

Supplementary Materials: A Time-Dependent Random State Approach for Large-scale Density Functional Calculations

Weiying Zhou¹ and Shengjun Yuan^{1,*}

¹*Key Laboratory of Artificial Micro- and Nano-structures of Ministry of Education and School of Physics and Technology, Wuhan University, Wuhan 430072, China*

1. Higher-order Finite-difference Pseudopotential Method

Within the non-relativistic Kohn–Sham DFT, the ground state of a system of N_e electrons subject to an external potential can be obtained by solving a set of one-particle equations, the Kohn–Sham equations (atomic units will be used throughout):

$$\left[-\frac{\nabla^2}{2} + V_{KS}[\rho(\mathbf{r})]\right]\varphi_i(\mathbf{r}) = \varepsilon_i\varphi_i(\mathbf{r}) \quad (1)$$

where Kohn-Sham potential $V_{KS}[\rho(\mathbf{r})]$ is usually divided as:

$$V_{KS}[\rho(\mathbf{r})] = V_{ext}[\rho(\mathbf{r})] + V_H[\rho(\mathbf{r})] + V_{xc}[\rho(\mathbf{r})] \quad (2)$$

where V_{ext} is the external potential, V_H is the Hartree potential, and V_{xc} is the exchange and correlation potential. In this paper, we implement real-space finite-element methods, resulting in $V_{KS}[\rho(\mathbf{r})] = V_{KS}(\mathbf{r})$.

In our letter, we impose a simple, uniform orthogonal three-dimensional (3D) grid where the points are described in a finite domain by (x_i, y_j, z_k) [1]. Kinetic-energy operator can be described by high-order finite-element difference method [2],

$$\frac{\partial^2\varphi}{\partial x^2} = \sum_{n=-N_h}^{N_h} C_n\varphi(x_i + nh, y_j, z_k) + O(h^{2N_h+2}) \quad (3)$$

where h is the grid spacing and N_h is the order of finite-element difference. Expansion coefficients C_n for a uniform grid are given in Table. S1 [2].

TABLE S1: Expansion coefficients C_n for higher-order finite-difference expressions of the second derivative.

	C_i	$C_{i\pm 1}$	$C_{i\pm 2}$	$C_{i\pm 3}$	$C_{i\pm 4}$	$C_{i\pm 5}$	$C_{i\pm 6}$
$N_h = 1$	-2	1					
$N_h = 2$	$-\frac{5}{2}$	$\frac{4}{3}$	$-\frac{1}{12}$				
$N_h = 3$	$-\frac{49}{18}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$			
$N_h = 4$	$-\frac{205}{72}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$		
$N_h = 5$	$-\frac{5269}{1800}$	$\frac{5}{3}$	$-\frac{5}{21}$	$\frac{5}{126}$	$-\frac{5}{1008}$	$\frac{1}{3150}$	
$N_h = 6$	$-\frac{5369}{1800}$	$\frac{12}{7}$	$-\frac{15}{56}$	$\frac{10}{189}$	$-\frac{1}{112}$	$\frac{2}{1925}$	$-\frac{1}{16632}$

The Hartree energy density and potential are given by:

$$\varepsilon_H(\mathbf{r}) = \frac{1}{2} \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (4)$$

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (5)$$

The Hartree potential V_H could be obtained by solving Poisson's equation.

For the exchange-correlation part, we use local-density approximation (LDA):

$$E_{xc}[n] = \int d\mathbf{r} n(\mathbf{r})\varepsilon_{xc}(n(\mathbf{r})) \quad (6)$$

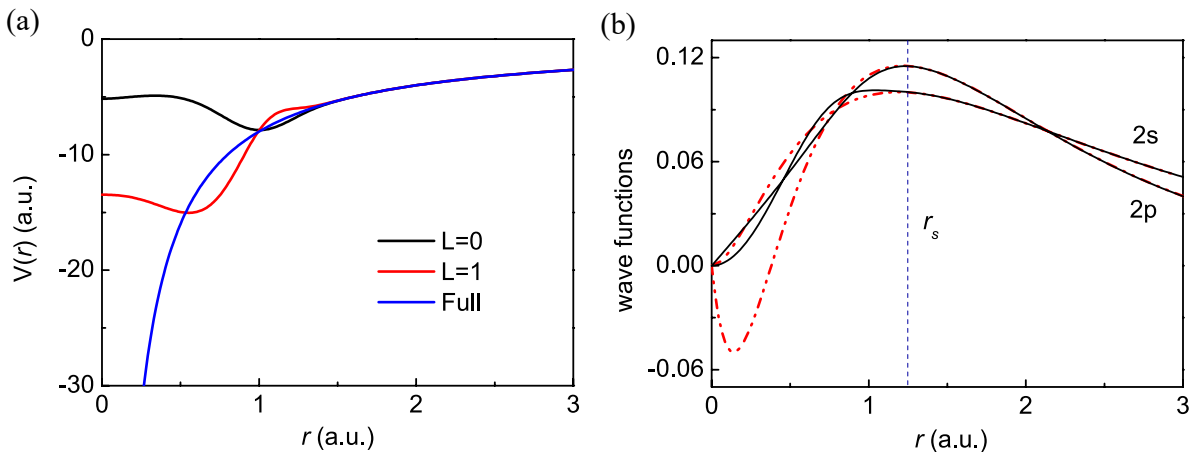


FIG. S1: (a) pseudopotential of carbon atom generated by program ATOM [5]. (b) atomic wavefunction (Red dashed curve) and pseudo-wavefunction (black solid curve) of carbon.

