Tucker-tensor algorithm for large-scale Kohn-Sham density functional theory calculations

Phani Motamarri and Vikram Gavini
Department of Mechanical Engineering, University of Michigan, Michigan 48109-2125, USA

Thomas Blesgen
Applied Mathematics, University of Applied Sciences Bingen, Berlinstr. 109, D-55411 Bingen, Germany
(Received 31 October 2015; revised manuscript received 7 January 2016; published 1 March 2016)

In this work, we propose a systematic way of computing a low-rank globally adapted localized Tucker-tensor basis for solving the Kohn-Sham density functional theory (DFT) problem. In every iteration of the self-consistent field procedure of the Kohn-Sham DFT problem, we construct an additive separable approximation of the Kohn-Sham Hamiltonian. The Tucker-tensor basis is chosen such as to span the tensor product of the one-dimensional eigenspaces corresponding to each of the spatially separable Hamiltonians, and the localized Tucker-tensor basis is constructed from localized representations of these one-dimensional eigenspaces. This Tucker-tensor basis forms a complete basis, and is naturally adapted to the Kohn-Sham Hamiltonian. Further, the locality of this basis in real-space allows us to exploit reduced-order scaling algorithms for the solution of the discrete Kohn-Sham eigenvalue problem. In particular, we use Chebyshev filtering to compute the eigenspace of the Kohn-Sham Hamiltonian, and evaluate nonorthogonal localized wave functions spanning the Chebyshev filtered space, all represented in the Tucker-tensor basis. We thereby compute the electron-density and other quantities of interest, using a Fermi-operator expansion of the Hamiltonian projected onto the subspace spanned by the nonorthogonal localized wave functions. Numerical results on benchmark examples involving pseudopotential calculations suggest an exponential convergence of the ground-state energy with the Tucker rank. Interestingly, the rank of the Tucker-tensor basis required to obtain chemical accuracy is found to be only weakly dependent on the system size, which results in close to linear-scaling complexity for Kohn-Sham DFT calculations for both insulating and metallic systems. A comparative study has revealed significant computational efficiencies afforded by the proposed Tucker-tensor approach in comparison to a plane-wave basis.

DOI: 10.1103/PhysRevB.93.125104

I. INTRODUCTION

Electronic structure calculations within the Kohn-Sham density functional theory (DFT) [1,2] have been very successful in providing significant insights into a wide range of materials properties over the past decade by enabling quantum-mechanically informed studies on ground-state materials properties. The Kohn-Sham approach to DFT is based on the key result by Hohenberg and Kohn [1] that the ground-state properties of a materials system can be described by a functional of electron density, which to date remains unknown. However, Kohn and Sham [2] addressed this challenge in an approximate sense by reducing the many-body problem of interacting electrons to an equivalent problem of noninteracting electrons in an effective mean field governed by the electron density. This effective single-electron formulation encompasses an unknown exchange-correlation term that includes the quantum-mechanical interaction between electrons, which is modeled in practice, and the widely used models have been successful in predicting a range of properties across various materials systems.

However, the computational complexity of traditional approaches of solving the Kohn-Sham problem scales as $\mathcal{O}(M\,N^2)$, where M denotes the number of basis functions and N specifies the system size (number of atoms or number of electrons). This enormous computational cost associated with Kohn-Sham DFT calculations, approaching cubic scaling as $M \propto N$, has restricted the size and complexity of accessible materials systems. Thus, to enable accurate large-scale DFT calculations, it is desirable to develop computational methods employing a systematically improvable and complete basis, but which is also effective as that it can accurately capture the electronic structure using a small number of basis functions (small M). In addition, it is also desirable to develop computational methods that exhibit reduced-order scaling with system size. To this end, this work develops an algorithm to construct an efficient, yet complete, basis that is systematically adapted to the Kohn-Sham Hamiltonian and combines this approach with reduced-order scaling methods for the solution of the Kohn-Sham problem to develop a computationally efficient methodology for large-scale Kohn-Sham DFT calculations.

Among the complete basis sets employed in DFT calculations, the plane-wave basis [3–5] is the most widely used, and is naturally suited for the computation of bulk-properties of materials. Although the plane-wave basis provides variational convergence in the ground-state energy with exponential convergence rate, the computations are restricted to periodic geometries with periodic boundary conditions. Furthermore, the plane-wave basis functions are extended in real space, and this limits the scalability of numerical implementations on parallel computing architectures. The other widely employed basis sets include the atomic-orbital-type basis functions [6–8], which are reduced-order basis functions that provide good accuracy with relatively few basis functions. However, these basis sets do not constitute a complete basis and may not offer systematic convergence for all materials systems. Also, in some cases, parallel scalability across processors is limited due to the nonlocality of these basis functions. Recent efforts

have also focused on developing adaptive reduced-order basis functions [9,10], which offers a promising direction to develop computationally efficient large-scale DFT calculations.

Over the past few decades, systematically improvable real-space techniques for DFT calculations have been an active area of research. Some notable developments include discretization techniques based on finite difference discretization [11,12], wavelet basis [13,14], and finite element basis [15–19]. Among the real-space techniques, the finite element basis—a piecewise polynomial basis—has desirable features such as admitting general geometries and boundary conditions, locality of the basis functions that supports development of reduced-order scaling methods via localization, and good parallel scalability. However, the number of basis functions M required to achieve chemical accuracy is usually larger in comparison to plane-wave basis and atomic-orbital basis. Thus it is desirable to develop a basis that is systematically improvable and complete such as plane waves, wavelets, or finite elements, has locality in real space such as wavelets and finite elements, is efficient such as atomic-orbital type basis, and exhibits good parallel scalability.

In addition to developing efficient basis functions, many efforts in the past decade have focused on developing algorithms for the solution of Kohn-Sham equations that have a reduced computational complexity. We refer to [20,21] for a comprehensive review of these methods. These methods usually exploit the locality [22] in representing the wave functions directly or indirectly, by either computing the single-electron density matrix (divide and conquer method [23–25], Fermi-operator expansion type techniques [26–30], density-matrix minimization [31,32], subspace projection type methods [33,34]), or representing the density matrix in terms of localized Wannier functions (Fermi-operator projection method [35,36], orbital minimization approach [37,38]). While these methods have been successful in achieving linear-scaling complexity for materials systems with a band gap, the computational complexity can deviate significantly from linear scaling for metallic systems with vanishing band gaps. The development of reduced-order scaling techniques which can handle both insulating and metallic systems in a unified framework is still an active area of research [26,27,29,30,34].

In this work, we exploit low-rank tensor-structured approximations [39,40] to develop a Tucker-tensor algorithm for solving the Kohn-Sham equations. This constitutes constructing a complete, yet efficient localized Tucker-tensor basis that is adapted to the Kohn-Sham Hamiltonian, and using subspace-projected localization techniques for the solution of Kohn-Sham equations in the Tucker-tensor basis. This work has been inspired by recent studies on *a posteriori* numerical analysis of the computed electronic structure of materials systems [41], which revealed that tensor-structured approximations based on canonical and Tucker type representations [42–44] can provide low-rank approximations to the electronic structure of a wide range of materials systems. Further, a recent study [45] has shown that the Tucker rank required to approximate the computed electronic structure of materials is only weakly dependent on the system size, thus providing a useful direction to exploit the low-rank Tucker approximation for developing reduced-order scaling algorithms for DFT calculations.

The key challenge in this work is to develop a systematic procedure for computing the Tucker-tensor basis adapted to the Kohn-Sham eigenvalue problem in order to efficiently represent the *a priori* unknown Kohn-Sham wave functions. To this end, for every self-consistent field (SCF) iteration of a DFT calculation, we compute a spatially additive separable approximation of the Kohn-Sham Hamiltonian and solve for the 1D-eigenfunctions of the separable one-dimensional Hamiltonians. Using a localization procedure [46], we construct a one-dimensional nonorthogonal localized basis spanning the eigenspaces of the corresponding one-dimensional Hamiltonians. We then construct the Tucker-tensor basis using the tensor product of these one-dimensional localized basis functions. The discrete Kohn-Sham eigenproblem is subsequently computed by projecting the continuous problem onto the space spanned by this Tucker-tensor basis, where all the operations are conducted using tensor-structured algorithms. The eigenspace corresponding to the occupied states of the discrete Kohn-Sham Hamiltonian is computed by Chebyshev filtering followed by the computation of nonorthogonal localized wave functions (represented in the Tucker-tensor basis) spanning the eigenspace. The relevant quantities such as the density matrix, the electron-density, and the band energy are computed via Fermi-operator expansion of the subspace-projected Hamiltonian onto the space spanned by the nonorthogonal localized wave functions.

The proposed Tucker-tensor approach constructs a localized tensor-structured basis adapted to the Kohn-Sham Hamiltonian in every SCF iteration and consequently deviates significantly from the usual fixed spatial basis sets currently employed in DFT calculations. The complexity estimates suggest that the proposed algorithm scales linearly with system size if the discretized matrices in the localized Tucker-tensor basis and the localized wave functions are sufficiently sparse (realized for large-scale materials systems). Even in the case where the sparsity is not realized, like relatively smaller materials systems, reduced-order scaling with system size is obtained if the Tucker-rank remains only weakly dependent on the system size.

In order to assess the accuracy and performance of the proposed Tucker-tensor algorithm, we conduct benchmark pseudopotential calculations on both metallic and insulating systems. In all our benchmark studies, we observe an exponential convergence in the ground-state energy with the Tucker rank. Further, we find that the number of Tucker-tensor basis functions required to obtain chemical accuracy grows sublinearly with the system size, both for metallic and insulating systems. Interestingly, the Tucker rank, and hence the number of Tucker-tensor basis functions, was insensitive to significant perturbations in the electronic structure—such as those resulting from introducing random vacancies in a nanocluster. The computational time for these benchmark calculations suggests a close to linear-scaling complexity with respect to the system size for both metallic and insulating systems, which is closely related to the sublinear dependence on the number of Tucker-tensor basis functions with the system size. In the limit of very large system sizes, the required number of Tucker-tensor basis functions will scale linearly with system size. However, a sufficient increase in the system size renders the matrices involved in the proposed

algorithm sparse, owing to the locality in the Tucker-tensor basis and the wave functions, which in turn results in a linear-scaling computational complexity of the proposed approach. A comparative study of the proposed approach on modest-size benchmark calculations suggests that the number of Tucker-tensor basis functions required to achieve chemical accuracy is about five times lower than the number of plane-wave basis functions, and offers about a three to fourfold computational speedup compared to plane-wave discretizations.

The remainder of this article is organized as follows. We begin by recalling some fundamentals of tensor-structured techniques in Sec. II, followed by the real-space formulation of the Kohn-Sham density functional theory in Sec. III. We then discuss the proposed Tucker-tensor algorithm for Kohn-Sham DFT in Sec. IV followed by the numerical results on benchmark problems in Sec. V. We conclude with an outlook on future developments in Sec. VI.

II. LOW-RANK TENSOR APPROXIMATIONS

Tensors, when represented efficiently by a small number of parameters, have significant advantages in terms of reducing the storage and computational costs in a variety of applications. For convenience, we recall here some fundamental concepts of the tensor-structured methods and refer to [39,42–44] for a comprehensive review. For convenience, we restrict our presentation here to tensors of order three.

Let A be a real-valued tensor of order three,

$$A = \left(a_{i_1 i_2 i_3}\right) \in \mathbb{V}, \tag{1}$$

where $(i_1,i_2,i_3) \in \times_{k=1}^{3} I_k$ with nonempty finite index sets I_1, I_2, $I_3 \subset \mathbb{N}$, and $\mathbb{V} := \times_{k=1}^{3} \mathbb{V}_k$ with $\mathbb{V}_k := \mathbb{R}^{|I_k|}$.

