System-Specific Separable Basis Based on Tucker Decomposition: Application to Density Functional Calculations

Jeheon Woo, Woo Youn Kim,* and Sunghwan Choi*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: For fast density functional calculations, a suitable basis that can accurately represent the orbitals within a reasonable number of dimensions is essential. Here, we propose a new type of basis constructed from Tucker decomposition of a finite-difference (FD) Hamiltonian matrix, which is intended to reflect the system information implied in the Hamiltonian matrix and satisfies orthonormality and separability conditions. By introducing the system-specific separable basis, the computation time for FD density functional calculations for seven two- and three-dimensional periodic systems was reduced by a factor of 2−71 times, while the errors in both the atomization energy per atom and the band gap were limited to less than 0.1 eV. The accuracy and speed of the density functional calculations with the proposed basis can be systematically controlled by adjusting the rank size of Tucker decomposition.
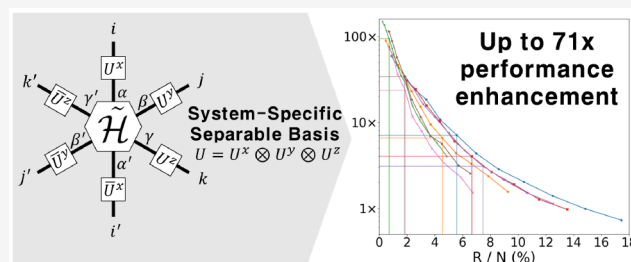
INTRODUCTION

Numerical methods for replacing partial differential equations with finite-dimensional algebraic equations are essential for the electronic structure calculations of molecular or solid systems.[1,2] The atom-centered and plane-wave basis methods are frequently used for nonperiodic and periodic systems, respectively.[3,4] Real-space methods are potentially competitive with the aforementioned methods because of their flexibility and computational simplicity.[5] However, they have not yet been widely adopted for electronic structure calculations. Discretization of the simulation domain results in a large dimension for the Hamiltonian and orbitals. To mitigate the increase in memory usage and computational costs due to the large dimension, tensor decomposition techniques can be applied to real-space methods.[6−9]

Tensor decomposition techniques are not limited to real-space methods.[10] They have been actively investigated to accelerate various quantum chemistry methods that require a large amount of computational and memory resources (e.g., a perturbation method,[10,11] coupled cluster theory,[12−14] and full basis representation methods[15]). For density functional calculations, tensor decomposition techniques using real-space methods have been studied because orbital values on a rectangular grid can be represented as an order-3 tensor. Solala et al. applied tensor decomposition to density functional calculations to minimize memory load.[6] Their results show that the Tucker decomposition method can successfully compress the orbitals represented on a cubic grid from the results of bubbles and the cube numerical framework, which is a variation of real-space methods. Tensor decomposition can be used to compress orbitals on a three-dimensional (3D) grid

and build a basis for a self-consistent field (SCF) procedure. Gavini et al. proposed a Tucker tensor basis derived from a separable approximation of the Hamiltonian. It effectively reduces the dimensions of a Kohn−Sham (KS) Hamiltonian matrix originally represented on an equidistant finite-element grid.[8]

The separability of the basis is an important property to reduce the computational costs of many operations. A typical separable basis can be obtained from the simple product of three 1D functions (e.g., a Gaussian function). To impose system information on such a basis, a 1D function that reflects the system information on a general polyatomic structure needs to be developed. By contrast, a Hamiltonian matrix that implicitly includes all system information can be easily constructed on a real-space grid. Using the Hamiltonian on a real-space grid is an attainable solution for imposing system information on a separable basis.

Herein, we propose a system-specific separable basis derived from a finite-difference (FD) KS Hamiltonian matrix and investigate its performance for 2D and 3D periodic structures. The resulting basis is constructed by reflecting information on the Hamiltonian of the system and is also separable along the axes of the spatial coordinates. These two features are common

to other types of Tucker tensor basis.[8] A key contribution of this work is that the basis is built directly from a finite-difference method, and its nonzero patterns are used in the projection process instead of introducing a separable Hamiltonian. In addition, the convergence of our basis is systematically controlled by increasing the rank size of Tucker decomposition. In the following, we briefly explain the mathematical background of our method, followed by the implementation details. We then discuss the performance of the proposed basis on density functional calculations for 2D and 3D periodic systems and demonstrate its advantages for reducing the computation time of density functional calculations.

■ METHOD

Tucker Representation and Higher-Order Singular-Value Decomposition. Here, we briefly introduce a Tucker representation and a higher-order singular-value decomposition (HOSVD) method for completeness. A more detailed explanation can be found in previous papers.[16,17]

The Tucker representation is used to represent an order-d tensor, $C \in \mathbb{C}^{N_1 \times N_2 \cdots \times N_d}$, as a contraction of a small order-d core tensor, $\tilde{C} \in \mathbb{C}^{r_1 \times r_2 \cdots r_d}$, and d unitary factor matrices, $U^{(n)} \in \mathbb{C}^{N_n \times r_n}$ ($n \in \{1, 2, ..., d\}$), where N_n and r_n are the dimensions of the nth axis for the original and core tensors, respectively.[16-19] Then, the Tucker decomposition can be written as

$$C_{i_1 i_2 \cdots i_d} \approx \sum_{j_1, j_2, \cdots j_d} \tilde{C}_{j_1 j_2 \cdots j_d} U^{(1)}_{i_1 j_1} U^{(2)}_{i_2 j_2} \cdots U^{(d)}_{i_d j_d} \tag{1}$$

The convergence of eq 1 is mathematically guaranteed as r_n approaches N_n.[10] However, the Tucker representation is frequently used to find a compact representation of a given tensor, which means $r_n < N_n$. This compact representation can reduce the computational complexity and memory consumption of tensor operations, thereby minimizing accuracy loss.[17]

An HOSVD method is the most common choice to find a set of $U^{(d)}$ and the corresponding C.[10,16] This is one multilinear extension of the matrix singular-value decomposition (SVD). In the HOSVD method, $U^{(n)}$ is obtained as the left singular vector of a factor-n flattened tensor $C_{(n)}$ ($\in \mathbb{C}^{N_n \times (N_1 \cdots N_{n-1} N_{n+1} \cdots N_d)}$). Because all $U^{(n)}$ from HOSVD are unitary, $\tilde{C}$ in eq 1 can be evaluated from the contraction of the original tensor with the obtained $U^{(n)}$, as follows:

$$\tilde{C}_{j_1 j_2 \cdots j_d} = \sum_{i_1, i_2, \cdots i_d} C_{i_1 i_2 \cdots i_d} (U^{(1)}_{i_1 j_1})^* (U^{(2)}_{i_2 j_2})^* \cdots (U^{(d)}_{i_d j_d})^* \tag{2}$$

where $(\cdot)^*$ indicates a complex conjugate.

To obtain a compact Tucker representation, the singular vectors of $C_{(n)}$ with large singular values are denoted as $U^{(n)}(\in \mathbb{C}^{N_n \times r_n})$. The compact core tensor is computed using consecutive tensor contractions, as shown in eq 2. Although the memory usage for $\tilde{C}$ and $U^{(n)}$ is much smaller than that for C, the key patterns of C can be recovered by contraction with $U^{(n)}$, as in a typical compact SVD.

