Chinese Physics Letters **40**, 027101 (2023) **Express Letter**

A Time-Dependent Random State Approach for Large-Scale Density Functional Calculations

Weiqing Zhou(周巍青) and Shengjun Yuan(袁声军)*

Key Laboratory of Artificial Micro- and Nano-structures of Ministry of Education and School of Physics and Technology, Wuhan University, Wuhan 430072, China

(Received 17 December 2022; accepted manuscript online 17 January 2023)

We develop a self-consistent first-principle method based on the density functional theory. Physical quantities such as the density of states, Fermi energy and electron density are obtained using a time-dependent random state method without diagonalization. The numerical error for calculating either global or local variables always scales as $1/\sqrt{SN_e}$, where N_e is the number of electrons and S is the number of random states, leading to a sublinear computational cost with the system size. In the limit of large systems, one random state could be enough to achieve reasonable accuracy. The accuracy and scaling properties of using the method are derived analytically and verified numerically in different condensed matter systems. Our time-dependent random state approach provides a powerful strategy for large-scale density functional calculations.

DOI: 10.1088/0256-307X/40/2/027101

First-principles calculation using the density functional theory (DFT) is one of the most powerful computational methods for multi-electron systems and contributes extensively to physics, chemistry, and materials science. DFT theorems prove that there is a one-to-one mapping between the ground-state wave function and the ground-state electron density.[1] In the mid 1960s, Kohn and Sham showed that the finding of the ground-state density could be determined by a set of single-electron equations (Kohn–Sham equations),[2] which is also known as the KS-DFT. However, the KS-DFT suffers from a size limitation caused by diagonalization, in which the computational cost exhibits a cubic scaling with the system size. Although many efforts, such as iterative diagonalization schemes,[3] preconditioned conjugate-gradient minimizations,[4–6] and the Car-Parrinello method,[7] have been taken to improvement of the scaling behavior for a relatively small system, it is still hard to handle systems of more than a few hundreds or thousands of atoms.

This size limitation has stimulated the development of linear-scaling DFT.[8–21] The first attempt can be traced back to the 'divide-and-conquer' method of Yang.[11] In 1992, Baroni and Giannozzi also proposed an algorithm that determines the electron density directly by using Green's function.[12] In 1993, the density-matrix minimization approach was proposed by Li, Numes, and Vanderbilt.[13] Following these strategies, many linear-scaling DFT codes have been developed.[10,15–20] The Chebyshev filter method is another successful attempt to reduce the size of the effective dimension of Hilbert space, but there are other non-linear factors dominated in large systems.[22] A linear-scaling algorithm using atomic orbitals (LCAO) basis sets[14] can be applied to suitable systems with clearly separated occupied and empty states.[21] Furthermore, the orbital-free DFT (OF-DFT)[23,24] is a linear-scaling approach that avoids complete diagonalization, but the kinetic energy density functionals have not yet reached a good accuracy for many elements.[25,26] Recently, a linear-scaled DFT is realized by using a stochastic technique in a trace formula,[9,27–30] in which the statistical error of calculating a global variable, such as the total energy (per electron), is reduced by the sample average from different random orbitals. For local quantities, such as the electron density, many stochastic samples are required to reach a reasonable accuracy.

In this Letter, we develop a self-consistent first-principle calculation method based on the DFT without any diagonalization of the Hamiltonian matrix. The physical quantities such as density of states (DOS), Fermi energy and real-space distribution of electron density are calculated using the so-called time-dependent random state (TDRS) method. We show that the numerical error of a global or local variable always scales as $1/\sqrt{SN_e}$, where N_e is the number of electrons and S is the number of random states. It leads to an overall sublinear scaling of the computational costs, and one needs fewer random states for larger systems. The method becomes extremely powerful for massive quantum systems, and a calculation using one random state is enough to achieve reasonable accuracy when $N_e \to \infty$. Our time-dependent random-states DFT (rsDFT) originates from the real-space TDRS method developed in the tight-binding calculations, with an extension from globe variables (such as density of states,[31] electronic and optical conductivities,[32] polarization and screening functions,[33] etc.), to a local variable of electron density. It is a general strategy for the local variable calculations and can be applied in the tight-binding model or other physical models as well.

