

Linear Weak Scalability of Density Functional Theory Calculations without Imposing Electron Localization

Marcel D. Fabian, Ben Shpiro, and Roi Baer*



Cite This: *J. Chem. Theory Comput.* 2022, 18, 2162–2170



Read Online

ACCESS |



Metrics & More

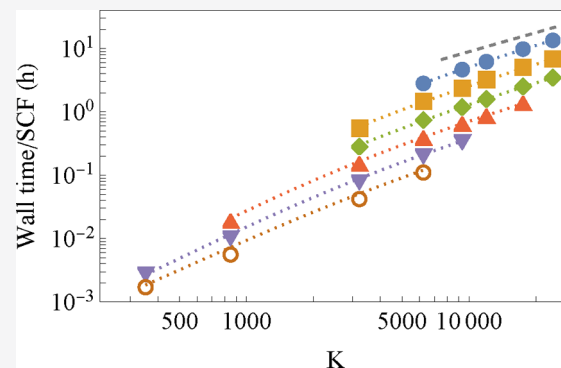


Article Recommendations



Supporting Information

ABSTRACT: Linear scaling density functional theory (DFT) approaches to the electronic structure of materials are often based on the tendency of electrons to localize in large atomic and molecular systems. However, in many cases of actual interest, such as semiconductor nanocrystals, system sizes can reach a substantial extension before significant electron localization sets in, causing a considerable deviation from linear scaling. Herein, we address this class of systems by developing a massively parallel DFT approach which does not rely on electron localization and is formally quadratic scaling yet enables highly efficient linear wall-time complexity in the weak scalability regime. The method extends from the stochastic DFT approach described in Fabian et al. (*WIRES: Comp. Mol. Sci.* 2019, e1412) but is entirely deterministic. It uses standard quantum chemical atom-centered Gaussian basis sets to represent the electronic wave functions combined with Cartesian real-space grids for some operators and enables a fast solver for the Poisson equation. Our main conclusion is that when a processor-abundant high-performance computing (HPC) infrastructure is available, this type of approach has the potential to allow the study of large systems in regimes where quantum confinement or electron delocalization prevents linear scaling.



1. INTRODUCTION

In the past few decades, supercomputers' massive number-crunching power, measured in floating-point operations per second (FLOPS), has grown a million-fold¹ and is currently pushing toward the exaflop (10^{18} FLOPS) realm. Combining this new technology with electronic structure calculations can revolutionize computational materials science and biochemistry, provided we complement it with algorithms that can efficiently exploit its massively parallel-based infrastructure.

One of the key questions then becomes how to quantify the efficiency of a certain algorithm on a massively parallel machine. A crucial measure in this regard is the *speedup*, which we define as the ratio

$$S(W, M) \equiv \frac{T_1(W)}{T_M(W)} \quad (1.1)$$

between the wall-times, $T_1(W)$, for executing a given computational work W using a single processor and $T_M(W)$ for its execution using M processors working in parallel. In operational regimes where the speedup is nearly proportional to M , i.e., $S = \mathbb{E} \times M$, there is a clear advantage in using a parallel multiprocessor approach where \mathbb{E} is the efficiency, with $\mathbb{E} = 1$ being ideal.

The efficient use of parallel computing was discussed by Amdahl in his seminal paper,² in which he identified in W an inherently serial (subscript s) and parallelizable (subscript p)

part, $W = W_s + W_p$. He assumed that the execution wall-time is independent of M for completing W_s and decreases linearly with M for W_p . Amdahl defined the *serial fraction* as $s_A = \frac{T_1(W_s)}{T_1(W)}$, measured on a single processor machine for a given job independent of M . With this definition, the speedup can be expressed as $S_A(W, M) = \left(s_A + \frac{1-s_A}{M}\right)^{-1}$ (Amdahl's law, also called *strong scalability*) and saturates once M exceeds the value of $1/s_A$.

Gustafson pointed out^{3,4} that in real-world usage the definition for the serial fraction should depend on M , because one does not generally take a fixed-sized problem, as Amdahl did, but rather scales the workload W with the available computing power. He then defined the serial fraction $s_G = \frac{T_M(W_s)}{T_M(W)}$ as measured on the M -processor system and showed that the speedup can be expressed as $S_G(M) = s_G + M(1 - s_G)$ (Gustafson's law, also called *weak*

Received: August 17, 2021

Published: March 26, 2022



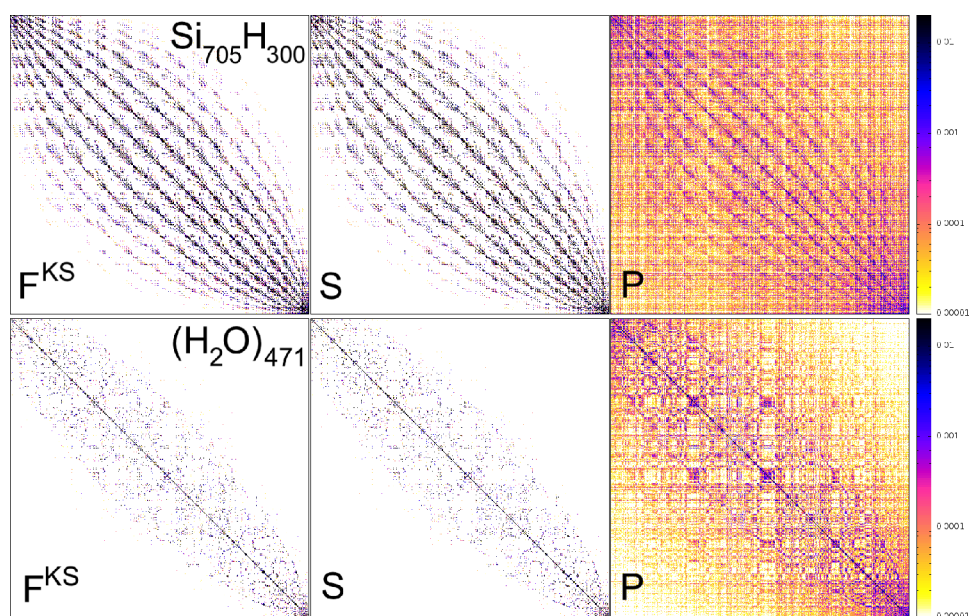


Figure 1. Fock (F^{KS}), overlap (S), and DM (P) matrices of $\text{Si}_{705}\text{H}_{300}$ and $(\text{H}_2\text{O})_{471}$ cluster (calculated within the LDA) are shown using a color-coded plot. The basis set in both systems is similar in size, with $K \approx 6000$. For clearer inspection of the sparsity pattern, the rows and columns are permuted so as to achieve a minimum bandwidth around the diagonal. We applied the MinimumBandWidthOrdering command of Mathematica³⁴ to the atomic proximity matrix $D_{AB} = \Theta(R_0 - R_{AB})$ (where R_{AB} is the distance between any pair of atoms A, B and $R_0 = 10a_0$ is the proximity distance), giving a permutation which is then used to order the atom-centered basis functions.

scalability), enabling linear speed up which does not inherently saturate as M increases.

