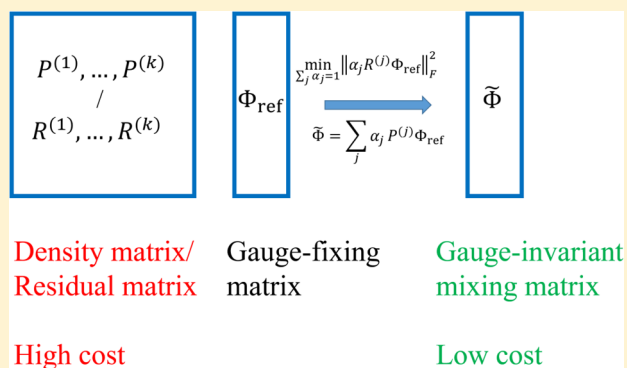# Projected Commutator DIIS Method for Accelerating Hybrid Functional Electronic Structure Calculations

Wei Hu,*[†] Lin Lin,*[†,‡] and Chao Yang*[†]

[†]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
[‡]Department of Mathematics, University of California, Berkeley, California 94720, United States

ABSTRACT: The commutator direct inversion of the iterative subspace (commutator DIIS or C-DIIS) method developed by Pulay is an efficient and the most widely used scheme in quantum chemistry to accelerate the convergence of self-consistent field (SCF) iterations in Hartree−Fock theory and Kohn−Sham density functional theory. The C-DIIS method requires the explicit storage of the density matrix, the Fock matrix, and the commutator matrix. Hence, the method can only be used for systems with a relatively small basis set, such as the Gaussian basis set. We develop a new method that enables the C-DIIS method to be efficiently employed in electronic structure calculations with a large basis set such as planewaves for the first time. The key ingredient is the projection of both the density matrix and the commutator matrix to an auxiliary matrix called the gauge-fixing matrix. The resulting projected commutator-DIIS method (PC-DIIS) only operates on matrices of the same dimension as that consists of Kohn−Sham orbitals. The cost of the method is comparable to that of standard charge mixing schemes used in large basis set calculations. The PC-DIIS method is gauge-invariant, which guarantees that its performance is invariant with respect to any unitary transformation of the Kohn−Sham orbitals. We demonstrate that the PC-DIIS method can be viewed as an extension of an iterative eigensolver for nonlinear problems. We use the PC-DIIS method for accelerating Kohn−Sham density functional theory calculations with hybrid exchange−correlation functionals, and demonstrate its superior performance compared to the commonly used nested two-level SCF iteration procedure. Furthermore, we demonstrate that in the context of *ab initio* molecular dynamics (MD) simulation with hybrid functionals one can extrapolate the gauge-fixing matrix to achieve the goal of extrapolating the entire density matrix implicitly along the MD trajectory. Numerical results indicate that the new method significantly reduces the number of SCF iterations per MD step, compared to the commonly used strategy of extrapolating the electron density.

Density matrix/ Residual matrix

$P^{(1)}, ..., P^{(k)}$ / $R^{(1)}, ..., R^{(k)}$

High cost

Φ_{ref}

Gauge-fixing matrix

$\min_{\sum_j \alpha_j = 1} \left\| \sum_j \alpha_j R^{(j)} \Phi_{ref} \right\|_F^2$

$\widetilde{\Phi} = \sum_j \alpha_j P^{(j)} \Phi_{ref}$

$\widetilde{\Phi}$

Gauge-invariant mixing matrix

Low cost

1. INTRODUCTION

In electronic structure calculations based on the Hartree−Fock (HF) theory and the Kohn−Sham density functional theory (KSDFT), the density matrix needs to be computed self-consistently. For quantum chemistry software packages based on Gaussian orbitals or localized atomic orbitals, the most widely used numerical scheme to achieve self-consistency is the commutator direct inversion of the iterative subspace (commutator DIIS or C-DIIS) method.[1] Despite the fact that there is no known theoretical guarantee for the global convergence of the C-DIIS method and the fact that the method may sometimes fail when the initial guess is far away from the true solution, numerous numerical results indicate that the method often converges rapidly at least when the initial guess is close to the true solution.[2−4] However, the C-DIIS method requires the explicit storage of the density matrix, the Fock matrix, and the commutator matrix. Hence, the method is limited to systems discretized using a small basis set such as Gaussian orbitals. For large basis sets such as planewaves, it is prohibitively expensive to store the full Fock matrix or the

density matrix, and the C-DIIS method is not directly applicable. Instead, in large basis set calculations, the most commonly used method is to only perform mixing on local quantities such as the electron density or the local potential, which correspond to the diagonal elements of the density matrix and the Fock matrix in the real space representation, respectively. When combined with a good preconditioner,[5,6] the density mixing and potential mixing schemes can also be highly efficient for KSDFT calculations with local and semilocal exchange−correlation functionals, such as the local density approximation (LDA),[7−9] the generalized gradient approximation (GGA),[10−12] and meta-GGA functionals.[13−15]

Hybrid exchange−correlation functionals, such as B3LYP,[16] PBE0,[17] and HSE,[18,19] are known to be more reliable in producing high fidelity results for electronic structure calculations. Hybrid functionals include a fraction of the Fock exchange operator, which not only depends on the electron

density but also the off-diagonal elements of the density matrix. This leads to a significant increase in the computational cost compared to calculations with semilocal functionals. For calculations performed in large basis sets such as planewaves, an iterative diagonalization procedure is used solve the Kohn−Sham equations. The diagonalization procedure requires multiplying the exchange operator with the occupied orbitals in each iteration. These multiplications alone often constitutes more than 95% of the overall computational time in a conventional approach. Many numerical methods have been developed to reduce the cost of such multiplication operations. Notably, linear scaling methods[20−25] construct a sparse approximation to the exchange operator and have been developed and applied to large systems with substantial band gaps. Recently, we have developed the adaptively compressed exchange operator (ACE) formulation[26,27] and the interpolative separable density fitting (ISDF) method[28,29] to reduce the cost associated with the exchange operator. Unlike linear scaling methods, one notable feature of ACE and ISDF is that these methods are insensitive to the energy band gap of the target system and hence can be applied to insulating, semiconducting, and even metallic systems.

Here, we are concerned with yet another difficulty in hybrid functional calculations related to the SCF iterations, which is somewhat orthogonal to the difficulty introduced by the large computational cost associated with the multiplication of the exchange operator. Note that self-consistency must be achieved for the entire density matrix. For algorithms that use a large basis set to discretize the problem, it is not practical to store and update the density matrix directly. The commonly used approach to achieve self-consistency is a nested two-level SCF iteration procedure that consists of an inner SCF loop and an outer SCF loop. In the inner SCF loop, the density matrix and hence the Fock exchange operator are fixed (implicitly by fixing the Kohn−Sham orbitals), and the SCF iteration is only performed for the charge density. Once the error of the inner iteration reaches below certain tolerance level, the density matrix is updated using a fixed point iteration in the outer iteration, which is carried out implicitly through a fixed point iteration for all the Kohn−Sham orbitals. The use of the two nested SCF loops allows charge mixing schemes such as Anderson mixing[30] and Kerker preconditioner[5] to be performed to avoid the "charge sloshing" phenomenon.[6,31] However, it also significantly increases the number of iterations for hybrid functional calculations to converge. There are other methods for accelerating the SCF iteration, such as the E-DIIS method.[32] Besides the SCF iteration procedure, HF and KSDFT calculations can also be performed through direct minimization of the orbitals without storing the density matrix, such as the direct optimization method[31,33,34] and the Car−Parrinello method[35] with a damped dynamics. The advantage of these methods is that one can use the same algorithmic structure for semilocal and hybrid functional calculations. However, the convergence rate of these methods may be slower when compared to SCF iteration methods such as C-DIIS.