The most accurate formulae for the exchange-correlation functional were obtained by fitting the QMC results for the Jellium model. Various parameterizations are available. We use one of the most popular choices proposed by Vosko-Wilk [3]. From the Jellium model, the local part of the exchange is given by :

$$\begin{aligned} \varepsilon_x &= -\frac{3}{4} \left(\frac{3}{2\pi} \right)^{2/3} \frac{1}{r_s}, \\ V_x &= \left(\frac{3}{2\pi} \right)^{2/3} \frac{1}{r_s}. \end{aligned} \quad (7)$$

For $V_{ext}(r)$, we use full ionic potential $-\frac{Z}{r}$ for the cases of single atoms. For other systems, we implement a pseudopotential operator to reduce the computational demand. We use the projection scheme of the pseudopotential operator suggested by Kleinman and Bylander [4]:

$$V_{ps}(r) = \sum_{a=1}^N \left[V_{a,ps}^{loc}(r) + \frac{|\Delta V_{a,ps}^l(r) \varphi_{lm}^a\rangle \langle \Delta V_{a,ps}^l(r) \varphi_{lm}^a(r)|}{\langle \varphi_{lm}^a(r) | \Delta V_{a,ps}^l(r) | \varphi_{lm}^a(r) \rangle} \right] \quad (8)$$

where the total pseudopotential can be divided into non-local and local part $\Delta V_{a,ps}^l(\mathbf{r}) \equiv V_{a,ps}^l(\mathbf{r}) - V_{a,ps}^{loc}(\mathbf{r})$. $V_{a,ps}^{loc}$ is the local part with specific angular momentum l component of atom a , which differs from zero only in the region smaller than the cutoff radius $r < r_c$. φ_{lm}^a is the atomic pseudo wave function with lm quantum angular momentum numbers. It is worth noticing that the pseudopotential operator only needs to be calculated once at the very beginning since it only depends on the atomic configuration. Taking the carbon atom as an example, we use the program ATOM [5] to generate its pseudopotential. The type of pseudopotential is chosen as local density approximation (LDA) [6] and plotted in Fig. S1. In Fig. S1 (b), we show the atomic pseudo-wave-function, and indeed it is the same as a full-potential wave-function in the range of $r > r_s$ where r_s is the cutoff radius. The construction of charge density has been described in the main context.

2. Another Fermi-Dirac Filter

In this part, we add some detailed discussion of the methods used in rsDFT. First, we construct a random superposition state in a uniform real-space grid as an initial state,

$$|\varphi_0\rangle = \sum_{i=1}^N c_i |\mathbf{r}_i\rangle, \quad (9)$$

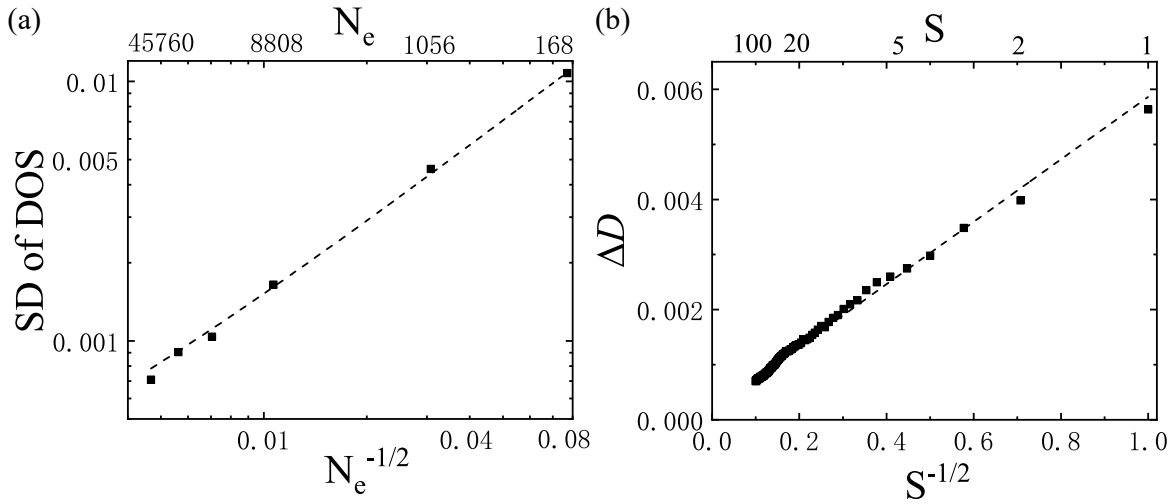


FIG. S2: The statistical error of calculated DOS as a function of the number of electrons for graphite nanocrystals (a) or the number of random states for C_{540} (b). In (a), the standard deviation of the DOS spectrum $\int_{-\infty}^{\infty} |D(\varepsilon) - \langle D(\varepsilon) \rangle| d\varepsilon$ is calculated based on the results from 500 individual random states, where $\langle D(\varepsilon) \rangle$ is the mean value of $\{D_k(\varepsilon)\}$ with $k = 1, 2, \dots, 500$. In (b), the error is defined as $\Delta D \equiv \int_{-\infty}^{\infty} |D(\varepsilon) - D_{Diag}(\varepsilon)| d\varepsilon$, where $D_{Diag}(\varepsilon)$ is the result obtained from the diagonalization, and each point is averaged from 100 groups of S random states.