The simplest decomposition of a given tensor is the *canonical decomposition* [44], given by a linear combination of rank-1 tensors

$$A \approx A^{(R)} = \sum_{i=1}^{R} c_i v_i^{(1)} \otimes v_i^{(2)} \otimes v_i^{(3)}, \tag{2}$$

where $\{v_i^{(k)}\}_{i=1}^{R}$ is a set of orthonormal vectors for $k = 1,2,3$. The parameter R in the decomposition is called the *canonical rank* of the tensor approximation. The storage cost of the tensor A in the canonical representation is $\mathcal{O}(Rn)$, where $n := \max_{k=1,2,3} |I_k|$ denotes the univariate grid size. However, the computation of this decomposition is a NP-hard and ill-posed problem [47]. Fast and stable algorithms for reducing arbitrary full-size tensors to the canonical format with controlled accuracy are lacking.

On the other hand, robust algorithms for the representation of the full-size tensors in the rank-structured *Tucker-tensor* format are available, and thus this is the preferred tensor-structured format in this work. The rank (r_1,r_2,r_3)-Tucker representation (approximation) of A is given by

$$A^{(\mathbf{r})} = \sum_{l_1=1}^{r_1} \sum_{l_2=1}^{r_2} \sum_{l_3=1}^{r_3} \beta_{l_1 l_2 l_3} v_{l_1}^{(1)} \otimes v_{l_2}^{(2)} \otimes v_{l_3}^{(3)}. \tag{3}$$

In Eq. (3), for each $k \in \{1,2,3\}$, $\{v_{l_k}^{(k)}\}_{1 \leqslant l_k \leqslant r_k}$ constitutes an orthonormal basis of $\mathbb{T}_k := \mathrm{span}_{1 \leqslant l_k \leqslant r_k} v_{l_k}^{(k)}$ with $\dim \mathbb{T}_k = r_k$. The coefficients tensor $\beta := (\beta_{l_1 l_2 l_3}) \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ is called the *core tensor*. As shown in Ref. [41], the Tucker approximation error of the electronic structure of molecular systems decays exponentially with increasing *Tucker rank* $r := \max_{k=1,2,3} r_k$. Further, the overall storage cost of $A^{(\mathbf{r})}$ is bounded by $r^3 + 3rn$. Since usually $r \ll n$, this leads to an impressive data compression [41,45]. Furthermore, $A^{(\mathbf{r})}$ can be computed from A by a minimization procedure,

$$A^{(\mathbf{r})} := \operatorname*{argmin}_{\mathcal{A} \in \mathcal{T}_{\mathbf{r}}} \|\mathcal{A} - A\|_F^2, \tag{4}$$

where $\|A\|_F = \sqrt{\mathrm{tr}(A^T A)}$ is the Frobenius norm. One method for solving the minimization problem in Eq. (4) is the alternating least squares (ALS) algorithm [44], and we refer to [40,43] for other algorithms.

The Tucker-tensor approximation discussed above becomes unattractive in higher dimensions due to the exponentially growing memory requirements for storing the core tensor when dealing with larger dimensions. This has motivated alternative tensor-structured formats like tensor trains (TT) [48,49], wherein a d-dimensional tensor $\boldsymbol{A} = (a_{i_1 i_2 i_3 \ldots i_d})$ is approximated as

$$\boldsymbol{A} \approx \sum_{\alpha_1,\alpha_2,\cdots,\alpha_{d-1},\alpha_d} G^{(1)}_{i_1 \alpha_1} G^{(2)}_{\alpha_1 i_2 \alpha_2} \ldots G^{(d)}_{\alpha_d i_d}, \tag{5}$$

where auxiliary indices α_k vary from 1 to r_k and r_k are called compression ranks or simply TT ranks. The basic arithmetic and storage involved in the TT approach is linear in dimension d and polynomial in $r = \max_k r_k$. We also note that more-general tensor decomposition approaches like the hierarchical tensor representation [50–52] and tree tensor network states [53,54] have been proposed to reduce the computational complexity and storage costs of the tensor-structured representations.

In this work, as explained in Sec. IV, we focus on developing a methodology to compute a Tucker-tensor basis that effectively represents the single-electron wave functions spanning the occupied eigenspace of the Kohn-Sham Hamiltonian. We restrict ourselves in this work to the Tucker-tensor format since the single-electron wave functions are functions in a three-dimensional space where the Tucker-tensor format is efficient. Furthermore, the underlying representation of the Tucker-tensor format provides a convenient way of computing the Galerkin projection of the continuous Kohn-Sham problem into the computed Tucker-tensor basis as discussed subsequently.

III. THE KOHN-SHAM DENSITY FUNCTIONAL THEORY

In Kohn-Sham density functional theory (DFT) [2,55], the variational problem of evaluating the ground-state properties of a given materials system consisting of N_e electrons and N_a atomic nuclei located at $\mathbf{R} = (\mathbf{R}_j)_{1 \leqslant j \leqslant N_a}$ is equivalent to solving the nonlinear eigenvalue problem for $N > N_e/2$ smallest eigenvalues

$$\left(-\tfrac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\varrho,\mathbf{R})\right)\psi_i = \epsilon_i \psi_i, \qquad i = 1,2,\ldots,N, \tag{6}$$

where ϵ_i and ψ_i denote the eigenvalues and the corresponding eigenfunctions (canonical single particle wave functions) of the Hamiltonian, respectively. In the present work, for the sake of simplicity, we discuss the formulation in a nonperiodic

setting restricting ourselves to spin-independent Hamiltonians. However, the present discussion as well as the ideas proposed subsequently can easily be generalized to periodic or semiperiodic materials systems and spin-dependent Hamiltonians.

The electron density—a central quantity in DFT—at any spatial point $\mathbf{x} = (x_1, x_2, x_3)$ in terms of the canonical wave functions is given by

$$\varrho(\mathbf{x}) = 2\sum_{i=1}^{N} f(\epsilon_i, \mu)|\psi_i(\mathbf{x})|^2, \tag{7}$$

where $f(\epsilon, \mu) \in [0,1]$ is the orbital occupancy function, and μ represents the Fermi energy which is computed from the constraint that the total number of electrons in the system is N_e. In ground-state DFT calculations, it is common to represent f by the Fermi distribution $f(\epsilon, \mu) = 1/(1 + \exp[(\epsilon - \mu)/\sigma])$, which tends to a Heaviside function as the parameter $\sigma \searrow 0$.

In Eq. (6), the effective single-electron potential in the Hamiltonian is given by

$$\begin{aligned} V_{\text{eff}}(\varrho, \mathbf{R}) &:= \frac{\delta E_{\text{xc}}}{\delta \varrho} + \frac{\delta E_{\text{H}}}{\delta \varrho} + V_{\text{ext}}(\mathbf{R}) \\ &= V_{\text{xc}}(\varrho) + V_{\text{H}}(\varrho) + V_{\text{ext}}(\mathbf{R}). \end{aligned} \tag{8}$$

In the above, E_{xc} represents the exchange-correlation energy that accounts for quantum-mechanical interactions between electrons, and we adopt the widely used local density approximation (LDA) [56,57]. The Hartree energy, E_{H}, represents the classical electrostatic interaction energy between the electrons and is given by

$$E_{\text{H}}(\varrho) := \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\varrho(\mathbf{x}')\varrho(\mathbf{x})}{|\mathbf{x} - \mathbf{x}'|}\, d\mathbf{x}'\, d\mathbf{x} = \frac{1}{2}\int_{\mathbb{R}^3} V_{\text{H}}(\varrho)\varrho(\mathbf{x})\, d\mathbf{x}. \tag{9}$$

Finally, $V_{\text{ext}}(\mathbf{R})$ denotes the external electrostatic potential corresponding to the nuclear charges. In this work, we adopt the commonly used pseudopotential approach, where only the valence-electron wave functions are computed. The pseudopotential, which provides the effective nuclear electrostatic potential $V_{\text{ext}}(\mathbf{R})$ for the valence electrons, is commonly represented by the operator $\mathcal{V}_{\text{ext}} = \mathcal{V}_{\text{loc}} + \mathcal{V}_{\text{nl}}$, where $\mathcal{V}_{\text{loc}}$ is the local part and $\mathcal{V}_{\text{nl}}$ is its nonlocal part. Using the norm-conserving Troullier-Martins pseudopotentials [58] in the Kleinman-Bylander form [59], the action of these operators on a Kohn-Sham wave function in real space is given by

$$\begin{aligned} V_{\text{loc}}(\mathbf{x}, \mathbf{R})\psi(\mathbf{x}) &:= \sum_{J=1}^{N_a} V_{\text{loc}}^J(\mathbf{x} - \mathbf{R}_J)\psi(\mathbf{x}), \\ V_{\text{nl}}(\mathbf{x}, \mathbf{R})\psi(\mathbf{x}) &:= \sum_{J=1}^{N_a}\sum_{lm} C_{lm}^J \varphi_{lm}^J(\mathbf{x} - \mathbf{R}_J)\Delta V_l^J(\mathbf{x} - \mathbf{R}_J), \end{aligned} \tag{10}$$

where

$$\Delta V_l^J(\mathbf{x} - \mathbf{R}_J) := V_l^J(\mathbf{x} - \mathbf{R}_J) - V_{\text{loc}}^J(\mathbf{x} - \mathbf{R}_J),$$

$$C_{lm}^J := \frac{\int \varphi_{lm}^J(\mathbf{x} - \mathbf{R}_J)\Delta V_l^J(\mathbf{x} - \mathbf{R}_J)\psi(\mathbf{x})\, d\mathbf{x}}{\int \varphi_{lm}^J(\mathbf{x} - \mathbf{R}_J)\Delta V_l^J(\mathbf{x} - \mathbf{R}_J)\varphi_{lm}^J(\mathbf{x} - \mathbf{R}_J)\, d\mathbf{x}}.$$

In the above, $V_l^J(\mathbf{x} - \mathbf{R}_J)$ denotes the pseudopotential component of atom J corresponding to the azimuthal quantum number l, $V_{\text{loc}}^J(\mathbf{x} - \mathbf{R}_J)$ is the corresponding local potential, and $\varphi_{lm}^J(\mathbf{x} - \mathbf{R}_J)$ is the corresponding single-atom pseudo-wave-function with azimuthal quantum number l and magnetic quantum number m.

For given positions of nuclei, the system of equations corresponding to the Kohn-Sham eigenvalue problem is

$$\begin{aligned} &\mathcal{H}\psi_i = \epsilon_i \psi_i, \\ &2\sum_{i=1}^{N} f(\epsilon_i, \mu) = N_e, \quad \varrho(\mathbf{x}) = 2\sum_{i=1}^{N} f(\epsilon_i, \mu)|\psi_i(\mathbf{x})|^2, \end{aligned} \tag{11}$$

where

$$\mathcal{H} := \left(-\tfrac{1}{2}\nabla^2 + V_{\text{xc}}(\varrho) + V_{\text{H}}(\varrho) + V_{\text{loc}}(\mathbf{x}, \mathbf{R}) + V_{\text{nl}}(\mathbf{x}, \mathbf{R})\right). \tag{12}$$

As the Hamiltonian $\mathcal{H}$ depends on ϱ, which in turn is computed from the eigenfunctions of $\mathcal{H}$, the system of equations in Eq. (11) is solved by a self-consistent field (SCF) iteration in a suitable basis. Upon self-consistently solving the Kohn-Sham eigenvalue problem, the ground-state energy is given by

$$\begin{aligned} E_{\text{tot}} &= 2\sum_{i=1}^{N} f(\epsilon_i, \mu)\epsilon_i + E_{\text{xc}}(\varrho) - \int_{\mathbb{R}^3} V_{\text{xc}}(\varrho)\varrho\, d\mathbf{x} \\ &\quad - \frac{1}{2}\int_{\mathbb{R}^3} \varrho V_{\text{H}}(\varrho)\, d\mathbf{x} + \frac{1}{2}\sum_{\substack{I,J=1 \\ I \neq J}}^{N_a} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \end{aligned}$$

Therein, the last term on the right denotes the nuclear-nuclear repulsive energy E_{ZZ} with Z_I denoting the valence charge of the Ith nucleus.