In this paper, we develop a new method that enables the C-DIIS method to be efficiently employed in electronic structure calculations with a large basis set for the first time. As a result, the two level SCF loops can be efficiently reduced into a single loop, which significantly cuts down the total number of SCF iterations. Our method never requires storing the density matrix, Fock matrix, or the commutator. It operates on matrices with $O(N_e)$ columns, where N_e is the number of electrons. The

key idea is to project both the density matrix and the commutator to an auxiliary matrix called the gauge-fixing matrix. The resulting projected commutator DIIS method (PC-DIIS) is inherently gauge-invariant and hence numerically stable. We remark that the gauge-invariance property is what distinguishes the PC-DIIS method from other acceleration schemes such as the RMM-DIIS method.[6] The PC-DIIS method can also be viewed as a generalized formulation of the C-DIIS method with its residual weighted by the gauge-fixing matrix.[36] We demonstrate that the PC-DIIS method can be viewed as an extension of an iterative eigensolver for nonlinear problems. The PC-DIIS method is well suited for accelerating Hartree−Fock and hybrid functional Kohn−Sham density functional theory calculations. The PC-DIIS method can also be naturally combined with other recently developed techniques for accelerating hybrid functional calculations such as ACE and ISDF. Similar to ACE and ISDF, we find that the effectiveness of the PC-DIIS method is not sensitive to the size of the band gap.

In the context of *ab initio* molecular dynamics (AIMD) simulation, since it is not practical to store the entire density matrix, one commonly used strategy is to only extrapolate the electron density from one time step to another. This leads to a mismatch between the electron density and the density matrix defined by the Kohn−Sham orbitals. We demonstrate that in the PC-DIIS method, one can extrapolate the gauge-fixing matrix along the MD trajectory. The gauge-fixing matrix can be chosen to be smooth with respect to the time t and to carry equivalent information as in the density matrix at each time step. As a result, the density matrix is implicitly extrapolated along the MD trajectory. This procedure shares similarity with the predictor−corrector scheme for extrapolating Kohn−Sham orbitals.[37] Compared to the method that only extrapolates the electron density, we find that extrapolating the gauge-fixing matrix can lead to significant reduction of the number of SCF iterations per MD step.

We demonstrate the performance of the PC-DIIS method for a number of systems with insulating and metallic characters. We find that the PC-DIIS method can significantly reduce the number of iterations and hence the wall clock time. For instance, using the HSE06 hybrid functional, we can fully converge a bulk silicon system with 1000 atoms within 5 min of wall clock time.

The rest of the manuscript is organized as follows. We introduce the two level nested SCF method and the C-DIIS method in Sections 2 and 3, respectively. We present the new PC-DIIS method in Section 4 and demonstrate how to extrapolate the gauge-fixing matrix in the context of AIMD simulation in Section 5. Numerical results are presented in Section 6, followed by a conclusion and discussion in Section 7.

2. TWO-LEVEL NESTED SCF METHOD

The Hartree−Fock-like equations as appeared in the Hartree−Fock theory and the Kohn−Sham density functional theory with hybrid functionals are a set of nonlinear equations as follows:

$$H[P]\psi_i = \left(-\frac{1}{2}\Delta + V_{\text{ion}} + V_{\text{Hxc}}[P] + V_X[P]\right)\psi_i = \varepsilon_i\psi_i,$$

$$\int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, \quad P(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{N_e} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') \tag{1}$$

Here, $H[P]$ is called a Hamiltonian operator (also called a Fock operator). The eigenvalues $\{\varepsilon_i\}$ are ordered nondecreasingly, and N_e is the number of electrons (spin degeneracy omitted for simplicity). P is the density matrix associated with $\{\psi_i\}$, $i = 1, 2,$ ..., N_e. It is an orthogonal projector with an exact rank N_e. The diagonal entries of P give the electron density $\rho(\mathbf{r}) = P(\mathbf{r}, \mathbf{r})$. V_{ion} characterizes the electron−ion interaction in all-electron calculations. V_{Hxc} is a local operator that characterizes the Hartree and the exchange−correlation contributions modeled at a local or semilocal level. It typically depends only on the electron density $\rho(\mathbf{r})$. The exchange operator V_X is an integral operator with a kernel

$$V_X[P](\mathbf{r}, \mathbf{r}') = -P(\mathbf{r}, \mathbf{r}')K(\mathbf{r}, \mathbf{r}') \tag{2}$$

where $K(\mathbf{r}, \mathbf{r}')$ is an operator that accounts for the electron−electron interaction. For example, in the Hartree−Fock theory, $K(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ is the Coulomb operator. In screened exchange theories such as HSE,[18,19] K is a screened Coulomb operator $K(\mathbf{r}, \mathbf{r}') = \text{erfc}(\mu|\mathbf{r}-\mathbf{r}'|)/|\mathbf{r}-\mathbf{r}'|$.

Methods for solving the Hartree−Fock equation often use an iterative procedure (i.e., the SCF iteration) in which the density matrix P is updated until it is consistent with the Hamiltonian operator $H[P]$. When a large basis set such as the planewave basis set is used, it becomes prohibitively expensive to store the density matrix P. On the other hand, one cannot simply take the output Kohn−Sham orbitals from one SCF iteration and use them as the input Kohn−Sham orbitals for the next SCF iteration. This type of fixed point iteration is known to suffer from the "charge sloshing" problem.[6,31] In practice, the most commonly used method, such as the one implemented in the Quantum ESPRESSO software package,[38] uses a two-level nested SCF procedure. The motivation for using a two-level nested SCF procedure is to apply advanced charge mixing schemes to the electron density ρ in the inner iteration to mitigate charge sloshing and to use a fixed point iteration to update the Kohn−Sham orbitals and consequently the exchange potential in the outer iteration. The update of the exchange potential is more costly, even though its contribution to total energy is typically much smaller.

The two-level nested SCF method is summarized in Algorithm 1. In each outer iteration, the exchange operator $V_X[P]$ is updated. This is implicitly done by updating a set of orbitals $\{\varphi_i\}_{i=1}^{N_e}$ defining the density matrix as $P = \sum_{i=1}^{N_e}\varphi_i\varphi_i^*$. We remark that this set of orbitals may be different from the Kohn−Sham orbitals in the inner SCF iteration. The update is done through a fixed point iteration; i.e., $\{\varphi_i\}_{i=1}^{N_e}$ are given by the output Kohn−Sham orbitals in the previous outer iteration. In the inner SCF iteration, with the exchange operator fixed, the Hamiltonian H only depends on the electron density ρ. Charge mixing schemes for ρ can be performed in similar fashion to what is done in a standard KSDFT calculation without the exchange operator in the inner SCF iteration. Finally, within each inner iteration, with both P and ρ fixed, eq 1 becomes a linear eigenvalue problem and can be solved by an iterative eigensolver such as the Davidson method,[39] the LOBPCG method,[40] or the PPCG method.[41] The outer SCF

iteration continues until convergence is reached, which can be monitored, e.g., in terms of the change of the exchange energy.

Algorithm 1 Two-level nested SCF method for solving Hartree-Fock-like equations.