In the field of quantum chemistry, the Tucker representation has been used to accelerate the calculation of two-electron integrals of atom-centered basis functions or tensor contractions for higher-order methods.[10,12,20-22] Here, we build a

separable basis using the Tucker decomposition of the KS Hamiltonian in the FD representation.

Tucker Decomposition of the Finite-Difference Hamiltonian. For an $N_x \times N_y \times N_z$ Cartesian grid, the FD Hamiltonian matrix, H, is constructed on $\mathbb{C}^{(N_x N_y N_z) \times (N_x N_y N_z)}$. H can be reshaped into an order-6 tensor $\mathcal{H} \in \mathbb{C}^{N_x \times N_y \times N_z \times N_x \times N_y \times N_z}$, the elements of which are given by

$$\mathcal{H}_{ijki'j'k'} = H_{pq} \tag{3}$$

where p and q denote the indices of grid points whose x, y, and z indices are (i, j, k) and (i', j', k'), respectively.

To ensure that the factor matrices of the Hamiltonian matrix span low-lying orbitals well, we introduce a constant fictitious potential, which implies that $\mathcal{H}_{ijki'j'k'} = H_{pq} - V_{\text{const}}\delta_{pq}$. This constant potential shifts the eigenspectrum of the original Hamiltonian downward without changing the eigenvectors, so that the factor matrices from the decomposition of a Hamiltonian matrix with fictitious potentials are more likely to span low-lying orbitals of the Hamiltonian, which are physically meaningful. For simplicity, we do not denote the fictitious potential in this section. A more detailed explanation for the fictitious potential is provided in the Appendix.

Using the HOSVD method, $\mathcal{H}$ can be decomposed into the core tensor, $\tilde{\mathcal{H}}$, and the corresponding factor matrices, U:

$$\mathcal{H}_{ijki'j'k'} \approx \sum_{\alpha,\beta,\gamma,\alpha',\beta',\gamma'} \tilde{\mathcal{H}}_{\alpha\beta\gamma\alpha'\beta'\gamma'} U^{(1)}_{i\alpha} U^{(2)}_{j\beta} U^{(3)}_{k\gamma} U^{(4)}_{i'\alpha'} U^{(5)}_{j'\beta'} U^{(6)}_{k'\gamma'} \tag{4}$$

where α, β, γ, α', β', and γ' are the indices of $\tilde{\mathcal{H}}(\in \mathbb{C}^{r_1 \times r_2 \times r_3 \times r_4 \times r_5 \times r_6})$.

Owing to the Hermitian property of H, if $r_1 = r_4$, $r_2 = r_5$, and $r_3 = r_6$, the flattened Hamiltonian matrices and factor matrices satisfy the following relations:

$$H_{(1)} = (H_{(4)})^*, \quad H_{(2)} = (H_{(5)})^*, \quad H_{(3)} = (H_{(6)})^*,$$

$$U^{(1)}_{i\alpha} = (U^{(4)}_{i\alpha})^*, \quad U^{(2)}_{j\beta} = (U^{(5)}_{j\beta})^*, \quad U^{(3)}_{k\gamma} = (U^{(6)}_{k\gamma})^* \tag{5}$$

Hereafter, for convenience, we use U^x, U^y, U^z, $H_{(x)}$, $H_{(y)}$, and $H_{(z)}$ instead of $U^{(1)}$, $U^{(2)}$, $U^{(3)}$, $H_{(1)}$, $H_{(2)}$, and $H_{(3)}$, respectively. Similarly, the rank sizes of U^x, U^y, and U^z are denoted as r_x, r_y, and r_z, respectively.

However, we define the square matrix form of $\tilde{\mathcal{H}}$ as

$$\tilde{H}_{\mu\nu} = \tilde{\mathcal{H}}_{\alpha\beta\gamma\alpha'\beta'\gamma'} \tag{6}$$

where μ and ν are the indices of $\tilde{H}$ ($\in \mathbb{C}^{(r_x r_y r_z) \times (r_x r_y r_z)}$). From eqs 4 and 5, $\tilde{H}$ can be rewritten as

$$\tilde{H} = (U^x \otimes U^y \otimes U^z)^H H (U^x \otimes U^y \otimes U^z) \tag{7}$$

where $\otimes$ and $(\cdot)^H$ denote the Kronecker product and the conjugate transpose, respectively. Here, $\tilde{H}$ is the projection of H on separable basis vectors U ($:= U^x \otimes U^y \otimes U^z$). As discussed in the previous section, the convergence of U_x, U_y, and U_z to make both sides of eq 7 equal is mathematically guaranteed as r_x, r_y, and r_z reach N_x, N_y, and N_z, respectively. Therefore, it is guaranteed that $\tilde{H}$ becomes equal to H when its dimension $r_x \times r_y \times r_z$ becomes $N_x \times N_y \times N_z$.

Here, U is a set of separable basis vectors that can reduce the dimensions of the Hamiltonian from $N_x \times N_y \times N_z$ to $r_x \times r_y \times r_z$. In addition, U satisfies the orthonormality condition
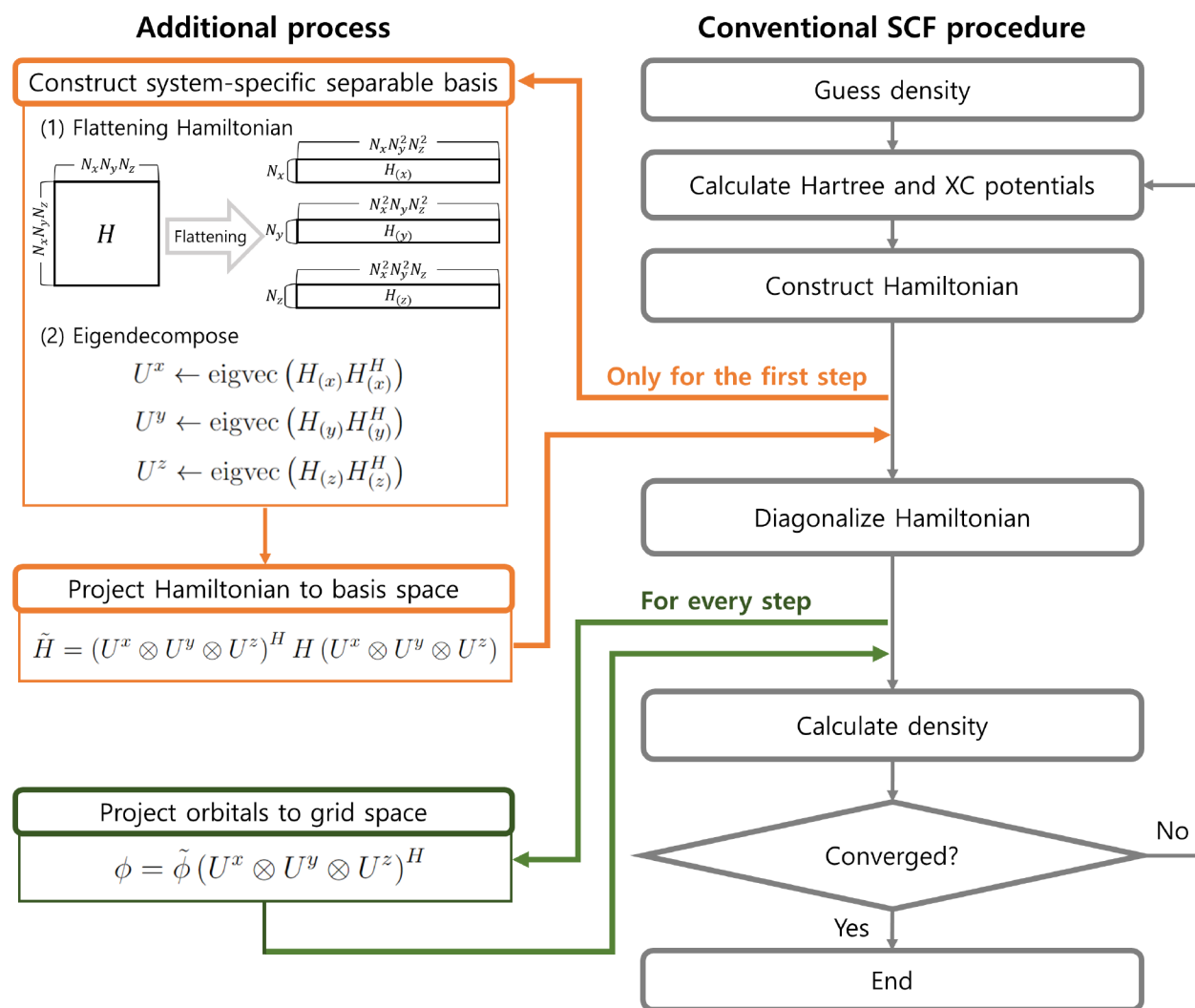
Figure 1. Schematic illustration of a conventional self-consistent field procedure (right panel) and the additional process introduced by the system-specific separable basis (left panel).