IV. TUCKER-TENSOR ALGORITHM FOR DFT

We now present a Tucker-tensor algorithm for the solution of the Kohn-Sham equations that has reduced computational complexity in comparison to conventional approaches. In every cycle of the SCF iteration, the proposed algorithm provides a prescription to compute a nonorthogonal locally adapted Tucker-tensor basis using a separable approximation of the Hamiltonian. The Kohn-Sham eigenvalue problem is subsequently solved by projecting the problem onto the span of this computed Tucker-tensor basis, and by computing the eigenspace corresponding to the occupied states using Chebyshev filtering techniques. Let $\varrho^{(n)}$ denote the input electron density to the n^{th} SCF iteration and $\mathcal{H}^n \equiv \mathcal{H}(\varrho^{(n)}(\mathbf{x}), \mathbf{R})$ be the corresponding Hamiltonian. The proposed Tucker-tensor algorithm consists of the following key steps with specific details discussed subsequently. (1) Construct a separable approximation of the Hamiltonian by using one of two proposed competing variational methods (outlined below),

$$\mathcal{H}_x + \mathcal{H}_y + \mathcal{H}_z \sim \mathcal{H}^n. \tag{13}$$

(2) Compute r_d one-dimensional eigenfunctions for $\mathcal{H}_x$, $\mathcal{H}_y$, $\mathcal{H}_z$ represented on a finite element grid, and subsequently

employ a localization procedure to evaluate nonorthogonal localized basis functions spanning the eigensubspaces in each spatial dimension. (3) Compute a nonorthogonal localized Tucker-tensor basis $\mathbb{T}^L := (T^L_{ijk})_{1\leqslant i,j,k\leqslant r_d}$ as the tensor-product of the one-dimensional localized basis functions of step 2. (4) Compute the projection $\mathcal{H}^n_h$ of $\mathcal{H}^n$ onto $\mathbb{T}^L$. (5) Employ Chebyshev filtering to compute the approximate occupied eigensubspace of $\mathcal{H}^n_h$ corresponding to the lower end of the eigenspectrum comprising of occupied states and a few unoccupied states above the Fermi energy. Subsequently, localize the Chebyshev filtered wave functions by utilizing a nonorthogonal localization procedure as described in Ref. [34]. (6) Project $\mathcal{H}^n_h$ onto the occupied eigensubspace of $\mathcal{H}^n_h$ represented by the localized Chebyshev filtered wave functions. Employ a Fermi-operator expansion of this subspace-projected Hamiltonian to compute the relevant quantities of interest such as the density matrix, the output electron-density and the ground-state energy. Then proceed with the SCF iteration.

We now begin to discuss various details of the proposed algorithm. Let the domain be cuboidal, i.e., $\Omega = \times_{k=1}^{3}\omega_k$ with one-dimensional bounded real sets ω_k, and enclose the compact support of the Kohn-Sham wave functions. We discretize ω_k by using isoparametric 1D finite elements, and represent functions on ω_k by using finite element basis functions—the piecewise polynomial functions constructed from the finite element discretization [60]. We denote by n_k (for $k = 1,2,3$) the dimension of the finite element space discretizing ω_k, or, in other words, the number of finite element basis functions in each spatial dimension k. In the present work, we use a higher-order finite element discretization with polynomial degree $p > 2$. We note that, while the ideas presented in this work are equally applicable to any basis, the choice of the finite element basis is motivated by the locality of the basis and its adaptive capability.

Given the input electron density to the n^{th} SCF iteration, $\varrho^{(n)}(\mathbf{x})$, we begin by computing the local effective potential on Ω,

$$V^{\text{loc}}_{\text{eff}}(\mathbf{x}) := V_{\text{xc}}(\varrho^{(n)}(\mathbf{x})) + V_{\text{H}}(\varrho^{(n)}(\mathbf{x})) + V_{\text{loc}}(\mathbf{x}). \tag{14}$$

We note that the evaluation of V_{H} [cf. Eq. (9)] involves the computation of a 3D convolution integral. To this end, for chosen rank $r_\varrho \in \mathbb{N}$ and $\mathbf{x}' = (x'_1,x'_2,x'_3)$, we first compute the rank-r_ϱ Tucker-tensor decomposition of the density $\varrho^{(n)}(\mathbf{x})$ as

$$\varrho^{(n)}(\mathbf{x}') \approx \sum_{i,j,k}^{r_\varrho} \sigma^{(n)}_{ijk}\varrho^{(n)}_i(x'_1)\varrho^{(n)}_j(x'_2)\varrho^{(n)}_k(x'_3). \tag{15}$$

Next, we approximate the kernel $|\mathbf{x}-\mathbf{x}'|^{-1}$ by a series of Gaussians (see Ref. [61], where also the values of α_p, β_p are tabulated), and obtain for a rank parameter $T \in \mathbb{N}$,

$$\frac{1}{|\mathbf{x}-\mathbf{x}'|} \approx \sum_{p=1}^{T} \alpha_p e^{-\beta_p(x_1-x'_1)^2} e^{-\beta_p(x_2-x'_2)^2} e^{-\beta_p(x_3-x'_3)^2}. \tag{16}$$

Thus the computation of $V_{\text{H}}(\varrho^{(n)})$ reduces to the computation of a series of 1D convolution integrals, as

$$\begin{aligned} V_{\text{H}}(\varrho^{(n)}(\mathbf{x})) &= \int_\Omega \frac{\varrho^{(n)}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,d\mathbf{x}' \\ &\approx \sum_{p=1}^{T}\alpha_p \sum_{i,j,k}^{r_\varrho}\sigma^{(n)}_{ijk}\Bigg[\int_{\omega_1}\varrho^{(n)}_i(x'_1)e^{-\beta_p(x_1-x'_1)^2}\,dx'_1 \\ &\quad\times\int_{\omega_2}\varrho^{(n)}_j(x'_2)e^{-\beta_p(x_2-x'_2)^2}dx'_2 \\ &\quad\times\int_{\omega_3}\varrho^{(n)}_k(x'_3)e^{-\beta_p(x_3-x'_3)^2}dx'_3\Bigg]. \end{aligned} \tag{17}$$

Upon evaluating V_H, we compute $V^{\text{loc}}_{\text{eff}}$ given by Eq. (14). Further, to aid the evaluation of terms arising in subsequent computations, we compute the rank-r_v Tucker-tensor decomposition of $V^{\text{loc}}_{\text{eff}}$, denoted by $\widehat{V^{\text{loc}}_{\text{eff}}}(\mathbf{x})$. For the same reason, by evaluating the rank-r_v Tucker-tensor decomposition of the atom-centered pseudopotential and pseudo-wave-function components, we compute the tensor-structured approximation of the nonlocal part of the pseudopotential operator and denote this by $\widehat{V_{\text{nl}}}(\mathbf{x},\mathbf{R})$.

A. Separable approximation of $\mathcal{H}^n$

We now explain step 1 of the Tucker-tensor algorithm in more detail and present two methods to compute the additive separable approximation of $\mathcal{H}^n$. One of the proposed methods is based on a rank-1 approximation of the eigenfunction corresponding to the lowest eigenvalue of the Kohn-Sham Hamiltonian, while the second method involves an additive separable approximation of the Kohn-Sham potential V_{eff}. While the first method is applicable to both local and nonlocal pseudopotentials, the latter is restricted to local pseudopotentials, only.

a. Method I. Rank-1 decomposition of wave functions. We start with the ansatz for the eigenfunction,

$$\psi(\mathbf{x}) := \prod_{k=1}^{3}\psi_k(x_k), \tag{18}$$

and denote by X the function space of all one-time (weakly) differentiable rank-1 functions in Ω. The problem of computing the smallest eigenvalue of the Kohn-Sham Hamiltonian [Eq. (12)] in the function space X is equivalent to the variational problem

$$\min_{\psi\in X}\mathcal{L}(\psi), \tag{19}$$

with the Lagrangian

$$\begin{aligned} \mathcal{L}(\psi) :=& \frac{1}{2}\int_\Omega\Bigg[\sum_{p=1}^{3}|\partial_{x_p}\psi_p(x_p)|^2\prod_{\substack{q=1\\q\neq p}}^{3}\psi^2_q(x_q) \\ &+ 2\big(\widehat{V^{\text{loc}}_{\text{eff}}}(\mathbf{x})+\lambda\big)\prod_{k=1}^{3}\psi^2_k(x_k) \\ &+ 2\prod_{k=1}^{3}\psi_k(x_k)\widehat{V_{\text{nl}}}(\mathbf{x},\mathbf{R})\prod_{k=1}^{3}\psi_k(x_k)\Bigg]d\mathbf{x}. \end{aligned}$$

Here, λ is a Lagrange multiplier corresponding to the constraint

$$\prod_{k=1}^{3}\int_{\omega_k}\psi_k^2(x_k)\,dx_k = 1. \tag{20}$$

Minimizers of (19) satisfy the Euler-Lagrange equations $\frac{\delta\mathcal{L}(\psi)}{\delta\psi_k}=0$ for $k=1,2,3$. Hence the minimizers ψ_k are the solutions of the one-dimensional problems

$$\left[-\frac{1}{2}\frac{d^2}{dx_k^2}+\frac{V_k^{\rm loc}(x_k)}{m_k}+\frac{V_k^{\rm nl}(x_k)}{m_k}\right]\psi_k(x_k) = -\left(\lambda+\frac{a_k}{2m_k}\right)\psi_k(x_k), \tag{21}$$

where we introduced the one-dimensional quantities

$$V_k^{\rm loc}(x_k) := \int_{\hat{\omega}_k}\widehat{V_{\rm eff}^{\rm loc}}(\mathbf{x})\prod_{\substack{j=1\\ j\neq k}}^{3}\psi_j^2(x_j)\,d\hat{\mathbf{x}}_k,$$

$$m_k := \int_{\hat{\omega}_k}\prod_{\substack{j=1\\ j\neq k}}^{3}\psi_j^2(x_j)\,d\hat{\mathbf{x}}_k,$$

$$a_k := \int_{\hat{\omega}_k}\sum_{\substack{p,q=1\\ p\neq q;\,p,q\neq k}}^{3}\left|\partial_{x_p}\psi_p(x_p)\right|^2\psi_q(x_q)\,d\hat{\mathbf{x}}_k,$$

$$V_k^{\rm nl}(x_k)\psi_k(x_k) := \sum_{J=1}^{N_{\rm a}}\sum_{lm}\gamma_{lm}^J\,v_{lm}^J\,F_{lm}^{J,k}(x_k) \quad \text{with}$$

$$F_{lm}^{J,k}(x_k) := \int_{\hat{\omega}_k}\widehat{\varphi_{lm}^J}(\mathbf{x}-\mathbf{R}_J)\widehat{\Delta V_l^J}(\mathbf{x}-\mathbf{R}_J)\prod_{\substack{j=1\\ j\neq k}}^{3}\psi_j(x_j)\,d\hat{\mathbf{x}}_k,$$

$$\gamma_{lm}^J := \int_{\omega_k}F_{lm}^{J,k}(x_k)\psi_k(x_k)\,dx_k,$$

$$\left(v_{lm}^J\right)^{-1} := \int_{\Omega}\widehat{\varphi_{lm}^J}(\mathbf{x}-\mathbf{R}_J)\widehat{\Delta V_l^J}(\mathbf{x}-\mathbf{R}_J)\widehat{\varphi_{lm}^J}(\mathbf{x}-\mathbf{R}_J)\,d\mathbf{x},$$

with notations

$$d\hat{\mathbf{x}}_1 := dx_2\,dx_3, \quad d\hat{\mathbf{x}}_2 := dx_1\,dx_3,$$
$$d\hat{\mathbf{x}}_3 := dx_1\,dx_2, \quad \hat{\omega}_k := \bigtimes_{\substack{j=1\\ j\neq k}}^{3}\omega_j.$$

In the above expressions, $\widehat{\Delta V_l^J}$ and $\widehat{\varphi_{lm}^J}$ denote the rank-r_v Tucker-tensor decomposition of ΔV_l^J and φ_{lm}^J, respectively. We note that the integrals involved in the above expressions reduce to a product of integrals in one dimension owing to the tensor-structured representation of all field quantities, thus rendering the computational complexity of evaluating these terms very low.