1: **while** Exchange energy is not converged **do**
2:    **while** Electron density ρ is not converged **do**
3:       Solve the linear eigenvalue problem $H\psi_i = \varepsilon_i\psi_i$ with any iterative eigensolver.
4:       Update $\rho^{\text{out}}(\mathbf{r}) \leftarrow \sum_{i=1}^{N_e}|\psi_i(\mathbf{r})|^2$.
5:       Update ρ using ρ^{out} and charge densities computed and saved from the previous iteration using a charge mixing scheme.
6:    **end while**
7:    Compute the exchange energy.
8:    Update $\{\varphi_i\}_{i=1}^{N_e} \leftarrow \{\psi_i\}_{i=1}^{N_e}$.
9: **end while**

---

3. COMMUTATOR DIIS METHOD

In the two-level nested SCF procedure, the density matrix is only updated implicitly in the outer iteration by a fixed point iteration of the Kohn−Sham orbitals. The overall convergence of this method is slower than an alternative scheme in which the density matrix and the Fock matrix are updated, e.g., by the commutator direct inversion of the iterative subspace method (commutator DIIS or C-DIIS) developed by Pulay.[1] The C-DIIS method, which requires saving density and Fock matrices for a few iterations, is often used in many hybrid DFT and Hartree−Fock calculations performed in a small basis sets. However, because it is too costly to compute and save the entire density and Fock matrices explicitly in each SCF for a calculation performed in a large basis set, the direct use of the C-DIIS method is not feasible.

Before we introduce an efficient way to perform C-DIIS in a hybrid DFT or Hartree−Fock calculation performed in a large basis, we first briefly introduce the standard C-DIIS method below. For simplicity, we assume the Kohn−Sham or Hartree−Fock equations are discretized by an orthonormal basis set. We define the residual R as the commutator between $H[P]$ and P, i.e.,

$$R[P] = H[P]P - PH[P] \tag{3}$$

The C-DIIS method is designed to minimize the residual within a subspace that contains the residuals associated with previous approximations to the Fock and density matrices. To be specific, if $H^{(k)}$ denotes the approximate Hamiltonian produced at step k, we define a new Hamiltonian $\tilde{H}^{(k+1)}$ at step $(k + 1)$ as a linear combination of $l + 1$ previous approximations to the Hamiltonian, i.e.,

$$\tilde{H}^{(k+1)} = \sum_{j=k-l}^{k} \alpha_j H^{(j)} \tag{4}$$

where α_j satisfies the constraint $\sum_{j=k-l}^{k} \alpha_j = 1$. Each Hamiltonian $H^{(j)}$ defines a density matrix $P^{(j)}$ through the solution of the linear eigenvalue problem (1). Before self-consistency is reached, the residual $R^{(j)} = R[P^{(j)}]$ defined by eq 3 is nonzero. However, when all the Hamiltonian matrices $\{H^{(j)}\}$ are close to the self-consistent Hamiltonian operator $H^{\star}$, it is reasonable to expect the residual associated with $\tilde{H}^{(k+1)}$ to be well approximated by $\tilde{R}^{(k+1)} \equiv \sum_{j=k-l}^{k}\alpha_j R^{(j)}$. The C-DIIS method determines $\{\alpha_j\}$ by minimizing $\tilde{R}^{(k+1)}$ with respect to α_j's, i.e., we solve the following constrained minimization problem in each C-DIIS iteration:

$$\min_{\{\alpha_j\}} \left\| \sum_{j=k-l}^{k} \alpha_j R^{(j)} \right\|_F^2, \quad \text{s.t.} \sum_{j=k-l}^{k} \alpha_j = 1 \tag{5}$$

Here, the Frobenius norm is defined as $\|A\|_F^2 = \text{Tr}[A*A]$. Although a vanishing commutator residual is only a necessary condition for reaching self-consistency, it has been found that, in the context of Hartree−Fock equation, the commutator corresponds to the gradient direction of the total energy with respect to P. Hence, the C-DIIS method can be interpreted as a quasi-Newton method.[42]

Note that the constraint in eq 5 can be eliminated by rewriting eq 4 as

$$\tilde{H}^{(k+1)} = H^{(k)} + \sum_{j=k-l+1}^{k} \beta_j (H^{(j-1)} - H^{(j)}) \tag{6}$$

The relation between eqs 4 and 6 can be seen from the following mapping between $\{\alpha_i\}$ and $\{\beta_j\}$ parameters, i.e.,

$$\alpha_j = \begin{cases} 1 - \beta_j, & j = k \\ \beta_{j+1} - \beta_j, & k - l + 1 \leq j \leq k - 1 \\ \beta_{j+1}, & j = k - l \end{cases} \tag{7}$$

Clearly, $\sum_{j=k-l}^{k} \alpha_j = 1$ is satisfied, and the new variables $\beta = (\beta_{k-l+1}, ..., \beta_k)^T$ become unconstrained. Define $Y^{(j)} = R^{(j-1)} - R^{(j)}$ for $k - l + 1 \leq j \leq k$, then the constraint minimization problem (5) becomes a unconstrained minimization problem

$$\min_{\{\beta_j\}} \left\| R^{(k)} + \sum_{j=k-l+1}^{k} \beta_j Y^{(j)} \right\|_F^2 \tag{8}$$

As a result, eq 8 has an analytic solution

$$\beta = -M^{-1}b \tag{9}$$

where the $l \times l$ matrix M and the vector b are defined as

$$M_{ij} = \text{Tr}[(Y^{(i)})^* Y^{(j)}], \quad b_j = \text{Tr}[(Y^{(j)})^* R^{(k)}] \tag{10}$$

respectively.

Using the solution (eq 9), we obtain $\tilde{H}^{(k+1)}$ via eq 6. This procedure is repeated until the residual $R[P]$ is sufficiently small. We remark that the solution in the form of eq 9 can also be interpreted as a variant of the Anderson acceleration method.[30]

4. PROJECTED COMMUTATOR DIIS

In this section, we introduce a new method that enables the C-DIIS method to be used for hybrid DFT or Hartree−Fock calculations performed within a large basis set. Since it is not possible to explicitly store or mix the density matrices in such calculations, it is tempting to perform the DIIS procedure on the $N_g \times N_e$ orbital matrix $\Psi = [\psi_1, ..., \psi_{N_e}]$, where N_g is the number of degrees of freedom to discretize each orbital (e.g., the number of planewaves) and N_e is the number of electrons or Kohn−Sham orbitals. However, one key difference between the density matrix and the orbital matrix is that the former is gauge invariant. That is, if we replace Ψ by ΨU where U is an $N_e \times N_e$ unitary gauge matrix, the density matrix $P = \Psi\Psi^* = \Psi U U^* \Psi^*$ does not change. Therefore, it is completely safe to combine two density matrices constructed from Ψ's that differ by a gauge transformation. It is important to note that in both the Hartree−Fock theory and KSDFT with semilocal and hybrid functionals, the total energy is invariant with respect to any gauge transformation.

However, since the orbital matrix Ψ is not gauge invariant, combining successive approximations to Ψ's that differ by a gauge transformation may hinder the convergence of the SCF iteration or may not be stable at all. This type of scenario may arise when eigenvalues of H are degenerate or nearly degenerate. A slight change in the potential in H from one SCF iteration to another may lead to an arbitrary rotation of the eigenvectors associated with these eigenvalues.[43] Consequently, a linear combination of the orbital matrices associated with these two consecutive SCF iterations is not expected to be effective.

To overcome this difficulty, we introduce an auxiliary orbital matrix Φ that spans the same subspace spanned by Ψ. This orbital matrix is obtained by applying the orthogonal projection operator associated with Ψ to a reference orbital matrix Φ_{ref} to be specified later in this section. That is, Φ is chosen to be

$$\Phi = P\Phi_{\text{ref}} = \Psi(\Psi^* \Phi_{\text{ref}}) \tag{11}$$

We require Φ_{ref} to be fixed throughout the entire SCF procedure. Note that Φ is invariant to any gauge transformation applied to Ψ. Therefore, the auxiliary orbital matrices obtained in successive SCF iterations can be safely combined to produce a better approximation to the desired invariant subspace.