Density of States and Fermi Energy. In the KS-DFT, the Fermi energy is determined by counting the number of occupied eigenstates, which are obtained from the diagonalization of the Hamiltonian. In the rsDFT,

*Corresponding author. Email: s.yuan@whu.edu.cn
© 2023 Chinese Physical Society and IOP Publishing Ltd

the Fermi energy is extracted by the integration of the DOS, which is calculated with the TDRS method without diagonalization:[31,34]

$$D(\varepsilon) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\varepsilon t}\langle\varphi|e^{-iHt}|\varphi\rangle dt, \tag{1}$$

where $|\varphi\rangle = \sum_i c_i|\boldsymbol{r_i}\rangle$ is a random state in the real space, and $\{c_i\}$ represents normalized random complex numbers. The state $|\varphi\rangle$ is also a random superposition state in the energy space and can be expressed as $|\varphi\rangle = \sum_n b_n|E_n\rangle$ with $b_n = \sum_i c_i a_i^*(E_n)$, $a_i(E_n) = \langle\mathbf{r_i}|E_n\rangle$. Thus, Eq. (1) becomes

$$D(\varepsilon) = \sum_{n=1}^{N} |b_n|^2\delta(\varepsilon - E_n), \tag{2}$$

where N is the dimension of the Hamiltonian. For a large but finite N, $|b_n| \to 1/\sqrt{N}$, the error of using $D(\varepsilon)$ to approximate the DOS vanish with $1/\sqrt{N}$.[31,34] As we normally use the same grid density, the dimension of the Hamiltonian N, determined by the number of grids, is linearly proportional to the number of atoms and the number of electrons. Thus the numerical error of calculating $D(\varepsilon)$ scales also with $1/\sqrt{N_e}$. The Fermi energy μ is determined by $N_e = \int_{-\infty}^{\mu} D(\varepsilon)d\varepsilon$. In the case that N_e is not enough to provide a desired accuracy, additional average of $D(\varepsilon)$ from different random states ($|\varphi_p\rangle = \sum_i c_{i,p}|\boldsymbol{r_i}\rangle$, where $p = 1, 2, \ldots, S$) can be introduced to reduce the statistical error. Then, according to the central limit theory, the overall error of calculating $D(\varepsilon)$ scales as $1/\sqrt{SN_e}$.[31,34] Numerically, the time-evolution operator e^{-iHt} can be decomposed using the Chebyshev polynomial method as discussed in Refs. [31,34], which is unconditionally stable and leads to a linear scaling on the system size as the Hamiltonian H is a spare matrix in the DFT. The energy resolution is determined by $1/N_t\tau$, where N_t is the number of time steps and τ is the time interval dt.

As a numerical check, we calculated the DOS of graphite nanocrystals and fullerene with a different number of carbon atoms. Here, the Kohn–Sham Hamiltonian is constructed with a given initial electron density $\rho(\boldsymbol{r})$ as

$$H = -\frac{\nabla^2}{2} + V_{\text{ext}}[\rho(\boldsymbol{r})] + V_{\text{H}}[\rho(\boldsymbol{r})] + V_{\text{xc}}[\rho(\boldsymbol{r})], \tag{3}$$

where $-\nabla^2/2$ is the kinetic energy, V_{ext} is the external potential, V_{H} is the Hartree potential, and V_{xc} is the exchange and correlation potential. The kinetic energy in the KS-Hamiltonian [Eq. (3)] is approximated by using the higher-order finite-difference expansion for the Laplacian operator in a uniform real-space grid.[35,36] The Hartree potential V_{H} is derived by solving the Poisson equation.[36] For exchange and correlation potential V_{xc}, we use the local density approximation.[37] The full ionic potential V_{ext} is effectively replaced by pseudo-potential in Kleinman–Bylander forms.[38]