The minimizing functions $\psi_k(x_k)$ obtained from the self-consistent solution of (21) are fixed to construct the one-dimensional potentials $V_k^{\rm loc}$ and $V_k^{\rm nl}$. The eigenfunctions of the resulting one-dimensional Hamiltonians in Eq. (21) are then used to construct the Tucker-tensor basis after localization, see Sec. IV B below.

b. Method II. Weighted residual minimization. In this method, which is restricted to local pseudopotentials, we construct an additive separable approximation of $V_{\rm eff}^{\rm loc}$ by solving the weighted residual minimization problem

$$\min_{V_k^{\rm loc}\in L^1(\omega_k)\,1\leqslant k\leqslant 3}\int_{\Omega}w(\mathbf{x})\left[\widehat{V_{\rm eff}^{\rm loc}}(\mathbf{x})-\sum_{l=1}^{3}V_l^{\rm loc}(x_l)\right]^2 d\mathbf{x}, \tag{22}$$

where $w(\mathbf{x})\in L^2(\Omega)$ represents a nonnegative weight function. We then construct the one-dimensional Hamiltonians for $k=1,2,3$ as $\mathcal{H}_k := -\frac{1}{2}\frac{d^2}{dx_k^2}+V_k^{\rm loc}(x_k)$ resulting in the one-dimensional eigenvalue problems

$$\mathcal{H}_k\xi_{k,i} = \varepsilon_{k,i}\xi_{k,i}\,. \tag{23}$$

The weight is chosen as $w(\mathbf{x}):=|\varrho^{(n)}(\mathbf{x})|^{\alpha}$ with $\alpha:=1$ to penalize the error in the separable approximation of $\widehat{V_{\rm eff}^{\rm loc}}(\mathbf{x})$ in the vicinity of atoms where the electron density is higher in comparison to the regions far away from the atoms.

B. Construction of a 3D Tucker-tensor basis $\mathbb{T}^L$

The methods outlined in Sec. IV A provide a systematic approach to constructing an additive separable approximation to the Kohn-Sham Hamiltonian. Solving the eigenvalue problems [Eq. (21) for method I or Eq. (23) for method II], we compute the eigenfunctions associated with the one-dimensional Hamiltonians in each spatial dimension. We remark that the one-dimensional eigenfunctions thus computed form a complete basis for admissible functions on each ω_k. In the discrete numerical setting, we compute r_{d_1}, r_{d_2}, r_{d_3} eigenfunctions corresponding to the lowest eigenvalues of the one-dimensional Hamiltonians in x_1, x_2, x_3 spatial directions, respectively. For the sake of notational simplicity in presenting our ideas, we assume $r_{d_1}=r_{d_2}=r_{d_3}=:r_d$. We denote by $(\xi_{k,i})_{1\leqslant i\leqslant r_d}$ the eigenfunctions in the direction k spanning the space $\mathbb{V}^{r_{d_k}}$ for $k=1,2,3$. The three-dimensional Tucker-tensor basis for the Kohn-Sham DFT problem can thus be constructed as a tensor product given by

$$\mathbb{T} := \{T_{abc}\}_{1\leqslant a,b,c\leqslant r_d} := \{\xi_{1,a}\xi_{2,b}\xi_{3,c}\}_{1\leqslant a,b,c\leqslant r_d}. \tag{24}$$

However, the eigenfunctions $(\xi_{k,i})_{1\leqslant i\leqslant r_d}$ have a global support on ω_k, thereby rendering the support of the corresponding three-dimensional Tucker-tensor basis global on Ω. The global nature of these functions results in dense matrices for the Kohn-Sham DFT problem, which is not desirable. To this end, we construct a localized representation of the Tucker-tensor basis $\{T_{abc}\}_{1\leqslant a,b,c\leqslant r_d}$ by localizing the 1D eigenfunctions $(\xi_{k,i})_{1\leqslant i\leqslant r_d}$ around the atomic locations in each of the spatial directions x_k for $k=1,2,3$. Various localization procedures employing nonorthogonal localized functions [37,62–64] have been proposed in the context of electronic structure calculations, which have better localizing properties than orthogonal functions. In the present work, we adopt the weighted L^2 localization technique proposed in E *et al.* [46] to construct the localized 1D basis-functions spanning the eigenspace $\mathbb{V}^{r_{d_k}}$ for $k=1,2,3$. However, we note that other localization procedures such as those proposed in Ref. [65] can also be used. We obtain the localized basis by solving for each $k=1,2,3$ the minimization problem

$$\operatorname*{argmin}_{\phi\in\mathbb{V}^{r_{d_k}},\,\|\phi\|=1}\int_{\omega_k}w(x_k)|\phi(x_k)|^2\,dx_k. \tag{25}$$

Here, $w(x_k) \geqslant 0$ is chosen to be a smooth weighting function of the form $|x_k - b_{I_k}|^2$, and b_{I_k} denotes a localization center. Such a choice of $w(x_k)$ minimizes the spread of the basis-functions from a localization center. In the present work, the localization center b_{I_k} is chosen to be the k^{th} direction atom-coordinate R_{I_k} corresponding to the I^{th} atom for $k = 1,2,3$. Let r_{I_k} denote the number of localized functions we desire to compute at every atom-coordinate R_{I_k}. Representing the localized function as

$$\phi(x_k) = \sum_{i=1}^{r_{d_k}} \alpha_i \xi_{k,i}(x_k) \in \mathbb{V}^{r_{d_k}}, \tag{26}$$

the minimization problem in Eq. (25) is equivalent to solving the generalized eigenvalue problem in each spatial direction k for the smallest r_{I_k} eigenvalues,

$$\mathbf{G}^{I_k}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}, \tag{27}$$

where for $i, j = 1, \ldots, r_{d_k}$

$$G^{I_k}_{ij} := \int_{\omega_k} \left|x_k - R_{I_k}\right|^2 \xi_{k,i}(x_k)\, \xi_{k,j}(x_k)\, dx_k\,. \tag{28}$$

In the present work, we choose r_{I_k} corresponding to the I^{th} atom such that $\sum_I r_{I_k} = r_{d_k}$. We note that we can rewrite $\mathbf{G}^{I_k}$ in Eq. (28) in matrix notation as

$$\mathbf{G}^{I_k} = \mathbf{L}_b^T \mathbf{K}_b^{I_k} \mathbf{L}_b, \tag{29}$$

where $(\cdot)^T$ is the matrix transpose, the columns of the matrix $\mathbf{L}_b$ correspond to the finite element nodal values of the eigenfunctions $\{\xi_{k,1}(x_k), \xi_{k,2}(x_k), \ldots, \xi_{k,r_{d_k}}(x_k)\}$, and

$$\left(\mathbf{K}_b^{I_k}\right)_{ij} := \int_{\omega_k} \left|x_k - R_{I_k}\right|^2 N_i(x_k) N_j(x_k)\, dx_k, \tag{30}$$

with N_i denoting the finite element basis function corresponding to node i.

Upon solving Eq. (27) for each I_k, we represent the computed localized one-dimensional functions spanning $\mathbb{V}^{r_{d_k}}$ by $(\phi_{k,i})_{1 \leqslant i \leqslant r_d}$. Thus the three-dimensional localized Tucker-tensor basis functions for solving the Kohn-Sham DFT problem are constructed to be

$$\mathbb{T}^L := \left\{T^L_{abc}\right\}_{1 \leqslant a,b,c \leqslant r_d} := \{\phi_{1,a}\phi_{2,b}\phi_{3,c}\}_{1 \leqslant a,b,c \leqslant r_d}. \tag{31}$$

In practice, we use a truncation tolerance to achieve a compact support for $(\phi_{k,i})_{1 \leqslant i \leqslant r_d}$, and consequently for $\{T^L_{abc}\}_{1 \leqslant a,b,c \leqslant r_d}$.

C. Discrete Kohn-Sham eigenvalue problem

The projection of $\mathcal{H}^n$ onto $\times_{k=1}^3 \mathbb{V}^{r_{d_k}}$, denoted by $\mathcal{H}^n_h$, expressed in the localized Tucker-tensor basis $\mathbb{T}^L$ is given by

$$\left(\mathcal{H}^n_h\right)_{(ijk),(abc)} := \sum_{p,q,r} \left\langle T^L_{ijk} \middle| T^L_{pqr} \right\rangle^{-1} \left\langle T^L_{pqr} \middle| \mathcal{H}^n \middle| T^L_{abc} \right\rangle. \tag{32}$$

We note that it is convenient to approximate the Kohn-Sham potential V_{eff} [Eq. (8)] using a Tucker-tensor decomposition, which reduces all integrals involved in Eq. (32) to products of one-dimensional integrals, and is used in the present work. The discrete Kohn-Sham eigenvalue problem in the localized Tucker-tensor basis is given by the non-Hermitian standard eigenvalue problem

$$\tilde{\mathbf{H}}\boldsymbol{\Psi}_i = \epsilon^h_i \boldsymbol{\Psi}_i, \tag{33}$$

with $\tilde{\mathbf{H}} := \mathbf{M}^{-1}\mathbf{H}$, where $\mathbf{H}$ denotes the discrete Hamiltonian matrix with matrix elements H_{IJ} and $\mathbf{M}$ denotes the overlap matrix arising because of the nonorthogonality of the localized Tucker-tensor basis functions with matrix elements M_{IJ} for subscripts $\text{I,J} \in \times_{k=1}^3 \{1, \ldots, r_{d_k}\}$. By ϵ^h_i we denote the i^{th} eigenvalue corresponding to the discrete eigenvector $\boldsymbol{\Psi}_i$ in Eq. (33) expressed in the localized Tucker-tensor basis $\mathbb{T}^L$. The matrix elements M_{IJ} and H_{IJ} are given by

$$M_{\text{IJ}} := \int_\Omega T^L_{\text{I}}(\mathbf{x})\, T^L_{\text{J}}(\mathbf{x})\, d\mathbf{x}, \tag{34}$$