These considerations can be applied to electronic structure calculations of extended systems in DFT codes that lower the cubic scaling by taking advantage of electron localization.^{5–25} For linear-scaling schemes, the Amdahl serial fraction $s_A = \frac{T_1(W_G)}{T_1(W)}$ is expected to be system-size independent (since both timings in the numerator and the denominator scale linearly with system size) while for codes of higher algorithmic complexity, s_A decreases as system size increases.²⁶ In a weak scalability analysis of the linear scaling codes, Gustafson's serial fraction $s_G = \frac{T_M(W_G)}{T_M(W)}$ is also expected to be system-size-independent (since both timings in the numerator and the denominator scale linearly with system size) and therefore take the form $s_G = \left(1 + \frac{M_0}{M}\right)^{-1}$, where M_0 is a constant (depending on the hardware and algorithm). For large M , the speedup saturates to $S_G(M) \rightarrow 1 + M_0$, but if M_0 is very large there is a sizable regime where $M \ll M_0$ and the s_G is essentially zero, so an ideal linear speedup emerges, as reported, for example, for the CONQUEST code,^{6,27} even up to $M = 200\,000$ cores on the Fujitsu-made K-computer. It is clear from the previous studies mentioned above that it is important to determine the strong and weak scalability properties of codes that can use massively parallel machines, because they are sensitive to many details concerning hardware, system size, algorithmic scaling, etc.

In this paper we develop an efficiently parallelizable, (semi)local DFT approach which offers quadratic scaling with system size and does not involve approximations derived from assuming electron localization. It combines several approaches, such as atom-centered Gaussian basis sets and real-space grids for providing the electrostatic and exchange–correlation energies (similar to SIESTA⁹ and CP2K/Quick-

step⁷) as well as Chebyshev expansion techniques for representing the density matrix.^{16,28–30} We describe the theory and implementation in section 2, where we also provide an illustration of the nonlocalized nature of electrons in the large benchmarking systems we use (see Figure 1). Next, we present the algorithmic complexity and the parallel strong/weak scalability properties of our approach in section 3, and finally, we summarize and discuss the conclusions in section 4.

2. METHOD

In our method, we work with standard quantum chemistry basis sets, composed of atom-centered local functions $\phi_\alpha(\mathbf{r})$, $\alpha = 1, \dots, K$. For calculating the necessary integrals, solving the Poisson equations, and generating the exchange–correlation potentials, we use a 3D Cartesian real-space grid of equidistant points spanning a simulation box, containing the system's atoms and electronic density. For this purpose, we developed an efficient method for evaluating the basis functions on a relevant set of grid points, outlined in Supporting Information A. Our method of combining basis functions and real-space grids is similar in spirit to those existing in the literature, such as SIESTA⁹ and CP2K/Quickstep,⁷ but differs in important details. Unlike SIESTA, we use standard nonorthogonal Gaussian basis sets, and unlike Quickstep, we represent the basis functions on the grid where all integrals are performed as summations. The first type of integral that we have to evaluate on the grid is the overlap matrix:

$$S_{\alpha\beta} = h^3 \sum_{\mathbf{g}} \phi_\alpha(\mathbf{r}_{\mathbf{g}}) \phi_\beta(\mathbf{r}_{\mathbf{g}}) \quad (2.1)$$

where $\mathbf{r}_{\mathbf{g}}$ are the grid points and h is the grid-spacing. Next, the kinetic energy integrals are evaluated as

$$T_{\alpha\beta} = \frac{1}{2} h^3 \sum_{\mathbf{g}} \nabla \phi_\alpha(\mathbf{r}_{\mathbf{g}}) \cdot \nabla \phi_\beta(\mathbf{r}_{\mathbf{g}}) \quad (2.2)$$

where the derivatives of the basis functions are calculated analytically and then placed on the grid (see Supporting Information A.2.3 for details). To avoid an excessive number of grid points, the equally spaced grid is complemented with norm-conserving pseudopotentials,³¹ representing the effects of the tightly bound core electrons (which are not treated explicitly) and taken into account in the KS Hamiltonian, represented by the Fock matrix

$$F^{\text{KS}} = T + V^{\text{NL}} + V^{\text{KS}} \quad (2.3)$$

where

$$V_{\alpha\beta}^{\text{NL}} = h^3 \sum_{\mathbf{g}} \phi_{\alpha}(\mathbf{r}_{\mathbf{g}}) \sum_{C \in \text{nuclei}} \hat{v}_{ni}^C \phi_{\beta}(\mathbf{r}_{\mathbf{g}}) \quad (2.4)$$

are the integrals for the nonlocal pseudopotential and

$$V_{\alpha\beta}^{\text{KS}} = h^3 \sum_{\mathbf{g}} \phi_{\alpha}(\mathbf{r}_{\mathbf{g}}) v_{\text{KS}}(\mathbf{r}_{\mathbf{g}}) \phi_{\beta}(\mathbf{r}_{\mathbf{g}}) \quad (2.5)$$

are the KS potential integrals, where

$$v_{\text{KS}}(\mathbf{r}_{\mathbf{g}}) = \sum_{C \in \text{nuclei}} v_{\text{loc}}^C(\mathbf{r}_{\mathbf{g}} - \mathbf{R}_C) + v_{\text{H}}[n](\mathbf{r}_{\mathbf{g}}) + v_{\text{xc}}[n](\mathbf{r}_{\mathbf{g}}) \quad (2.6)$$

In eq 2.6, $v_{\text{H}}[n](\mathbf{r}_{\mathbf{g}})$ is the Hartree potential on the grid which is evaluated directly from the grid representation of the electron density $n(\mathbf{r}_{\mathbf{g}})$ by a reciprocal space-based method for treating long-range interactions.³² The exchange–correlation potential $v_{\text{xc}}[n](\mathbf{r}_{\mathbf{g}})$ (within the local density approximation (LDA)) is also determined on the grid directly from the electron density. From the grid representation of the pseudopotentials^a we obtain the potential $v_{\text{loc}}^C(\mathbf{r}_{\mathbf{g}} - \mathbf{R}_C)$ appearing in eq 2.6 for nucleus C at position \mathbf{R}_C and, by grid integration, the matrix V^{NL} appearing in eq 2.3. All integral calculations are performed in parallel for different basis function pairs; for more details, see Supporting Information C.

