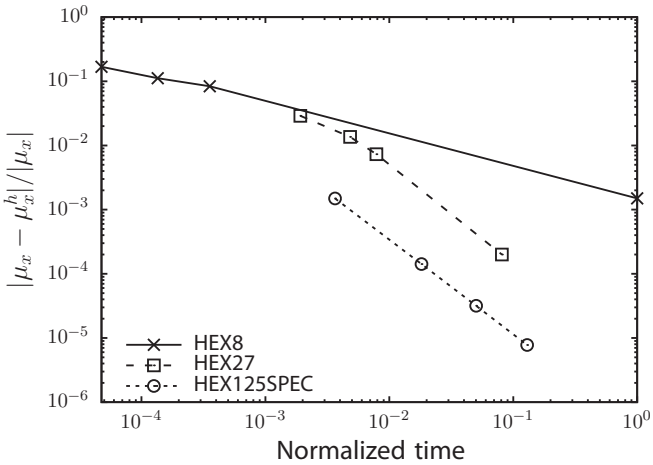

FIG. 7. Computational efficiency of various orders of finite elements for CH_4.

TABLE I. Simulation details for both pseudopotential (PSP) and all-electron (AE) benchmark systems: type of the electric field $\mathbf{E}_0(t)$; time step (Δt in a.u.); tolerance for Krylov subspace projection error [ϵ, cf. Eq. (38)]; total duration of simulation (T in fs).

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

[a] $\mathbf{E}_0(t) = \kappa e^{(t-t_0)^2/\omega^2}\hat{x}$, with $\kappa = 2 \times 10^{-5}$, $t_0 = 3.0$, and $\omega = 0.2$ (all in a.u.).
[b] $\mathbf{E}_0(t) = \kappa \sin^2(\pi/T)\sin(\omega t)\hat{x}$, with $\kappa = 0.01$, $\omega = 0.03$, $T = 5 \times (2\pi/\omega)$ (all in a.u.).

for other materials systems, in both pseudopotential and all-electron RT-TDDFT calculations. We use Al_2, Al_{13}, and Mg_2 as the benchmark metallic systems for pseudopotential calculations. Furthermore, we use buckminsterfullerene (C_{60}) as our benchmark insulating system for pseudopotential calculations. For the all-electron case, we use methane and benzene as our benchmark systems. Additionally, for the all-electron calculations we provide a comparison, in the absorption spectrum, with their pseudopotential counterparts. For all the above systems under consideration, except Mg_2, we use weak electric fields to excite them. For Mg_2, we use a strong laser pulse to study the efficacy of higher order finite elements for nonlinear response. Table I lists the important simulation parameters, for all the benchmark systems under consideration. We remark that the Δt as well as the tolerance for Krylov subspace projection error are chosen such that they achieve <10 meV accuracy in the excitation energies. For pseudopotential systems, we also provide comparison, wherever possible, against calculations based on a finite-difference discretization, by employing the same propagator (i.e., second-order Magnus) and simulation details (as listed in Table I). To this end, we use the OCTOPUS [30] software package to perform the finite-difference-based calculations.

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

size until it achieves <10 meV accuracy in the excitation energies (defined in Sec. VII C 1). The calculation based on the resulting OCTOPUS mesh is considered as the point of comparison (for both accuracy and efficiency) against the corresponding finite-element-based calculation. We add that, while dealing with uniform mesh, a typical workaround to the large domain requirement is to use a smaller domain with absorbing boundary conditions. Hence, to better assess the effects of absorbing boundary conditions, we employ them in finite-difference-based calculations for some of the benchmark systems discussed below.

1. Pseudopotential calculations: Al_2

We consider an aluminum dimer (Al_2) of bond length 4.74 a.u. In order to generate a suitable mesh, we use an adaptive HEX64SPEC finite-elements discretization that follows the coarsening rate obtained from Eq. (34) and is commensurate with an accuracy of 10 meV in the ground-state energy per atom. We use a cubical domain of length 60 a.u. to ensure that the wave functions decay to zero, and thereby avoid any reflection effects. We excite the ground state using the simulation parameters listed in Table I. We use the Fourier transform of the dipole moment to obtain the dynamic polarizability $\alpha_{a,b}(\omega)$, where a is the index of the electric field's polarization direction and b is the index of the measurement direction of the dipole. Subsequently, we obtain the absorption spectrum (dipole strength function) $S(\omega)$, given by $S(\omega) = \frac{2\omega}{3\pi}\mathrm{Tr}[\mathrm{Im}[\boldsymbol{\alpha}(\omega)]]$. The peaks in the absorption spectrum correspond to the excitation energies. We also assess the performance of higher order finite elements by comparing against the finite-difference scheme of OCTOPUS [30]. In order to highlight the effects of domain size for the finite-difference mesh, we use three cubical domains of sizes 38, 46, and 52 a.u., all with a grid spacing of 0.2 a.u. Furthermore, to understand the effect of absorbing boundary conditions, we perform an additional finite-difference calculation on the 38-a.u. mesh with a negative imaginary potential (NIP) near the boundaries. In particular, we use a potential of the form

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

with $\eta = 0.4$, $L = 18.0$, and $\Delta L = 1.0$ (all in atomic units). For clarity, we refer to the four finite-difference calculations, namely, with domain size 52 a.u., with domain size 46 a.u., with domain size 38 a.u., and with domain size 38 a.u. along with NIP absorbing boundary condition as FD-52, FD-46, FD-38, and FD-38-ABS, respectively. We use the same simulation details, namely, time step, duration of propagation, choice of propagator, and tolerance for Krylov subspace, as used for the finite-element case. Figure 8 compares absorption spectrum obtained from finite elements against finite difference. We have used a Gaussian window of the form $g(t) = e^{-\alpha t^2}$, with $\alpha = 0.005$ a.u., in the Fourier transform of the dipole moment to artificially broaden the peaks. As is evident from the figure, we get good agreement with the finite-difference-based results for FD-46 and FD-52. The finite-difference calculation with domain size 38 a.u., with and without the absorbing boundary condition, provides qualitatively different results with two peaks around 5 eV. We attribute this discrepancy to possible

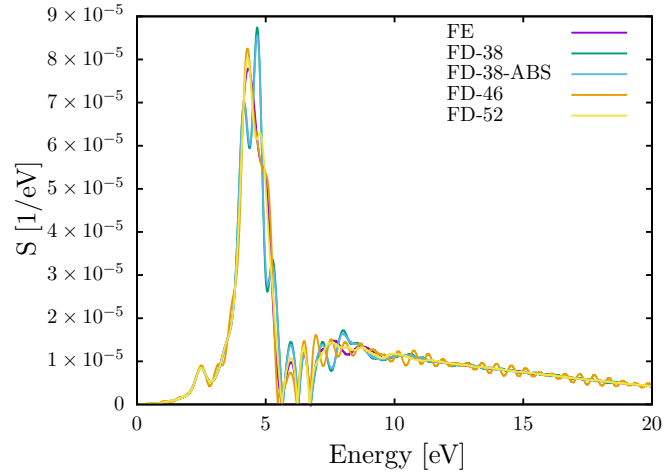

FIG. 8. Absorption spectra for Al_2.

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 II compares the first two excitation peaks, the degrees of freedom and the total computational time for the finite-element- and the finite-difference-based (46-a.u. domain size) 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 resolution of the mesh, as opposed to a uniform mesh in finite difference. We underline this by noting that while finite difference requires over 12×10^6 degrees of freedom, the finite elements require only 31 411 degrees of freedom to attain similar accuracies. Finally, comparing FD-46 and FD-52 according to the FE results, we observe that although we achieve convergence in the excitation energies by 46 a.u., the convergence of the peak values requires a larger domain.

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

TABLE II. 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

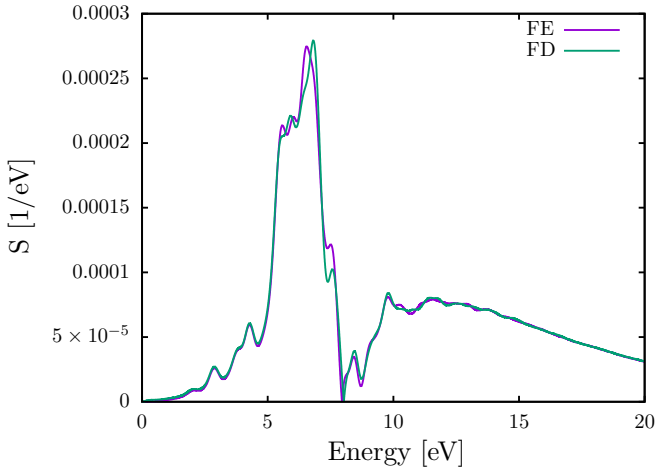

FIG. 9. Absorption spectra for Al_{13}.

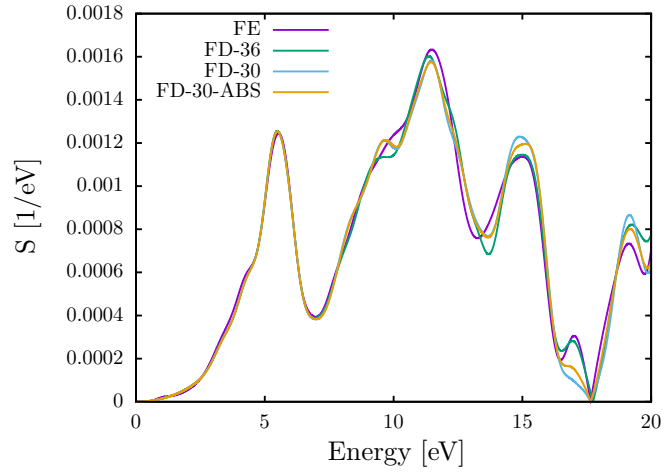

FIG. 10. Absorption spectra of buckminsterfullerene.

of length 70 a.u. to avoid reflection effects. We excite the system from its ground state using the parameters listed in Table I. 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. Figure 9 compares absorption spectrum obtained from finite elements against finite difference. We have used the same Gaussian window as in the case of Al_2. As is evident from the figure, the peaks for both finite element and finite difference are in good agreement. Table III compares the first two excitation peaks, degrees of freedom, and the total computational time for the finite-element- and the finite-difference-based calculations. Both the methods agree to within 10 meV in the first two excitation energies. In terms of computational efficiency, the finite elements attain an ∼8-fold savings in the computational time against finite difference, once again, attributed to the fewer degrees of freedom in finite elements owing to adaptive resolution of the mesh. In particular, the finite elements afford ∼30-fold fewer degrees of freedom as compared to finite difference.

