JCTC Journal of Chemical Theory and Computation

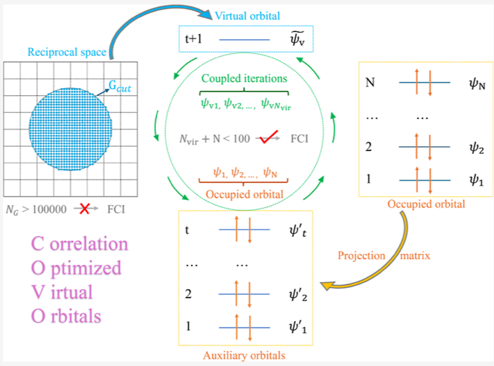
pubs.acs.org/JCTC Article

Improved Correlation Optimized Virtual Orbital Algorithm for Plane-Wave Full Configuration Interaction Calculations

Mingyu Qiu, Zhenlin Zhang, Zhiyuan Zhang, Yexuan Lin, Yingzhou Li,* Jinlong Yang,* and Wei Hu*

Cite This: https://doi.org/10.1021/acs.jctc.5c00586 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Full configuration interaction (FCI) calculations have historically faced significant challenges in dealing with periodic systems. The plane-wave basis sets are valued for their efficiency and broad applicability in various computational physics and chemistry simulations. Because of their natural periodicity, the plane-wave basis sets offer a potential solution to this problem. Moreover, FCI can address the limitations of widely used methods, such as density functional theory (DFT) with plane-wave basis sets, in accurately describing strongly correlated systems. However, the large basis set nature of the plane-wave makes them unsuitable for direct application in FCI calculations. To address this challenge, we propose an improved algorithm based on the correlation-optimized virtual orbital (COVOS) framework. By incorporating rotational matrices to enhance the active space dimension and optimizing orbitals through iterative coupled processes, we successfully compress the extensive plane-wave basis set into a manageable number of virtual orbitals suitable for FCI calculations while retaining most of the original basis set characteristics. We apply this method to supercell calculations and potential energy curves of periodic metallic systems. To further validate our approach, we test it on nonperiodic small molecular systems and compare the results with those obtained from DFT, second-order Møller−Plesset perturbation theory (MP2), random phase approximation (RPA), FCI calculations using the 6−31G or cc-pVDZ basis sets, and the original COVOS algorithm. The improved COVOS framework demonstrates significant advantages in convergence and correlation description over the original method. Furthermore, we observe metal divergence issues in MP2 calculations for certain metallic systems and note that RPA may overestimate the correlation energy of such systems. These findings underscore the importance of achieving FCI calculations with plane-wave basis sets.

1. INTRODUCTION

In quantum chemistry, the FCI theory is recognized as the most rigorous and precise approach.[1] FCI constructs and solves the Schrödinger equation for the entire electronic configuration space within a given basis, fully capturing electron correlation effects. Consequently, it is regarded as the "exact solution" in computational chemistry.[2] Traditional FCI methods are primarily based on atomic orbital expansions using Gaussian or atomic natural basis sets, which are highly localized in molecular systems. However, these basis sets suffer from significant basis set errors in periodic systems and struggle to describe electron delocalization and long-range correlations.[3]

The plane-wave method, which is based on periodic basis sets and incorporating algorithms such as fast Fourier transform (FFT),[4] naturally addresses boundary condition problems in periodic systems while avoiding the biases introduced by the choice of basis sets in traditional atomic orbital approaches.[5] This method is particularly well-suited for simulating crystalline and surface systems in condensed matter physics and materials science, and it has become a critical component of electronic structure methods such as DFT.[6,7] The development of plane-wave condensed matter calculations has brought about revolutionary changes to the fields of chemistry and materials science. Its precision and efficiency have enabled the simulation of complex chemical systems.

Compared to traditional basis sets in the FCI method, plane-wave basis sets offer several distinct advantages:

1 Controllable basis set error: the plane-wave basis set, systematically expanded through wavevector truncation,[8] ensures strict convergence and avoids nonphysical errors introduced by basis set selection.
2 Compatibility with periodic systems: plane-wave basis sets are highly compatible with periodic boundary conditions (PBC),[9] making them ideal in describing the electronic structure of crystals and surfaces.
3 Uniformity and delocalization: plane-wave basis sets are uniformly distributed in space and effectively describe electron delocalization, which is particularly crucial for

Received: April 14, 2025
Revised: June 6, 2025
Accepted: June 6, 2025

studying metallic systems and periodic strongly correlated materials where electron localization and long-range interactions coexist.[10] For systems with highly localized orbitals, such as transition metal oxides or actinides, plane waves are also supplemented with techniques such as hybrid functionals to capture strong correlation effects.

Despite these advantages, research on many-body correlation theories within the plane-wave framework has focused mainly on approximate methods, such as low-order many-body perturbation theory (MPn), RPA,[11] coupled cluster theory (CC), and the widely used DFT. However, the application of MPn and CC methods within the plane-wave framework remains limited. Both MP2[12−14] and CC[15−17] methods often rely on transforming Gaussian basis sets into grid-based representations resembling plane-wave approaches,[18] rather than utilizing native plane-wave implementations. While these methods are computationally efficient, they often face challenges in handling strongly correlated systems and high-precision computations, failing to achieve the systematically accurate results provided by FCI.

Thus, developing FCI algorithms within the plane-wave framework represents both a technical innovation and a computational breakthrough. This integration can fully leverage the strengths of the plane-wave method, address the limitations of traditional FCI in periodic systems, and provide novel tools for studying strongly correlated periodic systems. Specifically, this approach aims to address the following key scientific questions:

1. Achieving high-precision descriptions of electron correlation effects in periodic systems. Furthermore, FCI will serve as a benchmark[19] for assessing the accuracy of other approximate methods, such as perturbative algorithms.
2. Systematically investigating quantum phenomena in strongly correlated systems,[20] including the spin dynamics and exchange interactions in magnetic materials, as well as the pairing mechanisms explored within the plane-wave FCI framework and unconventional superconductivity in correlated electron systems such as high-temperature superconductors[21] and heavy Fermion compounds.

However, a significant challenge in performing FCI calculations with plane-wave basis sets lies in the overwhelmingly large number of basis functions.[22] For the computationally expensive FCI method, the excessive basis size becomes a critical bottleneck.[23,24] To address this, it is necessary to select a small number of orbitals with dominant contributions from the extensive basis set. Several approaches have been proposed to tackle this issue, with the following strategies. Considering that there are currently few optimization methods for FCI calculations based on plane-wave basis sets, this presentation also introduces some methods for other basis sets for reference.

1. Basis set compression techniques: basis set reduction strategies generally fall into two complementary approaches: projection operator methods optimize the plane-wave basis through subspace construction, where techniques like subspace diagonalization[25] iteratively refine Hamiltonian subspaces to capture essential system features while maintaining computational tractability. Mathematical compression techniques address basis redundancy through algebraic decomposition, exemplified by Cholesky factorization of overlap matrices and tensor hypercontraction (THC),[26,27] which, respectively, achieve dimensionality reduction via matrix factorization and low-rank tensor representations of electron integrals.
2. Multiresolution analysis: while plane-wave basis sets offer uniform spatial resolution, multiresolution methods employ hierarchical basis functions or dynamic grid refinement to resolve localized features. The multiwavelet approach utilizes wavelet bases to selectively capture regions with sharp electron density gradients, whereas adaptive mesh refinement (AMR) dynamically adjusts grid resolution based on local system properties.[28] These methods sacrifice the inherent periodicity of plane waves for enhanced efficiency in modeling molecular systems with strong electronic singularities, thereby constituting a complementary strategy to plane-wave approaches in multiscale simulations.
3. Truncation methods: these approaches systematically discard less significant orbitals or basis functions based on their contribution to the overall system. For example, Natural Orbital Truncation[29] removes orbitals with low occupation numbers, and Canonical Virtual Space Truncation[30] limits the virtual orbital space to simplify excited-state calculations.

While conventional truncation schemes focus on static orbital selection criteria, recent advances emphasize correlation-driven optimization of the virtual space. The COVOS[31,32] algorithm is an approach for generating optimized virtual orbital basis sets in plane-wave systems. Unlike other plane-wave-based optimization methods, which focus on fitting single-electron eigenvalue spectra or band structures,[33,34] COVOS optimizes small selective configuration interaction (CI) problems to construct virtual orbitals.

This work improves upon the COVOS algorithm from an algorithmic perspective and extends its applicability to larger computational systems, including supercell expansion and dissociation energy calculations. The goal is to bridge the current gap between the plane-wave basis set and accurate many-body theories in condensed matter calculations, advancing the application of FCI algorithms within plane-wave frameworks and providing theoretical support and computational tools for addressing complex periodic and strongly correlated systems.

In Section 2, we provide an overview of the COVOS algorithm and describe the improvements made; we introduced a rotation matrix to project the occupied orbitals into an active space of specific size, and based on this, we performed coupled iterations for the virtual orbitals and the rotation matrix. Section 3 presents the computation and results; the results show that the virtual orbitals obtained by the improved COVOS algorithm are significantly better than those from the original COVOS, and the algorithm has achieved the expected effects in various tests. Additionally, comparisons with other algorithms have demonstrated the benchmark role of FCI. Section 4 summarizes the findings and discusses prospects for future work.