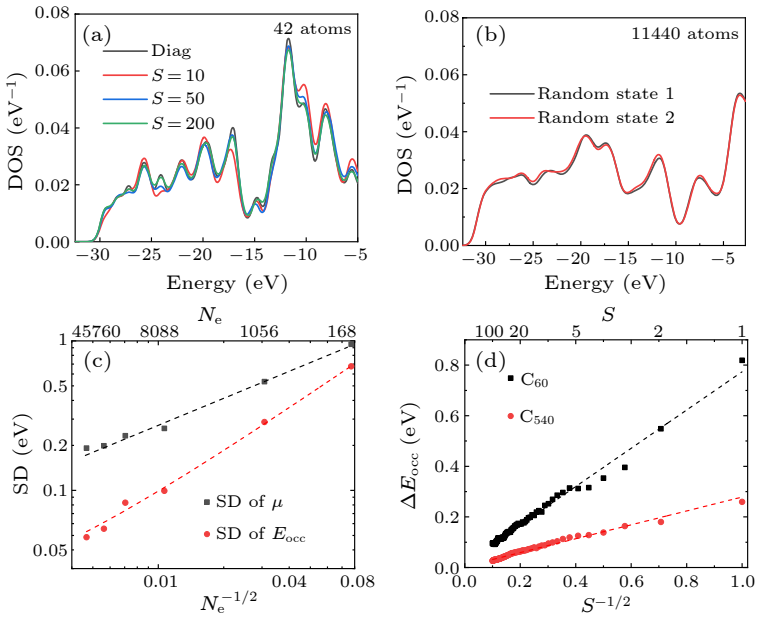

Fig. 1. (a) The DOS of a graphite nanocrystal with 42 atoms, calculated by exact diagonalization or using the TDRS method averaged with a different number (S) of random samples. (b) The DOS of a graphite nanocrystal with 11440 atoms was calculated using the TDRS method without random sample averaging. The different colors represent different initial random states. (c) The standard deviations (SD) of μ and $E_{\text{occ}}(\mu)$ for graphite nanocrystals with different size (N_e), obtained from the statistical analysis of results from 500 individual random states. (d) The error of $E_{\text{occ}}(\mu)$, with respect to the exact diagonalization, as a function of the number of random states (S) used in TDRS for fullerenes C_{60} and C_{540}. Here, each point in (d) is averaged from 100 groups of S random states.

In Fig. 1(a), we plot the DOS of a graphite nanocrystal with 42 carbon atoms, showing that the TDRS results with more random samples match better to the value from the diagonalization. For large graphite nanocrystals, such as the one with 11440 atoms shown in Fig. 1(b), the TDRS results obtained from two individual random states are

quite close to each other. As a quantitative measurement of the statistical error, the standard deviations (SDs) of results using only one random state are collected in Fig. 1(c). Here, in each case we considered 500 different random states and plotted the SDs of μ and $E_{\rm occ}(\mu)$ for graphite nanocrystals with different sizes, where $E_{\rm occ}(\mu)$ is the occupied energy defined as $E_{\rm occ}(\mu) = \int_{-\infty}^{\mu} \varepsilon D(\varepsilon) d\varepsilon$. It is clear that the SDs of μ and $E_{\rm occ}(\mu)$ reduce significantly when there are more electrons, and approach to zero with an error scales as $1/\sqrt{N_{\rm e}}$. This indicates that for very large systems, it is unnecessary to have additional averages with different random states, and thus using one random initial state is enough to provide a converged result in TDRS. On the other hand, for finite systems such as the fullerenes C_{60} and C_{540} shown in Fig. 1(d), the accuracy of using TDRS can be improved by the average using more random states and the result converges to the exact value (from diagonalization) with an error scale as $1/\sqrt{S}$. In total, our numerical results prove that the statistical error of using the TDRS method to calculate DOS or related variables scales as $1/\sqrt{SN_{\rm e}}$, just as we expected (also see the error analysis of DOS in the Supplementary Materials).

Electron Density. In the KS-DFT, the electron density is constructed by the superposition of occupied KS orbitals $|E_n\rangle$, namely $\rho(\boldsymbol{r}) = \sum_n f(E_n)||E_n(\boldsymbol{r})\rangle|^2$, where f is the Fermi–Dirac function. In the rsDFT, the knowledge of $|E_n\rangle$ is absent as we do not perform any diagonalization, but $\rho(\boldsymbol{r})$ will be obtained in a different way. In the basis of real-space grid, $\{|\boldsymbol{r_i}\rangle\}$, the wave functions of KS orbitals can be expressed as $|E_n\rangle = \sum_{i=1}^{N} a_i(E_n)|\boldsymbol{r_i}\rangle$, where N is the total number of grid points, i.e., the dimension of the Hamiltonian. We consider a random state $|\varphi\rangle$ in the real space, as the one used in Eq. (1), which is also a random superposition state in the energy space. Thus, a superposition of all occupied states, with a given Fermi energy μ and temperature T, can be constructed by applying a Fermi–Dirac (FD) filter on the random state as $|\varphi\rangle_{\rm FD} \equiv \sqrt{f(H)}|\varphi\rangle = \sum b_n \sqrt{f(E_n)}|E_n\rangle$, where $f(H) = 1/(e^{(H-\mu)/k_{\rm B}T} + 1)$ is the Fermi–Dirac operator. The intensity of state vector $|\varphi\rangle_{\rm FD}$ at grid $\boldsymbol{r_j}$ can be expressed as

$$\begin{aligned} \||\varphi\rangle_{\rm FD}(\boldsymbol{r_j})|^2 &= \sum_n |b_n|^2 f(E_n)|a_j(E_n)|^2 \\ &+ \sum_{m\neq n} b_m^* b_n f^{\frac{1}{2}}(E_m) f^{\frac{1}{2}}(E_n) a_j^*(E_m) a_j(E_n). \end{aligned} \tag{4}$$

As we see in the calculation of DOS, for a large but finite N, $|b_n| \to 1/\sqrt{N}$,[39] thus the first term in Eq. (4) converges to $\rho(\boldsymbol{r_j})/N$, where $\rho(\boldsymbol{r_j}) = \sum_n f(E_n)|a_j(E_n)|^2$ is exactly the electron density at grid $\boldsymbol{r}_j$. However, the value of this term is of the order of $O(N_{\rm e}/N^2)$, which is about $N_{\rm e}$ times smaller than the second term $\sim O(N_{\rm e}^2/N^2)$. This means that the second term in Eq. (4) becomes dominant when increasing the system size. To approximate $\rho(\boldsymbol{r_j})$ by using $\||\varphi\rangle_{\rm FD}(\boldsymbol{r_j})|^2$, one needs to reduce the second term significantly. This can be realized by using a time-dependent approach in the following way.