The electron density on the grid is formally defined as

$$n(\mathbf{r}_{\mathbf{g}}) = 2 \sum_{\alpha, \beta} P_{\alpha\beta} \phi_{\alpha}(\mathbf{r}_{\mathbf{g}}) \phi_{\beta}(\mathbf{r}_{\mathbf{g}}) \quad (2.7)$$

where P is the density matrix (DM) and the factor of 2 comes from integration over spin degrees of freedom. The DM must obey an electron-conserving criterion, namely that the integral over all grid points evaluates to the total number of electrons in the system: $h^3 \sum_{\mathbf{g}} n(\mathbf{r}_{\mathbf{g}}) = N_e$. Indeed, performing this integral and using eqs 2.1 and 2.7 we find

$$N_e = 2\text{Tr}[PS] \quad (2.8)$$

This relation is part of a more general requirement, that the Kohn–Sham eigenstates are populated according to the Fermi–Dirac function $f_{\text{FD}}(\epsilon) = \frac{1}{1 + e^{\beta(\epsilon - \mu)}}$ where ϵ is the corresponding energy eigenvalue. For the DM, this condition can be satisfied by defining²¹

$$P = f_{\text{FD}}(S^{-1}F^{\text{KS}})S^{-1} \quad (2.9)$$

For finite-temperature DFT, β is the inverse temperature and μ is the chemical potential. For ground-state calculations, β obeys $\beta(\epsilon_{\text{L}} - \epsilon_{\text{H}}) \gg 1$, where ϵ_{L} (ϵ_{H}) is the Kohn–Sham eigenvalue of the lowest unoccupied (highest occupied) molecular orbital. The chemical potential in the Fermi–

Dirac function is adjusted to reproduce the systems' number of electrons N_e through eq 2.8.

The use of atom-centered local basis functions allows for sparsity in the basic matrices F^{KS} and S , as illustrated in Figure 1 for two systems of similar size but different chemical nature, a 2.5 nm (diameter) semiconductor nanocrystal $\text{Si}_{705}\text{H}_{300}$ and a 3 nm water cluster $(\text{H}_2\text{O})_{471}$. For the matrix representation in Figure 1, we have ordered the atoms (and the basis functions associated with them) in a way that takes into account their spatial proximity (near atoms tend to have similar indices). Therefore, it is clear by mere inspection that F^{KS} and S have a relatively small spatial range and are therefore quiet sparse. Our approach makes an effort to exploit this property by using sparse matrix algebra. Despite the spatial locality of F^{KS} and S , P in these large systems is highly nonlocal, expressing the physical fact that the electronic coherence in these systems is long ranged. For the silicon system, this fits our intuition, namely that silicon is by nature a semiconductor, with properties which are close to those of metals. Although water is a large band gap system, it is known that under LDA it exhibits very small HOMO–LUMO gaps^{35–38} (see also Figure 2).

The various expectation values of relevant observables (i.e., operators in the grid representation) can be expressed as trace operations:

$$\langle \hat{O} \rangle = 2\text{Tr}[PO] \quad (2.10)$$

where

$$O_{\alpha\beta} = h^3 \sum_{\mathbf{g}} \phi_{\alpha}(\mathbf{r}_{\mathbf{g}}) \hat{O} \phi_{\beta}(\mathbf{r}_{\mathbf{g}}) \quad (2.11)$$

is the matrix representation of the one-body operator \hat{O} in the atomic basis. In order to expedite the calculation we need to parallelize the computational work, and this can be done by representing the trace operations as a sum over unit column vectors u_{α} (with coordinates $(u_{\alpha})_{\beta} = \delta_{\alpha\beta}$ i.e., zeros in all positions except at α), computed column by column:

$$\langle \hat{O} \rangle = 2 \sum_{\alpha} u_{\alpha}^T O P u_{\alpha} \quad (2.12)$$

For achieving this, we treat the DM as an operator; that is, we devise a linear-scaling method for applying it to the column vector u_{α} based on eq 2.9: $P u_{\alpha} = f_{\text{FD}}(S^{-1}F^{\text{KS}})S^{-1}u_{\alpha}$. The operation $S^{-1}u_{\alpha}$ is performed by the linear-scaling preconditioned conjugate-gradient approach involving repeated application of the sparse overlap matrix S on column vectors.^b The operation of $f_{\text{FD}}(S^{-1}F^{\text{KS}})$ on the column vector $S^{-1}u_{\alpha}$ employs a Chebyshev expansion^{28,40} of the function $f_{\text{FD}}(\epsilon)$, which results in repeated applications of the operator $S^{-1}F^{\text{KS}}$ to column vectors. Details are described in Supporting Information B. The entire procedure can be readily distributed over several processors in parallel, each commissioned with a distinct set of u_{α} column vectors. This calculation method has the additional benefit that it avoids storage of the nonsparse DM. We discuss the algorithmic complexity of the approach, as well as its weak and strong scalability in section 3.

Equations 2.1–2.9 and the techniques of their application discussed above form a series of nonlinear equations that must be solved together, to give the self-consistent-field (SCF) solution. The procedure is iterative and uses the direct inversion of the iterative subspace (DIIS) convergence acceleration method.⁴¹ Once converged, various expectation

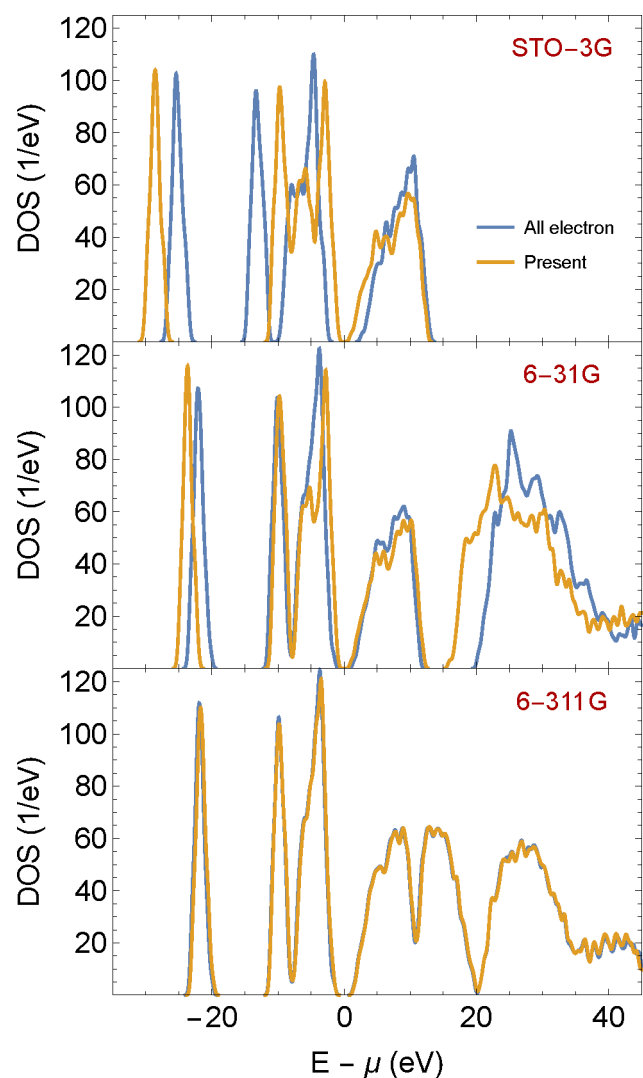


Figure 2. Density of state (DOS) as a function of energy, shifted by the chemical potential μ , for a water cluster $(\text{H}_2\text{O})_{100}$ in the LDA at a fixed geometry. The three panels compare all-electron calculations (performed by Q-Chem⁴³) with valence-electron-only calculations (performed by the present approach, using pseudopotentials). Each panel presents the results for different Gaussian basis sets, from single- to triple- ζ quality (STO-3G, 6-31G, and 6-311G). Both calculations use the eigenvalues ϵ_n of the converged KS Hamiltonian to obtain the DOS function $\rho_{\text{DOS}}(\epsilon) = 2 \times \frac{1}{\sqrt{2\pi}\sigma} \sum_n e^{-(\epsilon-\epsilon_n)^2/2\sigma^2}$ where $\sigma = 0.01E_h$. The calculation in the present approach used $\beta = 100E_h^{-1}$ with a real-space grid of spacing $\Delta x = 0.33 a_0$.