because U^x, U^y, and U^z are orthonormal matrices. If U spans physically meaningful eigenstates of the original Hamiltonian well, we only need to diagonalize $\tilde{H}$, which has a smaller dimension than that of H. In addition, U denotes a set of numerical basis vectors that are never explicitly constructed. Owing to its separability, operations with U can be replaced by operations with three small matrices, U^x, U^y, and U^z. Therefore, the memory requirement for U is not $N_x \times N_y \times N_z \times r_x \times r_y \times r_z$ but rather $N_x \times r_x + N_y \times r_y + N_z \times r_z$.

To evaluate $\tilde{H}$, instead of directly evaluating the right-hand side of eq 7, we project three terms of the Hamiltonian matrix separately: kinetic energy, local potential, and nonlocal potential terms. Owing to the properties of U and the nonzero patterns of the three terms, the evaluation of $\tilde{H}$ can be efficiently performed. A further explanation of the projection of the Hamiltonian matrix is described in the Appendix.

Unlike typical basis functions that use a predetermined formula, our U reflects the system information (e.g., relative positions of atoms, phase factors, and cell size) because it is constructed from the decomposition of the Hamiltonian matrix that includes system information. Hereafter, we name U the system-specific separable basis vector and investigate its applicability to density functional calculations.

Before we discuss the performance of the system-specific separable basis in a density functional calculation, we plot its overall process in Figure 1. The right and left panels of Figure 1 represent the conventional SCF procedure and the additional process, respectively. In the system-specific basis calculation, the basis vectors are constructed using the eigendecomposition of $H_{(x)}H_{(x)}^H$, $H_{(y)}H_{(y)}^H$, and $H_{(z)}H_{(z)}^H$. The eigenvectors of $H_{(x)}H_{(x)}^H$, $H_{(y)}H_{(y)}^H$, and $H_{(z)}H_{(z)}^H$ are identical with the left singular vectors of $H_{(x)}$, $H_{(y)}$, and $H_{(z)}$, respectively. To avoid the SVD of a large sparse matrix, we perform eigendecomposition instead of SVD.

After basis construction, the original Hamiltonian is projected to the obtained basis space. The eigenvalues and eigenvectors of the projected Hamiltonian are then computed using a typical matrix diagonalization method. The orbital, ϕ_p, for computing the density, ρ, is evaluated as $\widetilde{\phi_p}(U^x \otimes U^y \otimes U^z)^H$, where ϕ_p and $\widetilde{\phi_p}$ are the pth eigenvectors of H and $\tilde{H}$, respectively. We note that ϕ_p satisfies the orthonormality condition because both $\widetilde{\phi_p}$ and U are orthonormal. The Hartree and the exchange-correlation (XC) potentials for the obtained ρ are evaluated in the same way as the ordinary FD calculation.
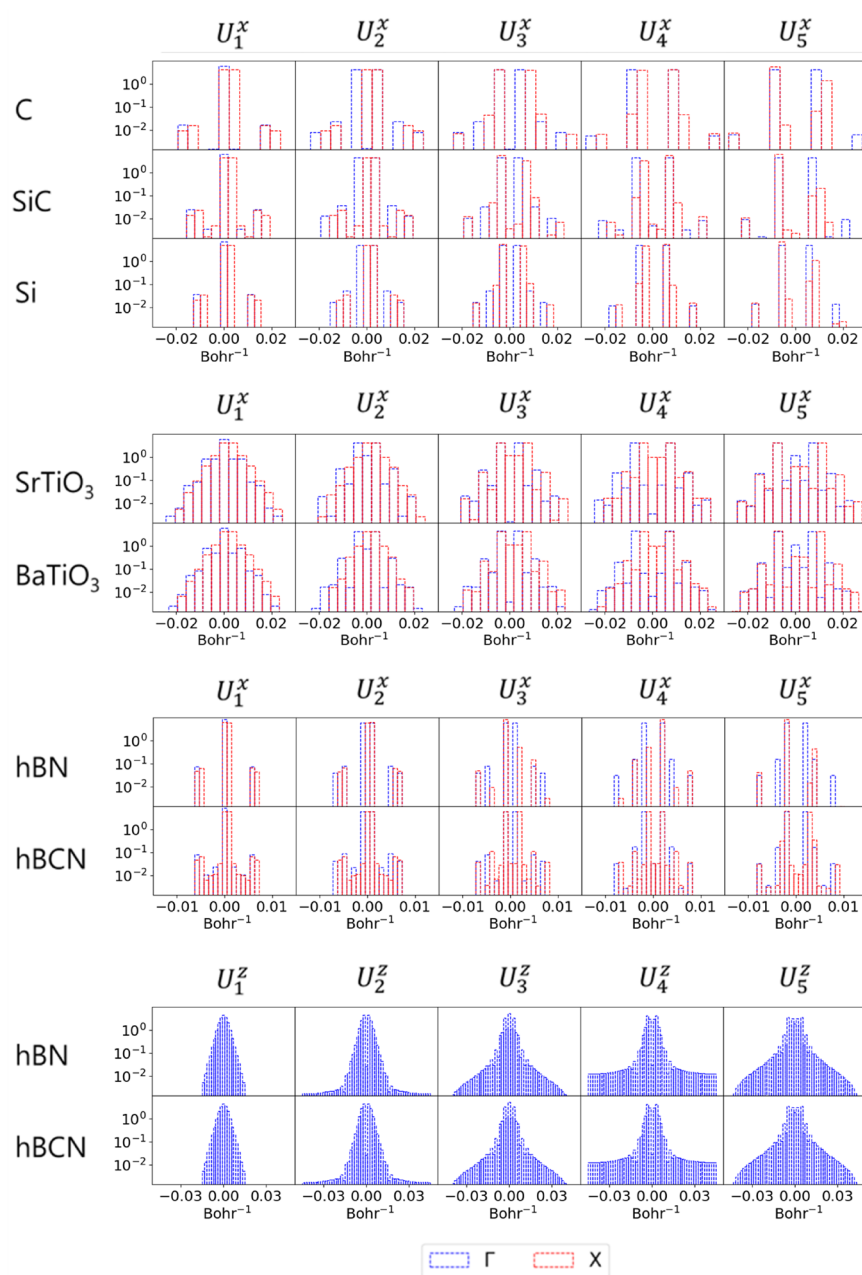
Figure 2. Results of 1D Fourier transform of the first five vectors of factor matrices from the initial Hamiltonian of the seven different systems: C, SiC, Si, $SrTiO_3$, $BaTiO_3$, hBN, and hBCN. The blue and red dashed lines indicate basis vectors from the Γ- and X-points, respectively. U^x and U^z denote the factor matrices of the x- and z-axis, respectively, where the z-axis is a nonperiodic axis of hBN and hBCN.