3. Pseudopotential calculations: Buckminsterfullerene

In this example, we consider the buckminsterfullerene molecule comprising of 60 carbon atoms (240 electrons) packed into the shape of a buckyball. As with Al_2, we use an adaptive HEX64SPEC finite-elements discretization that follows the coarsening rate obtained from Eq. (34) and is commensurate with an accuracy of 10 meV in the ground-state energy per atom. We use a cubical domain of length 50 a.u. to eliminate any reflection effects from the boundaries. We use the simulation parameters listed in Table I to excite the system from its ground state. As with previous cases, we also assess the performance of higher order finite elements by comparing against the 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. VII C 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 the finite-element-based and the three different finite-difference-based calculations. We have used the same Gaussian window of the form $g(t) = e^{-\alpha t^2}$, with $\alpha = 0.01$ a.u., to artificially broaden the peaks. As is evident from the figure, there is good agreement between the finite element and FD-36 for all the excitation peaks. On the other hand, while FD-30 and FD-30-ABS have good agreement with finite elements for the first two peaks, they differ for the rest, possibly because of reflection effects. Furthermore, comparing FD-30 and FD-30-ABS, we remark that the use of NIP-based absorbing boundary condition did not improve the absorption spectrum. This, once again, indicates that one cannot always dispense with the need for a larger domain by, solely, using absorbing boundary conditions. Table IV compares the first two excitation peaks, degrees of freedom, and the computational time for finite elements against that of FD-36. Both finite-element-based and FD-36-based results match within 30 meV in the first two peaks. Furthermore, the excitation energies are also in good agreement with results presented in Ref. [78] (the first two excitation peaks, as we estimate from the absorption spectrum reported in Ref. [78], are ∼5.6 eV and ∼11.5 eV, respectively.). In terms of computational efficiency, finite elements attain a ∼3-fold speedup over FD-36. This higher efficiency of the finite elements is,

TABLE III. 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

TABLE IV. 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	5200
FD	5.476	11.439	13 997 521	15 361

once again, attributed to a ∼9-fold fewer degrees of freedom required by the finite elements against that of finite difference.

4. Pseudopotential calculations: Mg_2

In this example, we study the higher harmonic generation in a magnesium dimer with bond length of 4.74 a.u. Unlike the previous examples, we use a strong laser pulse to excite the system from its ground state (see Table I for the simulation details). We use an adaptive HEX125SPEC mesh with the coarsening rate determined by Eq. (34). Furthermore, we use a cubical domain of length 100 a.u. to eliminate any reflection effects from the boundaries. We obtain the dipole power spectrum $P(\omega)$ of the system by taking the imaginary part of the Fourier transform of the acceleration of the dipole moment $\mu(t)$. To elaborate, $P(w) = \mathrm{Im}(\int_0^T e^{-i\omega t} \frac{d^2}{dt^2}\mu(t)\,dt)$. 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 Fig. 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 13th harmonic, which corroborates well with the plateau phenomenon, typically observed in experiments [79]. We emphasize that despite the large domain size used in this calculation, we require only ∼60 000 basis functions. This underlines the efficacy of higher order finite elements for even nonlinear regime in RT-TDDFT.

5. All-electron calculations: Methane (CH_4)

We now examine the competence of higher order finite elements for all-electron RT-TDDFT calculations by

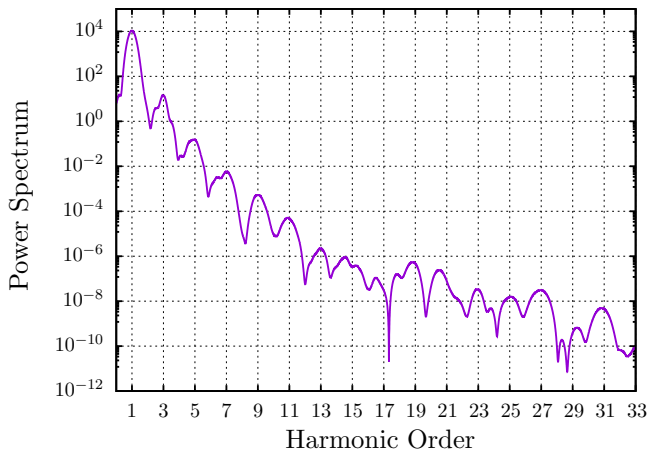

FIG. 11. Dipole power spectrum of Mg_2.

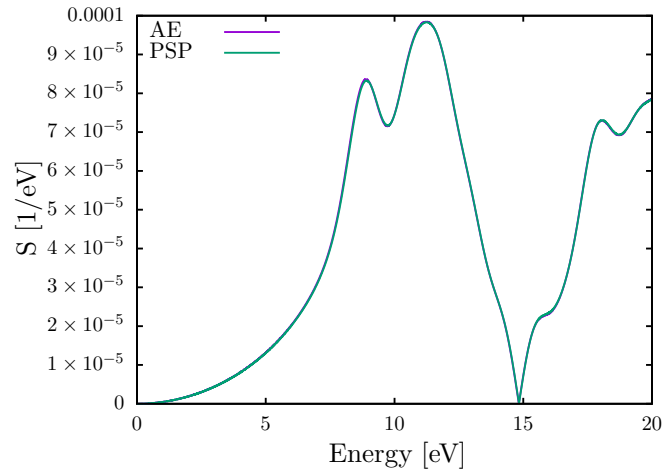

FIG. 12. Absorption spectra of methane.

providing a comparative study with its pseudopotential counterpart. In this example, we consider a methane molecule with the same geometry as described in Sec. VII A 2. We use HEX64SPEC and HEX125SPEC elements for the pseudopotential and all-electron case, respectively. For both all-electron and pseudopotential cases, we use the same mesh adaption strategy as used in all previous examples. For both the meshes, we use a large cubical domain of length 40 a.u., so as to eliminate reflection from the boundaries. Both the systems are excited from their respective ground states using the simulation details listed in Table I. The absorption spectra for both the calculations are shown in Fig. 12. We used the same Gaussian window as in the case of buckminsterfullerene (see Sec. VII C 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. In Table V 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 ∼100× 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) increase in $\|\psi_\alpha^h\|_{H^1(\Omega)}$, which, in turn, warrants a smaller time step [cf. Eq. (35)] as well as a larger Krylov subspace to achieve the prescribed accuracy. In particular, for the case of all-electron

TABLE V. 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

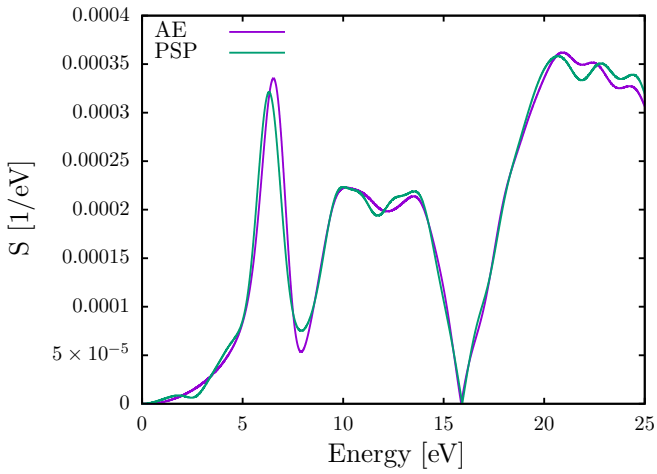

FIG. 13. Absorption spectra of benzene.

methane, we required ∼4× degrees of freedom and ∼10× 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. VII A 1). Moreover, one can mitigate the need of a refined mesh for the all-electron calculation by using enrichment ideas, wherein the standard (classical) finite-element basis is augmented with numerical atom-centered functions [57,80–82]. This idea has successfully attained 100–300× speedup over the standard (classical) finite elements for ground-state DFT calculations [57], and can be extended to RT-TDDFT to further the capabilities of finite elements.

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. VII C 5). The simulation details, for both the cases, are listed in Table I. Figure 13 compares the absorption spectra from the all-electron and pseudopotential calculations. Both the spectra compare well with the results presented in Ref. [59], in terms of first two excitation peaks (the first two excitation peaks, as we estimate from the absorption spectrum reported in Ref. [59], are ∼6.6 eV and ∼10 eV, respectively). We remark that while there is qualitative agreement between the pseudpotential and all-electron calculations, quantitatively the predictions from all-electron and pseudopotential calculations differ. In particular, the first two excitation peaks (see Table VI) differ up to ∼0.2 eV. As in both the pseudopotential and all-electron calculations, the spatiotemporal discretization parameters have been chosen to be commensurate with an accuracy of ∼10 meV in the excitation energies; we attribute this larger discrepancy to the pseudopotential approximation. This suggests that one should 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, allow for a robust tool for such transferability studies.