values such as charges and multipoles, density of states, and polarizability can be calculated, as well as forces on the nuclei,⁴² which can be used for structure optimization.

In order to check and validate the implementation of the algorithm outlined above, we show in Figure 2 the density of states (DOS) shifted for the chemical potential μ of a cluster of 100 water molecules, obtained with our program, and with the all-electron calculation performed in the commercially available quantum chemistry program Q-Chem.⁴³ Our code used $\beta = 100E_h^{-1}$, but we tested also larger values of β to ascertain that the results are visibly identical. We made comparisons using three different basis sets, ranging from single- to triple- ζ quality (STO-3G, 6-31G, and 6-311G). To complement the picture, we also give the frontier orbital

energies, band gaps, and chemical potentials corresponding to these calculations in Table 1.

Table 1. Comparison of Frontier Energy Levels, the Band Gap $\epsilon_g = \epsilon_L - \epsilon_H$, and the Chemical Potential $\mu = (\epsilon_H + \epsilon_L)/2$ for the DOS Calculations of Figure 2

basis	method	ϵ_H	ϵ_L	ϵ_g	μ
STO-3G	present	-4.5	-2.6	1.9	-3.6
	all-electron	1.3	4.5	3.1	2.9
	SBKJJC	-5.3	-1.6	3.7	-3.4
6-31G	present	-3.5	-1.9	1.5	-2.7
	all-electron	-3.5	-2.2	1.4	-2.8
6-311G	present	-4.7	-3.3	1.4	-4.0
	all-electron	-4.2	-2.7	1.4	-3.4

Looking at the shifted DOS, both the results of Q-Chem and those of the present code converge to indistinguishable values close to that of the all-electron highest-quality basis calculation. This validates our present code's calculations, even though a small shift still exists between the chemical potentials (-0.6 eV), as seen in Table 1. It is noteworthy that the DOS in our code is less sensitive to basis set quality than the all-electron code, where for the smallest STO-3G basis set the all-electron calculations deviate strongly from the converged basis set values, showing a large (6.3 eV) shift and a band gap which is more than a factor two too large. The stability of our calculations in comparison to Q-Chem can be attributed to the use of the norm-conserving pseudopotentials. Indeed, in Supporting Information F we show that effective core potentials stabilize the Q-Chem small basis set calculations as well.

An additional validation of our approach can be found in the Supporting Information G, where we compare the potential energy surface of the H_2 molecule calculated with both our code and Q-Chem and where we show the influence of the grid spacing on the accuracy of the calculation. Overall, the approximations that we employ lead to a systematic difference of $\sim 0.2\%$ in the electronic energy when compared with Q-Chem for most of the examined distance range (and maximally $\sim 0.4\%$) and a small corrugation which appears when the gridpoint spacing is larger than the width of the smallest Gaussian primitive. The relative errors in the electronic energy, and the fact that they are mostly a rigid shift, lead to deviance of the order of 0.05 eV in the bond energy, much smaller than typical 6-311G basis set errors.⁴⁴

3. SCALING PROPERTIES OF THE METHOD

In this section we study the method's algorithmic complexity and analyze the speedup achievable by parallelization in terms of strong and weak scalability.

3.1. Algorithmic Complexity. To understand the algorithmic complexity of our method, we have to examine how each part of our code scales as we increase the system size K . Here we are especially interested in the asymptotic behavior, meaning that the program part with the largest scaling will determine the overall algorithmic complexity. Our entire SCF cycle, that is described in detail in Supporting Information A.3, includes different integral calculations, solving the Poisson equation, and calculating the density. The integral calculation is expected to scale linearly with system size K , i.e. $O(K)$, because the relevant matrices (F^{KS} , S) are expected to become sparse (see also Figure 1). The Poisson equation is

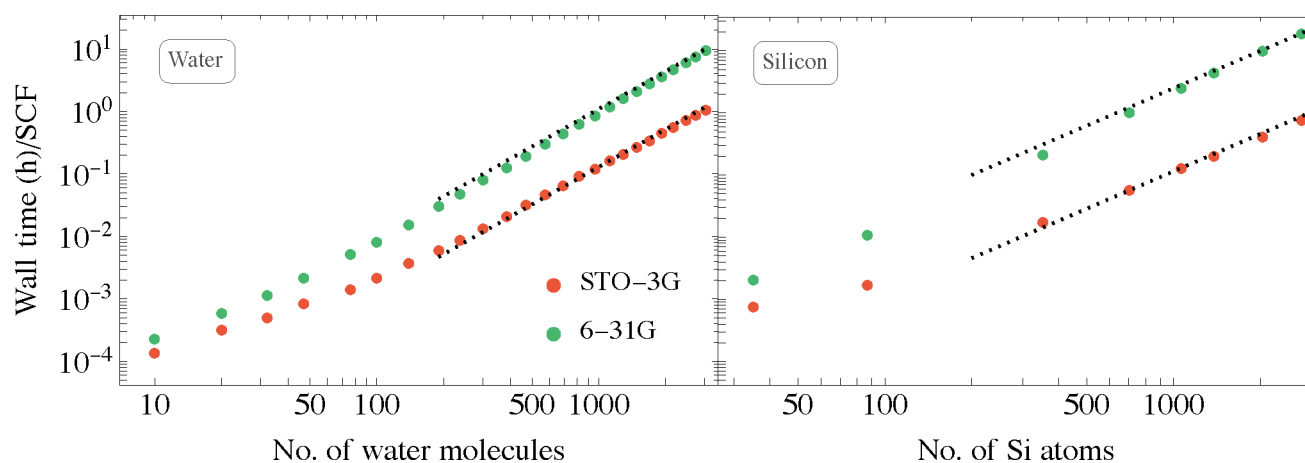


Figure 3. Wall-time as a function of system size, for the water clusters and the silicon nanocrystals, calculated using two basis sets within the LDA. The calculations were performed on eight Intel Xeon Gold 6132 CPU @ 2.60 GHz 755GB RAM (connected through Infiniband), using 112 cores for all systems. The dotted lines in the panels are guides to the eye with designated quadratic scaling. Fitting the function $t = Ax^n$ to the data at the larger time range, one obtains the exponent $n = 1.9$ (2.1) for both the water and the silicon systems in the STO-3G (6-31G) basis.