To accelerate the system-specific basis vector calculations, we introduce two approximations. The first approximation is using the fixed basis during the SCF loop. In other words, the basis is constructed in the first step of the SCF loop using the initial Hamiltonian matrix; it is then used in the subsequent SCF steps. Although the construction of the basis set is not computationally heavy, changes in the basis set at every SCF step reduce the speed of SCF convergence. For a fixed basis, only two local potential terms (Hartree and XC potentials) must be updated at each SCF step. Therefore, only the projection of the updated local potential is performed for each SCF step.

The second approximation is discarding the nonlocal pseudopotential in the basis construction. The errors

introduced by the two approximations are plotted in Figure S1. The approximations may induce errors up to 100 meV in both the atomization energy and the band gap; however, these deviations disappear when the rank size sufficiently converges, and the approximations lead to a ~6× increase in speed in all tested cases. Hereafter, all results are obtained using both approximations.

Implementation and Experiments. The construction of system-specific basis vectors and the projection of the Hamiltonian are performed using the Tucy package, which is written in C++ and has a Python interface. For density functional calculations, our Python package, called the grid-based open-source Python engine for large-scale simulation (GOSPEL), was used. GOSPEL supports FD calculations and
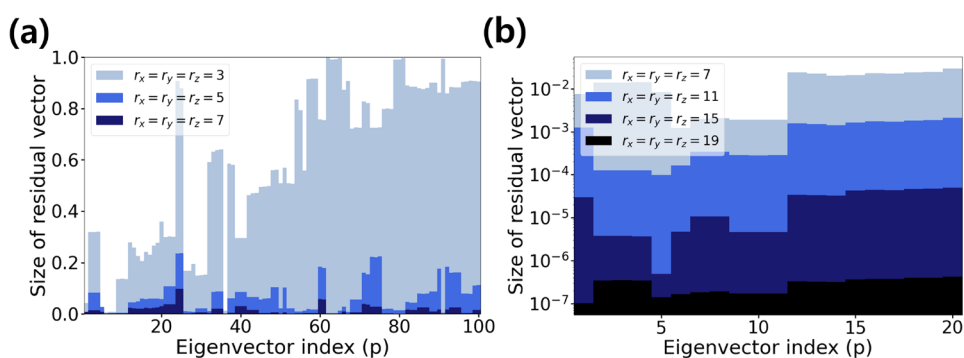
(a)   **(b)**

Figure 3. Size of the projection residual of (a) the first 100 orbitals and (b) the occupied orbitals of $SrTiO_3$ at the Γ-point. The residual is computed by the projection of the reference orbitals, ϕ_p, on the separable basis vectors, U, obtained from Tucker decomposition with different rank sizes.

system-specific separable basis calculations using the Tucy. In GOSPEL, the Hartree potential was evaluated using the interpolation scaling method, as in our previous studies.[23,24] An XC potential is evaluated from the experimental version of *libXC*.[25]

To assess the convergence and performance enhancement, both the reference FD and system-specific separable basis calculations were performed using the same systems. All computational options were used equally in both cases, and all calculations were performed using a single thread of an Intel Xeon Gold 6234 CPU. The PBE[26] functional was used for the XC functional, and optimized norm-conserving Vanderbilt[27] pseudopotentials were used. All 2D and 3D periodic structures were calculated using $(4 \times 4 \times 1)$ and $(4 \times 4 \times 4)$ k-point meshes, respectively. For the kinetic energy matrix, a 7-point FD matrix is used. SCF procedures end when the sum of the occupied band energies converges to less than 10^{-6} Hartree.

For iterative diagonalization for both typical FD and the projected Hamiltonian matrices, we use *LOBPCG* functions implemented in *scipy*,[28] a highly mature and optimized Python package for scientific computing. A compressed sparse row format is used for the FD Hamiltonian instead of a dense matrix format to compute the matrix-vector multiplications. Tucy and GOSPEL are freely available in their online git repositories (https://gitlab.com/jhwoo15/gospel and https://gitlab.com/jhwoo15/tucy, respectively).

Because the cell parameters do not exactly match the multiples of a given grid spacing, the actual grid spacing is set to have the closest value of the given spacing within a small difference (up to 0.1 bohr). Here, we denote the given grid spacing instead of the actual grid spacing in the paper for better readability. The actual grid spacing corresponding to each structure is listed in Table S1.

The performance of the system-specific separable basis is assessed for seven structures: three cubic diamond structures (C, SiC, and Si), two ABO_3 perovskites ($BaTiO_3$ and $SrTiO_3$), and two hexagonal 2D materials (hBN and hBCN). The atomic coordinates and cell parameters of the seven systems are presented in the third section of the Supporting Information. We used the atomization energy per atom for a fair comparison between systems with different numbers of atoms. Hereafter, we refer to the atomization energy per atom as the atomization energy.

■ RESULTS AND DISCUSSION

To confirm the dependency of the system-specific separable basis on the system information, we plot the results of the 1D Fourier transform of the first five vectors of the factor matrices constructed from the initial Hamiltonian at the Γ- (blue bars) and X- (red bars) points (see Figure 2). In the lowest panel, only Γ-point data is visible because only one k-point was sampled along the nonperiodic axis of the 2D hexagonal sheet structures.

First, the blue bars are symmetrically distributed in all cases because the basis vectors at the Γ-point are always real regardless of the structure. However, at the X-point, the basis vectors and Hamiltonian matrices are no longer symmetric because of the phase factor; therefore, the basis vectors are no longer symmetric in Fourier space. One interesting point related to the k-space is that the basis vectors at the X-point are not just shifts in the basis vector at the Γ-point. This indicates that the basis vectors at different points in the k-space are not the products of Γ-point basis vectors with the phase factor, and the basis vectors at each k-point are constructed in a way that reflects the overall Hamiltonian matrix.