$$H_{\text{IJ}} := \frac{1}{2}\int_\Omega \nabla T^L_{\text{I}}(\mathbf{x}) \cdot \nabla T^L_{\text{J}}(\mathbf{x}) d\mathbf{x} + \int_\Omega T^L_{\text{I}}(\mathbf{x}) \widehat{V^{\text{loc}}_{\text{eff}}}(\mathbf{x}) T^L_{\text{J}}(\mathbf{x}) d\mathbf{x} + \int_\Omega T^L_{\text{I}}(\mathbf{x}) \widehat{V_{\text{nl}}}(\mathbf{x},\mathbf{R}) T^L_{\text{J}}(\mathbf{x})\, d\mathbf{x}, \tag{35}$$

with $\widehat{V^{\text{loc}}_{\text{eff}}}$ and $\widehat{V_{\text{nl}}}$ denoting the rank-r_v Tucker-tensor decompositions of $V^{\text{loc}}_{\text{eff}}$ and V_{nl}, respectively. As a consequence of applying the Tucker-tensor decompositions $\widehat{V^{\text{loc}}_{\text{eff}}}$ and $\widehat{V_{\text{nl}}}$, the right-hand sides of (34) and (35) reduce to a tensor-structured format involving one-dimensional integrals. Thus the computational complexity associated with the computation of the discrete Hamiltonian and overlap matrix in Eqs. (34) and (35) is evaluated to be $\mathcal{O}(r_d^2\,\text{n}) + \mathcal{O}(r_d^6) + \mathcal{O}(r_d^2 r_v^3\,\text{n}) + \mathcal{O}(r_d^6 r_v^3)$, with $\text{n} := \max_k n_k$ relating to the number of nodes in the one-dimensional finite element mesh (univariate grid size). However, as we use a localized Tucker-tensor basis, by exploiting the locality in the basis functions, the computational complexity of evaluating the matrix elements reduces to $\mathcal{O}(c^{1/3}\text{n}) + \mathcal{O}(c) + \mathcal{O}(c^{1/3} r_v^3 \text{n}) + \mathcal{O}(c\, r_v^3)$, where c denotes the maximum number of nonzero entries in the matrices $\mathbf{H}$ and $\mathbf{M}$. Finally, the inverse overlap matrix $\mathbf{M}^{-1}$ involved in the computation of $\tilde{\mathbf{H}}$ is evaluated using a scaled third-order Newton-Schulz iteration [66].

D. Computation of the DFT ground-state energy

a. Chebyshev filtered subspace iteration. An approximation to the eigenspace of the discrete Kohn-Sham eigenproblem in Eq. (33), spanned by $N > N_e/2$ lowest eigenfunctions, is computed by using a Chebyshev-filtered subspace iteration (ChFSI) technique [67]. We refer to Refs. [19,34] for the details of its implementation in the context of finite element discretization. The ChFSI technique exploits the rapid growth of Chebyshev polynomials in $(-\infty, -1)$ to magnify the relevant eigenspectrum, and thereby providing an efficient approach for the solution of the Kohn-Sham eigenvalue problem.

In each iteration of the SCF procedure, the action of a Chebyshev filter on a given subspace is accomplished by the recursive construction of the Chebyshev polynomial of the discrete Hamiltonian together with its action on the subspace. This involves matrix-vector multiplications between the discretized Hamiltonian $\tilde{\mathbf{H}}$ and the vectors obtained during the course of the recursive iteration. We note that, if the discretized Hamiltonian is sufficiently sparse and the vectors obtained during the process of recursive iteration of the Chebyshev filtering procedure are sparse, the computational complexity of the relevant matrix-vector multiplications scales as $\mathcal{O}(N)$.

b. Localization and truncation. Developing a localized representation of the wave functions spanning the occupied eigenspace is one of the key ideas that has been exploited in developing reduced-order scaling algorithms [33,34], and is also employed here. We use the algorithm developed in Ref. [34] to construct a nonorthogonal localized basis of the subspace spanned by Chebyshev filtered wave functions. We recall the main ideas and present them in the context of the Tucker-tensor basis for the sake of completeness. The localized basis of the subspace spanned by the Chebyshev filtered wave functions, henceforth referred to as the localized Chebyshev filtered basis, is obtained by solving the generalized eigenvalue problem for the n_P smallest eigenvalues for every atom P,

$$\mathbf{W}^P\boldsymbol{\alpha} = \lambda\mathbf{S}\boldsymbol{\alpha}, \tag{36}$$

where for $l,m = 1,\ldots,N$

$$W^P_{lm} := \int_\Omega |\mathbf{x}-\mathbf{R}_P|^2\,\psi^f_l(\mathbf{x})\,\psi^f_m(\mathbf{x})\,d\mathbf{x}, \tag{37a}$$

$$S_{lm} := \int_\Omega \psi^f_l(\mathbf{x})\,\psi^f_m(\mathbf{x})\,d\mathbf{x}, \tag{37b}$$

and n_P denotes the number of localized functions we desire to compute at every atom centered at $\mathbf{R}_P = (R_{Px_1}, R_{Px_2}, R_{Px_3})$. The number n_P is chosen to be equal to the number of occupied single atom orbitals corresponding to the Pth atom; $\boldsymbol{\alpha}$ is a vector containing the coefficients corresponding to the linear combination of Chebyshev filtered functions $\{\psi^f_1(\mathbf{x}),\psi^f_2(\mathbf{x}),\ldots,\psi^f_N(\mathbf{x})\}$. The matrix $\mathbf{W}^P$ can be recast in matrix notation as

$$\mathbf{W}^P = \mathbf{L}^T\mathbf{K}^P\mathbf{L}, \tag{38}$$

where the columns of the matrix $\mathbf{L}$ correspond to the coefficients of the Chebyshev filtered wave functions expressed in Tucker-tensor basis, and with

$$K^P_{\mathrm{IJ}} := \int_\Omega |\mathbf{x}-\mathbf{R}_P|^2\,T^L_{\mathrm{I}}(\mathbf{x})T^L_{\mathrm{J}}(\mathbf{x})\,d\mathbf{x}. \tag{39}$$

Let $\mathbf{K}^0$ denote the matrix in Eq. (39) for a reference atom located at $\mathbf{R}_0$. We note that the matrix $\mathbf{K}^P$ for any P can be represented in terms of $\mathbf{K}^0$ as

$$\mathbf{K}^P = \mathbf{K}^0 + |\mathbf{R}_0-\mathbf{R}_P|^2\mathbf{M} + 2\sum_{k=1}^{3}\left(R_{0x_k}-R_{Px_k}\right)\mathbf{B}^{x_k}, \tag{40}$$

where

$$\mathbf{K}^0 := \left(\mathbf{K}^0_1,\mathbf{O}_2,\mathbf{O}_3\right) + \left(\mathbf{O}_1,\mathbf{K}^0_2,\mathbf{O}_3\right) + \left(\mathbf{O}_1,\mathbf{O}_2,\mathbf{K}^0_3\right),$$
$$\mathbf{M} := (\mathbf{O}_1,\mathbf{O}_2,\mathbf{O}_3)\ ,\ \mathbf{B}^{x_1} := (\mathbf{B}_1,\mathbf{O}_2,\mathbf{O}_3),$$
$$\mathbf{B}^{x_2} := (\mathbf{O}_1,\mathbf{B}_2,\mathbf{O}_3)\ ,\ \mathbf{B}^{x_3} := (\mathbf{O}_1,\mathbf{O}_2,\mathbf{B}_3),$$
$$\text{with the notation}\ \ (\mathbf{X},\mathbf{Y},\mathbf{Z}) := \mathbf{X}\otimes\mathbf{Y}\otimes\mathbf{Z},$$

and [with $\phi_{k,i}$ as in Eq. (31)]

$$\left(\mathbf{K}^0_k\right)_{ij} := \int_{\omega_k}\left(x_k - R_{0x_k}\right)^2\phi_{k,i}(x_k)\phi_{k,j}(x_k)\,dx_k,$$
$$(\mathbf{O}_k)_{ij} := \int_{\omega_k}\phi_{k,i}(x_k)\phi_{k,j}(x_k)\,dx_k, \tag{41}$$
$$(\mathbf{B}_k)_{ij} := \int_{\omega_k}\left(x_k - R_{0x_k}\right)\phi_{k,i}(x_k)\phi_{k,j}(x_k)\,dx_k$$

for $k = 1,2,3$. Thus $\mathbf{W}^P$, for any atom P, can be evaluated as a linear combination of five matrices independent of the atom P, where the integrals involved in each of the matrices can be evaluated as the product of one-dimensional integrals. We note that the matrices $\mathbf{K}^0_k$, $\mathbf{O}_k$, and $\mathbf{B}_k$ are sparse owing to the locality of the Tucker-tensor basis $\mathbb{T}^L$, thereby rendering $\mathbf{K}^P$ sparse. Further, we truncate the wave functions involved in the computation of $\mathbf{L}$ using a truncation tolerance, rendering $\mathbf{L}$ sparse. Thus the computational complexity involved in the construction of $\mathbf{W}^P$ for all atoms $P = 1,\ldots,N_{\mathrm{a}}$ scales as $\mathcal{O}(N)$. Using the eigenvectors $\boldsymbol{\alpha}$ from the solution of the eigenvalue problem in Eq. (36) for each atom P, the linear combination of the Chebyshev filtered vectors is computed to construct the nonorthogonal localized wave functions, which span the Chebyshev filtered space. We refer to these localized wave functions which span the Chebyshev filtered subspace as the localized Chebyshev filtered wave functions, and denote them in matrix form by $\boldsymbol{\Phi}_L$. In practice, we achieve compact support for these localized wave functions by introducing a truncation tolerance.

c. Computation of the electron-density. To compute the electron-density in a given self-consistent field iteration, we first evaluate the projection of the Hamiltonian onto the space spanned by the Chebyshev filtered wave functions represented in the basis of the localized Chebyshev filtered functions, which is given by $\mathbf{H}^\phi = \mathbf{S}^{-1}\boldsymbol{\Phi}_L^T\tilde{\mathbf{H}}\boldsymbol{\Phi}_L$ with $\mathbf{S} = \boldsymbol{\Phi}_L^T\mathbf{M}\boldsymbol{\Phi}_L$ [34]. Furthermore, $\mathbf{S}^{-1}$ can be computed in $\mathcal{O}(N)$ complexity if $\mathbf{S}$ and $\mathbf{S}^{-1}$ are exponentially localized [68]. If the discretized Hamiltonian $\tilde{\mathbf{H}}$ and the matrix $\boldsymbol{\Phi}_L$ are sparse with a bandwidth independent of N, $\mathbf{H}^\phi$ can be computed in $\mathcal{O}(N)$ complexity.