The columns of Φ are generally not orthogonal to each other. However, as long as the columns of Φ are not linearly dependent, both Ψ and Φ span the range of the density matrix P, which can also be written as

$$P = \Phi(\Phi^* \Phi)^{-1} \Phi^* \tag{12}$$

The new method we propose to accelerate the SCF iteration for hybrid DFT and HF calculations performed in a large basis set, which we call a projector commutator DIIS (PC-DIIS) method, constructs a new approximation to Φ in the kth SCF iteration by taking a linear combination of the auxiliary orbital matrices $\Phi^{(k-l)}, ..., \Phi^{(k)}$ obtained in $l+1$ previous iterations, i.e.

$$\tilde{\Phi}^{(k+1)} = \sum_{j=k-l}^{k} \alpha_j \Phi^{(j)} \tag{13}$$

The coefficients $\{\alpha_j\}$ in eq 13 are determined by minimizing the residual associated with $\tilde{\Phi}^{(k+1)}$, which, under the assumption that $\Phi^{(j)}$ are sufficiently close to the solution of the Kohn−Sham equations, is well approximated by $R \equiv \sum_{j=k-l}^{k} \alpha_j R_{\Phi^{(j)}}$, where the residual associated with an auxiliary orbital matrix Φ is defined by

$$R_\Phi = H[P]P\Phi_{\text{ref}} - PH[P]\Phi_{\text{ref}}$$
$$= (H[P]\Psi)(\Psi^* \Phi_{\text{ref}}) - \Psi((H[P]\Psi)^* \Phi_{\text{ref}}) \tag{14}$$

Note that evaluation of the residual in eq 14 only requires multiplying $H[P]$ with Ψ and the multiplications of matrices of sizes $N_g \times N_e$ and $N_e \times N_e$ only. These operations are already used in iterative methods for computing the desired eigenvectors Ψ of H. The PC-DIIS algorithm does not require $P, H[P]$ or R to be constructed or stored explicitly.

An interesting observation is that, if $\Phi_{\text{ref}} = \Psi$, then $\Psi^* H[P]\Psi$ is a diagonal matrix denoted by Λ. Consequently, the projected commutator takes the form

$$R_\Phi = H[P]\Psi - \Psi\Lambda \tag{15}$$

This expression coincides with the standard definition of the residual associated with an approximate eigenpair (Λ, Ψ).

Hence the PC-DIIS method can also be viewed as an extension of an iterative eigensolver for nonlinear problems.

Similar to the reformulation of the constrained minimization problem into an unconstrained minimization problem in the C-DIIS method presented in the previous section, the constrained minimization problem

$$\min_{\{\alpha_j\}} \left\| \sum_{j=k-l}^{k} \alpha_j R_\Phi^{(j)} \right\|_F^2, \quad \text{s.t.} \quad \sum_{j=k-l}^{k} \alpha_j = 1 \tag{16}$$

to be solved in the PC-DIIS method can also be reformulated as an unconstrained minimization problem. Using the same change of variable as that presented in Section 3, we can write

$$\tilde{\Phi}^{(k+1)} = \Phi^{(k)} + \sum_{j=k-l+1}^{k} \beta_j (\Phi^{(j-1)} - \Phi^{(j)}) \tag{17}$$

If we let $Y_{\Phi^{(j)}} = R_{\Phi^{(j-1)}} - R_{\Phi^{(j)}}$, the coefficients β_j's in eq 17 can be retrieved from the vector $\beta = -(M^\Phi)^{-1} b^\Phi$, where

$$M_{ij}^\Phi = \text{Tr}[(Y_{\Phi^{(i)}})^* Y_{\Phi^{(j)}}], \quad b_j^\Phi = \text{Tr}[(Y_{\Phi^{(j)}})^* R_{\Phi^{(k)}}] \tag{18}$$

Once $\tilde{\Phi}^{(k+1)}$ is obtained, a density matrix associated with this orbital matrix is implicitly defined through eq 12. This implicitly defined density matrix allows us to construct a new Hamiltonian from which a new set of Kohn–Sham orbitals $\Psi^{(k+1)}$ and auxiliary orbitals $\Phi^{(k+1)}$ can be computed.

We now discuss how to choose the gauge-fixing matrix Φ_{ref}. Note that in hybrid functional calculations the contribution from the exchange operator is relatively small. Hence, the density matrix associated with Kohn–Sham orbitals obtained from a DFT calculation that uses a local or semilocal exchange–correlation functional is already a good initial guess for the density matrix required in a hybrid functional calculation. Therefore, we may use these orbitals as Φ_{ref}. Compared to the two-level nested loop structure, the PC-DIIS method only requires one level of SCF iteration. The PC-DIIS method is summarized in Algorithm 2.

Algorithm 2 The PC-DIIS method for solving Hartree-Fock-like equations.

Input: Reference orbitals Φ_{ref}.
Output: Approximate solution $\Psi = \{\psi_i\}$, $i = 1, 2..., N_e$, to Eq. (1).

1: Construct the initial Hamitonian H and evaluate the exchange energy using Φ_{ref};
2: **while** Exchange energy is not converged **do**
3:     Solve the linear eigenvalue problem $H[P]\psi_i = \varepsilon_i \psi_i$ using an iterative eigensolver.
4:     Evaluate Φ, R_Φ according to Eq. (11), (14).
5:     Perform the DIIS procedure according to Eq. (17) to obtain the new $\tilde{\Phi}$ which implicitly defines a density matrix P via Eq. (12).
6:     Update the Hamiltonian $H[P]$.
7:     Compute the exchange energy.
8: **end while**

The discussion above is applicable when Ψ only contains the occupied orbitals. When Ψ also involves the unoccupied orbitals, we use the fact that the density matrix defining the Fock exchange operator only involves the occupied orbitals, and we only need to apply the PC-DIIS method to the occupied orbitals. We also remark that the PC-DIIS method is not yet applicable for finite temperature calculations with fractionally occupied orbitals, and this will be our future work.

5. WAVE FUNCTION EXTRAPOLATION IN *AB INITIO* MOLECULAR DYNAMICS

In *ab initio* molecular dynamics (AIMD) simulation, the electron density and the Kohn–Sham orbitals between consecutive MD steps are correlated. Hence, one can extrapolate electron density density and/or Kohn–Sham orbitals from previous MD steps to produce an initial guess for the new MD step. The simplest strategy is a linear extrapolation procedure. For instance, in KSDFT calculations with semilocal functionals, let $\rho(t - \delta t)$ and $\rho(t)$ be the electron density at time $t - \delta t$ and t, respectively. One can perform linear extrapolation

$$\rho^p = 2\rho(t) - \rho(t - \delta t) \tag{19}$$

and use ρ^p as the initial guess for the electron density at time $t + \delta t$. After the self-consistency is reached at time $t + \delta t$, we obtain the corrected electron density $\rho(t + \delta t)$. For Hartree–Fock-like equations, the Hamiltonian depends on the entire density matrix, and it is not sufficient to only extrapolate the electron density. Since it is prohibitively expensive to extrapolate the density matrix when a large basis set is used, one can only perform extrapolation on the Kohn–Sham orbitals. However, due to the arbitrariness in the choice of the gauge matrix, the Kohn–Sham orbitals $\Psi(t)$ may depend on the choice of the gauge in the eigensolver. In this case, $\Psi(t)$ is not even be continuous with respect to t.