Figure 2 also shows the structural dependency of the basis vectors. The ABO_3 structures have a common pattern. C and Si structures also share a similar pattern. However, the SiC structure has a different shape than those of other diamond structures. SiC is composed of two different elements; therefore, the nature of the covalent bonds in SiC is largely different than those of C and Si. Likewise, the basis vectors of two hexagonal sheet structures show different patterns along the x-axis, whereas the basis vectors along the z-axis show a similar trend. This implies that hBN and hBCN show different characteristics along the periodic axes but not along the nonperiodic axis. Although it is difficult to elucidate which system information changes a specific pattern in the basis vectors that we obtain, we can observe the structural and phase dependencies of the system-specific basis vectors.

We investigated whether the obtained basis vectors can properly span an orbital from the reference FD calculations. We projected the reference orbitals from the FD calculation of the $SrTiO_3$ on U and calculated their residuals. Figure 3 plots the sizes of the projection residual on U constructed with different rank sizes (r_x, r_y, and r_z). The large residual size means that the basis space does not sufficiently span the reference orbital. The tested reference orbitals were obtained from ordinary FD calculations of $SrTiO_3$. For good readability, we
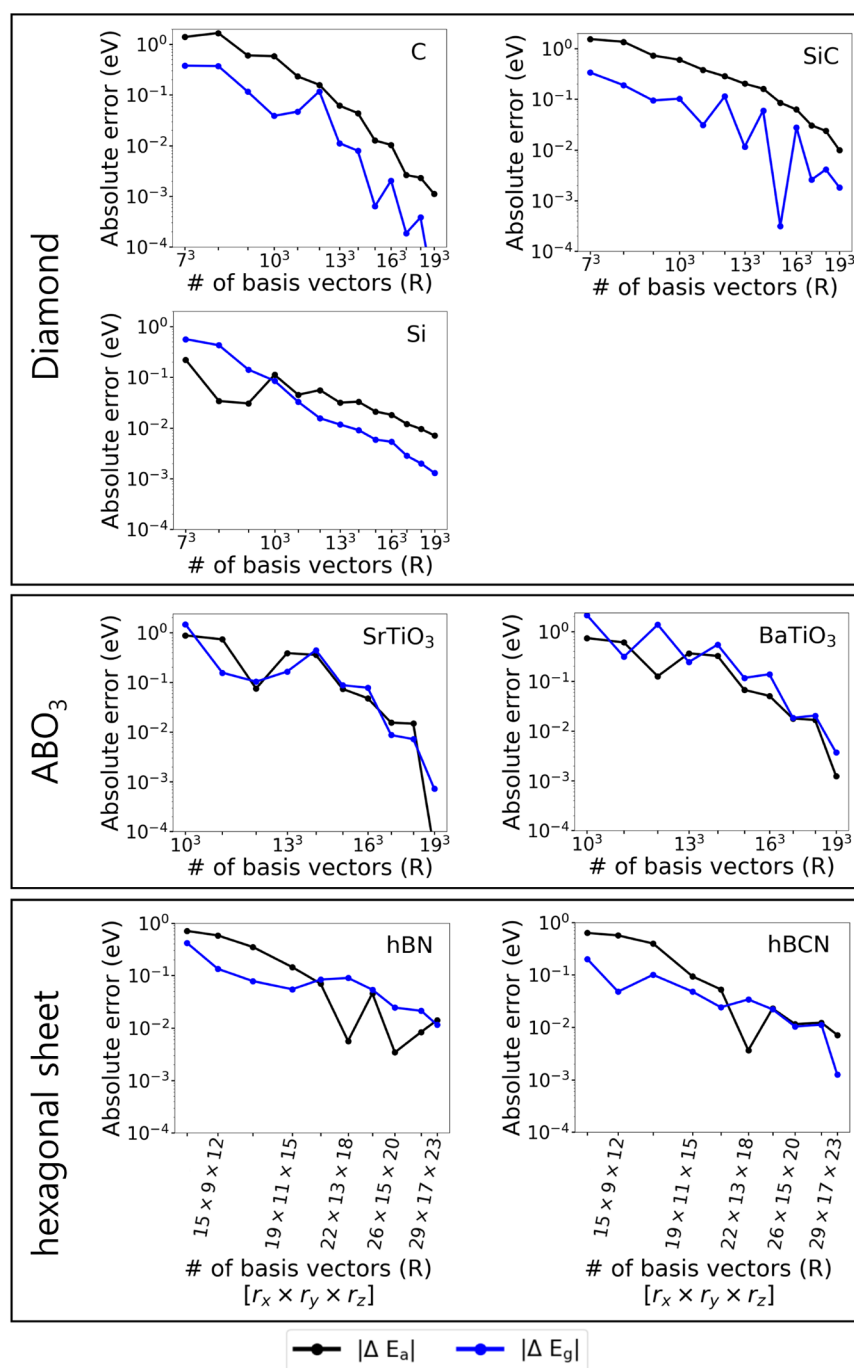
Figure 4. Convergence behavior of system-specific separable basis vectors as a function of the number of basis vectors, R, on a log–log scale. The blue and black lines indicate the absolute errors in the band gap and atomization energy per atom ($|\Delta E_g|$ and $|\Delta E_a|$), respectively. Both the system-specific separable basis and reference finite-difference calculations were performed with $h = 0.2$ bohr.

present the residual size of the first 100 orbitals and occupied orbitals in Figure 3a and 3b, respectively.

For small rank sizes, the basis vectors do not sufficiently cover the reference orbitals, but they span the orbitals better as the rank size increases (see Figure 3a). In addition, it was observed that the residual sizes for the low-lying orbitals do not always decrease first as the rank size increases, but those for high virtual orbitals slowly converge to zero. This indicates that the basis space spans the orbitals sufficiently, especially for low-lying orbitals.

To investigate the effects of system-specific separable basis vectors in the SCF procedure, we performed density functional

calculations for the tested structures. The computational details and results are summarized in Tables S3−S9.

Figure 4 shows the absolute errors in atomization energies, $|\Delta E_a|$ (black line), and band gaps, $|\Delta E_g|$ (blue line), as a function of the number of basis vectors, $R = r_x \times r_y \times r_z$. For the cubic diamond structures and ABO_3 structures, we sampled the same rank sizes for all three axes, whereas the rank size of the hexagonal sheet structures is proportional to the cell size along each axis. The same cell parameters and rank sizes were used for the single-atom calculations needed to calculate the atomization energies.
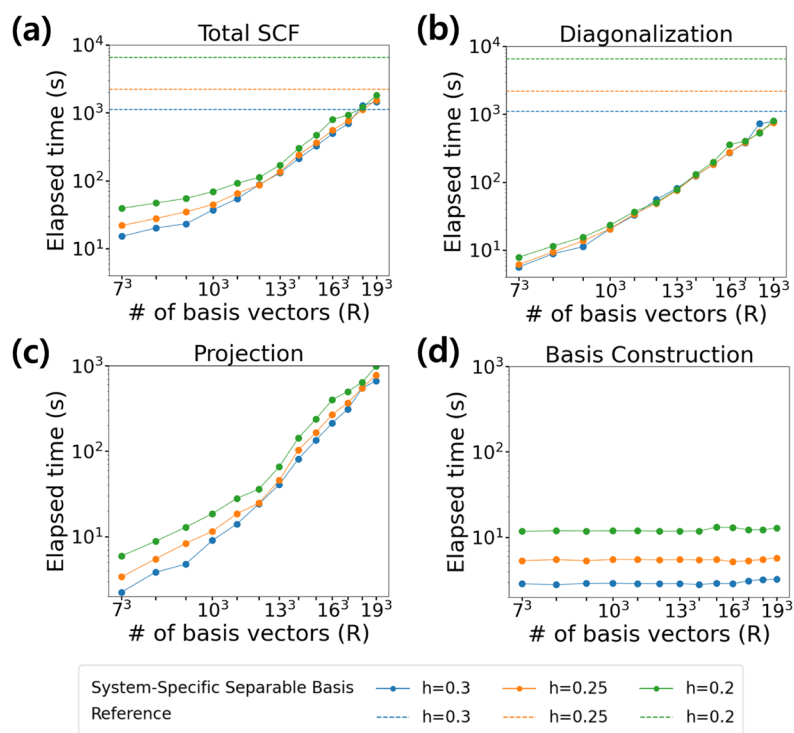
Figure 5. Elapsed times of density functional calculations according to the number of basis vectors on a log–log scale. Elapsed time for (a) total SCF calculation, (b) diagonalization, (c) projection of the Hamiltonian matrix, and (d) construction of the system-specific separable basis. The dashed lines in (a) and (b) represent the elapsed time for the reference finite-difference calculations for the same system (Si) and computational conditions. Blue, orange, and green lines represent the results for h of 0.3, 0.25, and 0.2 bohr, respectively.