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

Method	E_1	E_2	DOF	CPU hrs
AE	6.521	10.131	989 649	153 600
PSP	6.316	10.007	257 473	1 574

D. Scalability

Lastly, we demonstrate the parallel scalability (strong scaling) of the proposed finite-element basis in Fig. 14. We choose the buckminsterfullerene molecule containing ∼3.5 × 10^6 degrees of freedom (number of basis functions) as our fixed benchmark system and report the relative speedup with respect to the wall time on 24 processors. The use of any number of processors below 24 was unfeasible owing to the memory requirement posed by the system. As is evident from the figure, the scaling is in good agreement with the ideal linear scaling behavior up to 384 processors, at which we observe a parallel efficiency of 86.2%. However, we observe a deviation from linear scaling behavior at 768 processors with a parallel efficiency of 74.2%. This is attributed to the fact that, at 768 processors, the number of degrees of freedom possessed by each processor falls below 5000, which is low to achieve good parallel scalability.

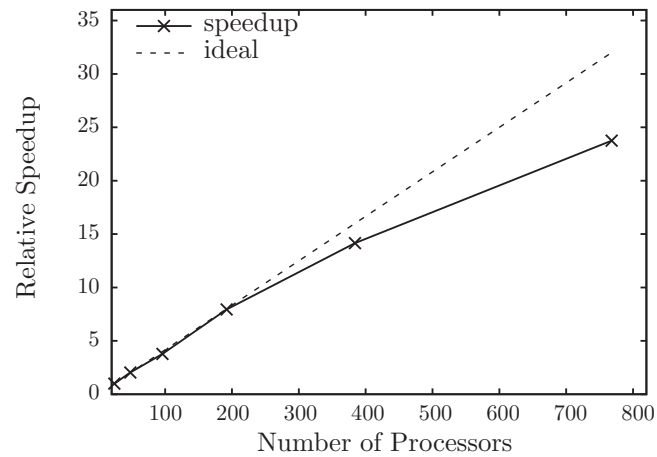

FIG. 14. Parallel scalability of the higher order finite-element implementation.

VIII. SUMMARY

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

We demonstrated the accuracy of the proposed approach through numerical convergence studies on both pseudopotential and all-electron systems, where we obtained close to optimal rates of convergence with respect to both spatial and temporal discretizations, as determined by our error estimates. The computational efficiency afforded by using higher order finite-element discretization was established, where a staggering (10–100)-fold speedup was obtained on benchmark systems by using a fourth-order finite element in comparison to linear and quadratic finite elements. Furthermore, we assessed the accuracy and efficiency afforded by our approach against the finite-difference-based method of the OCTOPUS software package, for pseudopotential calculations. Across all the benchmark systems considered, we obtained good agreement between the absorption spectrum evaluated using the finite-elements- and finite-difference-based (OCTOPUS) calculations. In terms of computational efficiency, we obtained (3–60)-fold speedup over finite difference, which is largely attributed to the adaptive spatial resolution afforded by our approach. We also demonstrated the efficacy of finite elements, especially its efficient handling of large domains, for nonlinear response by studying the higher harmonic generation under a strong electric field. We also demonstrated the competence of higher order finite elements for the all-electron RT-TDDFT calculations. This underscores the versatility of finite elements in handling both pseudopotential and all-electron calculations on an equal footing. Lastly, in terms of parallel scalability, we obtained good parallel efficiency up to 768 processors for a benchmark system comprising of the buckminsterfullerene molecule containing $\sim 3.5 \times 10^6$ basis functions.

Thus, the proposed approach offers a computationally efficient, systematically improvable, and scalable basis for RT-TDDFT calculations, applicable to both pseudopotential and all-electron cases. We remark that, for the all-electron case, the need for a highly refined mesh near the nuclei increases the computational cost, as observed from the reported studies. For systems with heavier atoms, the mesh requirements are expected to become even more exacting. However, this can be alleviated by employing an enriched finite-element basis [57], along with efficient numerical strategies to evaluate the inverse of the resultant overlap matrix as well as handle the increased quadrature costs [83]. These ideas have been successfully used for ground-state DFT [57,80–82], and its extension to RT-TDDFT is currently being investigated. Further, assessing the transferability of pseudopotentials for electron dynamics, enabled by the unified treatment of all-electron and pseudoptential calculations, is another interesting direction for future investigation.

ACKNOWLEDGMENTS

We are grateful for the support of Army Research Office through Grant No. 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 No. 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 No. 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 US 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 No. W911NF-18-1-0242.

APPENDIX: DERIVATION OF SPATIAL- AND TIME-DISCRETIZATION ERROR ESTIMATES

In this Appendix we provide the detailed derivation of the spatial- and time-discretization error estimates presented in the main text [i.e., Eqs. (28) and (29)].

1. Notations, assumptions, and preliminaries

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

$$L^2(\Omega_T) = \{f(\mathbf{r}, t) | ||f||_{L^2(\Omega)}(t) < \infty, \quad \forall\, t \in [0, T]\}, \tag{A1a}$$

$$H^1(\Omega_T) = \left\{f(\mathbf{r}, t) | f, \frac{\partial f}{\partial t}, Df \in L^2(\Omega_T)\right\}, \tag{A1b}$$

$$H_0^1(\Omega_T) = \{f(\mathbf{r}, t) | f \in H^1(\Omega_T), f(\mathbf{r}, t)|_{\partial\Omega} = 0, \quad \forall\, t \in [0, T]\}, \tag{A1c}$$

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)]

$$H^1_Z(\Omega_T) = \left\{ f(\mathbf{r},t) | f \in H^1(\Omega_T),\, f(\mathbf{r},t)|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|} \quad \forall\, t \in [0,T] \right\}, \tag{A2a}$$

$$H^1_{-Z}(\Omega) = \left\{ f(\mathbf{r}) | f \in H^1(\Omega),\, f(\mathbf{r})|_{\partial\Omega} = \sum_{I=1}^{N_a} \frac{-Z_I}{|\mathbf{r}-\mathbf{R}_I|} \right\}. \tag{A2b}$$

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 $(\ldots,\ldots)$, $||\ldots||_{L^p(\Omega)}$, and $||\ldots||_{H^1(\Omega)}$ are to be treated as time-dependent, unless otherwise specified.

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

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

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

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

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

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

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

(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

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

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

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

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

$$\left\| \rho_{\Psi_1} - \rho_{\Psi_2} \right\|_{L^1(\Omega)} \leqslant C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}, \tag{A3a}$$

$$\left\| \rho_{\Psi_1} - \rho_{\Psi_2} \right\|_{L^2(\Omega)} \leqslant C \sum_{\alpha=1}^{N_e} \|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{H^1(\Omega)}. \tag{A3b}$$

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

$$\||\mathbf{r}|^{-1} * \rho\|_{L^\infty(\Omega)} \leqslant C \||\mathbf{r}|^{-1}\|_{L^2(\Omega)} \|\rho\|_{L^2(\Omega)}, \tag{A4a}$$

$$\||\mathbf{r}|^{-1} * \rho\|_{L^2(\Omega)} \leqslant C \||\mathbf{r}|^{-1}\|_{L^2(\Omega)} \|\rho\|_{L^1(\Omega)}, \tag{A4b}$$

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

We now bound $\|V_{\mathrm{KS}}[\rho_{\Psi_1}]\psi_{1,\alpha} - V_{\mathrm{KS}}[\rho_{\Psi_2}]\psi_{2,\alpha}\|_{L^2(\Omega)}$, by decomposing V_{KS} into its Hartree (V_H), nuclear (V_N), exchange-correlation (V_{XC}), and field (V_{field}) components, and bounding each of the components. For the Hartree potential, we have, for $\forall\, v \in H^1_0(\Omega_T)$,

$$\left(V_H\left[\rho_{\Psi_1}\right]\psi_{1,\alpha} - V_H\left[\rho_{\Psi_2}\right]\psi_{2,\alpha}, v\right) = \left(V_H\left[\rho_{\Psi_1}\right](\psi_{1,\alpha} - \psi_{2,\alpha}), v\right) + \left(V_H\left[\rho_{\Psi_1} - \rho_{\Psi_2}\right]\psi_{2,\alpha}, v\right). \tag{A5}$$

Thus, using result of Eq. (A4b) along with the fact that $\psi_{2,\alpha} \in L^\infty(\Omega_T)$ (from assumption 1) and $V_H[\rho_{\Psi_1}] \in L^\infty(\Omega_T)$ [from Eq. (A4a)], it follows that

$$\left|\left(V_H\left[\rho_{\Psi_1}\right]\psi_{1,\alpha} - V_H\left[\rho_{\Psi_2}\right]\psi_{2,\alpha}, v\right)\right| \leqslant C\left(\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + \left\|\rho_{\Psi_1} - \rho_{\Psi_2}\right\|_{L^1(\Omega)}\|v\|_{L^2(\Omega)}\right). \tag{A6}$$

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

$$V_{\rm XC}\left[\rho_{\Psi_1}\right]\psi_{1,\alpha} - V_{\rm XC}\left[\rho_{\Psi_2}\right]\psi_{2,\alpha} = \left(V_{\rm XC}[\rho_X] + 2\chi_\alpha^2 V'_{\rm XC}[\rho_X]\right)(\psi_{1,\alpha} - \psi_{2,\alpha}), \tag{A7}$$

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

$$\begin{aligned}\left|\left(V_{\rm XC}\left[\rho_{\Psi_1}\right]\psi_{1,\alpha} - V_{\rm XC}\left[\rho_{\Psi_2}\right]\psi_{2,\alpha}, v\right)\right| &= \left|\left((V_{\rm XC}[\rho_X] + 2\chi_\alpha^2 V'_{\rm XC}[\rho_X])(\psi_{1,\alpha} - \psi_{2,\alpha}), v\right)\right| \\ &\leqslant \left\|V_{\rm XC}[\rho_X] + 2\chi_\alpha^2 V'_{\rm XC}[\rho_X]\right\|_{L^\infty(\Omega)}\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \\ &\leqslant C\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)},\end{aligned} \tag{A8}$$

where we have used the boundedness assumption on $V_{\rm XC}$ and $V'_{\rm XC}$ (assumption 5).