solved by a fast Fourier transform (FFT) which scales as $O(N_g \log N_g)$, where N_g are the grid points, expected to scale linearly with system size. This leaves only the density calculation, which is done according to eq 2.12. The application of the DM P to a column vector u_α expressed through a Chebyshev series, involves repeated applications of the operator $S^{-1}F^{KS}$ to the column vector $v = S^{-1}u_\alpha$ (see Supporting Information B for details). The length of the Chebyshev expansion, N_C , is independent of the system size K , and so the algorithmic complexity of the Pu_α operation is identical to that of one $S^{-1}F^{KS}v$ operation, namely, linear with K . There are a total of K different Pu_α operations (see eq 2.12), so that the overall algorithmic complexity of the method is asymptotically quadratic, i.e., $O(K^2)$. As the system size grows our algorithm could be modified to take advantage of the emerging sparsity of the DM, allowing for a K -independent complexity of each Pu_α operation. In such situations, one can expect an overall linear-scaling numerical complexity, i.e., $O(K)$. However, in the present paper, we focus on the broad class of systems which are very large but for which the DM has not yet localized. Hence, we are in the formally quadratic complexity regime.

To show that quadratic complexity is indeed what we achieve with this method, we plot in Figure 3 the wall-time per SCF cycle versus system size for water clusters (taken from <http://www.ergoscf.org/xyz/h2o.php>, accessed on 2022-03-05) and hydrogen-terminated silicon nanocrystals (we use a series of nanocrystals, starting from $\text{Si}_{35}\text{H}_{36}$ reaching $\text{Si}_{2785}\text{H}_{780}$; for details, see Supporting Information D), using STO-3G and the larger 6-31G basis sets. Going from the smaller to the larger basis set increases wall-time by a factor of 10–20. This result is a combination of several characteristics beyond the mere size of the basis set. For example, the magnitude of the Gaussian exponents of the basis set's primitives are relevant for the dimensioning of the grid. Higher-valued Gaussian exponents require a finer mesh and also increase the kinetic energy component of the Hamiltonian, which increases the Chebyshev expansion length. Smaller (diffuse) Gaussian exponents lead to larger grid windows (see also Supporting Information A.1.1) and hence an increase in overall grid size as well. Furthermore, the implementation of the linear scaling operation of S^{-1} , involving the incomplete Cholesky

decomposition and preconditioned conjugate gradients algorithms, is sensitive to the condition number of S , determined by near linear dependencies between basis functions. As seen in the figure, all cases show overall quadratic algorithmic complexity. It is worth noting that the small and intermediate sized systems in the figure exhibit a varying algorithmic complexity with system size associated with the interplay between linear complexity processes having a large prefactor and cubic stages because of the nonsparse nature of the Hamiltonian and overlap matrices.

3.2. Strong Scalability. In Figure 4 we study the strong scalability properties of our code, i.e., the scalability achievable when increasing the number of processors for a given task. We show in the figure the speedup and efficiency for a single SCF iteration of the $\text{Si}_{1379}\text{H}_{476}$ nanocrystal. Our definition for the speedup in eq 1.1 requires the knowledge of the elapsed wall time it takes a single processor (more accurately 1 core) to finish this nanocrystal calculation. Because of (human) time constraints we had to extrapolate this timing from a calculation on 36 cores on one single compute node by $T_1 = 36T_{36}$. The results can be analyzed in terms of the Amdahl law finding that the serial fraction is $s_A = 9 \times 10^{-5}$ showing a high degree of parallelization. Accordingly, the parallelization efficiency drops very slowly as the number of processors increases, with 96% efficiency even at $M = 500$ (see the inset in the top panel). We emphasize that this is achieved with a 10 Gb ethernet network communication. Potentially, the decay of efficiency may be slowed down by employing a faster communication solution. According to Amdahl's law, efficiency will drop to 0.5 when $M \approx \frac{1}{s_A} = 10^4$. In Supporting Information E we show results for a smaller system, $\text{Si}_{705}\text{H}_{300}$, where the Amdahl serial fraction is larger, $s_A = 2 \times 10^{-4}$, a system-size dependency due to the quadratic complexity of our method (see our discussion in section 1).

3.3. Weak Scalability. In this section we focus on the weak scalability properties of our method, namely how the wall time changes with system size K when the number of processors afforded to the calculation M grows in fixed proportion $r = K/M$. In the left panel of Figure 5 we present the wall-time T as a function of system size K for six series of runs we made with different fixed ratios ranging from $r = 4$ up

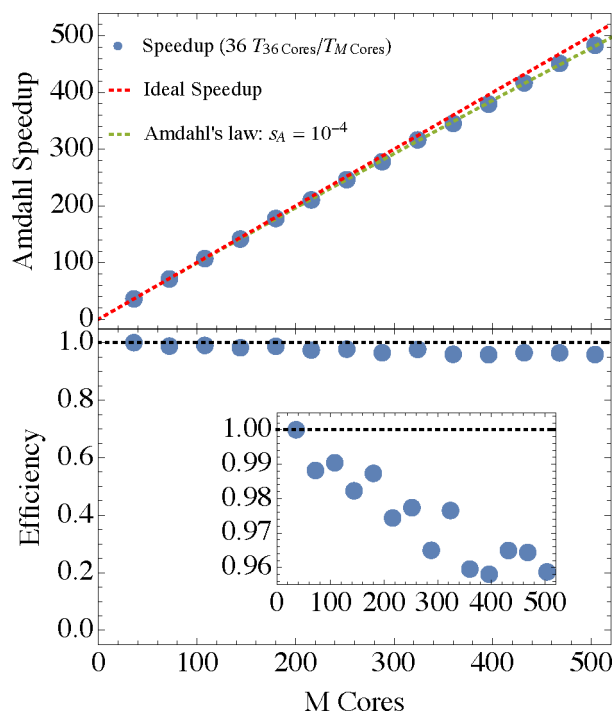


Figure 4. Strong scalability speedup analysis (upper panel) and efficiency $E(M) = S(M)/M$ (lower panel) for $\text{Si}_{1379}\text{H}_{476}$. The reference time for 1 processor for the speedup is extrapolated from $T_1 = 36T_{36}$. The inset in the lower panel enables a higher resolution of the efficiency regime close to unity. The calculation used the 6-31G basis set (11984 basis functions) within the LDA and were performed on several 2.60 GHz Intel Xeon Gold 6240 with 256 GB using 10 Gb Ethernet networking communications.