The atomization energy and the band gap do not converge monotonically with respect to R, whereas the monotonic convergence of the total energy is guaranteed by the variational principle, as shown in the 10th column in Tables S3–S9. In the test range of R, the systems show convergence within 0.1–10 meV for both $|\Delta E_a|$ and $|\Delta E_g|$. Elucidating the dependence of the error convergence on the structures is difficult with a few test cases. Nonetheless, we note that the system-specific separable basis converges well, even in systems with transition metals or a nonperiodic axis.

To investigate the speed of SCF calculations with a system-specific separable basis, we plot the elapsed times for the overall calculations and three major bottlenecks for Si calculations as a function of R in Figure 5. To check the dependence of the computation time on the grid spacing h, we also plotted the results with different h values. The blue, orange, and green lines indicate the results with h values of 0.3, 0.25, and 0.2 bohr, respectively. The dashed lines represent the elapsed time for the reference FD calculations. The elapsed time of the total SCF procedure with the system-specific separable basis increases as R increases but does not show a dramatic change with respect to h, as shown in Figure 5a. By contrast, the total elapsed time of the reference FD results increased rapidly as h decreased.

The increase in the elapsed time of the reference calculations originates from diagonalization, which is the primary bottleneck. As shown in Figure 5b, most of the elapsed time for the reference calculations is spent in diagonalization, and its cost is strongly dependent on the h values. For the case of a system-specific separable basis, the elapsed time of the diagonalization is independent of the choice of h and is much smaller than that of the reference cases. This is because the dimensions of the projected Hamiltonian matrix are determined not by the number of grid points, $N = N_x \times N_y \times N_z$, but by R which is much smaller than N.

The system-specific separable basis additionally induces the basis construction and projection processes. Figures 5c and d show the elapsed time for projection and basis construction, respectively. The computational time for basis construction relies on h values because we obtain U_x, U_y, and U_z from the direct diagonalization of small dense matrices, $H_{(x)}H_{(x)}^H, H_{(y)}H_{(y)}^H$, and $H_{(z)}H_{(z)}^H$. Despite the strong h dependence, the basis construction time occupies only a small part of the overall time. Contrary to the basis construction time, the cost of the projection depends on both h and R. The detailed computational complexities of the projections and basis construction are explained in the Appendix. The projection time increases as R increases and h decreases. However, large differences in the elapsed time of the projection as a function of h are not shown, except for a few small R cases. Hence, the total computational time for a system-specific basis calculation does not increase significantly as h decreases.

Although system-specific separable basis calculations require additional processes, they show excellent performance in diagonalization; thus, the overall computational costs are reduced in most cases. Here, we discuss only the results of Si, but we observed the same trend for other systems (see the sixth–ninth columns of Tables S3–S9).

Figure 6 summarizes the overall performance enhancement by the new basis with respect to the reference FD calculations as a function of R/N, when $h = 0.2$ bohr. Figure S3 shows the performance enhancement results with other h values. For all systems, smaller R/N values resulted in larger performance enhancement. This is because a smaller R/N implies a larger
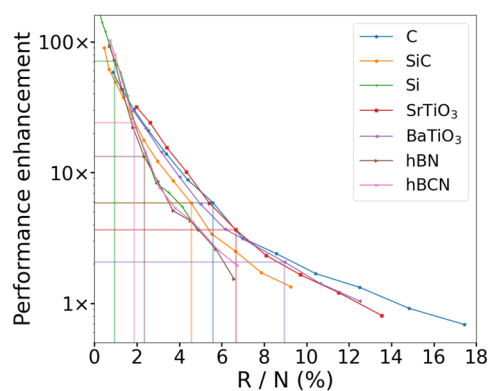
Figure 6. Performance enhancement by a system-specific separable basis as a function of the ratio of dimensions of the finite-difference and the projected Hamiltonian matrices, R/N. The intersections of the horizontal and vertical lines represent the smallest R/N cases, where both the atomization energy and the band gap have errors of less than 100 meV. For all calculations, with or without basis vectors, a grid spacing of 0.2 bohr was used.

reduction in the dimension of the Hamiltonian matrix, resulting in faster diagonalization. The elapsed time for each component for all calculations is included in Tables S3–S9. The intersections of the horizontal and vertical lines represent the smallest R/N case for each system, where both $|\Delta E_a|$ and $|\Delta E_g|$ were less than 100 meV. The intersections of the ABO_3 and hexagonal sheet structures were ~8% and ~2%, respectively. The diamond structures should have an R/N of ~5% to achieve a tolerance of 100 meV, except for Si. The N value for the Si system is greater than those of the other diamond structures because the cell volume of Si is larger than that of the others. In addition, the R for the error convergences was slightly smaller than that of the others. Therefore, the Si system showed significant performance enhancement.

To be more practical, a system-specific separable basis must achieve performance improvements with sufficiently high accuracy. The system-specific separable basis balances speed and accuracy by tuning R. A large R achieves high accuracy but simultaneously reduces the calculation speed, as shown in Figures 4 and 6. Table 1 summarizes the performance enhancements of the tested systems for 100, 50, and 25 meV tolerances for both types of errors. As shown in Figure 6, the use of large R/N to achieve high accuracy reduces the gains in computation speed. However, significant acceleration (2–14×) was achieved, even within a small tolerance value of 25 meV.

CONCLUSION

Here, we proposed a system-specific separable basis derived from Tucker decomposition of a finite-difference Hamiltonian and investigate its performance in density functional

calculations. We show that the new basis can successfully span low-lying orbitals and that their coverage is systematically improved by increasing the rank size of Tucker decomposition. The proposed basis dramatically reduces the dimensions of the Hamiltonian matrix and hence accelerates the diagonalization of the Hamiltonian matrix. We confirmed the properties of the basis vectors and measured the performance enhancements using seven selected systems. The system-specific separable basis achieved a 2−71× increase in computation speed with 100 meV tolerance for the band gap and the atomization energy. Higher accuracy can be achieved for all tested systems with a larger rank size but a lower gain in computation speed (e.g., 2−14× increase with a 25 meV tolerance). Here, we validated the performance of the system-specific separable basis only for density functional calculations. However, we expect it to be useful for higher-order quantum chemical methods because our basis benefits from the advantages of both the real-space method (e.g., fast numerical integrations and derivatives) and the typical basis function expansion method (e.g., low dimension of basis space and separability).