where N is the number of grid, $\{\mathbf{r}_i\}$ are the real space basis states, and $\{c_i\}$ are random complex numbers. Assuming that

$$|E_n\rangle = \sum_i^N a_i(E_n) |\mathbf{r}_i\rangle, \quad (10)$$

we have

$$\begin{aligned} |\varphi_0\rangle &= \sum_{i=1}^N c_i \sum_{n=1}^N |E_n\rangle \langle E_n | \mathbf{r}_i \rangle \\ &= \sum_{i=1}^N \sum_{n=1}^N c_i a_i^*(E_n) |E_n\rangle \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{n=1}^N c_i a_i^*(E_n) |\mathbf{r}_j\rangle \langle \mathbf{r}_j | E_n \rangle \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{n=1}^N c_i a_i^*(E_n) a_j(E_n) |\mathbf{r}_j\rangle. \end{aligned} \quad (11)$$

Now we consider another type of Dirac-Fermi filter different from the one introduced in the main text:

$$\begin{aligned} |\varphi\rangle_{fd} &\equiv f(H) |\varphi_0\rangle \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{n=1}^N c_i a_i^*(E_n) f(E_n) a_j(E_n) |\mathbf{r}_j\rangle \end{aligned} \quad (12)$$

In the inner product of $\langle \varphi_0 | \varphi \rangle_{fd}$ at grid \mathbf{r}_j can be calculated by using Eq. (11) and Eq. (12),

$$\begin{aligned}
\rho_{fd}(\mathbf{r}_j) &= {}_0 \langle \varphi | \mathbf{r}_j \rangle \langle \mathbf{r}_j | \varphi \rangle_{fd} \\
&= \sum_{i,i'=1}^N \sum_{n,m=1}^N c_i^* a_i(E_n) a_j^*(E_n) c_{i'} a_{i'}^*(E_m) f(E_m) a_j(E_m) \\
&= \sum_{i,i'=1}^N \sum_{n=m}^N c_i^* a_i(E_n) c_{i'} a_{i'}^*(E_n) f(E_n) |a_j(E_n)|^2 \\
&+ \sum_{i,i'=1}^N \sum_{n \neq m}^N c_i^* a_i(E_n) a_j^*(E_n) c_{i'} a_{i'}^*(E_m) f(E_m) a_j(E_m)
\end{aligned} \tag{13}$$

According to the central limit theorem, for a large but finite number (S) of the random states $|\varphi_p\rangle = \sum_i c_{i,p} |\mathbf{r}_i\rangle$, we have

$$\frac{1}{S} \sum_{p=1}^S c_{i,p} c_{i',p} = E(c^2) \delta_{i,i'} + O\left(\frac{1}{\sqrt{S}}\right). \tag{14}$$

Therefore, one proves that

$$\begin{aligned}
&\lim_{S \rightarrow \infty} \frac{1}{S} \sum_{p=1}^S \langle \varphi_p | \mathbf{r}_j \rangle \langle \mathbf{r}_j | \varphi_p \rangle_{fd} \\
&= \sum_{i=1}^N \sum_{n=1}^N E(|c|^2) f(E_n) |a_i(E_n)|^2 |a_j(E_n)|^2 \\
&+ \sum_{i=1}^N \sum_{n \neq m}^N E(|c|^2) f(E_m) a_i(E_n) a_i^*(E_m) a_j^*(E_n) a_j(E_m) \\
&= \sum_{i=1}^N |a_i(E_n)|^2 \sum_{n=1}^N E(|c|^2) f(E_n) |a_j(E_n)|^2 \\
&= \frac{1}{N} \sum_{n=1}^N f(E_n) |a_j(E_n)|^2 \\
&= \frac{1}{N} \rho_{diag}(\mathbf{x}_j)
\end{aligned} \tag{15}$$

here we used the normalization property of KS orbitals

$$\sum_{i=1}^N |a_i(E_n)|^2 = 1 \tag{16}$$

and the orthogonal property

$$\sum_{i=1}^N a_i(E_n) a_i^*(E_m) = 0 \tag{17}$$

for $m \neq n$. Eq. (15) indicates that

$$\rho_{fd}(\mathbf{r}_j) \equiv \frac{N}{S} \sum_{p=1}^S \langle \varphi_p | \mathbf{r}_j \rangle \langle \mathbf{r}_j | \varphi_p \rangle_{fd} \tag{18}$$

is an approximation of the charge density at \mathbf{r}_j with an error vanishes as $1/\sqrt{S}$, which can be verified in the zoom-in figure of Fig. S3 .

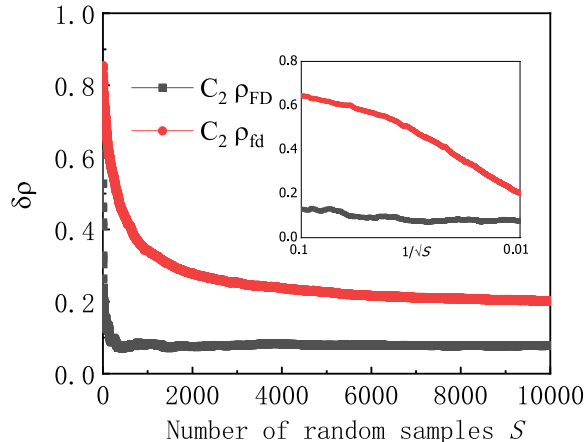


FIG. S3: The difference between the electron density obtained from KS-DFT ρ_{diag} and rsDFT as a function of the number of random samples S . In rsDFT, the results are obtained by using ρ_{FD} (Eq. (4) in the main context) or ρ_{fd} (Eq. (13)). Using $\sqrt{f(H)}$ instead of $f(H)$ in the Fermi-Dirac filter will significantly reduce the statistical error.