to $r = 120$ (in the actual calculation, r is the number of vectors u_α assigned to each processor (see eq 2.12)). The markers of each series fall on asymptotically straight lines in the log–log plot which appear parallel to the dark-dashed line, indicating a constant slope of 1. This confirms the claim of achieving linear-scaling wall-time in this regime of operation, where r is held constant. We would also like to examine the *speedup* in order

to determine the degree of efficiency of our calculation on the parallel machine. For calculating the speedup under our definition in eq 1.1 we need to be able to estimate the wall time $T_1(W)$, which for the large systems is not easily accessible because of (human) time constraints. Therefore, we developed the following model for the wall time, with which we will estimate the $M = 1$ wall times:

$$T(K, M) = \frac{\tau_2 K^3 / (K_0 + K)}{M} + \tau_1 K \log M \quad (3.1)$$

The first term on the right is the dominant parallelizable part of the calculation run on M processors (electron density calculation, see Supporting Information C for more information). For $K \gg K_0$ it exhibits quadratic scaling, while for $K \ll K_0$ the scaling is cubic because of insufficient sparsity of the Hamiltonian and overlap matrices for small K . The second term in eq 3.1 reflects the timing of the serial part of the calculation, dominated by the communication time needed for specific MPI functions (reduce and broadcast) and scales linearly with K and logarithmically with M .

Using the analytical model, the speedup can now be obtained by plugging eq 3.1 into eq 1.1, resulting in the following closed form expression:

$$S(r, M) = \frac{M}{\left(\frac{K_0}{rM} + 1\right) \frac{\tau_1}{\tau_2} \log M + 1} \quad (3.2)$$

From this equation, it can be seen that for asymptotically large values of M , the speedup approaches the limit $S(r, M) \rightarrow \frac{M}{\frac{\tau_1}{\tau_2} \log M + 1}$ and as long as r is not too small

$$r > \frac{\tau_1}{\tau_2} \log M \quad (3.3)$$

the speedup is close to ideal $S(r, M) \rightarrow M$.

We now fit our model to the calculation's timing results from the six constant- r series shown in the left panel of Figure 5 (a total of 32 data points). This leads to a best-fit set of parameters (in hours): $\tau_1 \rightarrow 5.16 \times 10^{-6}$ h,

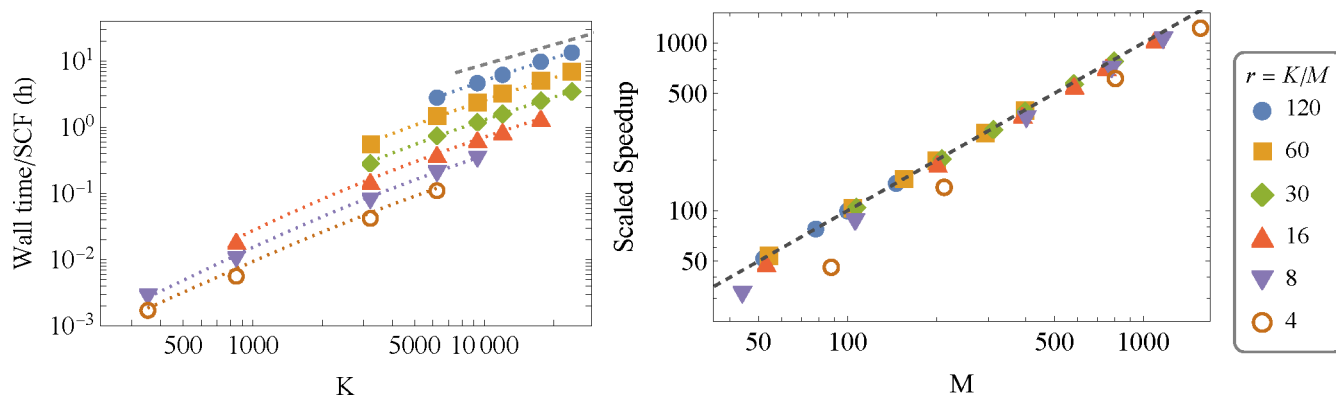


Figure 5. Weak scalability speedup analysis. Left: The wall-time for a single SCF cycle versus the number of basis set functions K in calculations given for several fixed values of r (K/M where M is the number of processors) on a series of eight hydrogen-terminated silicon nanocrystals (detailed in Supporting Information D) using the 6-31G basis set within the LDA. The black-dashed line is a guide to the eye showing linear scaling wall-time. The calculations were run on several Intel Xeon Gold 6240 CPUs @ 2.60 GHz 256GB RAM connected through a 10 Gb ethernet networking communications. We had access to at most 1584 cores; therefore, for $r = 4$ we could not treat systems greater than $K = 6420$, and similar though less stringent limitations appeared for $r = 8$ and 16. The colored dotted lines are the best-fit of the data to our model in eq 3.4. Right: The scaled speedup as a function of the number of processors M , calculated for the six values of r from eq 3.2 using the best-fit parameters of our model. The black-dashed line indicates the “perfect” speedup $S = M$.

$\tau_2 \rightarrow 5.63 \times 10^{-7}$ h, and $K_0 \rightarrow 2292.6$ for our model, and the resulting fit functions

$$\frac{T_{\text{fit}}(K; r)}{10^{-6} \text{ h}} = K \left(\frac{0.563 r}{\frac{2292.6}{K} + 1} + 5.16 \log(K/r) \right) \quad (3.4)$$

are plotted in the left panel of the figure as dotted colored lines, one for each value of r . It can be seen that these fit functions indeed reproduce the actual data (given as points) quite closely.

Having the best-fit parameters, let us now discuss the actual estimated values for the (scaled) speedup in the Gustafson sense. These estimates, based on eq 3.2, are plotted in the right panel of Figure 5. We see that for $r > 16$ the speedup is not too far from ideal, in accordance with the analysis presented above; however, as indicated in eq 3.3, the speedup is smaller when r decreases as is clearly visible for $r = 8$ and small M and for $r = 4$ for all values of M . However, even for these small r cases, the speedup is maintained as M increases and the calculation is still quite efficient.

4. SUMMARY AND CONCLUSIONS

In this paper we presented a parallelizable electronic structure approach to finite temperature density functional theory under (semi)local functionals using atom-centered Gaussian basis sets which offers linear wall-time complexity as a function of system size in the weak scalability regime. The inherent time complexity of the method is quadratic $O(K^2)$, as discussed in section 3.1, and it does not involve truncation of density matrix elements, characteristics of linear-scaling approaches.

Our trace-based calculation combined with Chebyshev expansions allows for efficient parallelization in the strong scalability sense, as shown in subsection 3.2. Because of the quadratic complexity, we found that the value of the Amdahl parameter was system-size dependent, with $s_A = 2 \times 10^{-4}$ for the $\text{Si}_{705}\text{H}_{300}$ system and $s_A = 9 \times 10^{-5}$ for $\text{Si}_{1379}\text{H}_{476}$. The overall weak scalability performance shows that linear scaling wall time is achievable, as demonstrated in section 3.3, and is highly efficient when the number of orbitals per processors r is not smaller than ~ 10 , and beyond that efficiency drops by a factor of ~ 1.5 .