■ APPENDIX

Negative Shift of the Hamiltonian Matrix

We introduced a fictitious potential to obtain singular vectors that well span low-lying orbitals. By the application of a large negative constant potential, the eigenspectrum of H is shifted down. This shift avoids a large positive eigenvalue of H, so the singular vectors that mainly span the high virtual orbitals are not selected in the compact HOSVD. In Figure S2, the results of representability tests with different sizes of the fictitious potentials are plotted. If the magnitude of the fictitious potential is sufficiently large, its value does not affect the representability of the basis vectors. We used −500 au as the fictitious potential for all calculations presented in this work.

Projecting the Hamiltonian Matrix on a System-Specific Separable Basis Space

A KS Hamiltonian matrix is composed of three terms: kinetic energy matrix, T, local potential matrix, V^{local}, and nonlocal potential matrix, V^{NL}. Each term has a specific nonzero pattern; therefore, we replace the projection of the Hamiltonian matrix, $\tilde{H}$, with the projection of each matrix using its nonzero pattern.

An FD kinetic energy matrix in order-6 tensor format, $\mathcal{T}$, can be represented as follows:

$$\mathcal{T}_{ijki'j'k'} = T^x_{ii'}\delta_{jj'}\delta_{kk'} + \delta_{ii'}T^y_{jj'}\delta_{kk'} + \delta_{ii'}\delta_{jj'}T^z_{kk'} \tag{8}$$

where T^n ($n \in \{x, y, z\}$) is the second-order FD matrix along the n-axis and δ is the Kronecker delta. Owing to the separability and orthonormality of U, the projection of the kinetic energy matrix on U, $\tilde{\mathcal{T}}$, was obtained by three 1D projections as follows:

Table 1. Performance Enhancements of the Tested Systems at Various Error Tolerances[a]

	diamond			ABO3		hexagonal sheet	
tolerance (meV)	C	SiC	Si	SrTiO₃	BaTiO₃	hBN	hBCN
100	5.88 (5.59%)	5.86 (4.56%)	71.43 (0.95%)	3.65 (6.66%)	2.07 (8.95%)	13.31 (2.34%)	24.17 (1.86%)
50	3.14 (6.98%)	2.51 (6.63%)	71.43 (0.95%)	1.65 (9.70%)	2.07 (8.95%)	4.27 (4.53%)	7.70 (3.06%)
25	2.39 (8.59%)	1.71 (7.87%)	14.07 (2.40%)	1.65 (9.70%)	2.07 (8.95%)	4.27 (4.53%)	5.36 (3.79%)

[a]The percentages in parentheses indicate the ratio of dimensions of the finite-difference and the projected Hamiltonian matrices, R/N.

$$\tilde{\mathcal{T}}_{\alpha\beta\gamma\alpha'\beta'\gamma'}$$
$$= \sum_{i,j,k,i',j',k'} \mathcal{T}_{ijki'j'k'}(U_{i\alpha}^x)^*(U_{j\beta}^y)^*(U_{k\gamma}^z)^*U_{i'\alpha'}^x U_{j'\beta'}^y U_{k'\gamma'}^z$$
$$= \sum_{i,i'} (U_{i\alpha}^x)^* T_{ii'}^x U_{i'\alpha'}^x + \sum_{j,j'} (U_{j\beta}^y)^* T_{jj'}^y U_{j'\beta'}^y$$
$$+ \sum_{k,k'} (U_{k\gamma}^z)^* T_{kk'}^z U_{k'\gamma'}^z \tag{9}$$

In eq 9, the projection of the kinetic energy matrix is calculated using the projections of the three 1D kinetic energy matrices. If we assume that $r_x = r_y = r_z = R^{1/3}$ and $N_x = N_y = N_z = N^{1/3}$, the computational complexity of evaluating each term of the second line in eq 9 becomes $O(R^{1/3}N^{2/3} + R^{2/3}N^{1/3})$.

In the FD representation, the local potential matrix, V^{local}, has nonzero values only on the diagonal, which means that $V_{pq}^{\text{local}} = V_p^{\text{local}}\delta_{pq}$, where p and q are the indices on a 3D grid. If p and q are decomposed into (i, j, k) and (i', j', k'), respectively, the local potential in tensor form becomes $\mathcal{V}_{ijki'j'k'}^{\text{local}} = V_{ijk}^{\text{local}}\delta_{ii'}\delta_{jj'}\delta_{kk'}$. Because of this nonzero pattern of the local potential, the 3D projection of $\mathcal{V}_{ijki'j'k'}^{\text{local}}$ is transformed as follows:

$$\tilde{\mathcal{V}}_{\alpha\beta\gamma\alpha'\beta'\gamma'}^{\text{local}}$$
$$= \sum_{i,j,k,i',j',k'} \mathcal{V}_{ijki'j'k'}^{\text{local}}(U_{i\alpha}^x)^*(U_{j\beta}^y)^*(U_{k\gamma}^z)^*U_{i'\alpha'}^x U_{j'\beta'}^y U_{k'\gamma'}^z$$
$$= \left(\sum_k (U_{k\gamma}^z)^* \left(\sum_j (U_{j\beta}^y)^* \left(\sum_i (U_{i\alpha}^x)^* \mathcal{V}_{ijk}^{\text{local}} U_{i\alpha'}^x \right) U_{j\beta'}^y \right) U_{k\gamma'}^z \right) \tag{10}$$

The first, second, and third parentheses in eq 10 represent the x-, y-, and z-axis projections, respectively. Each projection is performed with the contraction of an order-6 tensor with two matrices; therefore, it requires an 8-nested for-loop. However, using the nonzero pattern of $\mathcal{V}_{ijki'j'k'}$, the x-, y-, and z-axis projections are performed by 5-, 6-, and 7-nested for-loops, respectively. The computational complexities of the three projections are $O(R^{2/3}N)$, $O(R^{4/3}N^{2/3})$, and $O(R^2 N^{1/3})$, respectively. In our test, the last projection was a major bottleneck in the local potential projection.