Introduce the electron density $\rho_{\rm RS}(\boldsymbol{r})$ using a time-dependent RS approach as

$$\rho_{\rm RS}(\boldsymbol{r_j}) \equiv \frac{N}{2\pi} \int_{-\infty}^{\infty} |e^{-iHt}|\varphi\rangle_{\rm FD}(\boldsymbol{r_j})|^2 dt, \tag{5}$$

and one can prove that $\rho_{\rm RS}(\boldsymbol{r})$ converges to $\rho(\boldsymbol{r})$ in the limit of $N \to \infty$. Defining $\delta_\rho \equiv \sum_j |\rho_{\rm RS}(\boldsymbol{r_j}) - \rho(\boldsymbol{r_j})|/N_{\rm e}$ as a measurement error of electron density and using the property $\int_{-\infty}^{\infty} e^{-i(E_n - E_m)t} dt = 2\pi\delta(E_n - E_m)$, we have

$$\begin{aligned} \delta_\rho = \frac{1}{N_{\rm e}} \sum_j \Big| & \sum_n (N|b_n|^2 - 1) f(E_n)|a_j(E_n)|^2 \\ & + 2\pi N \sum_{m\neq n} b_m^* b_n f^{\frac{1}{2}}(E_m) f^{\frac{1}{2}}(E_n) a_j^*(E_m) a_j(E_n) \delta \\ & \cdot (E_n - E_m) \Big|. \end{aligned} \tag{6}$$

For a large but finite N, $|b_n| \to 1/\sqrt{N}$, both terms in Eq. (6) converge to zero, but the statistical errors of the first and second terms are $O(1/\sqrt{N})$ and $O(1/N)$, respectively. This indicates that in the limit of $N \to \infty$, $\rho_{\rm RS}(\boldsymbol{r})$ converges to $\rho(\boldsymbol{r})$, and δ_ρ is dominated by the first term and reduces to zero with a statistical error $\sim O(1/\sqrt{N})$.

The accuracy of using Eq. (5) to obtain the electron density can be further improved by averaging $\rho_{\rm RS}(\boldsymbol{r})$ from different initial random states. Similar to the calculation of DOS, consider a set of random states $|\varphi_p\rangle = \sum_i c_{i,p}|\boldsymbol{r_i}\rangle$, according to the central limit theorem, for a large but finite S, $\sum_{p=1}^{S} c_{i,p}c_{i',p}/S = E(|c|^2)\delta_{i,i'} + O(1/\sqrt{S})$,[34] where $E(|c|^2) \sim 1/N$ is the expectation value of $|c_i|^2$. As $b_n = \sum_{i=1}^{N} c_i a_i^*(E_n)$, using the normalization property $\sum_{i=1}^{N} |a_i(E_n)|^2 = 1$ and the orthogonal property $\sum_{i=1}^{N} a_i(E_n)a_i^*(E_m) = 0$ for $n \neq m$, we have $\sum_p b_{m,p}^* b_{n,p}/S = \delta_{m,n}/N + O(1/\sqrt{S})$. Thus, together with extra random states average, $\rho_{\rm RS}(\boldsymbol{r})$ is an accurate approximation of $\rho(\boldsymbol{r})$ with a statistical error δ_ρ scales as $1/\sqrt{SN_{\rm e}}$. This scaling behavior is indeed the same as the calculations of DOS given in Eq. (1).

Here, the time-evolution operator e^{-iHt} and the Fermi–Dirac filter $\sqrt{f(H)}$ are carried out numerically using the Chebyshev polynomials,[31] which is very efficient and accurate for sparse matrix H as we discussed in the part of DOS. In the Chebyshev decomposition of the Fermi–Dirac filter, the temperature must be finite, and we use $T = 10K$ in all the calculations. The method introduced here becomes more efficient at high temperatures, in which the ground state calculations also involve many states above the Fermi energy due to a nonzero occupation probability given by the Fermi–Dirac distribution. In the following, we show several examples and check the accuracy of obtaining electron density using the TDRS method introduced in Eq. (5).