2. PRELIMINARIES

We set N to represent the number of basis orbitals occupied by the ground state, corresponding to the total number of electrons in the system, and N_{vir} represents the size of the maximum virtual orbital space set; we can adjust N_{vir} based on the desired number of active orbitals for the final FCI calculation. Then we set ψ to represent the spatial orbital basis, and ψ_i $(i \le N)$ to represent the occupied orbital, and ψ_{vi} $(i \leq N_{\text{vir}})$ to represent virtual orbital; for ψ_{vi} that has not yet been optimized, we refer to it as $\tilde{\psi}_{vi}$. The general workflow of the original COVOS algorithm is as follows:

1. $n = 1$ is set.
2. The initial parameters $\tilde{\psi}_{vn}$ for the n-th virtual orbital ψ_{vn} are chosen. The occupied orbitals ψ_i and the initial guess for the virtual orbitals $\tilde{\psi}_{vn}$ can be derived from any self-consistent field orbitals. In this work, we specifically use the Kohn−Sham orbitals generated by KSSOLV.[35,36]
3. The virtual orbital $\tilde{\psi}_{vn}$ is interacted with the N occupied orbitals ψ_1, ψ_2, ..., ψ_N. Considering that involving all occupied orbitals in the virtual orbital optimization process can make gradient descent computationally expensive, we will select the outermost orbital ψ_N as the active orbital, along with the virtual orbital in the unoccupied states, to form a CI matrix as a smaller active space.
4. $\tilde{\psi}_{vn}$ is optimized using the gradient descent method to minimize the ground state energy obtained from the diagonalization of the CI matrix in step 3, ensuring that the orbital remains orthogonal to all occupied orbitals ψ_1, ψ_2, ..., ψ_N and the previously generated virtual orbitals ψ_{v1}, ψ_{v2}, ..., $\psi_{v(n-1)}$. If the optimization meets the convergence criteria, the optimization for ψ_{vn} is finalized; otherwise, Step 3 is revisited.
5. If $n < N_{\text{vir}}$, n is incremented by 1, and Step 2 is revisited. If $n = N_{\text{vir}}$, the generation of virtual orbitals is finalized.
6. A FCI calculation of the ground state energy is performed using all the occupied orbitals ψ_1, ψ_2, ..., ψ_N and the virtual orbitals ψ_{v1}, ψ_{v2}, ..., $\psi_{vN_{\text{vir}}}$.

During the gradient descent process, the gradient of each matrix element of the CI matrix, which is constructed for 2-electron system with 2 spatial orbitals $\tilde{\psi}_{vn}$ and ψ_N, is computed with respect to the virtual orbitals $\tilde{\psi}_{vn}$. The detailed derivation can be found in the Supporting Information part 1 and part 2.

In the absence of external fields, the one-electron integral takes the form

$$\langle p|q\rangle = \sum_{\mathbf{G}} \psi_p^*(\mathbf{G})\frac{|\mathbf{G}|^2}{2}\psi_q(\mathbf{G}) + (V_{\text{loc}}(\mathbf{G}) + V_{\text{nloc}}(\mathbf{G}))\rho_{pq}(\mathbf{G}) \tag{1}$$

where $\mathbf{G}$ is the reciprocal lattice vector, $\psi(\mathbf{G})$ are the expansion coefficients of the orbital in reciprocal space, $\frac{|\mathbf{G}|^2}{2}$ is the kinetic energy term in reciprocal space, and $\rho_{pq}(\mathbf{G})$ is the generalized density in reciprocal space, defined as

$$\rho_{pq}(\mathbf{G}) = \sum_{\mathbf{G}'} \psi_p^*(\mathbf{G}')\psi_q(\mathbf{G}' - \mathbf{G}) \tag{2}$$

The two-electron integral describes the Coulomb interaction between electrons, which can also be written as[32]

$$\langle pq|rs\rangle = \frac{\Omega}{2}\sum_{\mathbf{G}\neq 0} \rho_{pr}^*(\mathbf{G})\frac{4\pi}{|\mathbf{G}|^2}\rho_{qs}(\mathbf{G}) \tag{3}$$

It should be noted that when computing with a single Γ-point, the two-electron integrals can be expressed in a simplified form. However, when incorporating a full k-point mesh, the equations become significantly more complex due to the extended sampling of the Brillouin zone. This complexity arises from the need to carefully track phase relationships between Bloch functions at different k-points, enforce momentum conservation, and perform nontrivial summations over the k-point grid. These summations involve coupling wave functions from different points in reciprocal space, introducing intricate bookkeeping due to phase mismatches and effects analogous to what Harry F. King, Richard E. Stanton, Hojing Kim, Robert E. Wyatt, and Robert G. Parr referred to as noncoincidences, but generalized here to periodic boundary conditions. These noncoincidences reflect inherent structural mismatches between quantum states that do not align trivially across the symmetry-adapted k-space manifold. For further details, the reader is referred to the work of Bylaska, Hutter, Monkhorst, and others.[37−43]

The gradients with respect to ψ_j for all j of one-electron and two-electron integrals can be expressed as

$$\begin{aligned}\frac{\partial\langle p|q\rangle}{\partial\psi_j(\mathbf{G})} &= \delta_{pj}\left(\frac{|\mathbf{G}|^2}{2} + V_{\text{loc}}(\mathbf{G}) + V_{\text{nloc}}(\mathbf{G})\right)\psi_q(\mathbf{G}) \\ &\quad + \delta_{qj}\left(\frac{|\mathbf{G}|^2}{2} + V_{\text{loc}}(\mathbf{G}) + V_{\text{nloc}}(\mathbf{G})\right)\psi_p(\mathbf{G})\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{\partial\langle pq|rs\rangle}{\partial\psi_j(\mathbf{G})} &= \sum_{\mathbf{G}'} \delta_{pj}\psi_r(\mathbf{G} - \mathbf{G}')\rho_{qs}(\mathbf{G}')\frac{4\pi}{|\mathbf{G}'|^2} \\ &\quad + \delta_{qj}\psi_s(\mathbf{G} - \mathbf{G}')\rho_{pr}(\mathbf{G}')\frac{4\pi}{|\mathbf{G}'|^2} \\ &\quad + \delta_{rj}\psi_p(\mathbf{G} - \mathbf{G}')\rho_{qs}(\mathbf{G}')\frac{4\pi}{|\mathbf{G}'|^2} \\ &\quad + \delta_{sj}\psi_q(\mathbf{G} - \mathbf{G}')\rho_{pr}(\mathbf{G}')\frac{4\pi}{|\mathbf{G}'|^2}\end{aligned} \tag{5}$$

We denote the CI matrix as H_{CI}. When we perform gradient descent on the smallest eigenvalue of H_{CI}, first, we solve for the smallest eigenvalue E_{CI} and the corresponding eigenvector $\boldsymbol{x}$ of H_{CI}.

Subsequently, since the matrix elements of H_{CI} are entirely composed of one-electron and two-electron integrals, according to the Hellmann−Feynman theorem,[44] we can derive that

$$\begin{aligned}\frac{\partial E_{\text{CI}}}{\partial\psi_j(\mathbf{G})} &= \frac{\partial\boldsymbol{x}^T H_{\text{CI}}\boldsymbol{x}}{\partial\psi_j(\mathbf{G})} \\ &= \left\langle \psi_{\text{CI}}\left|\frac{\partial\hat{H}}{\partial\psi_j(\mathbf{G})}\right|\psi_{\text{CI}}\right\rangle \\ &= \sum_{pq} c_{pq}\frac{\partial\langle p|q\rangle}{\partial\psi_j(\mathbf{G})} + \sum_{pqrs} c_{pqrs}\frac{\partial\langle pq|rs\rangle}{\partial\psi_j(\mathbf{G})}\end{aligned} \tag{6}$$

where c_{pq} and c_{pqrs} are the single-electron and two-electron reduced density matrices of $\boldsymbol{x}^T H_{\text{CI}}\boldsymbol{x}$, abbreviated as 1-RDM and 2-RDM, $\hat{H}$ is the system Hamiltonian operator, and ψ_{CI} is the wave function corresponding to the smallest eigenvalue.

The pseudocode for the iterative process of ψ_{vn} in the original COVOS is given in Algorithm 1.

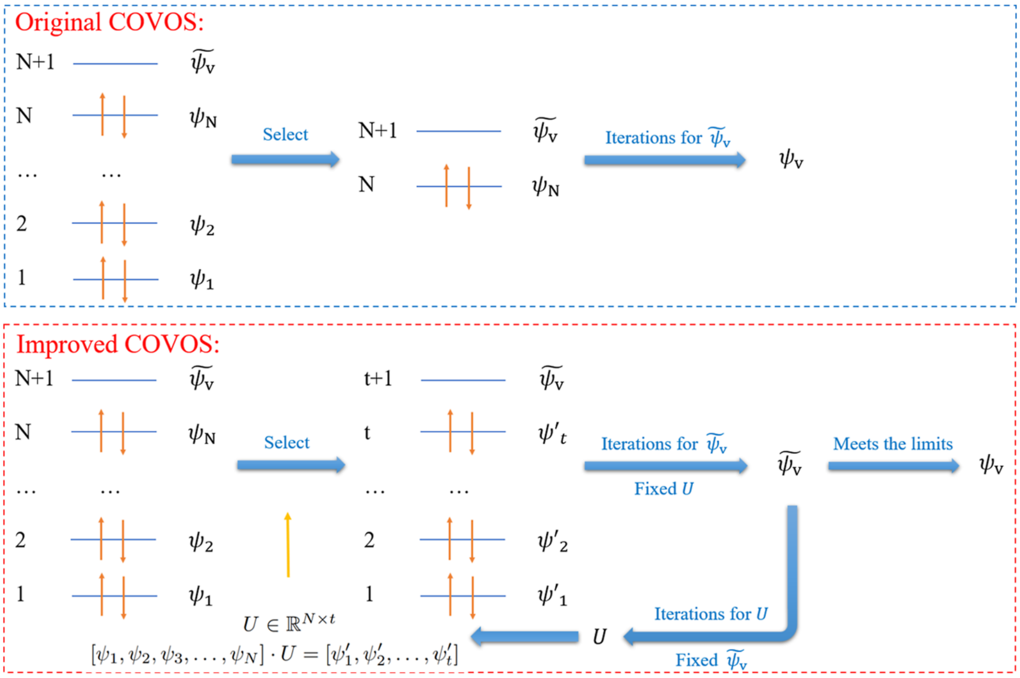


Figure 1. Overview diagrams of the original COVOS algorithm and the improved COVOS algorithm. We set N to represent the number of basis orbitals occupied by the ground state, corresponding to the total number of electrons in the system. Then we set ψ to represent the spatial orbital basis, and ψ_i $(i \leq N)$ to represent the occupied orbital, and ψ_v to represent virtual orbital; for ψ_v that has not yet been optimized, we refer to it as $\tilde{\psi}_v$, and U represents the projection matrix. Compared to the original COVOS algorithm, which only uses ψ_N to iterate $\tilde{\psi}_v$, we introduced a projection matrix U to generate auxiliary orbitals for the iterative calculation of $\tilde{\psi}_v$. The process involves continuously coupling the iteration of $\tilde{\psi}_v$ and U until convergence criteria are met.