Similarly, using the boundedness assumptions on $V_N^{\rm AE}$ (assumption 2), $V_N^{\rm PSP}$(assumptions 3 and 4), and $V_{\rm field}$ (assumption 6) it is easy to observe, $\forall\ v \in H_0^1(\Omega_T)$,

$$\left|\left(V_N^{\rm AE}\psi_{1,\alpha} - V_N^{\rm AE}\psi_{2,\alpha}, v\right)\right| \leqslant C\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}, \tag{A9}$$

$$\left|\left(V_N^{\rm PSP}\psi_{1,\alpha} - V_N^{\rm PSP}\psi_{2,\alpha}, v\right)\right| \leqslant C\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}, \tag{A10}$$

$$|(V_{\rm field}\psi_{1,\alpha} - V_{\rm field}\psi_{2,\alpha}, v)| \leqslant C\|\psi_{1,\alpha} - \psi_{2,\alpha}\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}. \tag{A11}$$

We now define the weak solution of the TDKS equation (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

$$i\left(\frac{\partial\psi_\alpha}{\partial t}, v\right) = \frac{1}{2}(\nabla\psi_\alpha, \nabla v) + (V_{\rm KS}[\rho]\psi_\alpha, v), \quad \forall\ v \in H_0^1(\Omega_T), \quad \text{and} \quad \forall\ t \in [0,T]. \tag{A12}$$

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 $V_N^{\rm AE}(\mathbf{r},\mathbf{R}) \in H^1_{-Z}(\Omega)$, satisfying

$$(\nabla V_H, \nabla v) = 4\pi(\rho, v), \quad \forall\ v \in H_0^1(\Omega_T), \quad \text{and} \quad \forall\ t \in [0,T], \tag{A13a}$$

$$\left(\nabla V_N^{\rm AE}, \nabla v\right) = 4\pi(b, v), \quad \forall\ v \in H_0^1(\Omega)\,. \tag{A13b}$$

2. Derivation of spatial-discretization 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 semidiscrete solution $\psi_\alpha^h(\mathbf{r},t)$, to Eq. (A12) 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

$$i\left(\frac{\partial\psi_\alpha^h}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla\psi_\alpha^h, \nabla v^h\right) + \left(V_{\rm KS}^h[\rho^h]\psi_\alpha^h, v^h\right), \quad \forall\ v^h \in X_0^{h,p} \times [0,T] \quad \text{and} \quad \forall\ t \in [0,T]\,, \tag{A14}$$

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

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

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

as an interpolation of the boundary conditions of Eq. (8) into $X^{h,p}$. This allows us to define the discrete counterpart of the weak solution described in Eq. (A13a) as $V_H^h[\rho^h](\mathbf{r},t) = V_{H,0}^h[\rho^h](\mathbf{r},t) + f^h(\mathbf{r};\mathbf{R})$, with $V_{H,0}^h[\rho^h](\mathbf{r},t) \in X_0^{h,p} \times [0,T]$, such that

$$\left(\nabla V_{H,0}^h, \nabla v^h\right) = 4\pi(\rho^h, v^h) - (\nabla f^h, \nabla v^h), \quad \forall\ v^h \in X_0^{h,p} \times [0,T], \quad \text{and} \quad \forall\ t \in [0,T]\,. \tag{A15}$$

Similarly, we define the discrete analog of the weak solution defined in Eq. (A13b) as $V_N^{{\rm AE},h}(\mathbf{r};\mathbf{R}) = V_{N,0}^{{\rm AE},h}(\mathbf{r};\mathbf{R}) - f^h(\mathbf{r};\mathbf{R})$, with $V_{N,0}^{{\rm AE},h}(\mathbf{r};\mathbf{R}) \in X_0^{h,p}$, such that

$$\left(\nabla V_{N,0}^{{\rm AE},h}, \nabla v^h\right) = 4\pi(b, v^h) + (\nabla f^h, \nabla v^h), \quad \forall\ v^h \in X_0^{h,p}. \tag{A16}$$

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

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:

$$(\nabla(\psi - \mathcal{P}_h\psi), \nabla v^h) = 0, \quad \forall\ \psi \in H_0^1(\Omega_T), \quad \forall\ v^h \in X_0^{h,p} \times [0,T], \quad \text{and} \quad \forall\ t \in [0,T]. \tag{A17}$$

This allows us to use some standard finite-element error estimates [84] to bound $\|\psi - \mathcal{P}_h\psi\|_{L^2(\Omega)}$.

In order to prove the bound of Eq. (28), we first present a general case with no assumptions on the initial orbitals $\psi_\alpha(\mathbf{r},0)$. We then present the special case of the initial orbitals being ground-state Kohn-Sham orbitals, as a corollary to the general case. Furthermore, we note that the an error estimate for $\|\psi_\alpha - \psi_\alpha^h\|_{H^1(\Omega)}$, in turn, requires an estimate for $\|\psi_\alpha - \psi_\alpha^h\|_{L^2(\Omega)}$. Therefore, in our subsequent analysis we report estimates for both $\|\psi_\alpha - \psi_\alpha^h\|_{L^2(\Omega)}$ and $\|\psi_\alpha - \psi_\alpha^h\|_{H^1(\Omega)}$). We emphasize that, although the numerical studies presented in this work have used hexagonal elements, the following results apply to other shapes of finite element and, hence, in our analysis we denote the mesh using the generic term "triangulation" [84]. In particular, we take a triangulation $\mathcal{T}^{h,p}$ of pth-order finite elements covering the domain Ω.

Proposition 1. Assuming uniqueness and existence of the solution to Eqs. (A12) and (A14), we obtain the following bounds on the finite-element semidiscrete approximation error to the Kohn-Sham orbitals:

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) \leqslant C_1 e^{C_2 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)$$
$$+ C_1 e^{C_2 t} t \sum_e h_e^{p+1}\left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + e^{C_2 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0), \tag{A18a}$$

$$\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leqslant C_3 e^{C_2 t}(t+1) \sum_e h_e^{p} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right)$$
$$+ C_3 e^{C_2 t} t \sum_e h_e^{p}\left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + C_3 e^{C_2 t} h_{\min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0), \tag{A18b}$$

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 $|\ldots|_{p,\Omega_e}$ is the seminorm in $H^p(\Omega_e)$. The arguments $s_{1,\alpha}$, $s_{2,\alpha}$, and s_3 are defined as

$$s_{1,\alpha} = \underset{0\leqslant s\leqslant t}{\arg\max} \|\psi_\alpha - \mathcal{P}_h\psi_\alpha\|_{L^2(\Omega)}(s), \quad s_{2,\alpha} = \underset{0\leqslant s\leqslant t}{\arg\max} \left\|\frac{\partial\psi_\alpha}{\partial t} - \mathcal{P}_h\frac{\partial\psi_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s), \quad \text{and}$$
$$s_3 = \underset{0\leqslant s\leqslant t}{\arg\max} \left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}(s). \tag{A19}$$

Proof. Taking $v = v^h \in X_0^{h,p} \times [0,T]$ in Eq. (A12) (continuous solution) and subtracting it from Eq. (A14) (semidiscrete solution), we get

$$i\left(\frac{\partial(\psi_\alpha - \psi_\alpha^h)}{\partial t}, v^h\right) = \frac{1}{2}\left(\nabla(\psi_\alpha - \psi_\alpha^h), \nabla v^h\right) + \left(V_{\rm KS}[\rho]\psi_\alpha - V_{\rm KS}^h[\rho^h]\psi_\alpha^h, v^h\right), \quad \forall\ v^h \in X_0^{h,p} \times [0,T]. \tag{A20}$$

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. (A20) as

$$i\left(\frac{\partial w_\alpha}{\partial t}, v^h\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, v^h\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla v^h\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla v^h\right) + \left(V_{\rm KS}[\rho]\psi_\alpha - V_{KS}^h[\rho^h]\psi_\alpha^h, v^h\right). \tag{A21}$$

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

$$i\left(\frac{\partial w_\alpha}{\partial t}, w_\alpha\right) = -i\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right) + \frac{1}{2}\left(\nabla u_\alpha, \nabla w_\alpha\right) + \frac{1}{2}\left(\nabla w_\alpha, \nabla w_\alpha\right) + \left(V_{\rm KS}[\rho]\psi_\alpha - V_{\rm KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right). \tag{A22}$$

Noting that

$$\frac{1}{2}\frac{d}{dt}\|w_\alpha\|_{L^2(\Omega)}^2 = \mathrm{Re}\left\{\left(\frac{\partial}{\partial t}w_\alpha, w_\alpha\right)\right\}, \tag{A23}$$

and comparing the imaginary parts of Eq. (A22), we have

$$\frac{1}{2}\frac{d}{dt}\|w_\alpha\|_{L^2(\Omega)}^2 = -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla u_\alpha, \nabla w_\alpha\right)\right\} + \frac{1}{2}\mathrm{Im}\left\{\left(\nabla w_\alpha, \nabla w_\alpha\right)\right\}$$
$$+ \mathrm{Im}\left\{\left(V_{\rm KS}[\rho]\psi_\alpha - V_{\rm KS}^h[\rho^h]\psi_\alpha^h, w_\alpha\right)\right\}. \tag{A24}$$

In the above equation, we note that $(\nabla u_\alpha, \nabla w_\alpha) = 0$, as a consequence of Eq. (A17). Furthermore, $(\nabla w_\alpha, \nabla w_\alpha)$ is real. Thus, Eq. (A24) simplifies to

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|w_\alpha\|^2_{L^2(\Omega)} &= -\mathrm{Re}\left\{\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right\} + \mathrm{Im}\left\{\left(V_{\mathrm{KS}}[\rho]\psi_\alpha - V^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right\} \\ &\leqslant \left|\left(\frac{\partial u_\alpha}{\partial t}, w_\alpha\right)\right| + \left|\left(V_{\mathrm{KS}}[\rho]\psi_\alpha - V^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right|. \end{aligned} \tag{A25}$$