Following [34], the electron-density is given by (cf. Eq. (60) in Ref. [34])

$$\varrho(\mathbf{x}) = 2\,\boldsymbol{T}^T(\mathbf{x})\,\mathbf{M}^{-1/2}\,\boldsymbol{\Phi}_L\,f(\mathbf{H}^\phi)\,\mathbf{S}^{-1}\,\boldsymbol{\Phi}_L^T\mathbf{M}^{-1/2}\boldsymbol{T}(\mathbf{x}), \tag{42}$$

where $\boldsymbol{T}^T(\mathbf{x}) = [T_1(\mathbf{x}), T_2(\mathbf{x}), T_3(\mathbf{x}), \ldots, T_{r_d^3}(\mathbf{x})]$ and

$$f(\mathbf{H}^\phi) = \frac{1}{1+\exp\left(\frac{\mathbf{H}^\phi-\mu}{\sigma}\right)}, \tag{43}$$

with μ being the chemical potential, $\sigma = k_BT$, and k_B the Boltzmann constant. A Chebyshev polynomial expansion is used to approximate $f(\mathbf{H}^\phi)$ in Eq. (43), and if $\mathbf{H}^\phi$ is sufficiently sparse, $f(\mathbf{H}^\phi)$ can be computed in $\mathcal{O}(N)$ complexity [27]. Furthermore, the computation of the Chebyshev polynomial expansion requires the evaluation of the Fermi energy μ, which is achieved by using the constraint

$$2\,\mathrm{tr}(f(\mathbf{H}^\phi)) = N_{\mathrm{e}}. \tag{44}$$

Here, N_{e} is the number of electrons in the given system. The Fermi energy can be efficiently computed with the methods described in Ref. [27], which scale as $\mathcal{O}(N)$. Finally, the band energy required in computing the total energy of the system is evaluated by

$$E_{\mathrm{band}} = 2\sum_{i=1}^{N} f(\epsilon_i,\mu)\epsilon_i = 2\,\mathrm{tr}(f(\mathbf{H}^\phi)\mathbf{H}^\phi). \tag{45}$$

V. NUMERICAL SIMULATIONS

In this section, we investigate the accuracy, performance, and scaling of the proposed Tucker-tensor algorithm for the solution of the Kohn-Sham equations. As benchmark examples we conduct pseudopotential calculations on nonperiodic three-dimensional materials systems representative of both metallic and insulating systems. The benchmark metallic systems considered include aluminum nanoclusters of various sizes: single aluminum atom, aluminum dimer, and nanoclusters containing $1 \times 1 \times 1$ (14 atoms), $2 \times 2 \times 2$ (63 atoms), $3 \times 3 \times 3$ (172 atoms), $4 \times 4 \times 4$ (365 atoms), and $5 \times 5 \times 5$ (666 atoms) face-centered-cubic (fcc) unit cells. The benchmark insulating systems include methane molecule and alkane chains C_8H_{18}, $C_{16}H_{34}$, and $C_{33}H_{68}$. The x, y, and z axes for the Tucker-tensor calculations are chosen as the eigendirections of the moment of inertia tensor of the atomic system computed using the atomic locations and atomic masses of the various elements in the materials system. This provides a systematic approach of orienting the axis to align with the atomistic system and improve the efficiency of the Tucker-tensor approach. In all our simulations, we choose the ranks r_ϱ, r_v, and the number T of terms in the expansion in Eq. (17), such that the approximation errors are higher-order compared to the discretization errors of the finite-dimensional Tucker-tensor basis in Eq. (24). In particular, we choose the ranks

$$r_\varrho = r_v := 45, \quad T := 35,$$

and the values of α_n and β_n are taken from Ref. [61]. Norm-conserving Troullier-Martins pseudopotentials [58] have been employed in the case of aluminum nanoclusters and alkane chains for investigating the performance of method I in the proposed Tucker-tensor algorithm, while bulk local pseudopotentials [70] are employed for conducting a comparative study between methods I and II. We use the n-stage Anderson mixing scheme [69] on the electron density for the self-consistent field iteration of the Kohn-Sham problem, and employed a stopping tolerance of 10^{-7} in the square of the L^2 norm of electron density difference between successive iterations. The Chebyshev filtered subspace iteration is used with a Chebyshev polynomial degree of 25 for the computation of the eigenspace associated with the occupied states. In our current Python implementation, all the matrices expressed in the Tucker-tensor basis are parallelized using MPI, and are executed on a parallel computing cluster with the following specifications: dual-socket eight-core Intel Core Sandybridge CPU nodes with 16 processors (cores) per node, 64 GB memory per node, and 40 Gbps Infiniband networking between all nodes for fast MPI communications. However, the ALS algorithm [Eq. (4)] employed in computing the Tucker-tensor decomposition of the three-dimensional fields, is not parallelized, thus requiring the various fields (ϱ, $V_{\rm H}$, $V_{\rm eff}$) on the tensor-structured grid to be stored locally on every compute node. This has limited the size of the materials systems considered in the present study.

The computational complexity of the proposed subspace projection algorithm relies on the locality of the Tucker-tensor basis, the locality of the localized Chebyshev filtered wave functions spanning the occupied space, and the dependence of the rank r_d on the system size. The truncation tolerances employed in the various stages of the algorithm determine the sparsity of the matrices in our formulation $(\tilde{\mathbf{H}}, \mathbf{H}^\phi, \mathbf{\Phi}_L, \mathbf{S}, \mathbf{W}^P)$. In the present study, we use dense data structures for all the matrices involved, since the truncation tolerances employed in our simulations resulted in matrices with fraction of nonzero entries greater than 2% for the materials systems studied. The overhead cost of using a sparse data-structure at these density fractions results in more computational inefficiencies than treating the matrices as dense matrices.

In the present work, we employ the recently developed real-space approach for Kohn-Sham DFT calculations using a higher-order finite element basis [19,34] to provide reference energies to assess the approximation errors in the ground-state energies obtained with the proposed Tucker-tensor approach. These reference energies are converged up to 0.1 meV in the

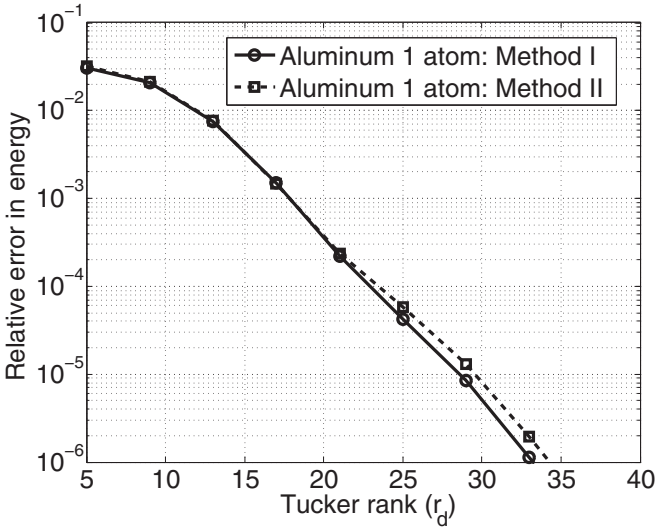

FIG. 1. Convergence of the ground-state energy with respect to the Tucker rank using local pseudopotential. Case study: aluminum atom.

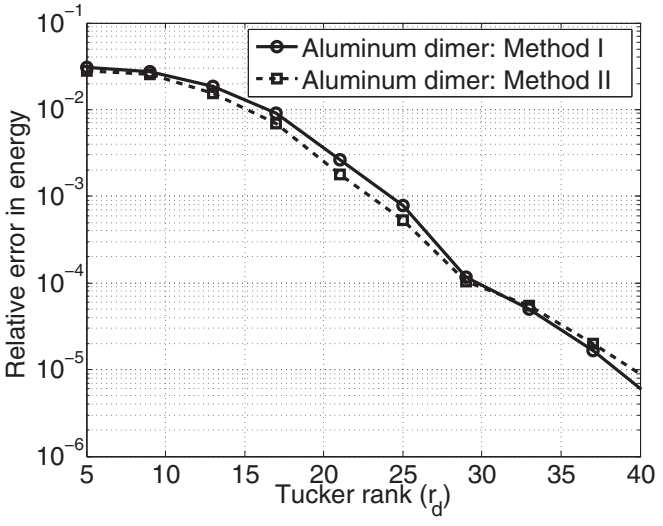

FIG. 2. Convergence of the ground-state energy with respect to the Tucker rank using local pseudopotential. Case study: aluminum dimer.

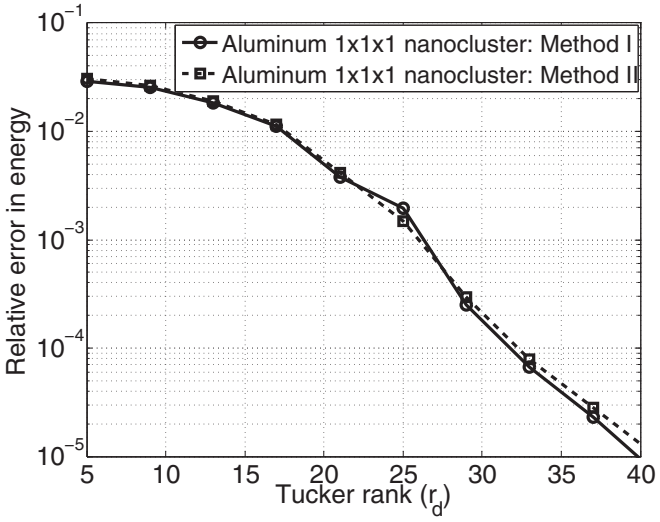

FIG. 3. Convergence of the ground-state energy with respect to the Tucker rank using local pseudopotential. Case study: Aluminum $1 \times 1 \times 1$ nanocluster

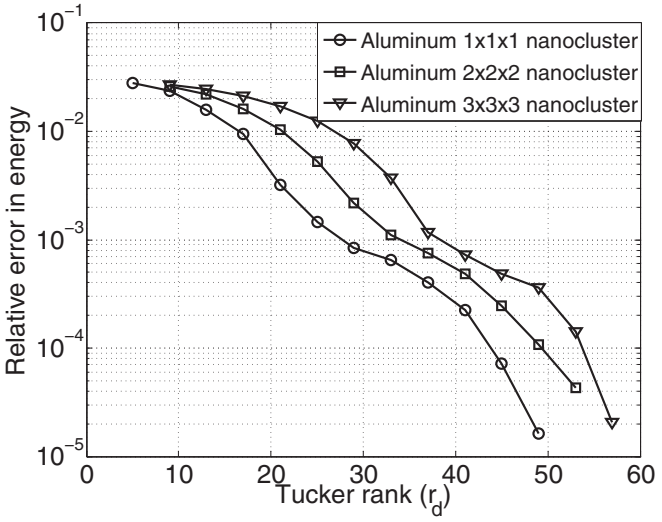

FIG. 5. Convergence of the ground-state energy with respect to the Tucker rank for nonlocal pseudopotential using method I. Case study: aluminum nanoclusters.

ground-state energy per atom with respect to discretization and other numerical parameters.

A. Metallic systems

We first conduct a comparative study between the two methods of constructing the separable Hamiltonian which were proposed in Sec. IV A. To this end, we employ bulk local pseudopotentials [70] to conduct simulations on three benchmark examples consisting of a single aluminum atom, aluminum dimer, and an aluminum nanocluster containing $1 \times 1 \times 1$ (14 atoms) fcc unit cell with a lattice constant of 7.45 a.u. For each of these benchmark systems, the relative error in ground-state energy is computed as a function of the Tucker rank r_d, and is plotted in Figs. 1–3. The results show that

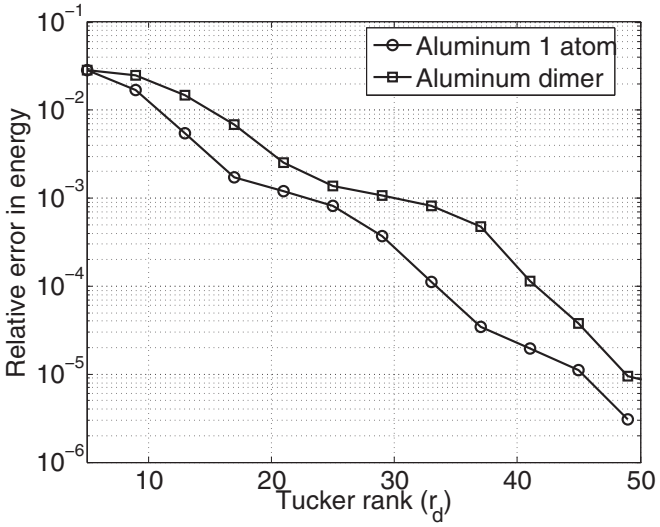

FIG. 4. Convergence of the ground-state energy with respect to the Tucker rank for nonlocal pseudopotential using method I. Case study: aluminum atom and dimer.

both methods of computing the separable Hamiltonian provide similar accuracies in the ground-state energies. Further, there is an exponential convergence in the ground-state energy for increasing Tucker ranks. We also note that the Tucker rank required to achieve chemical accuracy (∼5 meV in the ground-state energy per atom) is weakly dependent on the system size: ∼25 for single atom, ∼30 for dimer, and ∼32 for $1 \times 1 \times 1$ aluminum nanocluster.