However, the converge of $\rho_{fd}(\mathbf{r}_j)$ using the Fermi-Dirac filter $f(H)$ is slower than the one using double $\sqrt{f(H)}$ introduced in the main text. The reason is that, in Eq. (13), the sum in the second term ($n \neq m$) involves all unoccupied states associated with the index n , and their number is several orders larger than the number of occupied states because $N \gg N_e$. To overcome this difficulty, we introduce $|\varphi\rangle_{FD} = \sqrt{f(H)}|\varphi\rangle_0$, the one used in the main context. The main advantage of using $\rho_{FD}(\mathbf{r}_j)$ is that the sums in the second term ($n \neq m$) of Eq. (4) of the main context includes only occupied states, leading to a much faster convergence compared with Eq. (13) (see Fig. S3).

3. Chebyshev Polynomials Method

In the numerical calculation, the operators $\frac{1}{\sqrt{e^{\beta(H-\mu)}+1}}$ and e^{-iHt} are approximated by using the Chebyshev polynomial method. In general, a function $f(x)$ whose values are in the range $[-1,1]$ can be expressed as,

$$f(x) = \frac{1}{2}c_0T_0(x) + \sum_{k=1}^{\infty} c_kT_k(x) \quad (19)$$

where $T_k(x) = \cos(k \arccos x)$ and the coefficients c_k are

$$c_k = \frac{2}{\pi} \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} f(x) T_k(x) \quad (20)$$

if we let $x = \cos\theta$, then $T_k(x) = T_k(\cos\theta) = \cos k\theta$, and

$$\begin{aligned} c_k &= \frac{2}{\pi} \int_0^\pi f(\cos\theta) \cos k\theta d\theta \\ &= \text{Re} \left[\frac{2}{N} \sum_{n=0}^{N-1} f\left(\cos \frac{2\pi n}{N}\right) e^{2\pi ink/N} \right], \end{aligned} \quad (21)$$

which can be calculated by the fast Fourier transform (FFT). We normalize H such that $\tilde{H} = H/\|H\|$ has eigenvalues in the range $[-1,1]$ and put $\tilde{\beta} = \beta/\|\beta\|$. Then