Algorithm 1 The iterative process of ψ_{vn} in the original COVOS

```
**Input:** Initial orbitals $\tilde{\psi}_{vn}$, occupied orbitals $\psi_1, \psi_2, \ldots, \psi_N$ and the previously generated virtual orbitals $\psi_{v1}, \psi_{v2}, \ldots, \psi_{v(n-1)}$, learning rate $\alpha$, convergence threshold $\epsilon$
**Output:** Optimized orbitals $\psi_{vn}$
Construct a set $\mathcal{S} = [\psi_1, \psi_2, \ldots, \psi_N, \psi_{v1}, \psi_{v2}, \ldots, \psi_{v(n-1)}]$
Initialize iteration counter $k \leftarrow 0$, error $error \leftarrow \infty$
while $error > \epsilon$ do
    Orthogonalize and normalize $\tilde{\psi}^k_{vn}$ with respect to all orbitals in $\mathcal{S}$
    Construct $H_{CI}$ for 2-electron system with 2 spatial orbitals $\tilde{\psi}^k_{vn}$ and $\psi_N$
    Compute gradient $\frac{\partial E_{CI}}{\partial \tilde{\psi}^k_{vn}}$
    Update parameters: $\tilde{\psi}^{k+1}_{vn} \leftarrow \tilde{\psi}^k_{vn} - \alpha \frac{\partial E_{CI}}{\partial \tilde{\psi}^k_{vn}}$
    Compute new error: $error \leftarrow \|\tilde{\psi}^k_{vn} - \tilde{\psi}^{k+1}_{vn}\|^2$
    Increment iteration counter: $k \leftarrow k + 1$
end while
Orthogonalize and normalize $\tilde{\psi}_{vn}$ with respect to all orbitals in $\mathcal{S}$
return Optimized orbitals $\psi_{vn}$
```

3. OPTIMIZATION ALGORITHM

The original COVOS algorithm includes only the outermost occupied orbitals, designated as active orbitals,[32] in the optimization process. This approach significantly reduces the computational cost of the COVOS virtual orbital calculations in many-body systems. Simultaneously, approximating the outermost orbitals as the most likely to transition ensures the utility of the virtual orbitals. However, for many systems, the transitions of inner orbitals are also non-negligible,[45,46] and the original COVOS algorithm lacks representation for these orbital transitions.

To address this limitation, we propose a new projection method that incorporates inner orbitals into the optimization of virtual orbitals in COVOS, and we present our algorithm in Figure 1. This method is inspired by the treatment of the FCI basis in OptOrb.[23,47,48] The same strategy can be extended to orbital optimization. Specifically, we project the N occupied orbitals onto t $(t \leq N)$ auxiliary orbitals $[\psi'_1, \psi'_2, \cdots, \psi'_t]$ using a projection matrix $U \in \mathbb{R}^{N\times t}$, where t represents the adjustable size of the active space. These auxiliary orbitals are then used to optimize the virtual orbitals

$$[\psi_1, \psi_2, \psi_3, \ldots, \psi_N] \times U = [\psi'_1, \psi'_2, \ldots, \psi'_t] \tag{7}$$

The coefficients of the U matrix are determined by the energy difference generated when the initial parameters of virtual orbital $\tilde{\psi}_v$ interact with each occupied orbital under an FCI operation ψ_i for a 2-electron system with 2 spatial orbitals. For this system, this value corresponds to the difference between the FCI energy and the HF energy, $E_{FCI} - E_{HF}$. We can refer to it as the correlation strength[49] and name it $E_c(i, v)$. This difference reflects the weight of the transition effect of each occupied orbital relative to the Kohn–Sham orbital. In the final FCI calculation of the overall ground state energy, E_c represents the total correlation energy of the system.[50] We can simply make the first column of U proportional to $E_c(i, v)$, that means $U(i, 1) = E_c(i, v)$, while the other columns are generated by various functions. The U matrix always ensures that $U^T U = I$. This is to ensure the orthogonality and normalization of $\psi'_1, \psi'_2, \cdots, \psi'_t$.

The auxiliary orbitals obtained via the U matrix, combined with the virtual orbital, result in $N + 1$ orbitals before rotation and $t + 1$ orbitals after rotation: $[\psi'_1, \psi'_2, \ldots, \psi'_t, \tilde{\psi}_v]$. In actual calculations, we prefer to use U^{alg} for algorithm computation, which makes the computational logic clearer. These $t + 1$ orbitals can be expressed using a rotation matrix $U^{alg} \in \mathbb{R}^{(N+1)\times(t+1)}$ as

$$U^{alg} = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \tag{8}$$

$$[\psi_1, \psi_2, \psi_3, \ldots, \psi_N, \tilde{\psi}_v] \times U^{alg} = [\psi'_1, \psi'_2, \ldots, \psi'_t, \tilde{\psi}_v] \tag{9}$$

At this stage, the generated auxiliary orbitals and the virtual orbital are subjected to an FCI calculation involving $2t$ electrons and $t + 1$ orbitals. The gradient is then computed based on the FCI results, as previously described, to optimize the virtual orbital.

Since the introduction of rotation matrix U requires its optimization, a coupled iterative strategy is employed. For a given $\tilde{\psi}_{vn}$, the rotation matrix U is iterated to convergence to minimize the E_c with $t + 1$ orbitals and $2t$ electrons. Subsequently, U is kept fixed while the virtual orbital $\tilde{\psi}_{vn}$ is iteratively optimized to minimize the FCI energy for the same condition, this process continues until convergence. After the virtual orbital $\tilde{\psi}_{vn}$ is finalized through this iterative process, it is kept fixed while a new iteration is performed to calculate U. This coupled iteration proceeds alternately until convergence.

In summary, the improved COVOS algorithm has modified the previous steps 3 and 4. And we obtain the following novel COVOS algorithm.

1 The rotation matrix U is initialized.

2The virtual orbital $\tilde{\psi}_{vn}$ is interacted with the t auxiliary orbitals $\psi'_1, \psi'_2, ..., \psi'_t$, which originate from the interaction of the rotation matrix U with the occupied orbitals. That forms a CI matrix for a $2t$-electron system with $t + 1$ spatial orbitals. While keeping $\tilde{\psi}_{vn}$ unchanged, it is iterated to optimize the rotation matrix U to minimize E_c calculated from the $2t$-electron system with $t + 1$ spatial orbitals. The optimization method chosen is gradient descent, continuing until the convergence criteria for U are met. In actual calculations, we prefer to use U^{alg} for computation, which makes the computational logic clearer.

3 $\tilde{\psi}_{vn}$ is optimized using the gradient descent method to minimize the ground state energy obtained from the diagonalization of the CI matrix in step 2, ensuring that the orbital remains orthogonal to all occupied orbitals ψ_1, ψ_2, ..., ψ_N and the previously generated virtual orbitals $\psi_{\mathrm{v}1}$, $\psi_{\mathrm{v}2}$, ..., $\psi_{\mathrm{v}(n-1)}$. The process continues until the optimization met the convergence criteria, and the optimization for $\tilde{\psi}_{vn}$ is finalized.

4 If the coupled iterations reach the convergence criteria, the optimization of $\psi_{\mathrm{v}n}$ is terminated. Otherwise, Step 2 is revisited.

For a given rotation matrix U that produces U^{alg}, the one- and two-electron integrals in the iterative process can be expressed in terms of the one- and two-electron integrals of all occupied and virtual orbitals. Specifically

$$\langle p'|q'\rangle = \sum_{pq}^{N+1} U^{\mathrm{alg}}_{pp'} U^{\mathrm{alg}}_{qq'} \langle p|q\rangle \tag{10}$$

$$\langle p'q'|r's'\rangle = \sum_{pqrs}^{N+1} U^{\mathrm{alg}}_{pp'} U^{\mathrm{alg}}_{qq'} U^{\mathrm{alg}}_{rr'} U^{\mathrm{alg}}_{ss'} \langle pq|rs\rangle \tag{11}$$

where the $\langle pq|rs\rangle$ and $\langle p|q\rangle$ represents the one- and two-electron integrals before projection, while $\langle p'q'|r's'\rangle$ and $\langle p'|q'\rangle$ represents those after projection.

$$\begin{aligned}\frac{\partial E_{\mathrm{c}}}{\partial U} &= \frac{\partial \boldsymbol{x}^T H_{\mathrm{CI}} \boldsymbol{x}}{\partial U} - \frac{\partial E_{\mathrm{HF}}}{\partial U} \\ &= \left\langle \psi_{\mathrm{CI}} \left| \frac{\partial \hat{H}}{\partial U} \right| \psi_{\mathrm{CI}} \right\rangle - \frac{\partial E_{\mathrm{HF}}}{\partial U} \\ &= \sum_{p'q'} c_{p'q'} \frac{\partial \langle p'|q'\rangle}{\partial U} + \sum_{p'q'r's'} c_{p'q'r's'} \frac{\partial \langle p'q'|r's'\rangle}{\partial U}\end{aligned} \tag{12}$$

where $c_{p'q'}$ and $c_{p'q'r's'}$ are the 1-RDM and 2-RDM of $\boldsymbol{x}^T H_{\mathrm{CI}} \boldsymbol{x} - E_{\mathrm{HF}}$.

Using these expressions, $\frac{\partial E_{\mathrm{c}}}{\partial U}$ can be analytically computed based on the values of $\langle pq|rs\rangle$ and $\langle p|q\rangle$.

The pseudocode for the iterative process of $\psi_{\mathrm{v}n}$ in the improved COVOS is shown in Algorithm 2.

Through this approach, we achieve arbitrary extension of the active orbital space in COVOS and enhance the sampling of inner orbitals rather than limit the process to the outermost orbitals. This improvement has significant advantages in many-body systems, where it provides a substantial boost in computational accuracy, thereby enabling us to use fewer orbitals to save computational time for FCI.