Nonetheless, the density matrix $P(t)$ is smooth with respect to t. Our main observation is that if we can choose the time-dependent gauge-fixing matrix $\Phi_{\text{ref}}(t)$ that is smooth with respect to t, then the corresponding matrix $P(t)\Phi_{\text{ref}}(t)$ will also be smooth respect to t. In particular, according to eq 12, the information contained in $P(t)$ and $P(t)\Phi_{\text{ref}}(t)$ are equivalent. For each given time t, the only constraint in the choice of $\Phi_{\text{ref}}(t)$ is that $P(t)\Phi_{\text{ref}}(t)$ should have full column rank. Assuming we have already obtained full rank $\Phi_{\text{ref}}(t-\delta t)$ and $\Phi_{\text{ref}}(t)$, then we can perform linear extrapolation

$$\Phi_{\text{ref}}^p = 2\Phi_{\text{ref}}(t) - \Phi_{\text{ref}}(t - \delta t) \tag{20}$$

to obtain the predicted gauge-fixing matrix Φ_{ref}^p. It follows from eq 12 again that the density matrix associated with Φ_{ref}^p is defined by

$$P^p = \Phi_{\text{ref}}^p [(\Phi_{\text{ref}}^p)^* \Phi_{\text{ref}}^p]^{-1} (\Phi_{\text{ref}}^p)^* \tag{21}$$

We stress that we never explicitly construct the density matrix P^p but only implicitly use the matrix factors in eq 21 to update the electron density and the exchange operator. After self-consistency is reached at $t + \delta t$, we obtain the Kohn–Sham orbitals denoted by $\Psi(t + \delta t)$. This gives us the corrected gauge-fixing matrix at time $t + \delta t$ as

$$\begin{aligned} \Phi_{\text{ref}}(t + \delta t) &= P(t + \delta t)\Phi_{\text{ref}}^p \\ &= \Psi(t + \delta t)(\Psi^*(t + \delta t)\Phi_{\text{ref}}^p) \end{aligned} \tag{22}$$

which is clearly gauge invariant with respect to $\Psi(t + \delta t)$. Again due to eq 12, the $\Phi_{\text{ref}}(t + \delta t)$ and $P(t + \delta t)$ span the same space and the MD simulation can continue. Our numerical results indicate that the extrapolation of the density matrix by means of the gauge-fixing matrix can effectively reduce the number of SCF iterations in AIMD simulation.

6. NUMERICAL RESULTS

We demonstrate the accuracy and efficiency of the PC-DIIS method using the DGDFT (Discontinuous Galerkin Density Functional Theory) software package.[44−48] DGDFT is a massively parallel electronic structure software package designed for large scale DFT calculations involving up to tens of thousands of atoms. It includes a self-contained module called PWDFT for performing planewave-based electronic
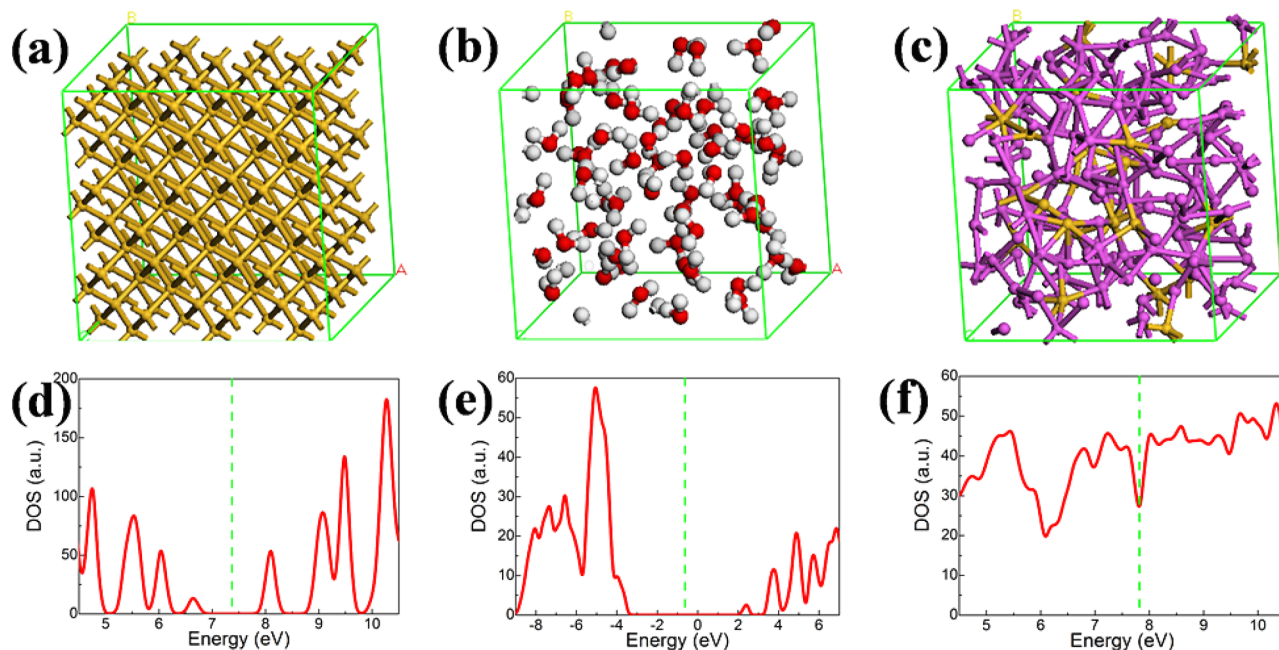
Figure 1. Atomic structures of (a) Si_{216}, (b) $(H_2O)_{64}$, and (c) $Al_{176}Si_{24}$. The white, red, yellow and pink balls denote hydrogen, oxygen, silicon, and aluminum atoms, respectively. The total densities of states (DOS) of semiconducting (d) Si_{216} and (e) $(H_2O)_{64}$, and (f) metallic $Al_{176}Si_{24}$. The Fermi levels are marked by green dotted lines.

Table 1. Differences between HF and Total Energies and Atomic Forces Computed by Nested Two-Level SCF Procedures and by the PC-DIIS Method[a]

Systems	E_{gap}	ΔE_{HF}	ΔE_{tot}	ΔF
Si_{64}	1.484561 (1.10×10^{-07})	1.56×10^{-09}	1.25×10^{-08}	4.45×10^{-06}
Si_{216}	1.449789 (3.00×10^{-08})	4.63×10^{-10}	4.62×10^{-09}	7.27×10^{-07}
Si_{512}	1.324900 (8.00×10^{-07})	1.95×10^{-09}	1.95×10^{-08}	1.56×10^{-05}
Si_{1000}	1.289140 (1.00×10^{-08})	1.00×10^{-09}	1.00×10^{-08}	5.30×10^{-07}
$(H_2O)_{64}$	5.991825 (5.00×10^{-07})	2.60×10^{-08}	1.04×10^{-07}	9.59×10^{-07}
$Al_{176}Si_{24}$	0.098631 (2.20×10^{-05})	5.50×10^{-09}	2.15×10^{-07}	4.10×10^{-06}

[a]The unit for HF and total energy difference is Hartree/atom. The unit atomic for force difference is Hartree/Bohr. Column 2 lists the energy gap in eV and the energy difference (in eV) is shown in the parentheses.

structure calculations (mostly for benchmarking and validation purposes). We implemented the PC-DIIS method in PWDFT. We use the Message Passing Interface (MPI) to handle data communication and the Hartwigsen-Goedecker-Hutter (HGH) norm-conserving pseudopotential.[49] All calculations use the HSE06 functional.[19] All calculations are carried out on the Edison systems at the National Energy Research Scientific Computing Center (NERSC). Each node consists of two Intel "Ivy Bridge" processors with 24 cores in total and 64 gigabyte (GB) of memory. Our implementation only uses MPI. The number of cores is equal to the number of MPI ranks used in the simulation.