We first consider fullerene with different numbers of carbon atoms and calculate the charge density for a given Hamiltonian using either the TDRS method or the standard diagonalization. In Fig. 2(a), we plot the statistical error δ_ρ as a function of the number of time steps N_t for C_{20}, C_{60}, C_{180}, C_{240}, and C_{540}, where the reference charge density ρ in each case is the one obtained from the diagonalization. We see that in all cases, δ_ρ drops rapidly when introducing the time evolution, and for a given evolution period (the same N_t), δ_ρ is always smaller in a larger sample. In Fig. 2(b), we plot the minimum value of δ_ρ as a function of the number of electrons $N_{\rm e}$ in the limit of $N_t \to \infty$. It shows clearly that δ_ρ scales exactly as $O(1/\sqrt{N_{\rm e}})$, as same as we predicted from Eq. (6). The

error bars are the standard deviations of electron density obtained from each individual random initial state, i.e., without any random sample averaging. In practice, it might be numerically expensive to perform a very long evolution, and thus one needs to consider using only finite or relatively small N_t. Thus we plot in Fig. 2(c) the similar results to Fig. 2(b) but with finite N_t. We see that: (1) in the absence of the time-dependent approach ($N_t = 0$), the value of δ_ρ keeps the same amplitude, independent of the system size; (2) when the time evolution is introduced, the value of δ_ρ starts to decrease when increasing the size of the sample (N_e). In Fig. 2(d), we consider the influence of sample average on the statistical error by plotting the value of δ_ρ as a function of sample number S. Here the reference charge density is the one obtained from C_{60} with exact diagonalization, and we see that the scaling of the error follows $1/\sqrt{S}$, for both cases with or without the time-dependent approach. The results presented in Fig. 2 verified numerically that the statistical error of calculating the charge density using the TDRS method scales as $1/\sqrt{SN_e}$, with the same scaling behavior as the calculation of DOS.

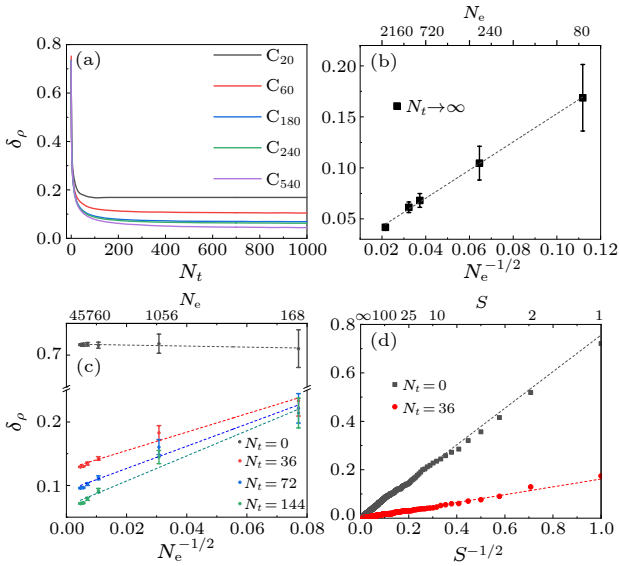

Fig. 2. (a) The statistical error δ_ρ as a function of the evolution time N_t for C_{20}, C_{60}, C_{180}, C_{240} and C_{540}, where $\rho_{\rm Ref}$ is the charge density obtained from diagonalization. The time step is $\tau = 64\pi$ and the number of random samples $S = 20$. (b) The values of δ_ρ in (a) in the limit of $N_t \to \infty$, plotted as a function of N_e. The error bars indicate the corresponding standard deviations. (c) The same as (b) but with finite N_t, where the reference charge density is the mean value averaged from 20 random samples with $N_t = 144$. $N_t = 0$ means no time-dependent approach is involved in calculating charge density. (d) The value of δ_ρ, plotted as a function of sample number S for C_{60}, without or with time-dependent approach ($\tau = 64\pi$).

Self-Consistent Iteration. Now, we consider the detailed self-consistent iterations in the rsDFT. Here, the numerical results obtained from the widely used commercial KS-DFT package VASP (Vienna *ab initio* simulation package) [40] are also presented as references.