$$f(\tilde{H}) = \sum_{k=0}^{\infty} c_k T_k(\tilde{H}) \quad (22)$$

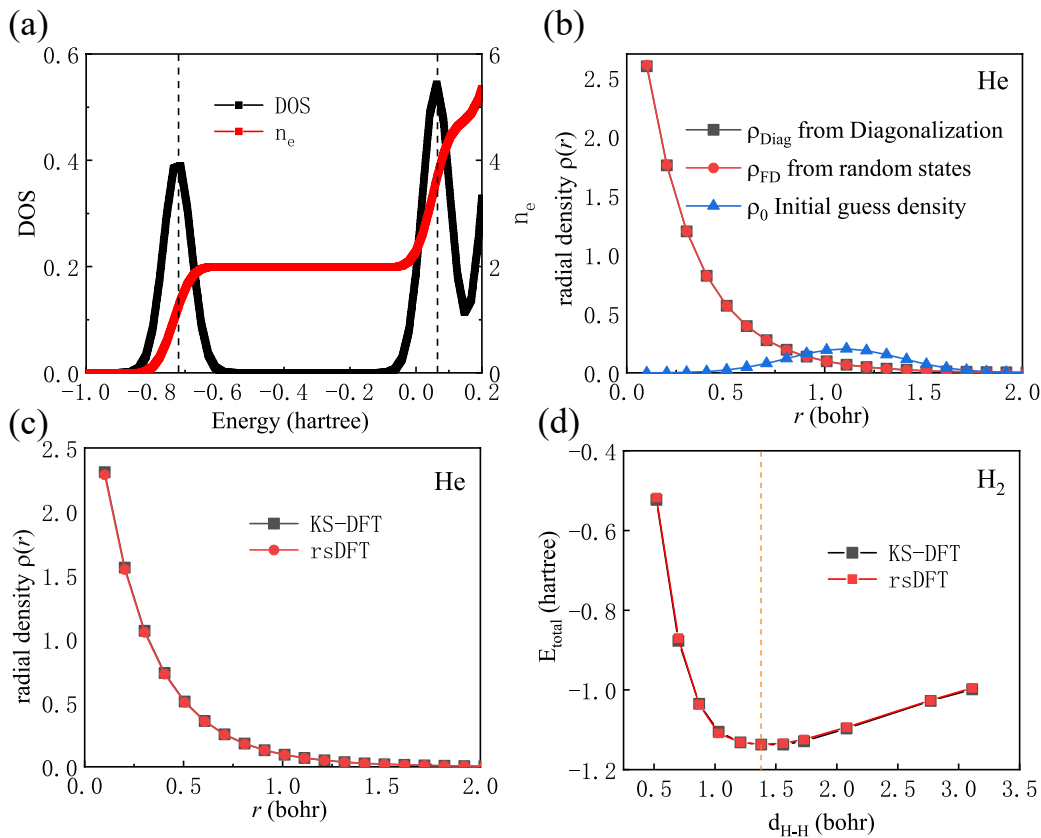


FIG. S4: (a) DOS and carrier density of one He atom calculated using the time-dependent random state method of one He atom. The vertical dashed lines indicate the energies obtained from the diagonalization. (b) The output electron density after one iteration using diagonalization (black) and Fermi-Dirac filter on random states (red) from the same input electron density (blue). For the He atom, due to the spherical symmetry, the electron density is only a function of distance (r) from the center of the atom. (c) The converged ground-state electron density of the He atom was obtained from KS-DFT and rsDFT. (d) The total energy as a function of bond length of molecular H_2 from obtained from KS-DFT and rsDFT, respectively.

where the Chebyshev polynomial $T_k(x)$ is the Chebyshev polynomial of the first kind. $T_k(x)$ obeys the following recurrence relation:

$$T_{k+1}(x) + T_{k-1}(x) = 2xT_k(x)$$

with

$$T_0(x) = 1, T_1(x) = x.$$

In Table. S2, we present the number of nonzero Bessel function (N_{Bessel}) as a function of time step τ , and the corresponding number of total matrix-vector operations $N_{\text{operations}}$ for the same propagation time $T = 1024\pi$. We can see a larger τ leads to fewer operations with the same total propagation time.

4. SINGLE ATOM

Let us consider a single Helium atom. In step (I), we construct a KS-Hamiltonian based on an initial electron density. In step (II), we obtain DOS $D(\varepsilon)$ by using the time-evolution method without the diagonalization of the Hamiltonian matrix and subsequently determine the Fermi level μ (see Fig. S4(a)).

As a comparison, the energies of KS orbitals from the diagonalization are also shown in Fig. S4(a), which agree very well with our results. In step (III), as there is only one occupied state, one can just use ρ_{FD} to obtain the electron

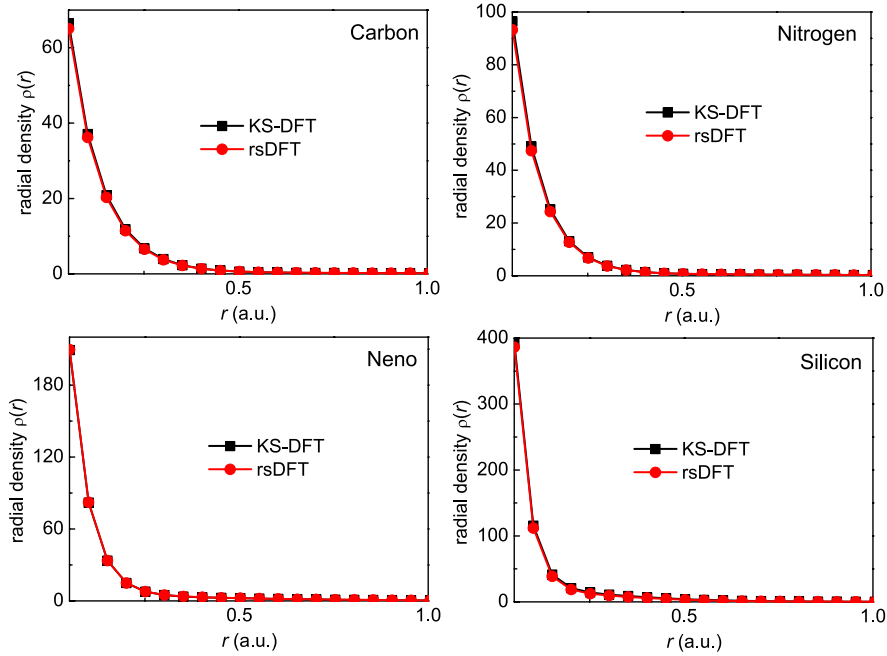


FIG. S5: Comparisons of ground-state charge density calculated by KS-DFT based on diagonalization and rsDFT of single atoms.

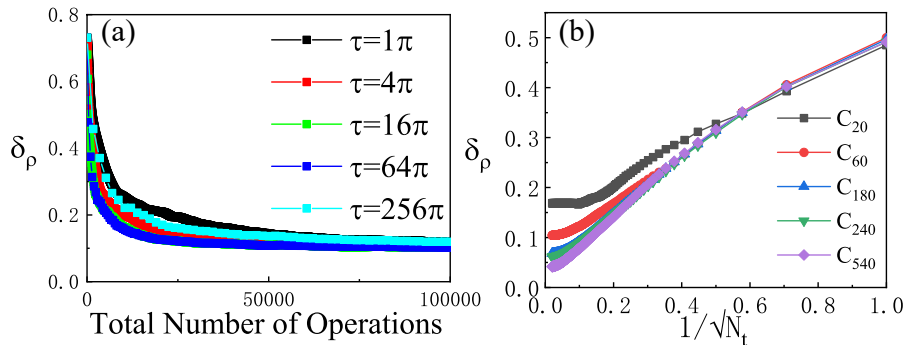


FIG. S6: (a) The statistical error δ_ρ as a function of total operations. Different colour indicates different time step τ in time evolution. (b) The statistical error δ_ρ as a function of $1/\sqrt{N_t}$ for different fullerenes, where $\tau = 64\pi$.