4. RESULTS AND DISCUSSION

4.1. Algorithm Convergence Test and Potential Energy Surface.

Since plane-wave systems are more suitable for periodic calculations, the Na_2 system serves as an excellent example for our study. Here, the Na_2 we use is the same as the Li_2 mentioned later and is actually chosen as a periodic system. Its primitive cell is a rectangular structure, with two sodium atoms located at the fractional coordinates (0,0,0) and $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ sites. The LiH system also meets periodic conditions, and its basic supercell contains four Li atoms and four H atoms. When calculating periodic systems within the plane-wave framework, it is only necessary to use the standard unit cell structure. However, for nonperiodic molecular systems, we often need to add vacuum layers to construct approximate nonperiodic boundary conditions. In addition, to reduce computational complexity, the pseudopotential files we use contain the frozen core approximation. In the results presented below, unless otherwise specified, the unit of energy is Hartree, and we will use "COVOSn" to represent FCI energy using n COVOS orbitals. We test only closed shells without considering spin. In most of the tests we conducted, since we set $t = 1$ (we will analyze the rationality of this approach in the subsequent calculations), the memory and time required for the COVOS optimization of orbitals can be negligible compared to the final FCI calculation, and the FCI solver is chosen as CDFCI.[51,52]

Algorithm 2 The iterative process of ψ_{vn} in the improved COVOS

```
Input: Initial orbitals $\tilde{\psi}_{vn}$, occupied orbitals $\psi_1, \psi_2, \ldots, \psi_N$ and the previously generated virtual orbitals $\psi_{v1}, \psi_{v2}, \ldots, \psi_{v(n-1)}$, learning rate $\alpha_1, \alpha_2$, convergence threshold $\epsilon, \epsilon_1, \epsilon_2$, the adjustable size $t$
Output: Optimized orbitals $\psi_{vn}$
Construct a set $\mathcal{S} = [\psi_1, \psi_2, \ldots, \psi_N, \psi_{v1}, \psi_{v2}, \ldots, \psi_{v(n-1)}]$
Initialize rotation matrix $U$, iteration counter $k \leftarrow 0$, error $error \leftarrow \infty$
while $error > \epsilon$ do
    Orthogonalize and normalize $\tilde{\psi}_{vn}^{k}$ with respect to all orbitals in $\mathcal{S}$
    Initialize iteration counter $k_1 \leftarrow 0$, error $error_1 \leftarrow \infty$
    while $error_1 > \epsilon_1$ do
        Compute $U^{\mathrm{alg},k_1}$ and auxiliary orbitals $\psi'$
        Construct a set $\mathcal{S}' = [\psi'_1, \psi'_2, \ldots, \psi'_t]$
        Construct $H_{\mathrm{CI}}$ for $2t$-electron system with $t+1$ spatial orbitals $\tilde{\psi}_{vn}$ and $S'$
        Compute gradient $\frac{\partial E_c}{\partial U^{k_1}}$
        Update parameters: $U^{k_1+1} \leftarrow U^{k_1} - \alpha_1 \frac{\partial E_c}{\partial U^{k_1}}$
        Orthogonalize and normalize $U^{k_1+1}$ itself
        Compute new error: $error_1 \leftarrow \|U^{k_1} - U^{k_1+1}\|^2$
        Increment iteration counter: $k_1 \leftarrow k_1 + 1$
    end while
    Initialize iteration counter $k_2 \leftarrow 0$, error $error_2 \leftarrow \infty$
    while $error_2 > \epsilon_2$ do
        Orthogonalize and normalize $\tilde{\psi}_{vn}^{k_2}$ with respect to all orbitals in $\mathcal{S}$
        Construct $H_{\mathrm{CI}}$ for $2t$-electron system with $t+1$ spatial orbitals $\tilde{\psi}_{vn}^{k_2}$ and $\mathcal{S}'$
        Compute gradient $\frac{\partial E_{\mathrm{CI}}}{\partial \tilde{\psi}_{vn}^{k_2}}$
        Update parameters: $\tilde{\psi}_{vn}^{k_2+1} \leftarrow \tilde{\psi}_{vn}^{k_2} - \alpha \frac{\partial E_{\mathrm{CI}}}{\partial \tilde{\psi}_{vn}^{k_2}}$
        Compute new error: $error_2 \leftarrow \|\tilde{\psi}_{vn}^{k_2} - \tilde{\psi}_{vn}^{k_2+1}\|^2$
        Increment iteration counter: $k_2 \leftarrow k_2 + 1$
    end while
    Compute new error: $error \leftarrow \|\tilde{\psi}_{vn}^{k} - \tilde{\psi}_{vn}^{k+1}\|^2$
    Increment iteration counter: $k \leftarrow k + 1$
end while
Orthogonalize and normalize $\tilde{\psi}_{vn}$ with respect to all orbitals in $\mathcal{S}$
return Optimized orbitals $\psi_{vn}$
```

First, we performed a full-basis FCI calculation (96 orbitals in total) for the Na_2 system on a coarse grid as a benchmark, which means we only used a plane-wave basis set of 96. We also evaluated the convergence of COVOS orbitals and compared them with Kohn−Sham (KS) orbitals generated by DFT,[53] as shown in Figure 2. It can be observed that COVOS achieves chemical accuracy rapidly with only 10 orbitals and that the FCI results based on the COVOS orbitals outperform those using KS orbitals for any number of orbitals. Both approaches eventually converge to the same result when the full orbital configuration is considered.

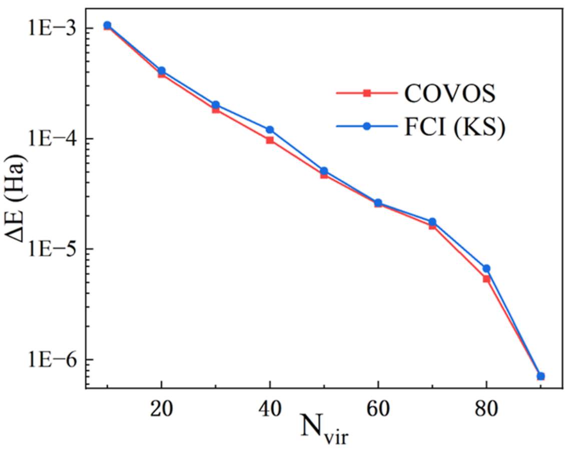


Figure 2. Error with respect to the actual energy (COVOS95) ΔE as a function of the number of virtual orbitals for Na_2, where N_{vir} is the number of selected virtual orbitals. The COVOS orbitals show stronger convergence than the KS orbitals, and both eventually agree when all orbitals are included.

For calculations on a regular grid, the results for different energy cutoffs (Ecut)[54] and numbers of orbitals are shown in Table 1. The energy values of the Na_2 unit cell converge around Ecut = 15. The virtual orbital count also converges to chemical accuracy at approximately 40 orbitals, as indicated by the trend. Compared with DFT results, the energy is significantly reduced. Remarkably, with an original basis dimension of approximately 20,000, the COVOS method achieves a superior correlation energy approximation with only 40 orbitals. Additionally, we calculate the energy variations with respect to interatomic distances near the equilibrium positions for Na_2 and Li_2, which is a part of calculating potential energy surface (PES), as well as the energy changes in the supercells. The equilibrium point results shown in Figure 5 demonstrate excellent symmetry along the equilibrium positions due to the periodic nature of Na_2 and Li_2. The trends of the three curves are almost identical, with the energies calculated by COVOS being overall lower than those from DFT, which is an expected outcome.

4.2. Improvement Validation. The electronic structure of HCl, comprising hydrogen's single electron and chlorine's seven valence electrons, is analyzed using four computational methods in Figure 3. From the perspective of molecular orbital theory, the improved accuracy across panels arises from advancements in treating electron correlation and orbital hybridization.

While DFT captures the overall electron density distribution, its reliance on approximate exchange−correlation functionals leads to slight overdelocalization in the bonding region between H and Cl, as well as insufficient resolution of chlorine's lone pairs. Incorporating FCI with KS orbitals improves electron correlation effects, yielding a sharper lone pair localization on chlorine. However, computational limitations restrict the basis set's completeness, leaving minor artifacts near the H−Cl bond. The evolution of the COVOS highlights pivotal advancements in orbital optimization strategies. In the original COVOS framework, valence-specific orbital optimization is achieved by iteratively refining virtual orbitals derived exclusively from the outermost valence orbitals. While this approach enhances electron density accumulation in the H−Cl covalent bond region compared to DFT or FCI(KS), its restriction to outer valence orbitals introduces residual asymmetry in the lone pair lobes of chlorine. This limitation likely stems from neglecting the interplay between the core and valence orbitals, which subtly modulates electron correlation and hybridization. The improved COVOS method addresses this by incorporating a holistic orbital optimization scheme that accounts for the superpositional effects of all orbitals—core and valence—during virtual orbital construction. By integrating dynamic correlation corrections and employing an expanded basis set, this refined

Table 1. Ground State Energy for Different Energy Cutoffs (Ecut) and Methods[a]

Ecut	DFT	HF	COVOS4	COVOS16	COVOS30	COVOS40	**COVOS70**
10	−0.619861	−0.614642	−0.634784	−0.637169	−0.637515	−0.637607	−0.637697
15	−0.619862	−0.614658	−0.634790	−0.637190	−0.637536	−0.637627	−0.637718
20	−0.619862	−0.614662	−0.634807	−0.637190	−0.637537	−0.637627	−0.637718

[a] HF represents the Hartree−Fock energy, using Hartree as the unit. COVOS*n* represent calculations using *n* COVOS orbitals. The energy values of the Na_2 unit cell converge around Ecut = 15 and 40 COVOS orbitals.

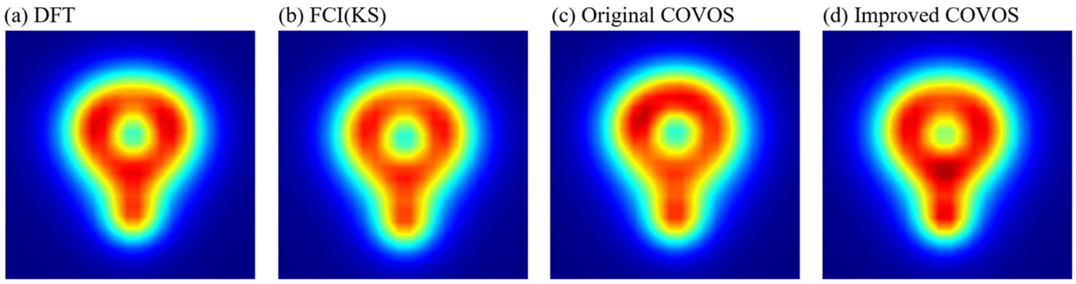


Figure 3. Electron density plots of the HCl molecule, reflecting contributions from the hydrogen atom (one electron) and chlorine's valence electrons (seven electrons), computed using four distinct methods: (a) DFT, (b) FCI with KS orbitals (FCI(KS)), (c) original COVOS, and (d) improved COVOS. Progressive refinement in electron density distribution is observed from left to right, reflecting enhanced accuracy in capturing covalent bonding and lone pair characteristics.