In this section, we demonstrate the performance of the PC-DIIS method for accelerating hybrid functional calculations by using six different systems. They consist of four bulk silicon systems (Si_{64}, Si_{216}, Si_{512}, and Si_{1000}),[26] a bulk water system with 64 molecules ($(H_2O)_{64}$), and a disordered silicon aluminum alloy system ($Al_{176}Si_{24}$)[47] as shown in Figure 1. Bulk silicon systems (Si_{64}, Si_{216}, Si_{512}, and Si_{1000}) and bulk water system ($(H_2O)_{64}$) are semiconducting with a relatively large energy gap $E_{gap} > 1.0$ eV, and the $Al_{176}Si_{24}$ system is metallic with a small energy gap $E_{gap} < 0.1$ eV. The density of states of these systems is given in Figure 1. All systems are closed shell systems, and the number of occupied bands is $N_{band} = N_e/2$. In

order to compute the energy gap in the systems, we also include two unoccupied bands in all calculations.

6.1. Accuracy. We first validate the accuracy of the PC-DIIS method compared to the two-level nested SCF procedure for hybrid functional calculations. In both cases, we use the adaptively compressed exchange (ACE)[26] formulation to accelerate calculations, which reduces the number of times the exchange operator is applied to Kohn−Sham orbitals without loss of accuracy. Table 1 shows the differences between the energy gaps, the HF energies, the total energies, and the atomic forces computed by PC-DIIS and a nested two-level SCF procedure, respectively. The HF and total energy differences as well as the difference in atomic forces are define by

$$\Delta E_{HF} = (E_{HF}^{PC-DIIS} - E_{HF}^{NESTED})/N_A$$

$$\Delta E = (E^{PC-DIIS} - E^{NESTED})/N_A$$

$$\Delta F = \max_I |F_I^{PC-DIIS} - F_I^{NESTED}|$$

where the superscript NESTED denotes quantities obtained from a nested two-level SCF procedure, and N_A is the total number of atoms and I is the atom index.

All calculations start from initial Kohn–Sham orbitals obtained from converged calculations using the PBE functional.[12] The kinetic energy cutoff E_{cut} is set to 10 Hartree for bulk silicon systems, 40 Hartree for the $Al_{176}Si_{24}$ system, and 60 Hartree for water system.

Our calculations indicate that the electronic properties obtained from the PC-DIIS method is fully comparable to that obtained from the nested two-level SCF procedure for both semiconducting and metallic systems. The remaining difference between the two methods is comparable to the residual error in the SCF iteration and can further be reduced with a tighter convergence criterion.

6.2. Efficiency. We demonstrate the efficiency of the PC-DIIS method by performing hybrid DFT calculations for a bulk silicon system with 1000 atoms ($N_{band} = 2000$) on 2000 computational cores. In both PC-DIIS and the nested two-level SCF procedure, we use the ACE-ISDF method[29] to accelerate hybrid functional calculations within each SCF iteration. A direct application of the exchange operator to $N_{band} = 2000$ occupied orbitals requires the solution of $N_{band}^2 = 4,000,000$ Poisson-like equations. The ISDF method compresses these equations into cN_{band} equations, where c is referred to as the rank parameter. Following ref 29, we choose the rank parameter $c = 6$, which reduces the number of Poisson-like equations to merely 12,000. The corresponding error in the energy and atomic force is below 10^{-3} Hartree/atom and 10^{-3} Hartree/Bohr, respectively. For comparison, we also report the wall clock time for conventional hybrid functional calculations, which uses the nested two-level method without the ACE-ISDF method. The results are summarized in Table 2.

Table 2. Comparison of Computational Efficiency Exhibited by PC-DIIS and Nested Two-Level SCF procedures with and without Using ACE-ISDF Technique on Si_{1000} System Using 2000 Computational Cores[a]

Outer SCF	PC-DIIS	NESTED (Two-level SCF procedure)			
		ACE-ISDF		Conventional	
iteration number	ACE-ISDF Time	#Inner	Time	#Inner	Time
1th	38.36	11	360.94	5	1962.32
second	39.14	8	145.48	4	1565.80
third	39.08	6	114.78	4	1566.30
fourth	39.70	3	62.67	2	831.04
fifth	38.31	2	48.50	1	460.96
sixth	39.01	1	34.82	–	–

[a]We report the wall clock time (second) per outer iteration. For the nested two-level method method, we also report the number of inner SCF iterations for each outer SCF iteration.

The nested two-level SCF procedure requires six and five outer SCF iterations to converge regardless whether ACE-ISDF is used. A similar number of outer SCF iterations is required in the PC-DIIS method also. Hence, the total number of iterations from the PC-DIIS method is very comparable to the number of outer SCF iterations in the nested two-level SCF method. However, the nested two-level SCF method involves many inner SCF iterations especially at the beginning stage, which increases the computational time significantly. Figure 2 shows the relation between the residual of the HF energy in hybrid functional calculations with respect to the wall clock time, plotted on a logarithmic scale. The PC-DIIS method only takes 233.6 s to converge the entire simulation, which is significantly
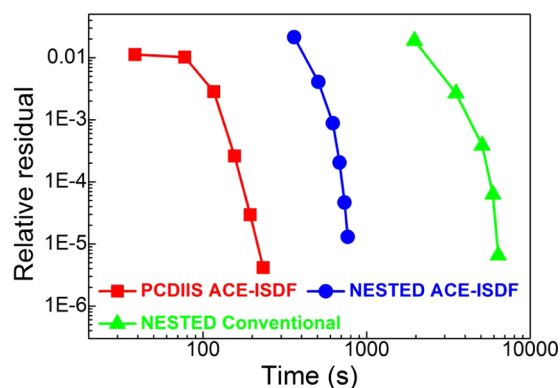
Figure 2. The relative residual of HF energy with respect to the wall clock time computed with one-level (PC-DIIS) and two-level (NESTED) SCF procedures for ACE-ISDF enabled and conventional hybrid functional calculations on the Si_{1000} system on 2000 cores.

faster than the two-level SCF procedure with ACE-ISDF (767.19 s) and without ACE-ISDF (6847.28 s), respectively.

6.3. AIMD. We demonstrate the performance of the PC-DIIS method in the context of AIMD simulation for the Si_{64} system under the NVE ensemble and the $(H_2O)_{64}$ system under the NVT ensemble, respectively.

Figure 3(a) shows the number of SCF iterations along AIMD trajectory for the Si_{64} system, using linear extrapolation of the electron density ("density") and linear extrapolation of the gauge-fixing matrix ("wave function"), respectively. In the former case, the initial guess for the density matrix at time $t + \delta t$ is given by the output density matrix at time t. In the latter case, the extrapolated gauge-fixing matrix Ψ_{ref}^p provides the initial guess for both the density matrix and the density in a consistent fashion at time $t + \delta t$. The convergence criterion is 10^{-6}, and the time step is 1.0 fs (fs). We find that the density extrapolation requires on average 14 iterations per MD step, while the wave function extrapolation only requires five iterations per MD step, respectively. We also observe that the variance of the number of iterations in the wave function extrapolation is significantly smaller.

In AIMD simulation under the NVE ensemble, the total energy is conserved, and hence, the error of the numerical scheme can also be measured in terms of the relative energy drift, defined as $\Delta E_{drift}(t + \delta t)) = (E_{tot}(t + \delta t)) - E_{tot}(t)) E_{tot}(t)$. Figure 3(b) reports the relative drift along the MD trajectory, with different choices of maximum number of SCF iterations per MD step in the PC-DIIS method. The linear extrapolation of the gauge-fixing matrix is used as the initial guess. We find that using a maximum of three SCF iterations already lead to small but noticeable linear drift in the total energy, while the energy drift becomes significantly smaller and stable when a maximum number of four SCF iterations is used per MD step.