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

$$\begin{aligned}\left(V_{\mathrm{KS}}[\rho]\psi_\alpha - V^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right) &= \left(V_{\mathrm{XC}}[\rho]\psi_\alpha - V_{\mathrm{XC}}[\rho^h]\psi^h_\alpha, w_\alpha\right) + \left(V_H[\rho]\psi_\alpha - V_H[\rho^h]\psi^h_\alpha, w_\alpha\right) \\ &\quad + \left(\left(V_H[\rho^h] - V^h_H[\rho^h]\right)\psi^h_\alpha, w_\alpha\right) + \left(V_N\psi_\alpha - V_N\psi^h_\alpha, w_\alpha\right) \\ &\quad + \left(V_N\psi^h_\alpha - V^h_N\psi^h_\alpha, w_\alpha\right) + \left(V_{\mathrm{field}}\psi_\alpha - V_{\mathrm{field}}\psi^h_\alpha, w_\alpha\right). \end{aligned} \tag{A26}$$

We note that the term $(V_N\psi^h_\alpha - V^h_N\psi^h_\alpha, w_\alpha)$, 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. (A6), (A8), (A9), (A10), and (A11), with $v = w_\alpha$, and using the fact that $\psi^h \in L^\infty(\Omega)$, it is straightforward to show that

$$\begin{aligned}\left|\left(V_{\mathrm{KS}}[\rho]\psi_\alpha - V^h_{\mathrm{KS}}[\rho^h]\psi^h_\alpha, w_\alpha\right)\right| &\leqslant C_0\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}\|w_\alpha\|_{L^2(\Omega)} + C_1\left(\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + \|\rho - \rho^h\|_{L^1(\Omega)}\right)\|w_\alpha\|_{L^2(\Omega)} \\ &\quad + C_2\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)}\|w_\alpha\|_{L^2(\Omega)} + C_3\left\|V_N - V^h_N\right\|_{L^2(\Omega)}\|w_\alpha\|_{L^2(\Omega)}. \end{aligned} \tag{A27}$$

Using the above result in Eq. (A25), we obtain

$$\begin{aligned}\frac{d}{dt}\|w_\alpha\|_{L^2(\Omega)} &\leqslant \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + C_1\left(\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} + \|\rho - \rho^h\|_{L^1(\Omega)}\right) \\ &\quad + C_2\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_3\left\|V_N - V^h_N\right\|_{L^2(\Omega)} \\ &\leqslant \left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_0\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)} \\ &\quad + C_2\|V_H[\rho^h] - V^h_H[\rho^h]\|_{L^2(\Omega)} + C_3\left\|V_N - V^h_N\right\|_{L^2(\Omega)} + C_4\sum_{\beta=1}^{N_e}\left\|\psi_\beta - \psi^h_\beta\right\|_{L^2(\Omega)}, \end{aligned} \tag{A28}$$

where we have used Eq. (A3a) in the second line to simplify the term involving $\|\rho - \rho^h\|_{L^1(\Omega)}$. Summing the above equation over all index α, we have

$$\begin{aligned}\frac{d}{dt}\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)} &\leqslant \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5\left\|\psi_\alpha - \psi^h_\alpha\right\|_{L^2(\Omega)}\right) + C_6\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_7\left\|V_N - V^h_N\right\|_{L^2(\Omega)} \\ &\leqslant \sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} + C_5\|u_\alpha\|_{L^2(\Omega)} + C_5\|w_\alpha\|_{L^2(\Omega)}\right) + C_6\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)} + C_7\left\|V_N - V^h_N\right\|_{L^2(\Omega)}, \end{aligned} \tag{A29}$$

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

$$\begin{aligned}\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) &\leqslant \sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(0) + C_5\int_0^t\sum_{\alpha=1}^{N_e}\|w_\alpha\|(s)\,ds + C_5\int_0^t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s) + \|u_\alpha\|_{L^2(\Omega)}(s)\right)ds \\ &\quad + C_8\int_0^t\left(\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)}(s) + \left\|V_N - V^h_N\right\|_{L^2(\Omega)}\right)ds. \end{aligned} \tag{A30}$$

Noting that $u_\alpha = \psi_\alpha - \mathcal{P}_h\psi_\alpha$, $\frac{\partial u_\alpha}{\partial t} = \frac{\partial \psi_\alpha}{\partial t} - \mathcal{P}_h\frac{\partial \psi_\alpha}{\partial t}$, and using the definitions of $s_{1,\alpha}$, $s_{2,\alpha}$, and s_3 [cf. Eq. (A19)], we can simplify the above equation as

$$\begin{aligned}\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) &\leqslant \sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(0) + C_5\int_0^t\sum_{\alpha=1}^{N_e}\|w_\alpha\|(s)\,ds + C_5 t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha})\right) \\ &\quad + C_8 t\left(\left\|V_H[\rho^h] - V^h_H[\rho^h]\right\|_{L^2(\Omega)}(s_3) + \left\|V_N - V^h_N\right\|_{L^2(\Omega)}\right). \end{aligned} \tag{A31}$$

Invoking the Grönwall's inequality on the above equation yields

$$\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) \leqslant e^{C_5 t}\left[\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(0) + C_5 t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha})\right)\right] + C_8 e^{C_5 t} t\left(\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}(s_3) + \left\|V_N - V_N^h\right\|_{L^2(\Omega)}\right). \tag{A32}$$

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

$$\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) \leqslant C_5 e^{C_5 t} t\sum_{\alpha=1}^{N_e}\left(\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)}(s_{2,\alpha}) + \|u_\alpha\|_{L^2(\Omega)}(s_{1,\alpha})\right) + C_8 e^{C_5 t} t\left(\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)}(s_3) + \left\|V_N - V_N^h\right\|_{L^2(\Omega)}\right) + e^{C_5 t}\sum_{\alpha=1}^{N_e}\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0). \tag{A33}$$

Bounds on the terms involving $\|u_\alpha\|_{L^2(\Omega)}$, $\|\frac{\partial u_\alpha}{\partial t}\|_{L^2(\Omega)}$, $\|V_H[\rho^h] - V_H^h[\rho^h]\|_{L^2(\Omega)}$, and $\|V_N - V_N^h\|_{L^2(\Omega)}$ can now be obtained using the Ceá's lemma [84], a standard finite-element error estimate. The Ceá's lemma, in simple terms, is stated as follows. Let $\phi \in H^1(\Omega_T)$ and $\phi^h \in V^h \subseteq X^{h,p}$. If $y = \phi - \phi^h$ satisfies the following Galerkin orthogonality condition,

$$(\nabla y, \nabla v^h)(t) = 0, \quad \forall\, v^h \in V^h \quad \text{and} \quad \forall\, t \in [0, T], \tag{A34}$$

then

$$\|y\|_{L^2(\Omega)} \leqslant C\sum_e h_e^{p+1}|\phi|_{p+1,\Omega_e} \quad \text{and} \tag{A35a}$$

$$\|y\|_{H^1(\Omega)} \leqslant C\sum_e h_e^{p}|\phi|_{p+1,\Omega_e}. \tag{A35b}$$

By definition of Ritz projection [Eq. (A17)], $y = u_\alpha = \psi_\alpha - \mathcal{P}_h\psi_\alpha$ satisfies Eq. (A34). Further, taking the time derivative of Eq. (A17), it is easy to verify that $y = \frac{\partial u_\alpha}{\partial t} = \frac{\partial \psi_\alpha}{\partial t} - \mathcal{P}_h\frac{\partial \psi_\alpha}{\partial t}$ also satisfies Eq. (A34). Thus, applying the Ceá's lemma [Eq. (A35a)] to u_α and $\frac{\partial u_\alpha}{\partial t}$ yields

$$\|u_\alpha\|_{L^2(\Omega)} \leqslant C\sum_e h_e^{p+1}|\psi_\alpha|_{p+1,\Omega_e} \tag{A36}$$

and

$$\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} \leqslant C\sum_e h_e^{p+1}\left|\frac{\partial \psi_\alpha}{\partial t}\right|_{p+1,\Omega_e}. \tag{A37}$$

We further simplify the above inequality, by using Eq. (1):

$$\begin{aligned}\left\|\frac{\partial u_\alpha}{\partial t}\right\|_{L^2(\Omega)} &\leqslant C\sum_e h_e^{p+1}\left|\frac{\partial \psi_\alpha}{\partial t}\right|_{p+1,\Omega_e} = C\sum_e h_e^{p+1}\left|-\frac{1}{2}\nabla^2\psi_\alpha + V_{\mathrm{KS}}[\rho]\psi_\alpha\right|_{p+1,\Omega_e}\\ &\leqslant C\sum_e h_e^{p+1}\left(|\psi_\alpha|_{p+3,\Omega_e} + |(V_H + V_N + V_{\mathrm{XC}} + V_{\mathrm{field}})\psi_\alpha|_{p+1,\Omega_e}\right)\\ &\leqslant C\sum_e h_e^{p+1}\left(|\psi_\alpha|_{p+3,\Omega_e} + |\psi_\alpha|_{p+1,\Omega_e}\right),\end{aligned} \tag{A38}$$

which follows from the definition of the $|\ldots|_{p+3}$ seminorm and the boundedness assumptions on V_N, V_{XC}, and V_{field} (assumptions 2–6). Lastly, it is straightforward to observe that both $y = V_H[\rho^h] - V_H^h[\rho^h]$ and $y = V_N - V_N^h$ satisfy Eq. (A34) [take the difference of Eqs. (A15) and (A13a) and Eqs. (A16) and (A13b), respectively). Thus, once again, applying the Ceá's lemma [Eq. (A35a)], we get

$$\left\|V_H[\rho^h] - V_H^h[\rho^h]\right\|_{L^2(\Omega)} \leqslant C\sum_e h_e^{p+1}|V_H[\rho^h]|_{p+1,\Omega_e}, \tag{A39a}$$

$$\left\|V_N - V_N^h\right\|_{L^2(\Omega)} \leqslant C\sum_e h_e^{p+1}|V_N|_{p+1,\Omega_e}. \tag{A39b}$$

Using Eqs. (A36), (A38), and (A39) in Eq. (A33), we have

$$\sum_{\alpha=1}^{N_e}\|w_\alpha\|_{L^2(\Omega)}(t) \leqslant C_9 e^{C_5 t} t\sum_e h_e^{p+1}\sum_{\alpha=1}^{N_e}\left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) + C_9 e^{C_5 t} t\sum_e h_e^{p+1}\left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + e^{C_5 t}\sum_{\alpha=1}^{N_e}\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0). \tag{A40}$$

Finally, expressing $\psi_\alpha - \psi_\alpha^h = w_\alpha + u_\alpha$ and using the result of Eq. (A36) in the above equation, we obtain

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) &\leqslant C \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} |\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) \\ &+ C_9 e^{C_5 t} t \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_9 e^{C_5 t} t \sum_e h_e^{p+1} \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \\ &\leqslant C_{10} e^{C_5 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_{10} e^{C_5 t} t \sum_e h_e^{p+1} \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + e^{C_5 t} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0). \end{aligned} \tag{A41}$$

This concludes the proof of Eq. (A18a).