In Fig. 3(a), we show the ground state DOS of C_{60} calculated using the rsDFT, standard KS-DFT (with diagonalization) and VASP, respectively. Pulay mix is adopted in the iteration to optimize the input density and accelerate the convergence. [41] In the rsDFT, we use 10 random samples in DOS calculations and 36 random samples with $N_t = 36$, $\tau = 64\pi$ in electron density calculations. The DOS values obtained from different approaches agree well with each other, indicating that (1) our DFT code based on diagonalization correctly reproduces the ground state from VASP, (2) the newly proposed rsDFT provides accurate results as these can be obtained from the standard KS-DFT and the diagonalization can be completely ruled out in the entire iteration process. In Fig. 3(b), we present more results of converged rsDFT calculations of fullerenes with different sizes, and show the total energy difference compared with the result from diagonalization in each iteration step. In general, the convergence to the ground state is more difficult in the larger system in the KS-DFT, requiring more iteration steps to reach the same accuracy for the total energy. In the rsDFT, however, the accuracy is increased automatically in larger systems due to the $1/\sqrt{SN_e}$ dependence of the statistical error, as shown clearly in converge of the total energy during the iterations in Fig. 3(b).

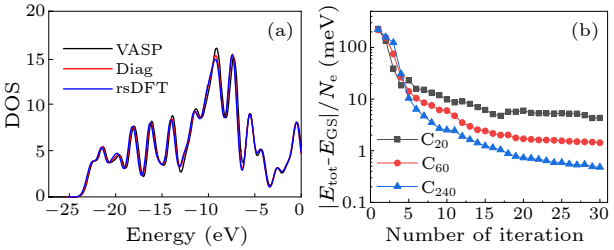

Fig. 3. (a) For C_{60}, the eigenvalue distributions of the ground state calculated by using the VASP, diagonalization and rsDFT, respectively. (b) The convergence of the total energy during the iteration for C_{20}, C_{60}, and C_{240}. Here the total energy of the ground state ($E_{\rm GS}$) is obtained from diagonalization.

Scaling Behavior. Lastly, we check the scaling behavior of the rsDFT. The non-diagonal elements of KS-Hamiltonian are from the kinetic energy and non-local pseudopotentials in Eq. (3), leading to a highly sparse Hamiltonian matrix due to the locality. The number of the nonzero elements in the matrix scales linearly with the number of atoms in the system. In the rsDFT, the basic and dominant calculations are the multiplications between the Hamiltonian matrix and a state vector, in which the number of operations scales linearly with the nonzero elements in the matrix. Therefore, the total computational load (CPU time) and memory cost of the rsDFT are linearly dependent on the dimension of the Hamiltonian, which is proportional to the number of atoms in the system. These linear-scaling behaviors are verified using graphite crystal, with up to 11440 carbon atoms (45760 electrons) in Figs. 4(a) and 4(b). As benchmark tests, we perform complete ground state calculations of graphite with 5040 and 11440 carbon atoms, respectively. The difference of input and output electron densities ($\Delta_\rho \equiv \sum_i |\rho_{\rm out}(\boldsymbol{r_i}) - \rho_{\rm in}(\boldsymbol{r_i})|/N$) as a function of iterative steps is plotted in Figs. 4(c) and 4(d). In all

cases, we used the same calculation parameters, such as the same real-space grid density, and the same accuracy for the Chebyshev decompositions of the time-evolution operator and the Fermi–Dirac filter. One should notice that, for 11440 carbon atoms, the total memory cost is only 16 GB. Due to the linear-scaling behavior, a self-consistent calculation of millions of atoms should be possible for computer clusters with Terabytes (TB) memory.

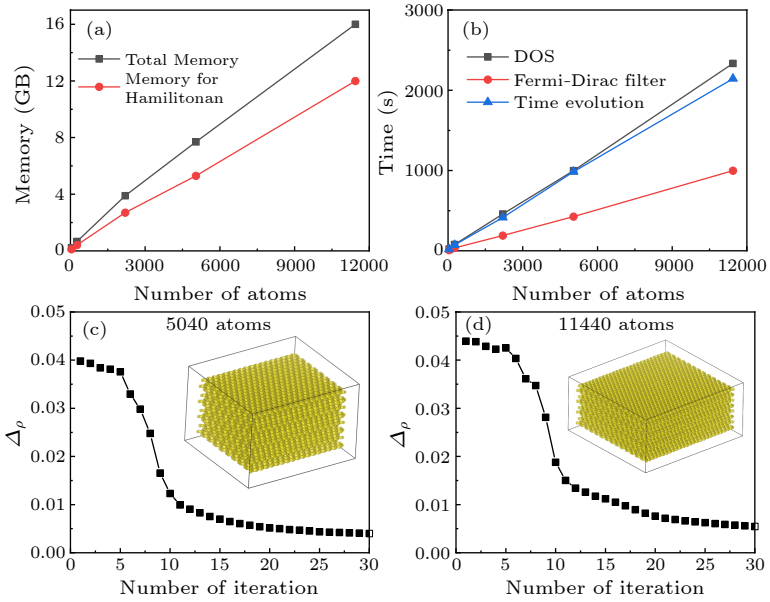

Fig. 4. [(a), (b)] The cost of memory and CPU time in the rsDFT for A-B stacked graphite with different numbers of carbon atoms. The black and red curves in (a) are the total memory cost and the memory used to store the sparse Hamiltonian matrix. The black, red and blue curves in (b) correspond to the time cost of one iteration for the calculations of DOS, Fermi–Dirac filter and time-evolution averaging, respectively. [(c), (d)] The difference of input and output electron density as a function of iteration steps for A-B stacked graphite with 5040 and 11440 atoms, respectively. The insets indicate the converged ground state density.