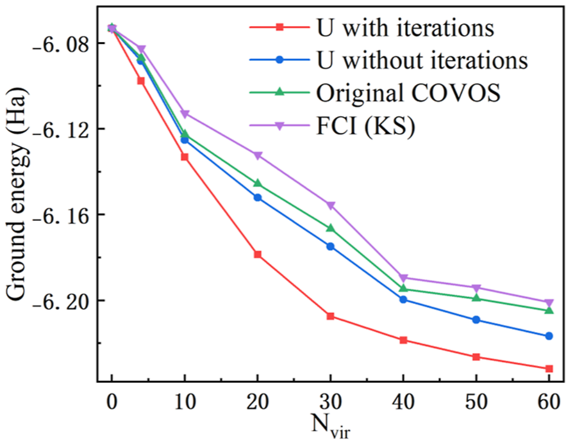


Figure 4. Comparison of the FCI energy for the SiH_4 system using virtual orbitals obtained through different methods is presented, where N_{vir} represents the number of virtual orbitals. The data in the figure indicate that the introduction of the rotation matrix U in the iteration process has a significant effect. The optimized COVOS basis set exhibits a smoother convergence curve and performs better in FCI calculations.

algorithm achieves a more balanced representation of electron density. The H−Cl bond region now exhibits a smoother density gradient, reflecting enhanced covalent character due to improved orbital overlap. Simultaneously, chlorine's lone pairs adopt symmetric, sp^3-hybridized geometries, a consequence of including core−valence polarization effects and interorbital correlation. These features align closely with theoretical predictions for heteronuclear diatomics. The contrast between the two COVOS variants underscores the importance of comprehensive orbital interaction modeling. While the original method prioritizes computational simplicity by focusing on valence orbitals, the improved version demonstrates that incorporating multiorbital superposition is critical for resolving subtle electronic features, such as lone pair symmetry and bond polarization. This advancement positions COVOS as a versatile tool for studying systems where both accuracy and efficiency are paramount, particularly in molecules with pronounced core−valence coupling.

When we select one auxiliary orbital, Figure 4 illustrates the role of the U matrix iteration in the COVOS algorithm. We compare the performance of the SiH_4 system under different scenarios: iterating the U matrix, not iterating the U matrix, the original COVOS algorithm, and the KS orbitals in FCI calculations. The results show that the virtual orbitals obtained through the coupled iterative active orbital approach significantly outperforms the initial COVOS algorithm. Even results calculated using only the initial rotation matrix U^0 exhibit substantial improvements compared with the original COVOS algorithm, particularly when the number of virtual orbitals increases.

4.3. Supercell Test and MP2 Comparison.

Supercell expansion increases the simulation system size, effectively approximating the behavior of the actual physical system.[55] This is particularly important for systems with long-range interactions, such as materials with surface adsorption, as it

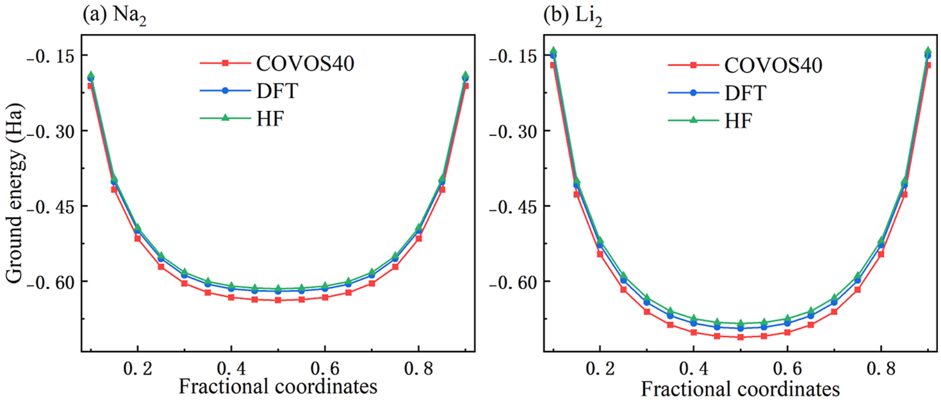

(a) Na_2
(b) Li_2

Figure 5. Ground state energy variations with interatomic distances for Na_2 and Li_2 in periodic conditions. (a) Na_2 system. (b) Li_2 system. We place one atom in the unit cell at the origin, and the other atom at the position on the diagonal where the fractional coordinate is located. When the fractional coordinate is $\frac{1}{2}$, the position is $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$. They demonstrate excellent symmetry along the equilibrium positions.

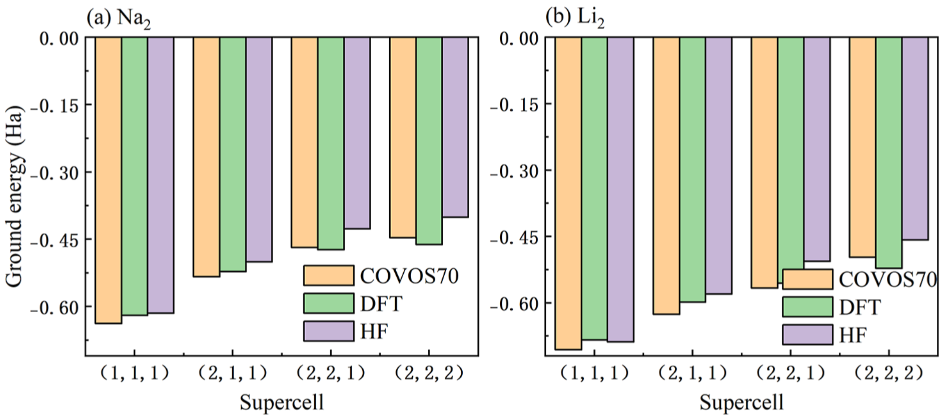


Figure 6. Ground state energy variations after supercell expansion for Na_2 and Li_2 in periodic conditions. (a) Na_2 system. (b) Li_2 system. The *x*-axis represents the degree of supercell expansion. Supercell calculations allow us to obtain more accurate results.

Table 2. Comparison of FCI Energy and MP2 Energy in Na_2 and LiH Systems, Using Hartree as the Unit[a]

Molecule	HF	COVOS4	COVOS16	COVOS30	COVOS40	COVOS50
Na_2(FCI)	−0.614658	−0.634790	−0.637190	−0.637536	−0.637627	−0.637668
Na_2(MP2)	−0.614658	−0.637221	−0.643007	−0.644422	−0.644815	−0.645020
LiH(FCI)	−3.071960	−3.110087	−3.158765	−3.194827	−3.202727	−3.211777
LiH(MP2)	−3.071960	−3.105716	−3.191571	−3.238124	−3.255190	−3.258975

[a] We found that, with the same orbitals, the energy obtained from MP2 calculations is lower than the FCI energy, indicating that MP2's correlation energy calculation is inaccurate.

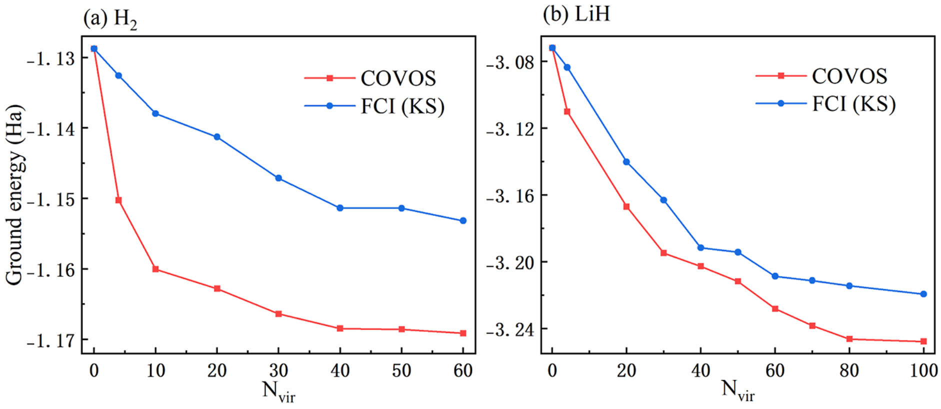


Figure 7. FCI energy variations for H_2 and LiH systems as a function of the number of COVOS and KS orbitals. (a) H_2 system: FCI energy plotted against the number of virtual orbitals N_{vir}. (b) LiH system: FCI energy plotted under the same conditions. The results show that COVOS orbitals perform significantly better than KS orbitals in FCI calculations.

ensures that the calculated properties more closely reflect those of the real-world system by minimizing artifacts introduced by periodic boundary conditions. In plane-wave basis set calculations, *k*-point sampling[56] significantly influences the electronic structure of the system. Expanding the supercell increases the density of sampling in the Brillouin zone, allowing for more accurate electronic structure calculations and thereby improving the computational precision. The energy variations in the supercells are illustrated in Figure 6.

Owing to the prohibitive computational cost of FCI, calculations on larger crystalline systems remain intractable. Consequently, this study focuses on small systems, such as Na_2 and Li_2, to ensure computational feasibility while retaining critical insights into the electronic structure behavior.

During the process of expanding the supercell, the unit cell energy continuously increases. This indicates that even under a periodic plane-wave basis set, using a Gamma-point unit cell calculation is inaccurate. Although the overall trend of the calculations remains consistent, the expansion of the supercell significantly increases the grid size, leading to a substantial increase in the number of plane-wave basis functions. Consequently, the 70 COVOS orbitals can no longer fully capture all of the correlation effects as effectively as they do in smaller supercells. This results in a weaker correlation energy. However, in terms of overall trends, the energies calculated by COVOS, DFT, and HF remain similar.