In order to derive estimates for $\sum_{\alpha=1}^{N_e} \|\psi_\alpha - \psi_\alpha^h\|_{H^1(\Omega)}(t)$, we use the inverse estimate [33] for $w_\alpha = (\mathcal{P}_h\psi_\alpha - \psi_\alpha^h) \in X_0^{h,p}$ to obtain

$$\|w_\alpha\|_{H^1(\Omega)}(t) \leqslant C h_{\min}^{-1} \|w_\alpha\|_{L^2(\Omega)}(t). \tag{A42}$$

Additionally, applying the Ceá's lemma [Eq. (A35b)] on $u_\alpha = (\psi_\alpha - \mathcal{P}_h\psi_\alpha)$, we have

$$\|u_\alpha\|_{H^1(\Omega)}(t) \leqslant C \sum_e h_e^p |\psi_\alpha|_{p+1,\Omega_e}(t). \tag{A43}$$

Combining Eqs. (A42) and (A43), we get

$$\sum_{\alpha}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) \leqslant \sum_{\alpha=1}^{N_e} (\|u_\alpha\|_{H^1(\Omega)}(t) + \|w_\alpha\|_{H^1(\Omega)}(t)) \leqslant C_{11} \sum_e h_e^p \sum_{\alpha=1}^{N_e} |\psi_\alpha|_{p+1,\Omega_e}(t) + C_{12} h_{\min}^{-1} \sum_{\alpha=1}^{N_e} \|w_\alpha\|_{L^2(\Omega)}(t). \tag{A44}$$

Finally, using the inequality obtained in Eq. (A40) in the above equation and using the fact that $h_e/h_{\min} \leqslant C$ for all the elements in $\mathcal{T}^{h,p}$ yields

$$\begin{aligned}\sum_{\alpha}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{H^1(\Omega)}(t) &\leqslant C_{11} \sum_e h_e^p \sum_{\alpha=1}^{N_e} |\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) \\ &+ C_{13} e^{C_5 t} t \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_{13} e^{C_5 t} t \sum_e h_e^p \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + C_{12} e^{C_5 t} h_{\min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \\ &\leqslant C_{14} e^{C_5 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_{14} e^{C_5 t} t \sum_e h_e^p \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right) + C_{14} e^{C_5 t} h_{\min}^{-1} \sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0). \end{aligned} \tag{A45}$$

This concludes the proof for Eq. (A18b). ■

Corollary 1. If the initial orbitals $\psi_\alpha(\mathbf{r}, 0)$ are obtained from a ground-state DFT calculation, wherein [48]

$$\left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(0) \leqslant C \sum_e h_e^{p+1} \left(|\psi_\alpha|_{p+1,\Omega_e} + |V_H[\rho^h]|_{p+1,\Omega_e} + |V_N|_{p+1,\Omega_e}\right), \tag{A46}$$

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

$$\begin{aligned}\sum_{\alpha=1}^{N_e} \left\|\psi_\alpha - \psi_\alpha^h\right\|_{L^2(\Omega)}(t) &\leqslant C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha})\right) \\ &+ C_1' e^{C_2 t}(t+1) \sum_e h_e^{p+1} \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e}\right), \end{aligned} \tag{A47a}$$

$$\sum_{\alpha}^{N_e} \left\| \psi_\alpha - \psi_\alpha^h \right\|_{H^1(\Omega)}(t) \leqslant C_3' e^{C_2 t}(t+1) \sum_e h_e^p \sum_{\alpha=1}^{N_e} \left(|\psi_\alpha|_{p+1,\Omega_e}(s_{1,\alpha}) + |\psi_\alpha|_{p+1,\Omega_e}(s_{2,\alpha}) + |\psi_\alpha|_{p+3,\Omega_e}(s_{2,\alpha}) \right)$$
$$+ C_3' e^{C_2 t}(t+1) \sum_e h_e^p \left(|V_H[\rho^h]|_{p+1,\Omega_e}(s_3) + |V_N|_{p+1,\Omega_e} \right). \tag{A47b}$$

The last equation concludes the proof of Eq. (28).

3. Derivation of time-discretization error estimate

Before proceeding to the proof for Eq. (29), we note that for an exponential operator of the form $e^{\boldsymbol{L}(t)}$, the partial derivative with respect to t is given by [68]

$$\frac{\partial}{\partial t} e^{\boldsymbol{L}(t)} = \mathrm{dexp}_{\boldsymbol{L}(t)}(\dot{\boldsymbol{L}}(t)) e^{\boldsymbol{L}(t)}, \tag{A48}$$

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

$$\mathrm{ad}_{\mathbf{X}}^k(\mathbf{Y}) = \mathrm{ad}_{\mathbf{X}}\left(\mathrm{ad}_{\mathbf{X}}^{k-1}(\mathbf{Y}) \right), \tag{A49}$$

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

We now present the proof for Eq. (29) in the following proposition. In the following analysis, we assume each time interval $[t_{n-1}, t_n]$ to be of length Δt. Moreover, for simpler terminology, we term $e^{\tilde{\mathbf{A}}_n}$ [cf. Eq. (25)] as the second-order Magnus propagator without explicitly spelling out the midpoint integration rule invoked in it.

Proposition 2. For a second-order Magnus propagator with a midpoint integration rule, we obtain the following bound for the time-discretization error in ψ_α^h:

$$\left\| \psi_\alpha^h(t_n) - \psi_\alpha^{h,n} \right\|_{L^2(\Omega)} \leqslant C(\Delta t)^2 t_n \max_{0 \leqslant t \leqslant t_n} \left\| \psi_\alpha^h(t) \right\|_{H^1(\Omega)}. \tag{A50}$$

Proof. To begin with, we introduce the following operators:

$$\mathbf{S}_0^k = e^{\mathbf{A}_k} e^{\mathbf{A}_{k-1}} \dots e^{\mathbf{A}_1} = \prod_{l=0}^{k-1} e^{\mathbf{A}_{k-l}} \quad \text{for } 0 < k \leqslant n\,, \quad \mathbf{S}_0^0 = I \tag{A51a}$$

$$\mathbf{R}_k^n = e^{\tilde{\mathbf{A}}_n} e^{\tilde{\mathbf{A}}_{n-1}} \dots e^{\tilde{\mathbf{A}}_{k+1}} = \prod_{l=0}^{n-k-1} e^{\tilde{\mathbf{A}}_{n-l}} \quad \text{for } 0 \leqslant k < n\,, \quad \mathbf{R}_n^n = I. \tag{A51b}$$

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

$$\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n = \left(\mathbf{R}_n^n \mathbf{S}_0^n - \mathbf{R}_0^n \mathbf{S}_0^0 \right) \bar{\boldsymbol{\psi}}_\alpha(0) = \sum_{k=1}^{n} \left(\mathbf{R}_k^n \mathbf{S}_0^k - \mathbf{R}_{k-1}^n \mathbf{S}_0^{k-1} \right) \bar{\boldsymbol{\psi}}_\alpha(0). \tag{A52}$$

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

$$\begin{aligned}
\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n &= \sum_{k=1}^{n} \left(\mathbf{R}_k^n \mathbf{S}_0^k - \mathbf{R}_{k-1}^n \mathbf{S}_0^{k-1} \right) \bar{\boldsymbol{\psi}}_\alpha(0) \\
&= \sum_{k=1}^{n} \left(\mathbf{R}_k^n e^{\mathbf{A}_k} \mathbf{S}_0^{k-1} - \mathbf{R}_k^n e^{\tilde{\mathbf{A}}_k} \mathbf{S}_0^{k-1} \right) \bar{\psi}_\alpha(0) \\
&= \sum_{k=1}^{n} \mathbf{R}_k^n (e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}) \mathbf{S}_0^{k-1} \bar{\boldsymbol{\psi}}_\alpha(0) \\
&= \sum_{k=1}^{n} \mathbf{R}_k^n (e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k}) \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}).
\end{aligned} \tag{A53}$$

Since $\mathbf{R}_k^n$ is a unitary operator, bounding $[\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n]$ reduces to finding the bound on $(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$. To this end, we extend the proof presented in Ref. [69] to the nonlinear case of the TDKS equations. To begin with, we split $(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ as

$$(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) = (e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) + (e^{\bar{\mathbf{A}}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}), \tag{A54}$$