Conclusion and Discussion. We developed a sublinear-scaled self-consistent first-principle calculation method, i.e., the rsDFT. We use a TDRS method to calculate the density of states and determine the Fermi energy. A Fermi–Dirac filter on a random state and, subsequently, a wave propagation according to the time-dependent Schrödinger equation are introduced to approximate the spatial distribution of the electron density. The accuracy can be improved by the average using different initial random states. The overall numerical error of either a global quantity or a local variable scales as $1/\sqrt{SN_\mathrm{e}}$, where N_e is the number of electrons and S is the number of random states. It leads to an overall sublinear scaling of the computational costs, as for larger systems, one needs fewer random states for the sample average. The method becomes extremely powerful when $N_\mathrm{e} \to \infty$, and a calculation using one random state is enough to achieve reasonable accuracy.

In the recently developed stochastic DFT (sDFT),[9,27–30] the physical quantities, such as the Fermi energy and electron density, are calculated using the trace formula of the stochastic technique without diagonalization, i.e., the trace of a variable is approximated by the average of its expectation values in stochastic orbitals. One of the main differences between rsDFT and sDFT is that all the variables in the rsDFT are obtained in a more deterministic way based on numerical solutions of the time-dependent Schrödinger equation, which is absent in the sDFT. In particular, the dominant noise in the electron density induced by occupied states after the Fermi–Dirac filter is dramatically reduced by the time-dependent approach in the rsDFT. A large number of stochastic orbitals (random samples) is required to reduce the statistical error of the electron density in the sDFT. To keep the same accuracy of calculating the local variables such as electron density, one cannot use fewer stochastic orbitals for larger systems, even in the limit of $N_\mathrm{e} \to \infty$.

In our earlier works, we have developed a so-called tight-binding propagation method (TBPM) for large-scale modelling of complex quantum systems. A direct extension of TBPM in the rsDFT is straightforward. For example, by using the TDRS-based method without diagonalization, one can calculate the electronic and optical conductivities,[42,43] polarization and screening functions,[32,44] diffusion coefficient and localization length,[33,44] quasieigenstate,[45] and many other applications as implemented in our homemade simulation package, TBPLaS.[46] The main advantage of the rsDFT and the other time-dependent methods mentioned above is that there is no diagonalization of the Hamiltonian matrix in the whole process, and the errors of these calculated variables all scale as $1/\sqrt{SN_\mathrm{e}}$, leading to an overall sublinear scaling on the computational costs. Furthermore, the atomic force can be calculated the same way as the traditional DFT or OF-DFT by using the formulation based on the electron density in real space.[47–50] It is also possible to extract the atomic force using the wave function of approximated ground states, which will be discussed in future work. The rsDFT provides a new possibility to

study large-scale systems from the first-principle calculations and can be used widely in physics, chemistry, biology and material science, with possible extension to large-scale TD-DFT and GW calculations.

Acknowledgements. S.Y. thanks Shiwu Gao, Hans De Raedt, Mikhail Katsnelson, Guodong Yu, and Yalei Zhang for many helpful discussions. This work was supported by the National Nature Science Foundation of China (Grant No. 11974263), and the Supercomputing Center of Wuhan University.