We next employ method I for computing the separable Hamiltonian while using the norm-conserving Troullier-Martins pseudopotentials [58] in the Kleinman-Bylander form [59]. The convergence of the ground-state energy with the Tucker rank is examined for the benchmark systems comprising of single aluminum atom, aluminum dimer, and aluminum nanoclusters containing $1 \times 1 \times 1$ (14 atoms), $2 \times 2 \times 2$ (63 atoms), and $3 \times 3 \times 3$ (172 atoms) fcc unit cells with a lattice constant of 7.45 a.u. Figures 4 and 5 show these results which indicate an exponential rate of convergence of the ground state energy with increasing Tucker rank. Furthermore, the number of basis functions, r_d^3, required to obtain chemical accuracy in the ground-state energy, for the case of nonlocal pseudopotentials, grows sublinearly with system size as $\mathcal{O}(N^{0.22})$ for the range of systems studied—with Tucker rank r_d being ∼33 for single atom, ∼41 for dimer,

TABLE I. Ground-state energies per atom (eV) for various sizes of aluminum nanoclusters computed with the proposed algorithm.

Al cluster	Tucker rank	Energy	Ref. energy
$1 \times 1 \times 1$	45	−55.80965	−55.81430
$2 \times 2 \times 2$	49	−56.45924	−56.46504
$3 \times 3 \times 3$	53	−56.69260	−56.69669
$4 \times 4 \times 4$	57	−56.80104	−56.80561
$4 \times 4 \times 4$ with 5 vacancies	57	−56.76531	−56.76964
$5 \times 5 \times 5$	60	−56.87367	−56.87822

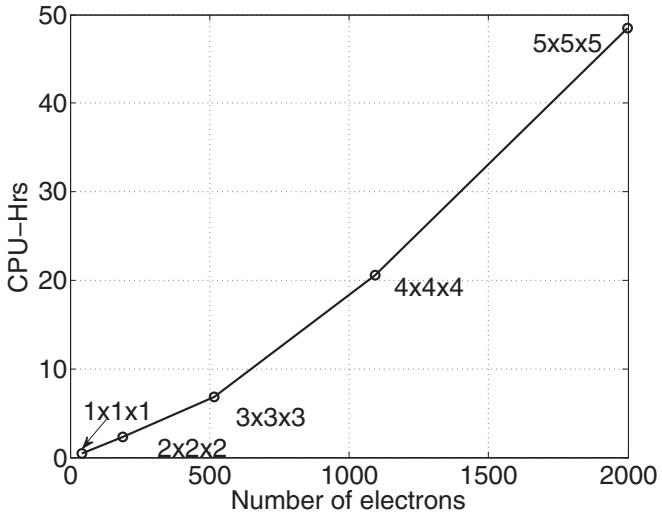

FIG. 6. Computational CPU-time per SCF iteration for $1 \times 1 \times 1$ to $5 \times 5 \times 5$ fcc Al nanoclusters.

and around 45, 49, and 53 for $1 \times 1 \times 1$, $2 \times 2 \times 2$, and $3 \times 3 \times 3$ aluminum nanoclusters, respectively. Moreover, we obtained ground-state energies within chemical accuracy for $4 \times 4 \times 4$ and $5 \times 5 \times 5$ nanoclusters using the Tucker-tensor basis with Tucker ranks of 57 and 60, respectively. We also introduced 5 random vacancies in the $4 \times 4 \times 4$ nanocluster and found that the ground-state energy within chemical accuracy is obtained with a Tucker basis of rank 57 even for this system. The ground-state energies computed with the proposed Tucker-tensor algorithm are tabulated in Table I, and are within chemical accuracy of the reference energies. This demonstrates the effectiveness of the computed Tucker-tensor basis in accurately representing the electronic structure of materials systems with varying sizes and complexity.

The computational CPU times per SCF iteration for each of these systems is plotted against the number of electrons in Fig. 6. All computational times reported in this study denote CPU times in hours (CPU time = number of cores $\times$ wall-clock time in hours). The scaling with the system size for the aluminum clusters is found to be $\mathcal{O}(N^{1.2})$. It is remarkable that we obtain close to linear-scaling complexity even for metallic systems with the proposed Tucker-tensor algorithm for the range of systems studied. Albeit using dense data structures in our computations, we obtain close to linear-scaling complexity due to the sublinear dependence

TABLE II. Comparison of the proposed Tucker-tensor approach with plane-wave basis for a $3 \times 3 \times 3$ FCC Al cluster. Reference ground-state energy for this system is -56.69669 eV per atom.

Type of basis set	Number of basis functions	Absolute error in energy per atom (meV)	Time (CPU hrs)
Plane-waves basis (cutoff energy 20 Ha; cell size 60 a.u.)	461,165	3.8	910
Tucker basis	148,877	4.1	360

TABLE III. Comparison of the proposed Tucker-tensor approach with plane-wave basis for a $5 \times 5 \times 5$ FCC Al cluster. Reference ground-state energy for this system is -56.87822 eV per atom.

Type of basis set	Number of basis functions	Absolute error in energy per atom (meV)	Time (CPU hrs)
Plane-waves basis (cutoff energy 20 Ha; cell size 80 a.u.)	1,093,421	4.3	8640
Tucker basis	216,000	4.6	2364

of the number of Tucker-tensor basis functions on the system size. We expect that in the limit of very large system sizes, the number of Tucker-tensor basis functions will grow linearly with the system size. However, the increase in system size renders the matrices involved in the proposed algorithm sparse, owing to the locality in the Tucker-tensor basis and the localized Chebyshev filtered wave functions. We note that the complexity estimates for the proposed Tucker-tensor algorithm (cf. Sec. IV) suggest linear-scaling complexity with system size for the case of sparse matrices. Thus we expect the close to linear-scaling computational complexity to also hold in the limit of large system sizes.

Tables II and III show the comparison of computational time and number of basis functions for the proposed algorithm using Tucker-tensor basis and plane-wave basis (ABINIT software [71]) for the computation of ground-state energy of $3 \times 3 \times 3$ and $5 \times 5 \times 5$ aluminum nanoclusters to within discretization error of less than 5 meV. The parameters used in the Tucker-tensor calculations (domain size, SCF mixing scheme and stopping tolerances) have also been used in the plane-wave calculations for a consistent comparison. These results show that the proposed Tucker-tensor approach requires a 3–5 times lower number of Tucker-tensor basis functions in comparison to the number of plane-wave basis functions. The computational times for the proposed methodology and the current nonoptimized implementation are also lower than the plane-wave implementation in ABINIT by a factor of 2.5 in the case of $3 \times 3 \times 3$ aluminum nanocluster and by a factor of 3.7 in the case of $5 \times 5 \times 5$ aluminum nanocluster. Further optimization of our in-house code may lead to more

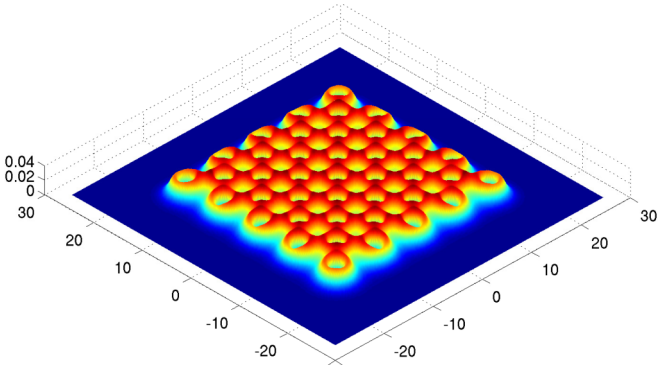

FIG. 7. Electron-density contours on the midplane of $4 \times 4 \times 4$ fcc nanoclusters.

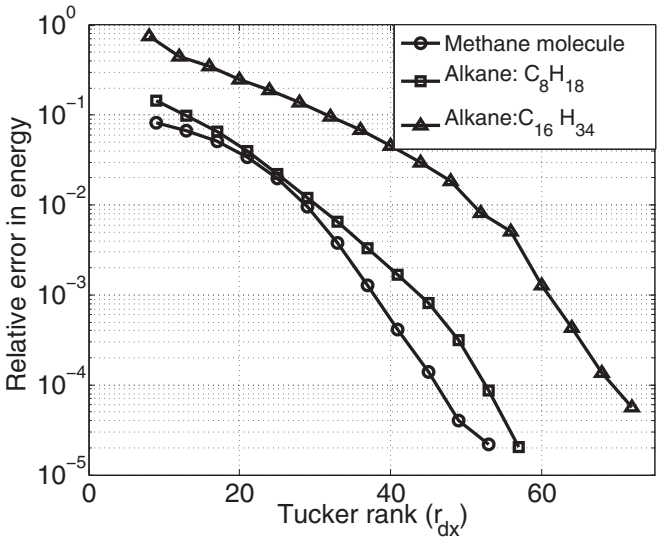

FIG. 8. Convergence of the ground-state energy with respect to the Tucker rank (r_{dx}) for the insulating benchmark systems.

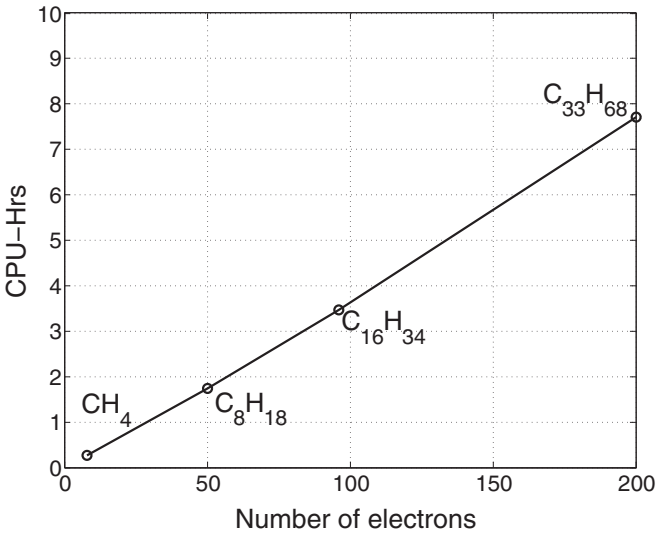

FIG. 9. Computational CPU time per SCF iteration for the insulating benchmark systems.

significant speedups than the factors reported here, and may provide significant savings in the computational times for large-scale DFT calculations. Figure 7 shows the electron-density contours on the mid-plane of a $4 \times 4 \times 4$ nanocluster computed with the proposed Tucker-tensor approach.