Our main conclusion is that this type of approach has the potential to be a useful and efficient tool for studying large systems in regimes where quantum confinement or electron delocalization prevents traditional linear-scaling to set in. Furthermore, for even larger systems, where electrons localize, we plan to enable linear scaling either through stochastic orbital methods²¹ or by exploiting directly the DM's finite range. While in this paper we were concerned mainly with the scalability of the density calculation, force evaluation, done after the density converges, is also an important goal, high on our list of future plans. We will follow our recent work developing a stochastic estimation of the exact energy derivative (Hellman–Feynman) forces.⁴² As shown in ref 45, these stochastic estimations lead to noisy forces that can be used only within Langevin dynamics. In our case, we expect that the deterministic evaluation of the exact derivatives will result in deterministic forces of sufficient quality to enable energy-conserving molecular dynamics simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00829>.

Details on our implementation of the Gaussian integral evaluation on the Cartesian grid, the Chebyshev expansion, our parallelization strategy, systems studied, density of states calculated with Q-Chem and effective core potentials, and the H_2 potential energy surface (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Roi Baer – Fritz Haber Research Center for Molecular Dynamics and the Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel; orcid.org/0000-0001-8432-1925; Email: roi.baer@huji.ac.il

Authors

Marcel D. Fabian – Fritz Haber Research Center for Molecular Dynamics and the Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
Ben Shpiro – Fritz Haber Research Center for Molecular Dynamics and the Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.1c00829>

Notes

The authors declare no competing financial interest. Binaries that support the findings of this study are available from the corresponding author upon reasonable request.

■ ACKNOWLEDGMENTS

We gratefully acknowledge support from the Israel Science Foundation grant 800/19. M.D.F. expresses special thanks to Roie Dann for helpful discussions.

■ ADDITIONAL NOTES

^aHere, we use the Kleinman–Bylander (KB) form,³³ which produces two types of operators, a nonlocal potential operator $\hat{v}_{\text{nl}}^C \equiv v_{\text{nl}}^C(r - R_C, r' - R_C)$ which is defined in a small sphere around each atomic core (R_C is the location of atom C) and a scalar potential $v_{\text{loc}}^C(r - R_C)$ containing the long-range electron-shielded nucleus Coulomb attraction.

^bWe use the incomplete Cholesky preconditioning³⁹ for the conjugated gradient approach implemented in the HSL-MI28 and MI21 codes, respectively, where HSL is a collection of FORTRAN codes for large-scale scientific computation (<http://www.hsl.rl.ac.uk/> accessed 2022-03-05).

■ REFERENCES

- Heldens, S.; Hijma, P.; Werkhoven, B. V.; Maassen, J.; Belloum, A. S. Z.; Van Nieuwpoort, R. V. The Landscape of Exascale Research: A Data-Driven Literature Analysis. *ACM Comput. Surv.* **2021**, *53*, 1–43.
- Amdahl, G. M. Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18–20, 1967 spring joint computer conference - AFIPS '67 (Spring)*, Atlantic City, NJ, 1967; p 483.