In the FCI calculations for the Na_2 and LiH systems, we also perform MP2 calculations. The results in Table 2 show that the MP2 energy is slightly lower than the FCI energy, which we attribute to the nonvariational nature of the MP2 method. Unlike variational methods that ensure an upper bound to the true ground-state energy, MP2 can yield lower energies due to the approximation involved in its perturbative approach.[57] Due to the presence of the Fermi surface, the density of electronic

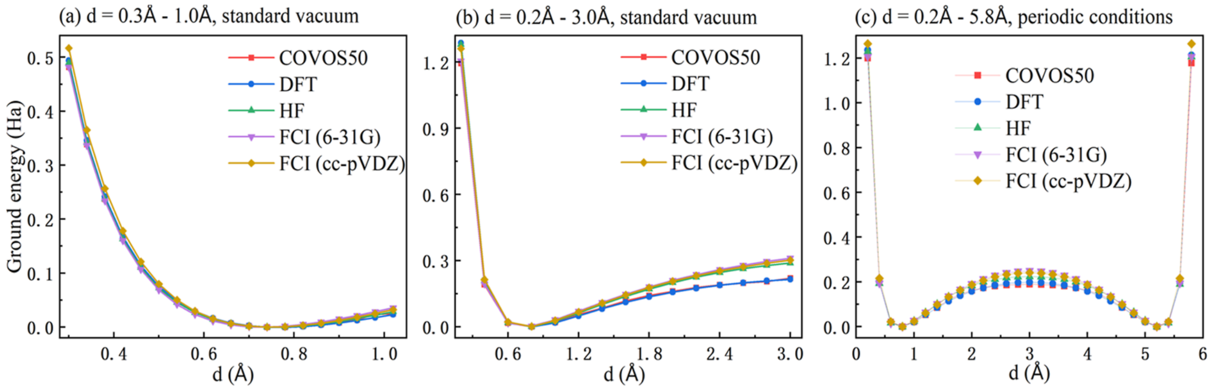


Figure 8. Dissociation energy curves for H_2 are shown, where d represents the interatomic distance of hydrogen atoms. (a) Short-range dissociation energy curves compared with the equilibrium position at $d = 0.74$ Å set as the zero-point energy. (b) Long-range dissociation energy curve comparison. (c) Reduced cell size tested for energy periodicity in a periodically ordered hydrogen chain system. Due to the systematic deviations in calculation results caused by the pseudopotentials used in the plane-wave basis set compared to Gaussian basis sets, we set the energy at $d = 0.74$ Å by each method and basis set as the energy zero point in (a−c).The figure shows that at smaller distances, both COVOS and DFT agree well with the commonly used FCI basis set. At larger distances, due to the periodicity of plane waves, unavoidable differences arise. When we reduce the supercell size and adjust the hydrogen atom spacing, we observed similar curves to those of Na and Li systems. The differences originate from the distinction between nonmetallic and metallic systems.

states near the Fermi surface is extremely high, leading to energy differences between occupied and unoccupied states approaching zero, which causes divergence in the MP2 method.

MP2 relies on perturbation theory, treating electron correlation effects as a second-order correction to the Hartree−Fock (HF) ground state. However, this approximation is inadequate for metallic systems due to the complex nature of electronic correlations and the unique behavior of delocalized electrons near the Fermi surface, which cannot be fully captured by the MP2 method.[13] The numerous nearly degenerate excited states near the Fermi surface cause the second-order perturbation corrections to be excessively amplified, resulting in significantly increased correlation energy values and an unexpectedly low total energy.

Compared with more accurate many-body methods, MP2 is only an approximation, and its perturbative corrections near the Fermi surface often exaggerate correlation effects. On the other hand, FCI can account for all electron correlations, providing a more accurate solution and avoiding the inaccuracies caused by divergence in MP2. Consequently, in metallic systems, the total energy calculated by MP2 may even be lower than the FCI energy, highlighting the limitations of MP2 in handling electron correlations in metallic or quasi-metallic systems.

4.4. Dissociation Energy Curves for H_2. For smaller molecular systems such as H_2, calculations with $C = 10$ Å and Ecut = 15 are shown in Figure 7, where we compare the results with those obtained using KS orbitals, where C is the length of the supercell. Since the initial guess of COVOS orbitals is derived from KS orbitals, this has an impact on the optimization process. The results indicate that the energy reduction achieved by COVOS orbitals is partially correlated with KS orbitals, but overall, COVOS orbitals demonstrate significantly superior performance, especially with fewer orbitals, which is critical for reducing the computational cost of FCI. With 60 orbitals, the FCI energy is nearly converged. Dissociation energy calculations for the H_2 system[58] are shown in Figure 8. Compared to the cc-pVDZ and 6−31G basis sets, at smaller distances (around 1.0 Å), the dissociation energy curves are nearly identical due to the sufficiently large vacuum layer in the periodic plane-wave system. At larger distances (1.0−3.0 Å), the periodicity implies the presence of equivalent hydrogen atoms at the supercell boundary, causing the plane-wave dissociation energy curve to exhibit slightly lower energies at long-range. In this range, the dissociation energy curve obtained with COVOS is nearly identical with the DFT results. To address this, we test a supercell size of 6.0 Å, calculating the energy variations with interatomic distances in a fully periodic system. The trends of the COVOS and DFT results remain consistent. In the periodic systems of Na_2 and H_2, the dissociation curves exhibit distinct shapes due to fundamental differences in their bonding characteristics. The U-shaped curve observed in Na_2 and Li_2 corresponds to metallic bonding, where the interaction between atoms is dominated by the delocalization of the valence electrons. This results in a balance between attractive and repulsive forces, leading to a stable minimum in the potential energy as a function of interatomic distance.[59] On the other hand, the W-shaped dissociation curve of H_2 reflects the nature of covalent bonding. When the interatomic distance exceeds the equilibrium bond length, the energy increase is primarily due to breaking of the covalent bond, which disrupts the shared electron pair that holds the two hydrogen atoms together. The relatively small peak in the W-shaped curve is a consequence of the periodic boundary conditions imposed on the system and not a result of dominant repulsive interactions. These distinctions highlight the fundamentally different bonding mechanisms in metallic systems like Na_2 and covalent systems like H_2.[60]

In Figure 7, we also perform calculations on the LiH system, obtaining results similar to those of H_2, where the FCI energy results approached convergence at 80 orbitals with the COVOS.

Notably, in this calculation, four occupied orbitals are projected onto one auxiliary orbital. When four virtual orbitals are used, the FCI energy obtained with COVOS orbitals optimized using all occupied orbitals was −3.12062, while the energy with one projected orbital was −3.11009. At 70 virtual orbitals, the FCI energy using all occupied orbitals was −3.23900, and with projection, it was −3.23826. This demonstrates that for larger systems with a high number of virtual orbitals, projection-based optimization can significantly improve computational efficiency during orbital optimization (by a factor of $O((N+1)^4)$) while maintaining the final FCI energy within chemical accuracy. Including more auxiliary orbitals in the iterations improves computational accuracy, particularly when the total number of virtual orbitals is small. However, increasing the number of auxiliary orbitals significantly raises the complexity of optimizing virtual orbitals. To save computational costs, we aim to include as few auxiliary orbitals as possible in the calculations.

4.5. Atomization Energy Test and RPA Comparison.

We also calculate the atomization energy for some molecular systems and compared them with DFT and experimental values,[61,62] as shown in Table 3.

Table 3. Comparison of the Atomization Energy Differences between FCI Calculations with COVOS Orbital, DFT, and Experiment, Using Hartree as the Unit[a]

molecule	DFT	COVOS70	experiment
C_2H_4	553.6	561.1	562.4
CS	195.6	175.6	171.2
CH_4	404.8	417.1	420.1
C_2H_6	693.6	698.0	717.2
Cl_2	69.0	48.7	58.0
SiH_4	303.1	302.7	322.0
HCl	99.9	100.6	106.4

[a]The experimental reference results are taken from.[61] The unit is kcal/mol. The results provided by COVOS are relatively close to the experimental values.

It should be emphasized that COVOS employs a plane-wave basis set, which is not inherently optimized for isolated molecular systems. The primary utility of applying COVOS to molecular systems resides in validating its precision under scenarios where testing on large periodic systems—such as extended solids or surfaces—faces significant computational challenges. This approach ensures rigorous benchmarking of the method's accuracy before scaling to complex periodic frameworks.

Overall, COVOS demonstrates higher accuracy than DFT, which is greatly related to the accurate calculation of correlation energy by the FCI algorithm. At the same time, for the ground state energy calculations of these molecular systems, the comparison of COVOS orbitals with DFT, KS orbitals, and the results obtained using KS orbitals for RPA calculations are shown in Table 4. We can compare and observe that the performance of the COVOS orbitals in FCI calculations is significantly superior to that of the KS orbitals. Using 70 KS orbitals, the RPA energy is found to be lower than the FCI energy, indicating that the RPA likely overestimates the correlation energy of the system.

Table 4. Comparison of the Energy Differences between DFT, FCI Calculations with COVOS Orbital and KS Orbital, and RPA, Using Hartree as the Unit[a]

molecule	DFT	COVOS70	FCI(KS70)	RPA(KS70)
C_2H_4	−13.6148	−13.6268	−13.4491	−13.5480
CS	−15.8023	−15.7703	−15.7163	−15.8184
CH_4	−8.0094	−8.0290	−7.9107	−7.9775
C_2H_6	−14.8358	−14.8429	−14.6528	−14.7718
Cl_2	−30.1217	−30.0893	−30.0025	−30.1283
SiH_4	−6.2392	−6.2385	−6.2085	−6.3067
HCl	−15.6640	−15.6651	−15.6049	−15.6795

[a]KS70 represent calculations using 70 KS orbitals. By comparing the FCI results of the KS orbitals with the RPA results, we found that RPA may overestimate the correlation energy of the system.

The RPA is a method derived from many-body perturbation theory and linear response theory. It computes the correlation energy by integrating over the screened Coulomb interaction within the framework of the Kohn−Sham orbitals. RPA is particularly effective for capturing long-range correlation effects, such as van der Waals interactions, and is free from self-interaction errors in the exchange−correlation energy.[63] However, its treatment of correlation is limited to the summation of ring diagrams.