We also apply the PC-DIIS method to perform AIMD simulations on liquid water system $(H_2O)_{64}$ at $T = 295$ K (Figure 4). We use a single level Nose-Hoover thermostat,[50,51] and the choice of mass of the Nose-Hoover thermostat is 85000 au. The MD time step size is 1.0 fs. After equilibrating the system starting from a prepared initial guess,[23] we perform the simulation for 2.0 ps to sample the radial distribution function. The average number of converged SCF iterations per MD step is 10. We compare the results from HSE06 and PBE functionals. In both cases, the van der Waals (VdW) interaction
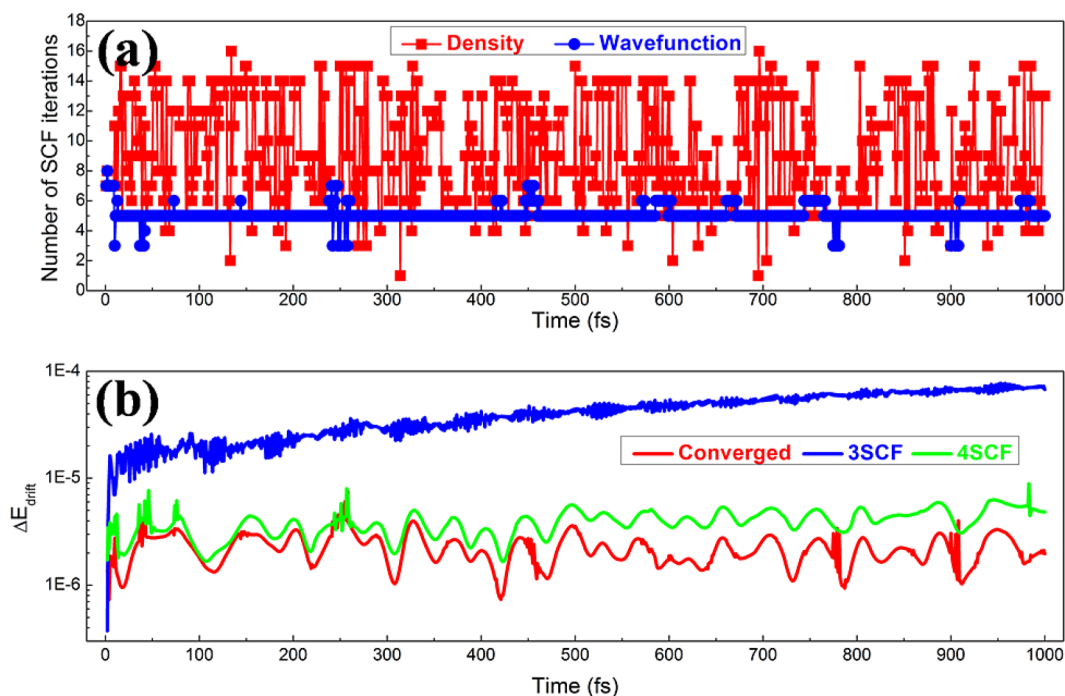
Figure 3. Comparison of AIMD simulations by using one-level (PC-DIIS) and two-level (NESTED) SCF procedures for ACE enabled hybrid functional calculations on the Si_{64} system. (a) The number of SCF iterations used per MD step using linear extrapolation of the density ("density") and linear extrapolation of the gauge-fixing matrix ("wave function"), respectively. (b) The relative energy drift by comparing different choices of maximum number of SCF iterations for the PC-DIIS method per MD step.
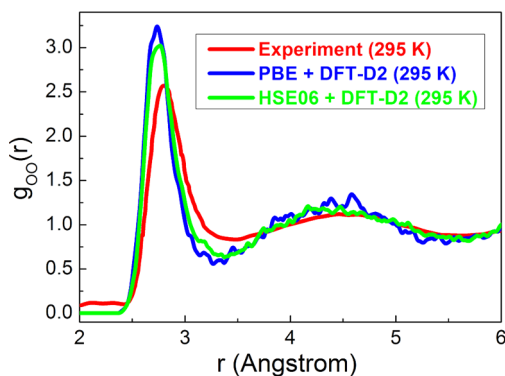
Figure 4. Oxygen−oxygen radial distribution functions $g_{OO}(r)$ of liquid water system $(H_2O)_{64}$ at $T = 295$ K obtained from AIMD simulations with semilocal GGA-PBE and hybrid HSE06 functional calculations with van der Waals correction (DFT-D2) compared to experimental measurement.

is modeled at the level of the DFT-D2 method.[52] We also benchmark our result with the experimental measurement from X-ray diffraction technique.[53] Although the second shell structure is not yet converged due to the relative short simulation length, the structure from the first shell is already clear. We observe that the PBE functional leads to over-structured radial distribution function, which is reduced by the HSE06 functional. This behavior is in quantitative agreement with previous hybrid functional DFT calculations,[23] where the remaining difference with respect to the experimental result can be to a large extent attributed to the nuclei quantum effects.[54]

7. CONCLUSION

We developed the projected commutator-DIIS (PC-DIIS) method for accelerating the self-consistent field (SCF) iteration

in Kohn−Sham density functional theory calculations with hybrid exchange−correlation functionals. The PC-DIIS method is particularly well suited in the context of a large basis set such as the planewave basis set, where it is prohibitively expensive to even store the density matrix and the Hamiltonian matrix. The key idea is to project both the density matrix and the commutator matrix to an auxiliary matrix called the gauge-fixing matrix Φ_{ref}. Then, we can extrapolate the projected matrices that are gauge-invariant and share the same dimension as the Kohn−Sham orbitals. This procedure also implicitly updates the entire density matrix and hence the Hamiltonian matrix in SCF iterations. Compared to the commonly used two-level nested SCF structure used for hybrid functional calculations, the PC-DIIS method only involves one SCF loop, without sacrificing the accuracy or the convergence rate. In the context of ab initio molecular dynamics (AIMD) simulation, the gauge-fixing matrix further provides an efficient and gauge-invariant way for implicit extrapolation of the density matrix. Numerical results indicate that the new extrapolation scheme significantly reduces the number of SCF iterations compared to the commonly used strategy of only extrapolating the electron density.

The PC-DIIS method can be directly extended along several directions. The two-level nested SCF structure is not only used in hybrid functional calculations but also in other contexts such as the DFT+U calculations.[55] Hence, the PC-DIIS method can be potentially useful for accelerating such calculations. For AIMD simulation, the new extrapolation scheme of the gauge-fixing matrix can be combined with time-reversible integrators such as the extended Lagrangian Born−Oppenhemier molecular dynamics (XL-BOMD) method[56] and the almost stable predictor−corrector (ASPC) method[57] to further reduce the number of SCF iterations. The freedom in the choice of Φ_{ref} allows one to consider choosing this matrix to be a sparse

matrix, so that the PC-DIIS method can be used in the context of linear scaling methods. Finally, the PC-DIIS method is not yet applicable to systems under finite temperature with fractionally occupied orbitals. We will explore these directions in the near future.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: whu@lbl.gov (W.H.).
*E-mail: linlin@math.berkeley.edu (L.L.).
*E-mail: cyang@lbl.gov (C.Y.).

ORCID
Wei Hu: 0000-0001-9629-2121
Lin Lin: 0000-0001-6860-9566

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Pulay, P. Improved SCF Convergence Acceleration. *J. Comput. Chem.* **1982**, *3*, 556−560.

(2) Rohwedder, T.; Schneider, R. An analysis for the DIIS acceleration method used in quantum chemistry calculations. *J. Math. Chem.* **2011**, *49*, 1889−1914.

(3) Høst, S.; Olsen, J.; Jansík, B.; Thøgersen, L.; Jørgensen, P.; Helgaker, T. The augmented Roothaan-Hall method for optimizing Hartree-Fock and Kohn-Sham density matrices. *J. Chem. Phys.* **2008**, *129*, 124106.