- (3) Gustafson, J. L. Reevaluating Amdahl's law. *Commun. ACM* **1988**, *31*, 532–533.
- (4) Gustafson, J. L.; Montry, G. R.; Benner, R. E. Development of parallel methods for a 1024-processor hypercube. *SIAM journal on Scientific and Statistical Computing* **1988**, *9*, 609–638.
- (5) Mohr, S.; Ratcliff, L. E.; Boulanger, P.; Genovese, L.; Caliste, D.; Deutsch, T.; Goedecker, S. Daubechies wavelets for linear scaling density functional theory. *J. Chem. Phys.* **2014**, *140*, 204110.
- (6) Nakata, A.; Baker, J. S.; Mujahed, S. Y.; Poulton, J. T. L.; Arapan, S.; Lin, J.; Raza, Z.; Yadav, S.; Truflandier, L.; Miyazaki, T.; Bowler, D. R. Large scale and linear scaling DFT with the CONQUEST code. *J. Chem. Phys.* **2020**, *152*, 164112.
- (7) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; Golze, D.; Wilhelm, J.; Chulkov, S.; Bani-Hashemian, M. H.; Weber, V.; Borštnik, U.; Taillefumier, M.; Jakobovits, A. S.; Lazzaro, A.; Pabst, H.; Müller, T.; Schade, R.; Guidon, M.; Andermatt, S.; Holmberg, N.; Schenter, G. K.; Hehn, A.; Bussy, A.; Belleflamme, F.; Tabacchi, G.; Glöß, A.; Lass, M.; Bethune, I.; Mundy, C. J.; Plessl, C.; Watkins, M.; VandeVondele, J.; Krack, M.; Hutter, J. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.
- (8) Olsen, J. M. H.; Reine, S.; Vahtras, O.; Kjellgren, E.; Reinholdt, P.; Dundas, K. O. H.; Li, X.; Cukras, J.; Ringholm, M.; Hedegård, E. D.; Remigio, R. D.; List, N. H.; Faber, R.; Tenorio, B. N. C.; Pedersen, T. B.; Rinkevicius, Z.; Sauer, S. P. A.; Mikkelsen, K. V.; Kongsted, J.; Coriani, S.; Ruud, K.; Helgaker, T.; Jensen, H. J. A.; Norman, P. Dalton Project: A Python platform for molecular- and electronic-structure simulations of complex systems. *J. Chem. Phys.* **2020**, *152*, 214115.
- (9) García, A.; Papior, N.; Akhtar, A.; Artacho, E.; Blum, V.; Bosoni, E.; Brandimarte, P.; Brandbyge, M.; Cerdá, J. I.; Corsetti, F.; Cuadrado, R.; Dikan, V.; Ferrer, J.; Gale, J.; García-Fernández, P.; García-Suárez, V. M.; García, S.; Huhs, G.; Illera, S.; Korytár, R.; Koval, P.; Lebedeva, I.; Lin, L.; López-Tarifa, P.; Mayo, S. G.; Mohr, S.; Ordejón, P.; Postnikov, A.; Pouillon, Y.; Pruneda, M.; Robles, R.; Sánchez-Portal, D.; Soler, J. M.; Ullah, R.; Yu, V.W.-z.; Junquera, J. SIESTA: Recent developments and applications. *J. Chem. Phys.* **2020**, *152*, 204108.
- (10) Prentice, J. C. A.; Aarons, J.; Womack, J. C.; Allen, A. E. A.; Andrinopoulos, L.; Anton, L.; Bell, R. A.; Bhandari, A.; Bramley, G. A.; Charlton, R. J.; Clements, R. J.; Cole, D. J.; Constantinescu, G.; Corsetti, F.; Dubois, S.M.-M.; Duff, K. K. B.; Escartín, J. M.; Greco, A.; Hill, Q.; Lee, L. P.; Linscott, E.; O'Regan, D. D.; Phipps, M. J. S.; Ratcliff, L. E.; Serrano, A. R.; Tait, E. W.; Teobaldi, G.; Vitale, V.; Yeung, N.; Zuehlsdorff, T. J.; Dziedzic, J.; Haynes, P. D.; Hine, N. D. M.; Mostofi, A. A.; Payne, M. C.; Skylaris, C.-K. The ONETEP linear-scaling density functional theory program. *J. Chem. Phys.* **2020**, *152*, 174111.
- (11) Rudberg, E.; Rubensson, E. H.; Salek, P.; Kruchinina, A. Ergo: An open-source program for linear-scaling electronic structure calculations. *SoftwareX* **2018**, *7*, 107–111.
- (12) Goedecker, S. Linear scaling electronic structure methods. *Rev. Mod. Phys.* **1999**, *71*, 1085–1123.
- (13) Yang, W.; Lee, T.-S. A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules. *J. Chem. Phys.* **1995**, *103*, 5674–5678.
- (14) Galli, G. Large-scale electronic structure calculations using linear scaling methods. *Physica Status Solidi B-Basic Research* **2000**, *217*, 231–249.
- (15) Scuseria, G. E. Linear scaling density functional calculations with Gaussian orbitals. *J. Phys. Chem. A* **1999**, *103*, 4782–4790.
- (16) Baer, R.; Head-Gordon, M. Sparsity of the Density Matrix in Kohn-Sham Density Functional Theory and an Assessment of Linear System-Size Scaling Methods. *Phys. Rev. Lett.* **1997**, *79*, 3962–3965.
- (17) Baer, R.; Head-Gordon, M. Energy renormalization-group method for electronic structure of large systems. *Physical Review B-Condensed Matter* **1998**, *58*, 15296–15299.
- (18) Baer, R.; Neuhauser, D.; Rabani, E. Self-Averaging Stochastic Kohn-Sham Density-Functional Theory. *Phys. Rev. Lett.* **2013**, *111*, 106402.
- (19) Osei-Kuffuor, D.; Fattebert, J.-L. Accurate and Scalable O(N) Algorithm for First-Principles Molecular-Dynamics Computations on Large Parallel Computers. *Phys. Rev. Lett.* **2014**, *112*, No. 046401.
- (20) Cytter, Y.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic Density Functional Theory at Finite Temperatures. *Phys. Rev. B* **2018**, *97*, 115207.
- (21) Fabian, M. D.; Shpiro, B.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic density functional theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, *9*, No. e1412.
- (22) Li, W.; Chen, M.; Rabani, E.; Baer, R.; Neuhauser, D. Stochastic embedding DFT: Theory and application to p-nitroaniline in water. *J. Chem. Phys.* **2019**, *151*, 174115.
- (23) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Energy window stochastic density functional theory. *J. Chem. Phys.* **2019**, *151*, 114116.
- (24) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Overlapped embedded fragment stochastic density functional theory for covalently-bonded materials. *J. Chem. Phys.* **2019**, *150*, No. 034106.
- (25) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Stochastic density functional theory: Real- and energy-space fragmentation for noise reduction. *J. Chem. Phys.* **2021**, *154*, 204108.
- (26) Corsetti, F. Performance Analysis of Electronic Structure Codes on HPC Systems: A Case Study of SIESTA. *PLoS One* **2014**, *9*, No. e95390.
- (27) Arita, M.; Arapan, S.; Bowler, D. R.; Miyazaki, T. Large-scale DFT simulations with a linear-scaling DFT code CONQUEST on K-computer. *Journal of Advanced Simulation in Science and Engineering* **2014**, *1*, 87–97.
- (28) Goedecker, S.; Colombo, L. Efficient Linear Scaling Algorithm for Tight-Binding Molecular Dynamics. *Phys. Rev. Lett.* **1994**, *73*, 122–125.
- (29) Goedecker, S.; Teter, M. Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals. *Phys. Rev. B* **1995**, *51*, 9455–9464.
- (30) Goedecker, S. Low Complexity Algorithms for Electronic Structure Calculations. *J. Comput. Phys.* **1995**, *118*, 261–268.
- (31) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993–2006.
- (32) Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* **1999**, *110*, 2810–2821.
- (33) Kleinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.
- (34) Wolfram. *Mathematica*, ver. 12.1.
- (35) Herbert, J. M.; Head-Gordon, M. Calculation of Electron Detachment Energies for Water Cluster Anions: An Appraisal of Electronic Structure Methods, with Application to $(\text{H}_2\text{O})_{20}^-$ and $(\text{H}_2\text{O})_{24}^-$. *J. Phys. Chem. A* **2005**, *109*, 5217–5229.
- (36) Rudberg, E. Difficulties in applying pure Kohn–Sham density functional theory electronic structure methods to protein molecules. *J. Phys.: Condens. Matter* **2012**, *24*, No. 072202.
- (37) Lever, G.; Cole, D. J.; Hine, N. D. M.; Haynes, P. D.; Payne, M. C. Electrostatic considerations affecting the calculated HOMO–LUMO gap in protein molecules. *J. Phys.: Condens. Matter* **2013**, *25*, 152101.
- (38) Sosa Vazquez, X. A.; Isborn, C. M. Size-dependent error of the density functional theory ionization potential in vacuum and solution. *J. Chem. Phys.* **2015**, *143*, 244105.
- (39) Scott, J.; Tüma, M. HSL_MI28: An Efficient and Robust Limited-Memory Incomplete Cholesky Factorization Code. *ACM Trans. Math. Softw.* **2014**, *40*, 1–19.
- (40) Baer, R.; Head-Gordon, M. Chebyshev expansion methods for electronic structure calculations on large molecular systems. *J. Chem. Phys.* **1997**, *107*, 10003–10013.
- (41) Pulay, P. Improved SCF convergence acceleration. *J. Comput. Chem.* **1982**, *3*, 556–560.

(42) Shpiro, B.; Fabian, M. D.; Rabani, E.; Baer, R. Forces from Stochastic Density Functional Theory under Nonorthogonal Atom-Centered Basis Sets. *J. Chem. Theory Comput.* **2022**, *18* (3), 1458–1466.

(43) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kus, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; III, H.L. W.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; Jr, R.A. D.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stock, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; III, W. A. G.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; III, H. F. S.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Voorhis, T. V.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(44) Jensen, S. R.; Saha, S.; Flores-Livas, J. A.; Huhn, W.; Blum, V.; Goedecker, S.; Frediani, L. The Elephant in the Room of Density Functional Theory Calculations. *J. Phys. Chem. Lett.* **2017**, *8*, 1449.

(45) Arnon, E.; Rabani, E.; Neuhauser, D.; Baer, R. Equilibrium configurations of large nanostructures using the embedded saturated-fragments stochastic density functional theory. *J. Chem. Phys.* **2017**, *146*, 224111.