The observed overestimation of the correlation energy by RPA can be attributed to several factors. First, the reliance on KS orbitals, which are approximate solutions to the electronic structure, can lead to inaccuracies in describing the virtual states. The quality of these orbitals significantly influences the computed RPA correlation energy. Second, RPA's inability to account for beyond-ring-diagram interactions inherently limits its accuracy, particularly in systems with strong correlation effects. Finally, the frequency-dependent integration required in RPA calculations is prone to numerical approximations and truncations, which can further increase the estimated correlation energy.

5. CONCLUSION

In this study, we investigate the problem of applying a plane-wave basis set to FCI calculations, aiming to achieve a higher FCI accuracy with as few virtual orbitals as possible. To this end, we improve the COVOS algorithm by expanding the active space optimization process to generate more effective virtual orbitals, and they perform excellently in all types of calculations. During this process, we primarily address two challenges. The first challenge is ensuring that the selected active space is effective. For this purpose, we choose E_c as the convergence criterion and adopt a gradient descent approach to iteratively select the active space. This is coupled with the virtual orbital iteration process in an iterative manner. The second challenge is balancing the computational cost and accuracy. While a larger active space can improve orbital selection, we found that for many-body calculations, using an active space with only one occupied orbital yields satisfactory results. This is because, when the number of virtual orbitals is large, the iterative active space steps account for transitions involving all occupied orbitals, fully leveraging the properties of these orbitals—something that the original COVOS algorithm could not achieve. This allows us to optimize more dominantly contributing orbitals without significantly increasing the computational complexity. Based on this approach, we conduct larger-scale calculations and compare the performance with other methods to explore the limits of plane-wave FCI calculations.

In the numerical experiments, we first perform FCI energy calculations for coarse systems to study the convergence behavior of COVOS orbitals with respect to the number of orbitals. We compare the performance of the new COVOS orbitals with that of the original COVOS and KS orbitals.

As visualized in the electron density slice plots, the improved COVOS achieves superior orbital localization, which aligns with classical chemical motifs (e.g., Lewis structures). This synergy enhances interpretability for bonding, charge transfer, and lone pairs, directly bridging computational accuracy with chemical intuition—critical for education and AI-driven molecular design.

https://doi.org/10.1021/acs.jctc.5c00586

The numerical results demonstrate the effectiveness of our work, as FCI energy calculations using our new COVOS orbitals show improvements exceeding chemical accuracy compared to the original COVOS and KS orbitals at the same number of orbitals. Therefore, even the initial choice of altering the active space, which barely increases the computational load, can significantly enhance the performance of the COVOS orbitals. Moreover, the strategy of coupled iterations has enabled us to obtain excellent virtual orbitals that are worth investigating. Subsequently, we examine periodic systems such as Na_2, Li_2, and LiH, computing FCI energies under various parameters, including a part of PES calculations and supercell expansion, and compare the results with DFT energies. In comparison with MP2 energies, we identify the divergence issue in periodic metallic systems. For such systems, MP2 fails to accurately calculate the total energy, yielding ground state energies that are lower than the true values. We also perform calculations on molecular systems such as H_2, SiH_4, and so on, including dissociation energy calculations for H_2, which are compared with results from cc-pVDZ, 6−31G, and DFT methods, validating the correctness of our calculations. Finally, we conduct calculations related to atomization energies, and the results indicate that COVOS's calculations of atomization energies are closer to experimental values compared to DFT. When comparing energies with RPA, we found that RPA may overestimate the correlation energy of the system. Sometimes we need to be more cautious about the results of RPA, which also emphasize the benchmark role of FCI.

In conclusion, the improved COVOS algorithm can optimize better-performing virtual orbitals, enabling us to use the FCI method with plane-wave basis sets to explore periodic system problems more effectively.

The COVOS algorithm presents several potential research directions. In the current implementation, we only consider the Γ-point, while one of the most prominent features of plane waves is the ability to perform calculations equivalent to periodic supercell expansion by incorporating *k*-points. For FCI, supercell expansion doubles both the number of occupied electrons and the total number of orbitals, while the Hamiltonian matrix dimension for a closed-shell system scales as $(C_N^m)^2$, where N is the number of spatial orbitals, and m is the number of electrons divided by two. This exponential increase in computational complexity makes supercell expansion in FCI infeasible. However, *k*-points can circumvent this issue by providing more precise sampling in the reciprocal space to obtain accurate results for periodic systems. Therefore, developing a COVOS algorithm with a *k*-point is an important future research direction.

Additionally, for iterative orbitals and rotation matrices, the current algorithm's use of coupled iteration may not be the most efficient approach. The criteria used to evaluate the orbital and active space performance could also be improved. Optimizing the iterative process could further enhance the performance of the COVOS orbitals. Beyond this, other aspects of wave function-based methods with plane waves, such as the impact of different singularity treatments on FCI energy, the choice of pseudopotentials, and other corrections, are also worthy of investigation.

Finally, we can also explore the role of the COVOS algorithm in other theories. For example, it can provide more accurate calculations for the impurity solver in dynamical mean field theory (DMFT)[64] or optimize the orbitals used in algorithms such as RPA and MP2 with similar ideas, thereby facilitating a wider range of research directions.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.5c00586.

> Construction of the Hamiltonian under plane-wave basis sets, derivation of one-electron integrals and two-electron integrals with their gradient formulas, and description of the singularity treatment scheme for two-electron integrals in computational implementation (PDF)

AUTHOR INFORMATION

Corresponding Authors

Yingzhou Li − *School of Mathematical Sciences, Fudan University, Shanghai 200433, China;* orcid.org/0000-0003-1852-3750; Email: yingzhouli@fudan.edu.cn

Jinlong Yang − *State Key Laboratory of Precision and Intelligent Chemistry, and Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, China;* orcid.org/0000-0002-5651-5340; Email: jlyang@ustc.edu.cn

Wei Hu − *School of Future Technology, Hefei National Research Center for Physical Sciences at the Microscale, and Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230026, China;* orcid.org/0000-0001-9629-2121; Email: whuustc@ustc.edu.cn

Authors

Mingyu Qiu − *School of Future Technology, Hefei National Research Center for Physical Sciences at the Microscale, and Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230026, China*

Zhenlin Zhang − *School of Future Technology, Hefei National Research Center for Physical Sciences at the Microscale, and Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230026, China*

Zhiyuan Zhang − *School of Future Technology, Hefei National Research Center for Physical Sciences at the Microscale, and Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230026, China*

Yexuan Lin − *School of Future Technology, Hefei National Research Center for Physical Sciences at the Microscale, and Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230026, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.5c00586

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is partly supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0450101), the Innovation Program for Quantum Science and Technology (2021ZD0303306), the National Natural Science Foundation of China (22288201, 22173093, 21688102), the Anhui Provincial Key Research and Development Program (2022a05020052), the National Key Research and Development Program of China (2016YFA0200604, 2021YFB0300600), the Network Information Project of

Chinese Academy of Sciences (CASWX2021SF-0103), the Hefei National Laboratory for Physical Sciences at the Microscale (KF2020003), the National Natural Science Foundation of China (22403024), the Anhui Provincial Natural Science Foundation (2308085QB52), the NSFC under grant 12271109, STCSM under grant 22TQ017 and 24DP2600100, and SIMIS under grant SIMIS-ID-2024-(CN). The authors thank the Hefei Advanced Computing Center, the Supercomputing Center of Chinese Academy of Sciences (Xiandao-1), the Supercomputing Center of USTC, and the National Supercomputing Center in Wuxi, Tianjin, Shanghai, and Guangzhou for the computational resources.

■ REFERENCES

(1) Eriksen, J. J. The shape of full configuration interaction to come. *J. Phys. Chem. Lett.* **2021**, *12*, 418−432.

(2) Olsen, J.; Jørgensen, P.; Simons, J. Passing the one-billion limit in full configuration-interaction (FCI) calculations. *Chem. Phys. Lett.* **1990**, *169*, 463−472.

(3) Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; d'Arco, P.; et al. CRYSTAL14: A program for the ab initio investigation of crystalline solids. *Int. J. Quantum Chem.* **2014**, *114*, 1287−1317.

(4) Schwarz, K.; Sideris, M.; Forsberg, R. The use of FFT techniques in physical geodesy. *Geophys. J. Int.* **1990**, *100*, 485−514.

(5) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169.

(6) Bylaska, E. J. *Annu. Rep. Comput. Chem.* **2017**, *13*, 185−228.

(7) Fonseca Guerra, C.; Snijders, J.; Te Velde, G. t.; Baerends, E. J. Towards an order-N DFT method. *Theor. Chem. Acc.* **1998**, *99*, 391−403.

(8) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T.; Joannopoulos, a. J. Iterative minimization techniques for ab initio total-energy calculations: Molecular dynamics and conjugate gradients. *Rev. Mod. Phys.* **1992**, *64*, 1045.

(9) Shepard, C.; Zhou, R.; Yost, D. C.; Yao, Y.; Kanai, Y. Simulating electronic excitation and dynamics with real-time propagation approach to TDDFT within plane-wave pseudopotential formulation. *J. Chem. Phys.* **2021**, *155*, 100901.

(10) Jollet, F.; Jomard, G.; Amadon, B.; Crocombette, J.; Torumba, D. Hybrid functional for correlated electrons in the projector augmented-wave formalism: Study of multiple minima for actinide oxides. *Phys. Rev. B* **2009**, *80*, 235109.

(11) Zhang, Z.; Yin, X.; Hu, W.; Yang, J. Machine learning k-means clustering of interpolative separable density fitting algorithm for accurate and efficient cubic-scaling exact exchange plus random phase approximation within plane waves. *J. Chem. Theory Comput.* **2024**, *20*, 1944−1961.

(12) Hubbard, J. The description of collective motions in terms of many-body perturbation theory. *Proc. R. Soc. London, A* **1957**, *240*, 539−560.

(13) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153*, 503−506.

(14) Rybkin, V. V.; VandeVondele, J. Spin-unrestricted second-order Møller−Plesset (MP2) forces for the condensed phase: From molecular radicals to F-centers in solids. *J. Chem. Theory Comput.* **2016**, *12*, 2214−2223.