(4) Garza, A. J.; Scuseria, G. E. Comparison of self-consistent field convergence acceleration techniques. *J. Chem. Phys.* **2012**, *137*, 054110.

(5) Kerker, G. P. Efficient Iteration Scheme for Self-Consistent Pseudopotential Calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1981**, *23*, 3082.

(6) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169.

(7) Ceperley, D. M.; Alder, B. J. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.* **1980**, *45*, 566.

(8) Perdew, J. P.; Zunger, A. Self-Interaction Correction to Density-Functional Approximations for Many-Electron Systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1981**, *23*, 5048.

(9) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 1703.

(10) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.

(11) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.

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

(13) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(14) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(15) Sun, J.; Perdew, J. P.; Ruzsinszky, A. Semilocal Density Functional Obeying a Strongly Tightened Bound for Exchange. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 685−689.

(16) Becke, A. D. A New Mixing of Hartree-Fock and Local Density Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372.

(17) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105*, 9982.

(18) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118*, 8207.

(19) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]. *J. Chem. Phys.* **2006**, *124*, 219906.

(20) Goedecker, S. Linear Scaling Electronic Structure Methods. *Rev. Mod. Phys.* **1999**, *71*, 1085.

(21) Bowler, D. R.; Miyazaki, T. $O(N)$ methods in Electronic Structure Calculations. *Rep. Prog. Phys.* **2012**, *75*, 036503.

(22) Guidon, M.; Hutter, J.; VandeVondele, J. Auxiliary Density Matrix Methods for Hartree - Fock Exchange Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2348−2364.

(23) DiStasio, R. A., Jr.; Santra, B.; Li, Z.; Wu, X.; Car, R. The Individual and Collective effects of Exact Exchange and Dispersion Interactions on the Ab Initio Structure of Liquid Water. *J. Chem. Phys.* **2014**, *141*, 084502.

(24) Dawson, W.; Gygi, F. Performance and Accuracy of Recursive Subspace Bisection for Hybrid DFT Calculations in Inhomogeneous Systems. *J. Chem. Theory Comput.* **2015**, *11*, 4655−4663.

(25) Damle, A.; Lin, L.; Ying, L. Compressed Representation of Kohn-Sham Orbitals via Selected Columns of the Density Matrix. *J. Chem. Theory Comput.* **2015**, *11*, 1463−1469.

(26) Lin, L. Adaptively Compressed Exchange Operator. *J. Chem. Theory Comput.* **2016**, *12*, 2242−2249.

(27) Hu, W.; Lin, L.; Banerjee, A. S.; Vecharynski, E.; Yang, C. Adaptively Compressed Exchange Operator for Large-Scale Hybrid Density Functional Calculations with Applications to the Adsorption of Water on Silicene. *J. Chem. Theory Comput.* **2017**, *13*, 1188−1198.

(28) Lu, J.; Ying, L. Compression of the electron repulsion integral tensor in tensor hypercontraction format with cubic scaling cost. *J. Comput. Phys.* **2015**, *302*, 329.

(29) Hu, W.; Lin, L.; Yang, C. Interpolative separable density fitting decomposition for accelerating hybrid density functional calculations with applications to defects in silicon. *J. Chem. Theory Comput.*, in press.

(30) Anderson, D. G. Iterative Procedures for Nonlinear Integral Equations. *J. Assoc. Comput. Mach.* **1965**, *12*, 547.

(31) Arias, T.; Payne, M.; Joannopoulos, J. Ab initio molecular dynamics: Analytically continued energy functionals and insights into iterative solutions. *Phys. Rev. Lett.* **1992**, *69*, 1077−1080.

(32) Kudin, K. N.; Scuseria, G. E.; Cancès, E. A black-box self-consistent field convergence algorithm: One step closer. *J. Chem. Phys.* **2002**, *116*, 8255−8261.

(33) Yang, C.; Meza, J. C.; Wang, L.-W. A trust region direct constrained minimization algorithm for the Kohn-Sham equation. *SIAM J. Sci. Comput.* **2007**, *29*, 1854−1875.

(34) Wen, Z.; Yin, W. *unpublished.*

(35) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.

(36) Garza, A. J.; Scuseria, G. E. On the equivalence of LIST and DIIS methods for convergence acceleration. *J. Chem. Phys.* **2015**, *142*, 164104.

(37) Kühne, T.; Krack, M.; Mohamed, F.; Parrinello, M. Efficient and accurate Car-Parrinello-like approach to Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.* **2007**, *98*, 1−4.

(38) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A. D.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

(39) Davidson, E. R. The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *J. Comput. Phys.* **1975**, *17*, 87−94.

(40) Knyazev, A. V. Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method. *SIAM J. Sci. Comput.* **2001**, *23*, 517−541.

(41) Vecharynski, E.; Yang, C.; Pask, J. E. A Projected Preconditioned Conjugate Gradient Algorithm for Computing Many Extreme Eigenpairs of a Hermitian Matrix. *J. Comput. Phys.* **2015**, *290*, 73−89.

(42) Kudin, K. N.; Scuseria, G. E. Converging self-consistent field equations in quantum chemistry - recent achievements and remaining challenges. *ESAIM: Math. Modell. Numer. Anal.* **2007**, *41*, 281−296.

(43) Golub, G. H., Van Loan, C. F. *Matrix computations*, 4th ed.; Johns Hopkins University Press: Baltimore, 2013.

(44) Lin, L.; Lu, J.; Ying, L.; E, W. Adaptive Local Basis Set for KohnCSham Density Functional Theory in a Discontinuous Galerkin Framework I: Total Energy Calculation. *J. Comput. Phys.* **2012**, *231*, 2140−2154.

(45) Hu, W.; Lin, L.; Yang, C. DGDFT: A Massively Parallel Method for Large Scale Density Functional Theory Calculations. *J. Chem. Phys.* **2015**, *143*, 124110.

(46) Hu, W.; Lin, L.; Yang, C. Edge Reconstruction in Armchair Phosphorene Nanoribbons Revealed by Discontinuous Galerkin Density Functional Theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31397−31404.

(47) Banerjee, A. S.; Lin, L.; Hu, W.; Yang, C.; Pask, J. E. Chebyshev Polynomial Filtered Subspace Iteration in the Discontinuous Galerkin Method for Large-Scale Electronic Structure Calculations. *J. Chem. Phys.* **2016**, *145*, 154101.

(48) Zhang, G.; Lin, L.; Hu, W.; Yang, C.; Pask, J. E. Adaptive Local Basis Set for Kohn-Sham Density Functional Theory in a Discontinuous Galerkin Framework II: Force, Vibration, and Molecular Dynamics Calculations. *J. Comput. Phys.* **2017**, *335*, 426−443.

(49) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 3641.

(50) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511.

(51) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695.

(52) Grimme, S. Semiempirical GGA-type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.

(53) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. Benchmark Oxygen-Oxygen Pair-Distribution Function of Ambient Water from X-ray Diffraction Measurements with a Wide Q-range. *J. Chem. Phys.* **2013**, *138*, 074506.

(54) Morrone, J.; Car, R. Nuclear quantum effects in water. *Phys. Rev. Lett.* **2008**, *101*, 017801.

(55) Liechtenstein, A. I.; Anisimov, V. I.; Zaanen, J. Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *52*, R5467−R5470.

(56) Niklasson, A. M. N. Extended Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.* **2008**, *100*, 123004.

(57) Kolafa, J. Time-Reversible Always Stable Predictor-Corrector Method for Molecular Dynamics of Polarizable Molecules. *J. Comput. Chem.* **2004**, *25*, 335−342.