TABLE S2: The number of nonzero Bessel function (N_{Bessel}) as a function of time step τ , and the corresponding number of total matrix-vector operations $N_{operations}$ for the same certain propagation time $T = 1024\pi$.

τ	N_{Bessel}	N_t	$N_{operations}$
π	20	1024	20480
2π	27	512	13824
4π	38	256	9728
8π	56	128	7168
16π	89	64	5695
32π	149	32	4768
64π	261	16	4176
128π	478	8	3824
256π	899	4	3596
512π	1727	2	3454

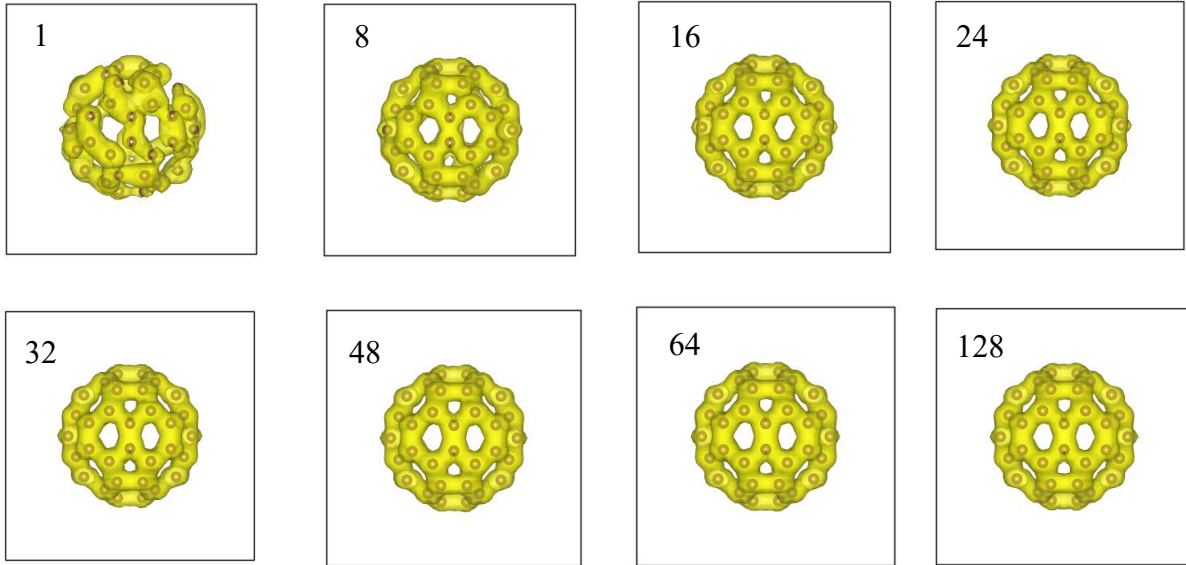
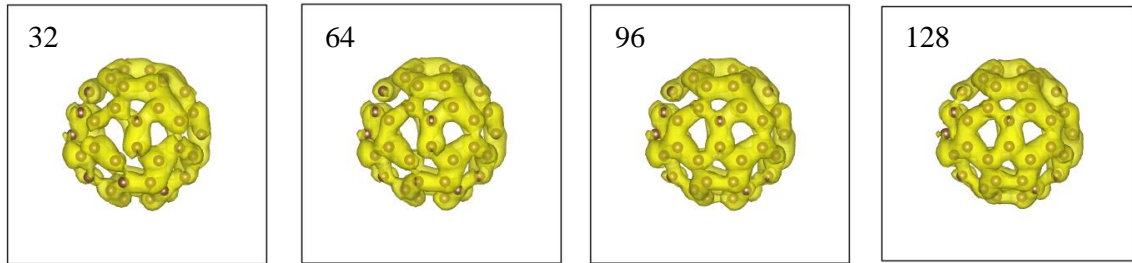


FIG. S7: The charge density of C_{60} calculated from average over a different number of random samples without time evolution.

(a) time intervals $dt = \pi$



(b) time intervals $dt = 32\pi$

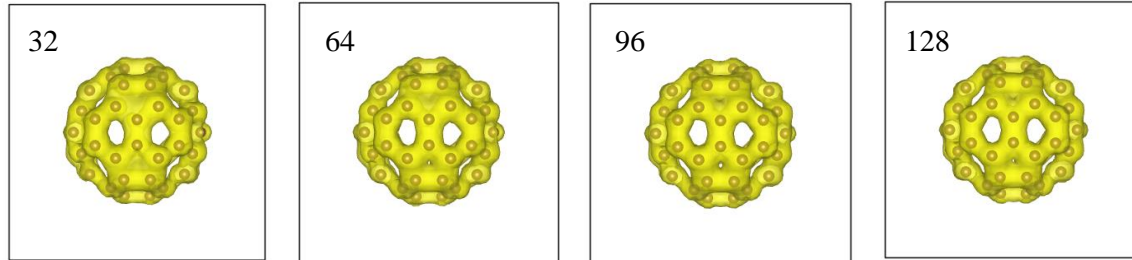


FIG. S8: The charge density of C_{60} calculated from one random state averaging over from time evolution for 32, 64, 96, 128 steps with $dt=\pi$ (a) and $dt=32\pi$ (b).