(15) Gan, Z.; Grant, D. J.; Harrison, R. J.; Dixon, D. A. The lowest energy states of the group-IIIA−group-VA heteronuclear diatomics: BN, BP, AlN, and AlP from full configuration interaction calculations. *J. Chem. Phys.* **2006**, *125*, 124311.

(16) Coester, F.; Kümmel, H. Short-range correlations in nuclear wave functions. *Nucl. Phys.* **1960**, *17*, 477−485.

(17) Bishop, R. F.; Kümmel, H. G. The coupled cluster method. *Phys. Today* **1987**, *40*, 52−60.

(18) Booth, G. H.; Tsatsoulis, T.; Chan, G. K.; Grüneis, A. From plane waves to local Gaussians for the simulation of correlated periodic systems. *J. Chem. Phys.* **2016**, *145*, 084111.

(19) Bauschlicher, Jr C. W.; Taylor, P. R. Benchmark full configuration-interaction calculations on H2O, F, and F-. *J. Chem. Phys.* **1986**, *85*, 2779−2783.

(20) Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *J. Chem. Phys.* **2017**, *146*, 224104.

(21) Dagotto, E. Correlated electrons in high-temperature superconductors. *Rev. Mod. Phys.* **1994**, *66*, 763.

(22) Rontani, M.; Cavazzoni, C.; Bellucci, D.; Goldoni, G. Full configuration interaction approach to the few-electron problem in artificial atoms. *J. Chem. Phys.* **2006**, *124*, 124102.

(23) Li, Y.; Lu, J. Optimal orbital selection for full configuration interaction (OptOrbFCI): pursuing the basis set limit under a budget. *J. Chem. Theory Comput.* **2020**, *16*, 6207−6221.

(24) Eriksen, J. J.; Lipparini, F.; Gauss, J. Virtual orbital many-body expansions: A possible route towards the full configuration interaction limit. *J. Phys. Chem. Lett.* **2017**, *8*, 4633−4639.

(25) Hutter, J.; Lüthi, H. P.; Parrinello, M. Electronic structure optimization in plane-wave-based density functional calculations by direct inversion in the iterative subspace. *Comput. Mater. Sci.* **1994**, *2*, 244−248.

(26) Fales, B. S.; Seritan, S.; Settje, N. F.; Levine, B. G.; Koch, H.; Martínez, T. J. Large-scale electron correlation calculations: Rank-reduced full configuration interaction. *J. Chem. Theory Comput.* **2018**, *14*, 4139−4150.

(27) Lee, J.; Berry, D. W.; Gidney, C.; Huggins, W. J.; McClean, J. R.; Wiebe, N.; Babbush, R. Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX Quantum* **2021**, *2*, 030305.

(28) Berger, M. J.; Colella, P. Local adaptive mesh refinement for shock hydrodynamics. *J. Comput. Phys.* **1989**, *82*, 64−84.

(29) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

(30) Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* **1999**, *110*, 2810−2821.

(31) Bylaska, E. J.; Song, D.; Bauman, N. P.; Kowalski, K.; Claudino, D.; Humble, T. S. Quantum solvers for plane-wave hamiltonians: Abridging virtual spaces through the optimization of pairwise correlations. *Front. Chem.* **2021**, *9*, 603019.

(32) Song, D.; Bauman, N. P.; Prawiroatmodjo, G.; Peng, B.; Granade, C.; Rosso, K. M.; Low, G. H.; Roetteler, M.; Kowalski, K.; Bylaska, E. J. Periodic plane-wave electronic structure calculations on quantum computers. *Mater. Theory* **2023**, *7*, 2.

(33) Prendergast, D.; Louie, S. G. Bloch-state-based interpolation: An efficient generalization of the Shirley approach to interpolating electronic structure. *Phys. Rev. B* **2009**, *80*, 235126.

(34) Shirley, E. L. Optimal basis sets for detailed Brillouin-zone integrations. *Phys. Rev. B* **1996**, *54*, 16464.

(35) Yang, C.; Meza, J. C.; Lee, B.; Wang, L.-W. KSSOLV-a MATLAB toolbox for solving the Kohn-Sham equations. *ACM Trans. Math Software* **2009**, *36*, 1−35.

(36) Jiao, S.; Zhang, Z.; Wu, K.; Wan, L.; Ma, H.; Li, J.; Chen, S.; Qin, X.; Liu, J.; Ding, Z.; et al. KSSOLV 2.0: An efficient MATLAB toolbox for solving the Kohn-Sham equations with plane-wave basis set. *Comput. Phys. Commun.* **2022**, *279*, 108424.

(37) Bylaska, E. J.; Tsemekhman, K.; Baden, S. B.; Weare, J. H.; Jonsson, H. Parallel implementation of γ-point pseudopotential plane-wave DFT with exact exchange. *J. Comput. Chem.* **2011**, *32*, 54−69.

(38) Bylaska, E. J.; Waters, K.; Hermes, E. D.; Zádor, J.; Rosso, K. M. A Filon-like integration strategy for calculating exact exchange in periodic boundary conditions: a plane-wave DFT implementation. *Mater. Theory* **2020**, *4*, 3.

(39) Bylaska, E. J.; Rosso, K. Corresponding orbitals derived from periodic bloch states for electron transfer calculations of transition metal oxides. *J. Chem. Theory Comput.* **2018**, *14*, 4416−4426.

(40) Bylaska, E. J.; Taylor, P. R.; Kawai, R.; Weare, J. H. LDA predictions of C20 isomerizations: Neutral and charged species. *J. Phys. Chem.* **1996**, *100*, 6966−6972.

(41) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188.

(42) Oddershede, J.; Kumar, L.; Monkhorst, H. J. On the hartree-fock limit for metallic hydrogen with plane wave basis. *Int. J. Quantum Chem.* **1974**, *8*, 447−450.

(43) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.

(44) Stanton, R. Hellmann-Feynman theorem and correlation energies. *J. Chem. Phys.* **1962**, *36*, 1298−1300.

(45) Von Barth, U.; Gelatt, C. Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations. *Phys. Rev. B* **1980**, *21*, 2222.

(46) Kiejna, A.; Kresse, G.; Rogal, J.; De Sarkar, A.; Reuter, K.; Scheffler, M. Comparison of the full-potential and frozen-core approximation approaches to density-functional calculations of surfaces. *Phys. Rev. B* **2006**, *73*, 035404.

(47) Bierman, J.; Li, Y.; Lu, J. Qubit Count Reduction by Orthogonally Constrained Orbital Optimization for Variational Quantum Excited-State Solvers. *J. Chem. Theory Comput.* **2024**, *20*, 3131−3143.

(48) Bierman, J.; Li, Y.; Lu, J. Improving the accuracy of variational quantum eigensolvers with fewer qubits using orbital optimization. *J. Chem. Theory Comput.* **2023**, *19*, 790−798.

(49) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.

(50) Damour, Y.; Véril, M.; Kossoski, F.; Caffarel, M.; Jacquemin, D.; Scemama, A.; Loos, P.-F. Accurate full configuration interaction correlation energy estimates for five-and six-membered rings. *J. Chem. Phys.* **2021**, *155*, 134104.

(51) Wang, Z.; Li, Y.; Lu, J. Coordinate descent full configuration interaction. *J. Chem. Theory Comput.* **2019**, *15*, 3558−3569.

(52) Wang, Z.; Zhang, Z.; Lu, J.; Li, Y. Coordinate Descent Full Configuration Interaction for Excited States. *J. Chem. Theory Comput.* **2023**, *19*, 7731−7739.

(53) Stowasser, R.; Hoffmann, R. What do the Kohn- Sham orbitals and eigenvalues mean? *J. Am. Chem. Soc.* **1999**, *121*, 3414−3420.

(54) Rozzi, C. A.; Varsano, D.; Marini, A.; Gross, E. K.; Rubio, A. Exact Coulomb cutoff technique for supercell calculations. *Phys. Rev. B* **2006**, *73*, 205119.

(55) Ghosh, G.; Van de Walle, A.; Asta, M. First-principles calculations of the structural and thermodynamic properties of bcc, fcc and hcp solid solutions in the Al−TM (TM= Ti, Zr and Hf) systems: a comparison of cluster expansion and supercell methods. *Acta Mater.* **2008**, *56*, 3202−3221.

(56) Partoens, B.; Peeters, F. From graphene to graphite: Electronic structure around the K point. *Phys. Rev. B* **2006**, *74*, 075404.

(57) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. Regularized second-order Møller−Plesset theory: A more accurate alternative to conventional MP2 for noncovalent interactions and transition metal thermochemistry for the same computational cost. *J. Phys. Chem. Lett.* **2021**, *12*, 12084−12097.

(58) Balakrishnan, A.; Smith, V.; Stoicheff, B. Dissociation energy of the hydrogen molecule. *Phys. Rev. Lett.* **1992**, *68*, 2149.

(59) Soler, J. M.; Beltrán, M. R.; Michaelian, K.; Garzón, I. L.; Ordejón, P.; Sánchez-Portal, D.; Artacho, E. Metallic bonding and cluster structure. *Phys. Rev. B* **2000**, *61*, 5771.

(60) Abell, G. Empirical chemical pseudopotential theory of molecular and metallic bonding. *Phys. Rev. B* **1985**, *31*, 6184.

(61) Ren, X.; Rinke, P.; Blum, V.; Wieferink, J.; Tkatchenko, A.; Sanfilippo, A.; Reuter, K.; Scheffler, M. Resolution-of-identity approach to Hartree−Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys.* **2012**, *14*, 053020.

(62) Feller, D.; Peterson, K. A. Re-examination of atomization energies for the Gaussian-2 set of molecules. *J. Chem. Phys.* **1999**, *110*, 8384−8396.

(63) Ren, X.; Rinke, P.; Joas, C.; Scheffler, M. Random-phase approximation and its applications in computational chemistry and materials science. *J. Mater. Sci.* **2012**, *47*, 7447−7471.

(64) Mejuto-Zaera, C.; Tubman, N. M.; Whaley, K. B. Dynamical mean field theory simulations with the adaptive sampling configuration interaction method. *Phys. Rev. B* **2019**, *100*, 125165.