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. (A54), we introduce the following auxiliary function,

$$\boldsymbol{\xi}_\alpha^k(t) = e^{\mathbf{B}_k(t)}\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}), \quad \forall\, t \in [t_{k-1}, t_k] \tag{A55}$$

where $\mathbf{B}_k(t) = \int_{t_{k-1}}^{t} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau$. We remark that $\boldsymbol{\xi}_\alpha^k(t)$ denotes the time evolution of $\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ using the truncated Magnus expansion, in the time interval $[t_{k-1}, t_k]$. Differentiating the above equation and using the result of Eq. (A48) gives

$$\dot{\boldsymbol{\xi}}_\alpha^k(t) = \mathrm{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))e^{\mathbf{B}_k(t)}\boldsymbol{\xi}_\alpha^k(t_{k-1}) = -i\bar{\mathbf{G}}_k(t)\boldsymbol{\xi}_\alpha^k(t), \quad \forall\, t \in [t_{k-1}, t_k] \tag{A56}$$

where $\bar{\mathbf{G}}_k(t) = i\,\mathrm{dexp}_{\mathbf{B}_k(t)}(\dot{\mathbf{B}}_k(t))$. 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. We now introduce the function $\boldsymbol{\gamma}_\alpha^k(t) = \bar{\boldsymbol{\psi}}_\alpha(t) - \boldsymbol{\xi}_\alpha^k(t)$, $\forall\, t \in [t_{k-1}, t_k]$. It is important to note that

$$\boldsymbol{\gamma}_\alpha^k(t_k) = \bar{\boldsymbol{\psi}}_\alpha(t_k) - \boldsymbol{\xi}_\alpha^k(t_k) = (e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1}), \tag{A57}$$

where the second equality follows from the definition of $\boldsymbol{\xi}_\alpha^k$ [Eq. (A55)] and the fact that $\mathbf{B}_k(t_k) = \int_{t_{k-1}}^{t_k} -i\bar{\mathbf{H}}[\rho(\tau)]d\tau = \bar{\mathbf{A}}_k$. Thus, the problem of bounding $(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})$ [the first term in Eq. (A54)] reduces to bounding $\boldsymbol{\gamma}_\alpha^k(t_k)$. To this end, we proceed, by first expressing the time derivative of $\boldsymbol{\gamma}_\alpha^k$ as

$$\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i[\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)]\bar{\boldsymbol{\psi}}_\alpha(t), \quad \forall\, t \in [t_{k-1}, t_k] \tag{A58}$$

which follows from Eqs. (21) and (A56). Now, taking the dot product with $\boldsymbol{\gamma}_\alpha^k(t)^\dagger$ on both sides yields

$$\boldsymbol{\gamma}_\alpha^k(t)^\dagger\dot{\boldsymbol{\gamma}}_\alpha^k(t) = -i\boldsymbol{\gamma}_\alpha^k(t)^\dagger\bar{\mathbf{G}}_k(t)\boldsymbol{\gamma}_\alpha^k(t) - i\boldsymbol{\gamma}_\alpha^k(t)^\dagger[\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)]\bar{\boldsymbol{\psi}}_\alpha(t). \tag{A59}$$

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

$$\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\|^2 = \mathrm{Im}\left\{\boldsymbol{\gamma}_\alpha^k(t)^\dagger[\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)]\bar{\boldsymbol{\psi}}_\alpha(t)\right\}. \tag{A60}$$

Consequently,

$$\frac{d}{dt}\left\|\boldsymbol{\gamma}_\alpha^k\right\| \leqslant \|[\bar{\mathbf{H}}(t) - \bar{\mathbf{G}}_k(t)]\bar{\boldsymbol{\psi}}_\alpha(t)\|. \tag{A61}$$

Time integrating the above equation yields

$$\left\|\boldsymbol{\gamma}_\alpha^k\right\|(t_k) = \left\|\bar{\boldsymbol{\psi}}_\alpha(t_k) - \boldsymbol{\xi}_\alpha^k(t_k)\right\| = \|(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\| \leqslant \int_{t_{k-1}}^{t_k} \|[\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau)\|d\tau, \tag{A62}$$

where we have used the result of Eq. (A57) along with the fact that $\|\boldsymbol{\gamma}_\alpha^k\|(t_{k-1}) = \bar{\boldsymbol{\psi}}_\alpha(t_{k-1}) - \boldsymbol{\xi}_\alpha^k(t_{k-1}) = 0$ [by the definition of $\boldsymbol{\xi}_\alpha^k(t)$, cf. Eq. (A55)]. Thus, the problem of bounding $\|(e^{\mathbf{A}_k} - e^{\bar{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\|$ further simplifies to finding a bound for $\int_{t_{k-1}}^{t_k} \|[\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau)\|d\tau$. To this end, we use the fact that $\bar{\mathbf{G}}_k(\tau) = i\,\mathrm{dexp}_{\mathbf{B}_k(\tau)}(\dot{\mathbf{B}}_k(\tau))$ and the definition of the operator $\mathrm{dexp}_{\mathbf{X}}(\mathbf{Y})$ to obtain

$$\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau) = -\frac{i}{2}[\mathbf{B}_k(\tau), \dot{\mathbf{B}}_k(\tau)] + \text{h.o.t.} = -\frac{i}{2}\int_{t_{k-1}}^{\tau}[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)]d\sigma + \text{h.o.t}, \tag{A63}$$

where h.o.t. stands for higher order terms.

In order to bound $[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)]$, we begin by rewriting $\bar{\mathbf{H}}$ in terms of $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$, i.e., its kinetic and Kohn-Sham potential components. To elaborate, $\bar{\mathbf{U}} = \mathbf{M}^{-1/2}\mathbf{U}\mathbf{M}^{-1/2}$ and $\bar{\mathbf{V}} = \mathbf{M}^{-1/2}\mathbf{V}\mathbf{M}^{-1/2}$, with $U_{jk} = \frac{1}{2}\int_\Omega \nabla N_j(\mathbf{r})\cdot\nabla N_k(\mathbf{r})d\mathbf{r}$ and $V_{jk} = \int_\Omega V_{\mathrm{KS}}^h[\rho^h](\mathbf{r}, t)N_j(\mathbf{r})N_k(\mathbf{r})\,d\mathbf{r}$. Noting that $\bar{\mathbf{U}}$ is time independent, we Taylor expand $\bar{\mathbf{H}}(\sigma)$ about τ to rewrite $[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)]$ as

$$[\bar{\mathbf{H}}(\tau), \bar{\mathbf{H}}(\sigma)] = [\bar{\mathbf{H}}(\tau), \bar{\mathbf{V}}'(\tau)](\sigma - \tau) + O[(\sigma - \tau)^2], \tag{A64}$$

where $\bar{\mathbf{V}}'(\tau) = \frac{d}{dt}(\bar{\mathbf{V}}(t))|_{\sigma=\tau}$. Thus, using the above relation in Eq. (A63) we get

$$[\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau) = \frac{i}{4}([\bar{\mathbf{H}}, \bar{\mathbf{V}}'(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau))(\tau - t_{k-1})^2 + O[(\tau - t_{k-1})^3]. \tag{A65}$$

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

$$\|[\bar{\mathbf{H}}(\tau) - \bar{\mathbf{G}}_k(\tau)]\bar{\boldsymbol{\psi}}_\alpha(\tau)\| \leqslant C(\tau - t_{k-1})^2\|\bar{\boldsymbol{\psi}}_\alpha(\tau)\|_{\bar{\mathbf{U}}} + O[(\tau - t_{k-1})^3]. \tag{A66}$$

Thus, substituting the above result into Eq. (A62) provides the following bound:

$$\|(e^{\mathbf{A}_k} - e^{\tilde{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\| \leqslant C(\Delta t)^3 \max_{t_{k-1}\leqslant t\leqslant t_k} \|\bar{\boldsymbol{\psi}}_\alpha(t)\|_{\bar{\mathbf{U}}}. \tag{A67}$$

This provides a bound for the first term (truncation error) on the right side of Eq. (A54). In order to bound the second term on the right side of Eq. (A54), i.e., the error due to midpoint quadrature rule, we begin with the following identity:

$$e^{\tilde{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k} = \int_0^1 \frac{d}{dx}(e^{(1-x)\tilde{\mathbf{A}}_k} e^{x\bar{\mathbf{A}}_k})\, dx = \int_0^1 e^{(1-x)\tilde{\mathbf{A}}_k}(\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k)e^{x\bar{\mathbf{A}}_k}\, dx. \tag{A68}$$

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 $|F - F_{1/2}| \leqslant C(b-a)^3 f''(\eta)$, for some $\eta \in [a, b]$. Thus, for the midpoint integration rule, $\|\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k\| \leqslant C(\Delta t)^3 \|\frac{d^2}{dt^2}(\bar{\mathbf{H}})|_{t'}\|$ 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

$$\|(e^{\tilde{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\| \leqslant \|(\bar{\mathbf{A}}_k - \tilde{\mathbf{A}}_k)\| \leqslant C(\Delta t)^3 \left\| \frac{d^2}{dt^2}(\bar{\mathbf{H}})|_{t'} \right\| \text{ for some } t' \in [t_{k-1}, t_k]. \tag{A69}$$

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

$$\|(e^{\tilde{\mathbf{A}}_k} - e^{\tilde{\mathbf{A}}_k})\bar{\boldsymbol{\psi}}_\alpha(t_{k-1})\| \leqslant C(\Delta t)^3. \tag{A70}$$

Thus, using the results of Eqs. (A67) and (A70) in Eq. (A53) along with the unitarity of the operators $\mathbf{R}_k^n$ yields

$$\left\|\bar{\boldsymbol{\psi}}_\alpha(t_n) - \bar{\boldsymbol{\psi}}_\alpha^n\right\| \leqslant C(\Delta t)^2 t_n \max_{t_0\leqslant t\leqslant t_n} \|\bar{\boldsymbol{\psi}}_\alpha(t)\|_{\bar{\mathbf{U}}}. \tag{A71}$$