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Chelikowsky J R, Jing X, Wu K, and Saad Y 1996 *Phys. Rev. B* **53** 12071
[4] Edelman A and Smith S T 1996 *BIT Numer. Math* **36** 494
[5] Teter M P, Payne M C, and Allan D C 1989 *Phys. Rev. B* **40** 12255
[6] Payne M C, Teter M P, Allan D C, Arias T A, and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
[7] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[8] Goedecker S 1999 *Rev. Mod. Phys.* **71** 1085
[9] Baer R, Neuhauser D, and Rabani E 2013 *Phys. Rev. Lett.* **111** 106402
[10] Jay L O, Kim H, Saad Y, and Chelikowsky J R 1999 *Comput. Phys. Commun.* **118** 21
[11] Yang W T 1991 *Phys. Rev. Lett.* **66** 1438
[12] Baroni S and Giannozzi P 1992 *EPL (Europhys. Lett.)* **17** 547
[13] Li X P, Nunes R W, and Vanderbilt D 1993 *Phys. Rev. B* **47** 10891
[14] Sánchez-Portal D, Ordejon P, Artacho E, and Soler J M 1997 *Int. J. Quantum Chem.* **65** 453
[15] Mohr S, Ratcliff L E, Genovese L, Caliste D, Boulanger P, Goedecker S, and Deutsch T 2015 *Phys. Chem. Chem. Phys.* **17** 31360
[16] Goringe C M, Hernández E, Gillan M J, and Bush I J 1997 *Comput. Phys. Commun.* **102** 1
[17] Hernández E, Gillan M J, and Goringe C M 1996 *Phys. Rev. B* **53** 7147
[18] Hine N D, Haynes P D, Mostofi A A, Skylaris C K, and Payne M C 2009 *Comput. Phys. Commun.* **180** 1041
[19] VandeVondele J, Krack M, Mohamed F, Parrinello M, Chassaing T, and Hutter J 2005 *Comput. Phys. Commun.* **167** 103
[20] Ghosh S and Suryanarayana P 2017 *Comput. Phys. Commun.* **216** 109
[21] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón P, and Sánchez-Portal D 2002 *J. Phys. Condens. Matter* **14** 2745
[22] Michaud-Rioux V, Zhang L, and Guo H 2016 *J. Comput. Phys.* **307** 593
[23] Wesolowski T A and Wang Y A 2013 *Recent Progress in Orbital-Free Density Functional Theory* (Singapore: World Scientific)
[24] Lignères V L and Carter E A 2005 *An introduction to Orbital-Free Density Functional Theory* (Berlin: Springer) p 137
[25] Zhou B, Ligneres V L, and Carter E A 2005 *J. Chem. Phys.* **122** 044103
[26] Wang Y A and Carter E A 2002 *Orbital-Free Kinetic-Energy Density Functional Theory* (Berlin: Springer) p 117
[27] Chen M, Baer R, Neuhauser D, and Rabani E 2019 *J. Chem. Phys.* **150** 034106
[28] Chen M, Baer R, Neuhauser D, and Rabani E 2019 *J. Chem. Phys.* **151** 114116
[29] Chen M, Baer R, Neuhauser D, and Rabani E 2021 *J. Chem. Phys.* **154** 204108
[30] White A J and Collins L A 2020 *Phys. Rev. Lett.* **125** 055002
[31] Yuan S, De Raedt H, and Katsnelson M I 2010 *Phys. Rev. B* **82** 115448
[32] Yuan S, Roldán R, and Katsnelson M I 2011 *Phys. Rev. B* **84** 035439
[33] Logemann R, Reijnders K J A, Tudorovskiy T, Katsnelson M I, and Yuan S 2015 *Phys. Rev. B* **91** 045420
[34] Hams A and De Raedt H 2000 *Phys. Rev. E* **62** 4365
[35] Chelikowsky J R, Troullier N, and Saad Y 1994 *Phys. Rev. Lett.* **72** 1240
[36] Chelikowsky J R, Troullier N, Wu K, and Saad Y 1994 *Phys. Rev. B* **50** 11355
[37] Vosko S H, Wilk L, and Nusair M 1980 *Can. J. Phys.* **58** 1200
[38] Kleinman L and Bylander D M 1982 *Phys. Rev. Lett.* **48** 1425
[39] Jin F, Willsch D, Willsch M, Lagemann H, Michielsen K, and De Raedt H 2021 *J. Phys. Soc. Jpn.* **90** 012001
[40] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
[41] Pulay P 1982 *J. Comput. Chem.* **3** 556
[42] Yuan S, Rösner M, Schulz A, Wehling T O, and Katsnelson M I 2015 *Phys. Rev. Lett.* **114** 047403
[43] Yuan S, De Raedt H, and Katsnelson M I 2010 *Phys. Rev. B* **82** 235409
[44] Yuan S, Wehling T O, Lichtenstein A I, and Katsnelson M I 2012 *Phys. Rev. Lett.* **109** 156601
[45] Shi H, Zhan Z, Qi Z *et al.* 2020 *Nat. Commun.* **11** 1
[46] Li Y, Zhan Z, Kuang X, Li Y, and Yuan S 2023 *Comput. Phys. Commun.* **285** 108632
[47] Shao X, Xu Q, Wang S, Lv J, Wang Y, and Ma Y 2018 *Comput. Phys. Commun.* **233** 78
[48] Ho G S, Lignères V L, and Carter E A 2008 *Comput. Phys. Commun.* **179** 839
[49] Golub P and Manzhos S 2020 *Comput. Phys. Commun.* **256** 107365
[50] Arnon E, Rabani E, Neuhauser D, and Baer R 2017 *J. Chem. Phys.* **146** 224111