Because all electron–electron interactions are represented by the local potential under the KS density functional theory, a nonlocal potential matrix, V^{NL}, is only from the pseudopotential. The most frequently used pseudopotential is the Kleinman–Bylander (KB)[29] formula given by

$$V_{pq}^{\text{NL}} = \sum_a \sum_{l,m} (P_p^{a,lm})^* D^{a,lm} P_q^{a,lm} \tag{11}$$

where P and D are a KB projector matrix and a coefficient matrix for each pair of angular momenta, respectively, and a, l, and m denote the indices for the atom, the angular momentum, and the magnetic momentum, respectively. Similar to eq 10, we consider the tensor form of a nonlocal potential term, $\mathcal{V}^{\text{NL}}$, and KB projectors, $\mathcal{P}_{ijk}^{a,lm}$ and $\mathcal{P}_{i'j'k'}^{a,lm}$). The projection of the nonlocal potential is computed as follows:

$$\tilde{\mathcal{V}}_{\alpha\beta\gamma\alpha'\beta'\gamma'}^{\text{NL}}$$
$$= \sum_{i,j,k,i',j',k'} \mathcal{V}_{ijki'j'k'}^{\text{NL}}(U_{i\alpha}^x)^*(U_{j\beta}^y)^*(U_{k\gamma}^z)^*U_{i'\alpha'}^x U_{j'\beta'}^y U_{k'\gamma'}^z$$
$$= \sum_a \sum_{l,m} (\tilde{\mathcal{P}}_{\alpha\beta\gamma}^{a,lm})^* D^{a,lm} \tilde{\mathcal{P}}_{\alpha'\beta'\gamma'}^{a,lm} \tag{12}$$

where $\tilde{P}_{\alpha'\beta'\gamma'}^{a,lm} = \sum_{i,j,k} \mathcal{P}_{ijk}^{a,lm} U_{i\alpha}^x U_{j\beta}^y U_{k\gamma}^z$ and $\tilde{\mathcal{P}}_{\alpha\beta\gamma}^{a,lm} = \sum_{i',j',k'} \mathcal{P}_{i'j'k'}^{a,lm} U_{i'\alpha'}^x U_{j'\beta'}^y U_{k'\gamma'}^z$. The computational complexity of $\mathcal{P}_{\alpha\beta\gamma}^{a,lm}$ was $O(NR^{1/3} + N^{2/3}R^{2/3} + N^{1/3}R)$. The evaluation of $\mathcal{P}_{\alpha\beta\gamma}^{a,lm}$ is performed only once in the overall calculation because the nonlocal term does not change during the SCF procedure.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c01263.

> Effect of approximations on the construction of basis vectors; changes in the residual according to the size of the fictitious potential; performance enhancement of the density functional calculations with other grid spacings as a function of R/N; actual grid spacings for each test system; elapsed times for each procedure and the corresponding errors; structural information on the tested systems (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

    Woo Youn Kim − *Department of Chemistry, KAIST, Yuseong-gu, Daejeon 34141, Republic of Korea;* Ⓘ orcid.org/0000-0001-7152-2111; Email: wooyoun@kaist.ac.kr

    Sunghwan Choi − *National Institute of Supercomputing and Networking, Korea Institute of Science and Technology Information, Daejeon 34141, Republic of Korea;* Ⓘ orcid.org/0000-0002-4330-7710; Email: sunghwanchoi@kisti.re.kr

Author

    Jeheon Woo − *Department of Chemistry, KAIST, Yuseong-gu, Daejeon 34141, Republic of Korea*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c01263

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.;Wiley: New York, 2006.

(2) Balazs, N.; Jensen, F. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; John Wiley & Sons, Inc.: New York, 2017; Chapter 3, pp 93−150.

(3) Davidson, E. R.; Feller, D. Basis Set Selection for Molecular Calculations. *Chem. Rev.* **1986**, *86*, 681−696.

(4) Kresse, G.; Furthmuller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.

(5) Beck, T. L. Real-space mesh techniques in density-functional theory. *Rev. Mod. Phys.* **2000**, *72*, 1041−1080.

(6) Solala, E.; Parkkinen, P.; Sundholm, D. Tensor decompositions for the bubbles and cube numerical framework. *Comput. Phys. Commun.* **2018**, *232*, 98−103.

(7) Motamarri, P.; Gavini, V.; Blesgen, T. Tucker-tensor algorithm for large-scale Kohn-Sham density functional theory calculations. *Phys. Rev. B* **2016**, *93*, 1−14.

(8) Lin, C. C.; Motamarri, P.; Gavini, V. Tensor-structured algorithm for reduced-order scaling large-scale Kohn-Sham density functional theory calculations. *npj Computational Materials* **2021**, *7*, 50.

(9) Blesgen, T.; Gavini, V.; Khoromskaia, V. Approximation of the electron density of Aluminium clusters in tensor-product format. *J. Comput. Phys.* **2012**, *231*, 2551−2564.

(10) Bell, F.; Lambrecht, D. S.; Head-Gordon, M. Higher order singular value decomposition in quantum chemistry. *Mol. Phys.* **2010**, *108*, 2759−2773.

(11) Benedikt, U.; Auer, A. A.; Espig, M.; Hackbusch, W. Tensor decomposition in post-Hartree-Fock methods. I. Two-electron integrals and MP2. *J. Chem. Phys.* **2011**, *134*, 054118.

(12) Lesiuk, M. Implementation of the Coupled-Cluster Method with Single, Double, and Triple Excitations using Tensor Decompositions. *J. Chem. Theory Comput.* **2020**, *16*, 453−467.

(13) Godtliebsen, I. H.; Thomsen, B.; Christiansen, O. Tensor Decomposition and Vibrational Coupled Cluster Theory. *J. Phys. Chem. A* **2013**, *117*, 7267−7279.

(14) Benedikt, U.; Böhm, K.-H.; Auer, A. A. Tensor decomposition in post-Hartree-Fock methods. II. CCD implementation. *J. Chem. Phys.* **2013**, *139*, 224101.

(15) Jerke, J.; Poirier, B. Two-body Schrödinger wave functions in a plane-wave basis via separation of dimensions. *J. Chem. Phys.* **2018**, *148*, 104101.

(16) De Lathauwer, L.; De Moor, B.; Vandewalle, J. A Multilinear Singular Value Decomposition. *SIAM Journal on Matrix Analysis and Applications* **2000**, *21*, 1253−1278.

(17) Kolda, T. G.; Bader, B. W. Tensor decompositions and applications. *SIAM Review* **2009**, *51*, 455−500.

(18) Khoromskaia, V.; Khoromskij, B. N. *Tensor Numerical Methods in Quantum Chemistry*; De Gruyter: Berlin, Germany, 2018.

(19) Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika* **1966**, *31*, 279−311.

(20) Khoromskaia, V.; Khoromskij, B. N. Tensor numerical methods in quantum chemistry: From Hartree-Fock to excitation energies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31491−31509.

(21) Schmitz, G.; Madsen, N. K.; Christiansen, O. Atomic-batched tensor decomposed two-electron repulsion integrals. *J. Chem. Phys.* **2017**, *146*, 134112.

(22) Schmitz, G.; Christiansen, O. Assessment of the overlap metric in the context of RI-MP2 and atomic batched tensor decomposed MP2. *Chem. Phys. Lett.* **2018**, *701*, 7−14.

(23) Kang, S.; Woo, J.; Kim, J.; Kim, H.; Kim, Y.; Lim, J.; Choi, S.; Kim, W. Y. ACE-Molecule: An open-source real-space quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 124110.

(24) Choi, S.; Kwon, O.-k.; Kim, J.; Kim, W. Y. Performance of heterogeneous computing with graphics processing unit and many integrated core for hartree potential calculations on a numerical grid. *J. Comput. Chem.* **2016**, *37*, 2193−2201.

(25) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1−5.

(26) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(27) Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B* **2013**, *88*, 085117.

(28) Virtanen, P.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.

(29) Kleinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425−1428.