density without time evolution. As a comparison, we also calculate the electron density ρ_{diag} based on the occupied KS orbital obtained from the diagonalization of KS-Hamiltonian, and plot together with ρ_{FD} in Fig. S4(b). We see that with only $S=10$ random samples, ρ_{FD} converges to ρ_{diag} with an error of $\Delta(\rho_{FD} - \rho_{diag}) = 8.06 \times 10^{-5}$, where $\Delta(\rho_{FD} - \rho_{diag}) \equiv \sum_{j=1}^N |\rho_{FD}(\mathbf{r}_j) - \rho_{diag}(\mathbf{r}_j)|/N$. In step (IV), we use ρ_{FD} as the new input electron density and perform the next iteration. The self-consistent iterations, including steps (I) to (IV), are continued until a threshold is reached. In our approach, since the space resolution (determined by N) is much larger than the energy resolution (determined by N_t), it is more accurate to use the electron density instead of the total energy to define the convergence

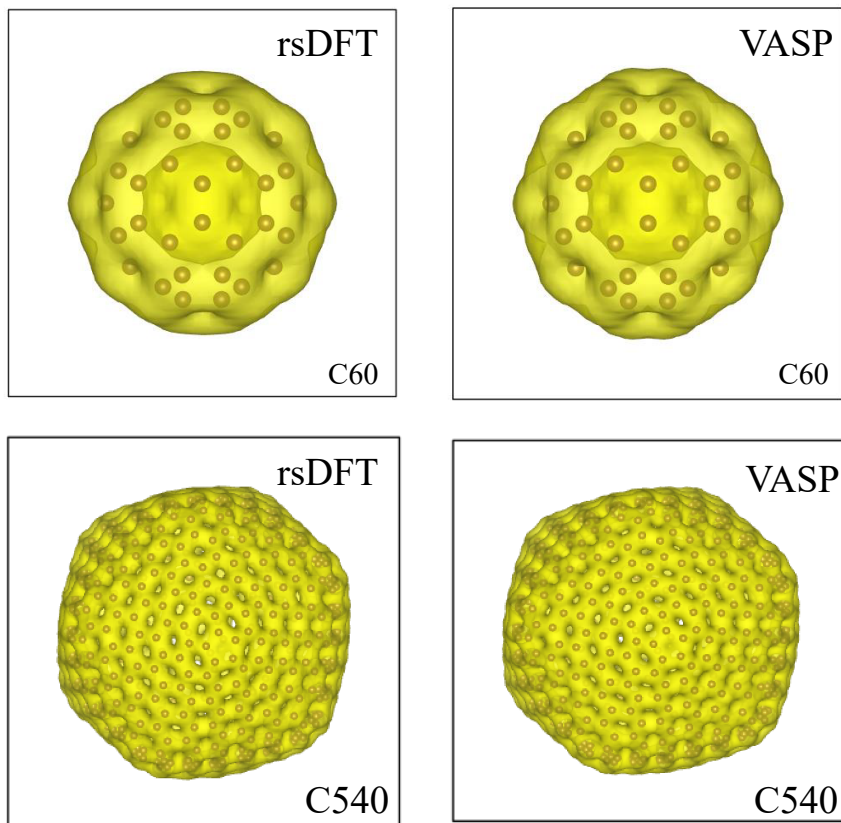


FIG. S9: The ground-state charge density calculated by rsDFT and VASP, respectively.

criterion. The ground-state electron density obtained from rsDFT without any diagonalization agrees well with the one from the common KS-DFT with diagonalization. Both are plotted in Fig. S4(c) for comparison. More examples of other single atoms can be found in Fig. S5.

4. MOLECULES

For molecular systems, we consider a diatomic model H_2 . The iterative calculations are similar to those of a single atom. Here we verify our approach by calculating the total energies for different H-H bond lengths and compare the results from our rsDFT approach and the common KS-DFT in Fig. S4(d). The two methods yield similar total energies for a given H-H bond length. The bond lengths in the ground state obtained from both ways are the same (74 pm), which agrees with the well-known result [7].

5. CLUSTERS

We extend our calculations to large atomic clusters of fullerenes C_{60} and C_{540} . In Fig. S7, we plot the ρ_{FD} averaging from up to 128 random states. As a comparison, we also present ρ_{RS} using only one random state, but different propagation time in Fig. S8(a) with $\tau = \pi$ and Fig. S8(b) with $\tau = 32\pi$. The real-space distribution of electron density in the ground state is visualized by VESTA [8] in Fig. S9. We use VASP (Vienna Ab initio Simulation Package) [9] to represent the standard KS-DFT method. VASP is a very efficient and widely used commercial KS-DFT package. The electron density distributions obtained from rsDFT and VASP are very similar.