B. Insulating systems

We consider three-dimensional alkane chains as our benchmark systems, including CH_4 (methane), C_8H_{18}, $C_{16}H_{34}$, and $C_{33}H_{68}$. We use norm-conserving Troullier-Martins pseudopotentials [58], and method I for computing the separable approximation of the Hamiltonian. We orient the alkane chains along the x direction and use C-C and C-H bond lengths to be 2.91018 and 2.0598 a.u. with the H-C-H and C-C-C bond angles taken to be 109.47^0. Figure 8 shows the convergence of the ground-state energy with increasing Tucker rank r_{dx}. For these simulations, we choose $r_{dy} = r_{dz} = 46$ for methane and $r_{dy} = r_{dz} = 55$ for C_8H_{18}, $C_{16}H_{34}$ and $C_{33}H_{68}$. In the case of alkane chains, the results indicate that the Tucker rank required to achieve chemical accuracy in the ground-state energy is—$r_{dx} \sim 46$ for CH_4, $r_{dx} \sim 55$ for C_8H_{18}, $r_{dx} \sim 68$ for $C_{16}H_{34}$, and $r_{dx} \sim 85$ for $C_{33}H_{68}$. Furthermore, the number of basis functions ($r_{dx}r_{dy}r_{dz}$) grows sublinearly with the system size as $\mathcal{O}(N^{0.3})$ for the range of systems studied. The computed ground-state energies with their Tucker ranks are tabulated in Table IV. The computational CPU times per SCF iteration for these systems plotted against the number of electrons are given in Fig. 9, and the scaling with system size is found to be $\mathcal{O}(N^{1.05})$. Figure 10 shows the electronic structure—isocontours of the electron density—of CH_4 and C_8H_{18}.

TABLE IV. Ground-state energies per atom (eV) for the various insulating systems computed using the proposed algorithm.

Cluster	Tucker rank(r_{dx})	Energy	Ref. energy
CH_4	46	−43.73506	−43.73892
C_8H_{18}	55	−58.77419	−58.77903
$C_{16}H_{34}$	68	−60.49686	−60.50081
$C_{33}H_{68}$	85	−61.43695	−61.44174

VI. SUMMARY

An algorithm for the solution of the Kohn-Sham problem is presented that exploits the low-rank approximation of the electronic-structure afforded by Tucker-tensor representations. A systematic procedure is developed for computing a localized Tucker-tensor basis adapted to the Kohn-Sham eigenvalue problem. To this end, in every iteration of the self-consistent

FIG. 10. Isocontours of the electron density of CH_4 and C_8H_{18} computed with the proposed Tucker-tensor DFT algorithm.

field procedure of the Kohn-Sham problem, a separable approximation of the Kohn-Sham Hamiltonian is constructed, and the localized Tucker-tensor basis is computed using the eigenfunctions of the separable Hamiltonians in each spatial dimension. The localized Tucker-tensor basis is subsequently used to solve the Kohn-Sham eigenvalue problem by using Chebyshev filtering and Fermi-operator expansion techniques to compute the occupied eigenspace and the electron-density. Numerical investigations on representative benchmark examples reveal an exponential convergence of the ground-state energy with respect to the Tucker rank. In addition, the Tucker rank required to obtain chemical accuracy in the computed ground-state energies is found to only weakly depend on the system size, with the number of Tucker-tensor basis functions exhibiting a sublinear dependence on the system size for the range of benchmark systems considered in this study. Our benchmark studies suggest further that the proposed algorithm exhibits a close to linear-scaling complexity with system size for both insulating and metallic systems. This reduced-order scaling is a result of combining the low-rank Tucker-tensor basis with localization techniques, and constitutes a promising direction for large-scale DFT calculations. A comparative numerical study for $3 \times 3 \times 3$ and $5 \times 5 \times 5$ aluminum nanoclusters as benchmark systems shows about a fivefold reduction in the number of basis functions and about a three to fourfold computational speedup for the current implementation of the proposed algorithm over the plane-wave implementation in ABINIT. We note that there is much scope for optimizing our current Python implementation, and thus the computational efficiency afforded by the proposed algorithm may potentially be much larger. Finally, in the present work, we used a serial version of the ALS algorithm to compute the Tucker-tensor decomposition of the three-dimensional fields, thus limiting the sizes of the materials systems to those systems where the data corresponding to all relevant three-dimensional fields is accommodated in the memory of a single compute node. Overcoming this limitation, and developing an efficient and scalable parallel implementation of all aspects of the proposed algorithm has the potential to enabling DFT calculations on system sizes not accessible heretofore.

ACKNOWLEDGMENTS

The authors gratefully acknowledge W. Hackbusch (MPI MIS, Leipzig) for the valuable comments and useful discussions. This work was performed in part under the auspices of Air Force Office of Scientific Research Grant No. FA9550-13-1-0113. V.G. also gratefully acknowledges the support of the Alexander von Humboldt Foundation through a research fellowship, and the hospitality of the Max-Planck Institute for Mathematics in the Sciences during the initial stages of this work. We also acknowledge Advanced Research Computing at the University of Michigan for providing the computing resources through the Flux computing platform.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[4] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Z. für Kristallogr. **220**, 567 (2005).
[5] X. Gonze, J. M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G. M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J. Y. Raty, and D. C. Allan, Comput. Mat. Science **25**, 478 (2002).
[6] W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys. **51**, 2657 (1969).
[7] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).
[8] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).
[9] E. Cancès, C. Le Bris, Y. Maday, and G. Turinici, J. Sci. Comput. **17**, 461 (2002).
[10] L. Lin, J. Lu, L. Yang, and Weinan E, J. Comput. Phys. **231**, 2140 (2012).
[11] J. R. Chelikowsky, N. Troullier, and Y. Saad, Phys. Rev. Lett. **72**, 1240 (1994).
[12] L. Kronik, A. Makmal, M. L. Tiago, M. M. G. Alemany, M. Jain, X. Huang, Y. Saad, and J. R. Chelikowsky, Phys. Status Solidi B **243**, 1063 (2006).
[13] K. Cho, T. A. Arias, J. D. Joannopoulos, and P. K. Lam, Phys. Rev. Lett. **71**, 1808 (1993).
[14] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, and J. F. Mèhaut, C. R. Mécanique **339**, 149 (2011).
[15] E. Tsuchida and M. Tsukada, Phys. Rev. B **52**, 5573 (1995).
[16] J. E. Pask, B. M. Klein, C. Y. Fong, and P. A. Sterne, Phys. Rev. B. **59**, 12352 (1999).
[17] J. E. Pask, B. M. Klein, P. A. Sterne, and C. Y. Fong, Comp. Phys. Comm. **135**, 1 (2001).
[18] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, and M. Ortiz, J. Mech. Phys. Solids **58**, 256 (2010).
[19] P. Motamarri, M. R. Nowak, K. Leiter, J. Knap, and V. Gavini, J. Comput. Phys. **253**, 308 (2013).
[20] S. Goedecker, Rev. Mod. Phy. **71**, 1085 (1999).
[21] D. R. Bowler and T. Miyazaki, Rep. Prog. Phys. **75**, 036503 (2012).
[22] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[23] W. Yang, J. Mol. Struc. Theochem. **255**, 461 (1992).
[24] T. Ozaki, Phys. Rev. B **74**, 245101 (2006).
[25] M. Barrault, E. Cancès, W. W. Hager, and C. L. Bris, J. Comput. Phys. **222**, 86 (2007).
[26] S. Goedecker and M. Teter, Phys. Rev. B **51**, 9455 (1995).
[27] R. Baer and M. Head-Gordon, J. Chem. Phys. **107**, 10003 (1997).
[28] L. Lin, J. Lu, L. Ying, and Weinan E, Chin. Ann. Math. **30**, 729 (2009).
[29] L. Lin, M. Chen, C. Yang, and L. He, J. Phys. Condens. Matter **25**, 295501 (2013).
[30] P. Suryanarayana, K. Bhattacharya, and M. Ortiz, J. Mech. Phys. Solids **61**, 38 (2013).

[31] X. P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).
[32] P. D. Haynes, C. K. Skylaris, A. A. Mostofi, and M. C. Payne, Phys. Stat. Sol.(b) **243**, 2489 (2006).
[33] C. J. Garcia-Cervera, J. Lu, Y. Xuan, and Weinan E, Phys. Rev. B **79**, 115110 (2009).
[34] P. Motamarri and V. Gavini, Phys. Rev. B **90**, 115127 (2014).
[35] O. F. Sankey, D. A. Drabold, and A. Gibson, Phys. Rev. B **50**, 1376 (1994).
[36] U. Stephan and D. A. Drabold, Phys. Rev. B **57**, 6391 (1998).
[37] F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993).
[38] W. Gao and Weinan E, Discrete Contin. Dyn. Sys. A **23**, 249 (2009).
[39] W. Hackbusch and B. N. Khoromskij, J. Complexity **23**, 697 (2007).
[40] B. N. Khoromski and V. Khoromskaia, SIAM J. Sci. Comput. **31**, 3002 (2009).
[41] B. N. Khoromskij, V. Khoromskaia, S. R. Chinnamsetty, and H.-J. Flad, J. Comp. Phys. **228**, 5749 (2009).
[42] L. R. Tucker, Psychometrika **31**, 279 (1966).
[43] L. de Lathauwer, B. de Moor, and J. Vandewalle, SIAM J. Matrix Anal. Appl. **21**, 1324 (2000).
[44] T. G. Kolda and B. W. Bader, SIAM Rev. **51**, 455 (2009).
[45] T. Blesgen, V. Gavini, and V. Khoromskaia, J. Comput. Phys. **231**, 2551 (2012).
[46] Weinan E, L. Tiejun, and L. Jianfeng, Proc. Natl. Acad. Sci. USA **107**, 1273 (2010).
[47] V. de Silva and L. H. Lim, SIAM J. Matrix Anal. & Appl. **30**, 1084 (2008).
[48] I. V. Oseledets and E. E. Tyrtyshnikov, SIAM J. Sci. Comput. **31**, 3744 (2009).
[49] I. Oseledets, SIAM. J. Sci. Comput. **33**, 2295 (2011).
[50] L. Grasedyck, SIAM J. Matrix Anal. Appl. **31**, 2029 (2010).
[51] W. Hackbusch and R. Schneider, *Tensor Spaces and Hierarchical Tensor Representations*, in Extraction of Quantifiable Information from Complex Systems Vol. 102 (Springer Verlag, New York, 2014), p. 237.
[52] W. Hackbusch and S. Kühn, J. Fourier Anal. Appl. **15**, 706 (2009).
[53] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).
[54] L. Grasedyck, D. Kressner, and C. Tobler, GAMM-Mitt **36**, 53 (2013).
[55] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2011).
[56] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
[57] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[58] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[59] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[60] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite-element Methods* (Springer, New York, 2002).
[61] D. Braess and W. Hackbusch, *On the Efficient Computation of High-Dimensional Integrals and the Approximation by Exponential Sums*, Multiscale, Nonlinear and Adaptive Approximation Vol. 39 (Springer, New York, 2009).
[62] P. W. Anderson, Phys. Rev. Lett. **21**, 13 (1968).
[63] J. Kim, F. Mauri, and G. Galli, Phys. Rev. B **52**, 1640 (1995).
[64] P. Ordejón, D. A. Drabold, M. P. Grumbach, and R. M. Martin, Phys. Rev. B **48**, 14646 (1993).
[65] V. Ozolins, R. Lai, R. Calflisch, and S. Osher, Proc. Natl. Acad. Sci. USA **110**, 18368 (2013).
[66] B. Jansik, S. Host, P. Jorgensen, J. Olsen, and T. Helgaker, J. Chem. Phys **126**, 124104 (2007).
[67] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, Phys. Rev. E **74**, 066704 (2006).
[68] E. Rubensson and P. Salek, J. Comput. Chem. **26**, 1628 (2005).
[69] D. G. Anderson, J. ACM **12**, 547 (1965).
[70] C. Huang and E. A. Carter, Phys. Chem. Chem. Phys. **10**, 7109 (2008).
[71] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, Ph. Ghosez, M. Giantomassi, S. Goedecker, D. R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini *et al.*, Comput. Phys. Commun. **180**, 2582 (2009).