Finally, noting that the coefficient vectors for the spatial fields $\psi_\alpha^h(\mathbf{r}, t_n)$ and $\psi_\alpha^{h,n}(\mathbf{r})$ are given by $\mathbf{M}^{-1/2}\bar{\boldsymbol{\psi}}_\alpha(t_n)$ and $\mathbf{M}^{-1/2}\bar{\boldsymbol{\psi}}_\alpha^n$, respectively, it is now trivial to arrive at Eq. (29) from the above equation. ■

[1] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3] H. Appel, E. K. U. Gross, and K. Burke, Phys. Rev. Lett. **90**, 043005 (2003).
[4] X. Gonze and J.-P. Vigneron, Phys. Rev. B **39**, 13120 (1989).
[5] S. J. A. van Gisbergen, J. G. Snijders, and E. J. Baerends, Phys. Rev. Lett. **78**, 3097 (1997).
[6] G. Stefanucci and C.-O. Almbladh, Europhys. Lett. **67**, 14 (2004).
[7] S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio, and E. K. U. Gross, Phys. Rev. B **72**, 035308 (2005).
[8] C. Jamorski Jödicke and H. P. Lüthi, J. Am. Chem. Soc. **125**, 252 (2003).
[9] T. Stein, L. Kronik, and R. Baer, J. Am. Chem. Soc. **131**, 2818 (2009).
[10] T. Burnus, M. A. L. Marques, and E. K. U. Gross, Phys. Rev. A **71**, 010501(R) (2005).
[11] X.-M. Tong and S.-I. Chu, Phys. Rev. A **57**, 452 (1998).
[12] X.-M. Tong and S.-I. Chu, Phys. Rev. A **64**, 013417 (2001).
[13] D. A. Telnov and S.-I. Chu, Phys. Rev. A **79**, 041401(R) (2009).
[14] M. E. Casida, in *Recent Advances In Density Functional Methods: (Part I)* (World Scientific, Singapore, 1995).
[15] M. Petersilka, U. J. Gossmann, and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).
[16] J. Theilhaber, Phys. Rev. B **46**, 12990 (1992).
[17] K. Yabana and G. F. Bertsch, Phys. Rev. B **54**, 4484 (1996).
[18] R. Baer and R. Gould, J. Chem. Phys. **114**, 3385 (2001).
[19] G. Vignale, Phys. Rev. Lett. **74**, 3233 (1995).
[20] N. T. Maitra, K. Burke, and C. Woodward, Phys. Rev. Lett. **89**, 023002 (2002).
[21] M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross, *Time-dependent Density Functional Theory* (Springer, Berlin, 2006).
[22] N. T. Maitra and K. Burke, Phys. Rev. A **63**, 042501 (2001).
[23] F. Gygi, IBM J. Res. Dev. **52**, 137 (2008).
[24] A. Schleife, E. W. Draeger, Y. Kanai, and A. A. Correa, J. Chem. Phys. **137**, 22A546 (2012).
[25] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, J. Phys.: Condens. Matter **14**, 2745 (2002).
[26] Y. Takimoto, F. D. Vila, and J. J. Rehr, J. Chem. Phys. **127**, 154114 (2007).
[27] M. Kuisma, A. Sakko, T. P. Rossi, A. H. Larsen, J. Enkovaara, L. Lehtovaara, and T. T. Rantala, Phys. Rev. B **91**, 115431 (2015).
[28] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).

[29] K. Lopata and N. Govind, J. Chem. Theory Comput. **7**, 1344 (2011).

[30] A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. A. L. Marques, E. K. U. Gross, and A. Rubio, Phys. Status Solidi B **243**, 2465 (2006).

[31] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Phys. Rev. B **71**, 035109 (2005).

[32] M. Walter, H. Häkkinen, L. Lehtovaara, M. Puska, J. Enkovaara, C. Rostgaard, and J. J. Mortensen, J. Chem. Phys. **128**, 244101 (2008).

[33] S. Brenner and R. Scott, *The Mathematical Theory of Finite Element Methods* (Springer, Berlin, 2007).

[34] T. J. R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis* (Dover, New York, 2012).

[35] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman *et al.*, J. Chem. Phys. **129**, 014109 (2008).

[36] B. Natarajan, L. Genovese, M. E. Casida, T. Deutsch, O. N. Burchak, C. Philouze, and M. Y. Balakirev, Chem. Phys. **402**, 29 (2012).

[37] S. R. White, J. W. Wilkins, and M. P. Teter, Phys. Rev. B **39**, 5819 (1989).

[38] E. Tsuchida and M. Tsukada, J. Phys. Soc. Jpn. **67**, 3844 (1998).

[39] J. E. Pask, B. M. Klein, C. Y. Fong, and P. A. Sterne, Phys. Rev. B **59**, 12352 (1999).

[40] J. E. Pask, B. M. Klein, P. A. Sterne, and C. Y. Fong, Comput. Phys. Commun. **135**, 1 (2001).

[41] J. E. Pask and P. A. Sterne, Model. Simul. Mater. Sci. Eng. **13**, R71 (2005).

[42] J.-L. Fattebert, R. D. Hornung, and A. M. Wissink, J. Comput. Phys. **223**, 759 (2007).

[43] D. Zhang, L. Shen, A. Zhou, and X.-G. Gong, Phys. Lett. A **372**, 5071 (2008).

[44] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, and M. Ortiz, J. Mech. Phys. Solids **58**, 256 (2010).

[45] J. Fang, X. Gao, and A. Zhou, J. Comput. Phys. **231**, 3166 (2012).

[46] G. Bao, G. Hu, and D. Liu, J. Comput. Phys. **231**, 4967 (2012).

[47] H. Chen, X. Gong, L. He, Z. Yang, and A. Zhou, Adv. Comput. Math. **38**, 225 (2013).

[48] P. Motamarri, M. R. Nowak, K. Leiter, J. Knap, and V. Gavini, J. Comput. Phys. **253**, 308 (2013).

[49] H. Chen, X. Dai, X. Gong, L. He, and A. Zhou, Multiscale Model. Simul. **12**, 1828 (2014).

[50] P. Motamarri and V. Gavini, Phys. Rev. B **97**, 165132 (2018).

[51] P. Motamarri, S. Das, S. Rudraraju, K. Ghosh, D. Davydov, and V. Gavini, Comput. Phys. Commun. (2019), doi:10.1016/j.cpc.2019.07.016

[52] S. Yamakawa and S.-a. Hyodo, Phys. Rev. B **71**, 035113 (2005).

[53] E. J. Bylaska, M. Holst, and J. H. Weare, J. Chem. Theory Comput. **5**, 937 (2009).

[54] L. Lehtovaara, V. Havu, and M. Puska, J. Chem. Phys. **131**, 054103 (2009).

[55] V. Schauer and C. Linder, J. Comput. Phys. **250**, 644 (2013).

[56] P. Motamarri and V. Gavini, Phys. Rev. B **90**, 115127 (2014).

[57] B. Kanungo and V. Gavini, Phys. Rev. B **95**, 035112 (2017).

[58] P. Motamarri, V. Gavini, K. Bhattacharya, and M. Ortiz, Phys. Rev. B **95**, 035111 (2017).

[59] L. Lehtovaara, V. Havu, and M. Puska, J. Chem. Phys. **135**, 154104 (2011).

[60] G. Bao, G. Hu, and D. Liu, J. Comput. Phys. **281**, 743 (2015).

[61] B. Hermansson and D. Yevick, Phys. Rev. B **33**, 7241 (1986).

[62] E. K. U. Gross and W. Kohn, Phys. Rev. Lett. **55**, 2850 (1985).

[63] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[64] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[65] P. Motamarri, M. Iyer, J. Knap, and V. Gavini, J. Comput. Phys. **231**, 6596 (2012).

[66] K. Bathe, *Finite Element Procedures* (Prentice Hall, Englewood Cliffs, NJ, 1996).

[67] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).

[68] S. Blanes, F. Casas, J. Oteo, and J. Ros, Phys. Rep. **470**, 151 (2009).

[69] M. Hochbruck and C. Lubich, SIAM J. Numer. Anal. **41**, 945 (2003).

[70] R. Radovitzky and M. Ortiz, Comput. Methods Appl. Mech. Eng. **172**, 203 (1999).

[71] J. P. Boyd, *Chebyshev and Fourier Spectral Methods* (Dover, New York, 2001).

[72] A. Castro, M. A. Marques, and A. Rubio, J. Chem. Phys. **121**, 3425 (2004).

[73] M. Hochbruck, C. Lubich, and H. Selhofer, SIAM J. Sci. Comput. **19**, 1552 (1998).

[74] C.-L. Cheng, J. S. Evans, and T. Van Voorhis, Phys. Rev. B **74**, 155112 (2006).

[75] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[76] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, J. Comput. Phys. **219**, 172 (2006).

[77] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, Phys. Rev. E **74**, 066704 (2006).

[78] Y. Kawashita, K. Yabana, M. Noda, K. Nobusada, and T. Nakatsukasa, J. Mol. Struct.: THEOCHEM **914**, 130 (2009).

[79] T. Brabec and F. Krausz, Rev. Mod. Phys. **72**, 545 (2000).

[80] J. Pask, N. Sukumar, M. Guney, and W. Hu, Partition-of-unity finite-element method for large scale quantum molecular dynamics on massively parallel computational platforms, Technical Report No. LLNL-TR-470692, Lawrence Livermore National Laboratory (LLNL), https://e-reports-ext.llnl.gov/pdf/471660.pdf.

[81] J. Pask, N. Sukumar, and S. Mousavi, Int. J. Multiscale Comput. Eng. **10**, 83 (2012).

[82] J. E. Pask and N. Sukumar, Extreme Mech. Lett. **11**, 8 (2017).

[83] S. Mousavi, J. Pask, and N. Sukumar, Int. J. Numer. Methods Eng. **91**, 343 (2012).

[84] P. Ciarlet, *The Finite Element Method for Elliptic Problems* (SIAM, Philadelphia, 2002).