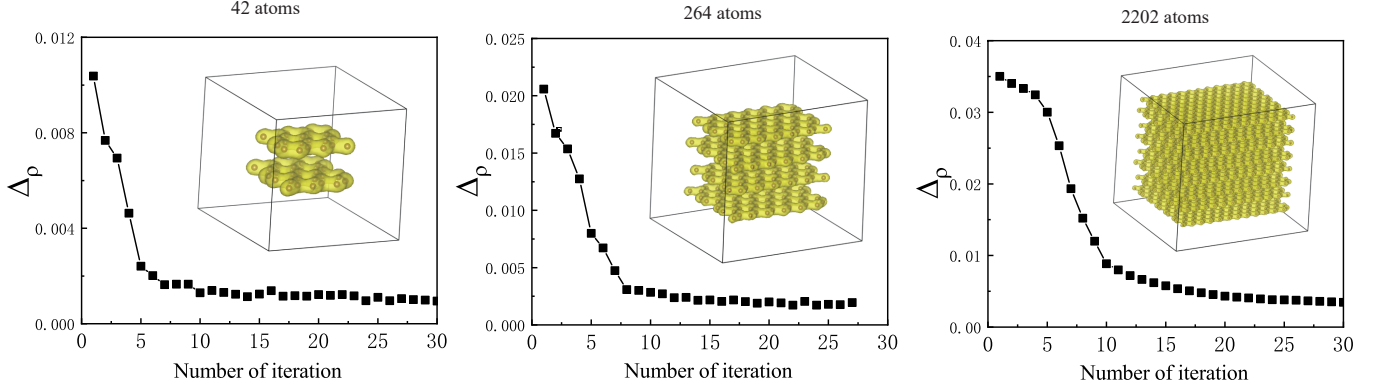


FIG. S10: $\Delta(\rho_{in} - \rho_{out})$ as a function of iterative steps for A-B stacked graphite calculated by rsDFT. The insets indicate the converged ground state densities.

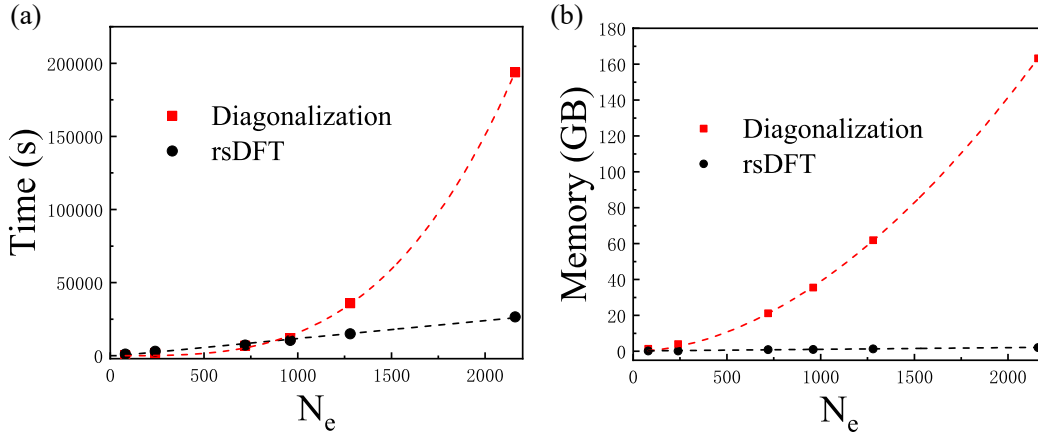


FIG. S11: Time cost of per iteration (a) and memory cost (b) during self-consistent calculations for fullerenes with different numbers of electrons. Black points refer to the traditional KS-DFT method with diagonalization, and red points refer to the rsDFT method without diagonalization. In rsDFT, we used 36 random samples for the average in the DOS and charge density calculations, the time step is $\tau = 64\pi$ and the number of time steps is $N_t = 36$.

6. CRYSTALS

More rsDFT calculations of graphite nanocrystals with different numbers of carbon atoms are plotted in Fig. S10.

7. CPU TIME AND MEMORY COST

To have a direct comparison of the CPU time and memory cost between the traditional KS-DFT (with diagonalization) and rsDFT (without diagonalization), we performed calculations for the fullerenes with different numbers of atoms (electrons) on a server with 40 CPU cores (2*Intel(R) Xeon(R) CPU Gold 6248). As shown in Fig. S11(a), if the system has less than ~ 1000 electrons, the traditional KS-DFT is much faster, but when the system size reaches ~ 1000 electrons, the rsDFT method becomes more efficient. The results in Fig. S11(a) also indicates that the time cost of rsDFT scales linearly with the system size, whereas the traditional KS-DFT scales approximately as $O(N_e^3)$. Although the accuracy of rsDFT and KS-DFT are not exactly the same, we estimate that rsDFT becomes more efficient when the system contains a few thousand or more electrons.

* s.yuan@whu.edu.cn

- [1] J. R. Chelikowsky, N. Troullier, and Y. Saad, *Phys. Rev. Lett.* **72**, 1240 (1994).
- [2] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, *Higher-order finite-difference pseudopotential method: An application to diatomic molecules*, *Phys. Rev. B* **50**, 11355 (1994).
- [3] S. H. Vosko, L. Wilk, and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, *Can. J. Phys.* **58**, 1200 (1980).
- [4] L. Kleinman and D. M. Bylander, *Efficacious Form for Model Pseudopotentials*, *Phys. Rev. Lett.* **48**, 1425 (1982).
- [5] ATOM, a program for DFT calculations in atoms and pseudopotential generation, distributed as part of the SIESTA software package. See <http://www.icmab.es/siesta/atom>.
- [6] D. R. Hamann, M. Schlüter, and C. Chiang, *Norm-Conserving Pseudopotentials*, *Phys. Rev. Lett.* **43**, 1494 (1979).
- [7] K.-P. Huber, *Molecular spectra and molecular structure: IV. Constants of diatomic molecules* (Springer Science & Business Media, 2013).
- [8] K. Momma and F. Izumi, *VESTA: a three-dimensional visualization system for electronic and structural analysis*, *J. Appl. Crystallogr.* **41**, 653 (2008).
- [9] G. Kresse and J. Furthmüller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*, *Comput. Mater. Sci.* **6**, 15 (1996).