JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

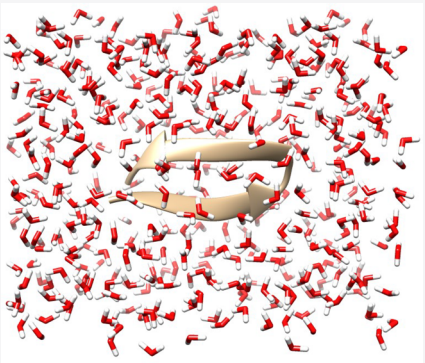
Forces from Stochastic Density Functional Theory under Nonorthogonal Atom-Centered Basis Sets

Ben Shpiro, Marcel David Fabian, Eran Rabani,* and Roi Baer*

Cite This: *J. Chem. Theory Comput.* 2022, 18, 1458−1466 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We develop a formalism for calculating forces on the nuclei within the linear-scaling stochastic density functional theory (sDFT) in a nonorthogonal atom-centered basis set representation (Fabian et al. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2019**, *9*, e1412, 10.1002/wcms.1412) and apply it to the Tryptophan Zipper 2 (Trp-zip2) peptide solvated in water. We use an embedded-fragment approach to reduce the statistical errors (fluctuation and systematic bias), where the entire peptide is the main fragment and the remaining 425 water molecules are grouped into small fragments. We analyze the magnitude of the statistical errors in the forces and find that the systematic bias is of the order of 0.065 eV/Å ($\sim 1.2 \times 10^{-3} E_h/a_0$) when 120 stochastic orbitals are used, independently of system size. This magnitude of bias is sufficiently small to ensure that the bond lengths estimated by stochastic DFT (within a Langevin molecular dynamics simulation) will deviate by less than 1% from those predicted by a deterministic calculation.



1. INTRODUCTION

Kohn−Sham density functional theory (KS-DFT) is often used for estimating the forces on the nuclei in *ab initio* molecular dynamics simulations, with which reliable predictions concerning structure and properties of molecules can be made. Despite the fact that it can be used to study extended molecular systems relevant to biomolecular chemistry and materials science,[1−4] the conventional applications are limited in size due to the cubic algorithm complexity. Therefore, several approaches to KS-DFT have been developed and are routinely used for treating such extended systems. These include linear-scaling approaches which rely on electron localization within the system's interior volume,[5−33] or the tight-binding DFT approach, which uses a very small basis set complemented by approximations calibrated with empirical data,[34−36] and the orbital-free DFT, which is applicable to relatively homogeneous systems.[37,38] The way many of the linear scaling approaches achieve their gentle algorithmic complexity involves imposing a sparse structure on the KS density matrix (DM) in a local real-space basis representation, effectively truncating the protruding elements. The rationale of this procedure relies on the electron localization which characterizes many large systems.[39] However, in metallic systems at low-temperature and for low band semiconductors, the localization length is very large, and such approaches are difficult to apply.[17]

In order to enable treatment of systems in which electron coherence is nonlocal, a different linear scaling approach was proposed and dubbed stochastic density functional theory (sDFT).[40] In sDFT, we use a sparse representation of the KS-DM which does not rely on truncation or modification of its elements. Instead, sDFT is based on the paradigm that the expectation values of the system observables can be regarded as random variables in a stochastic process with an expected value and a fluctuation. The fact that estimation of electronic structure quantities can be done by statistical sampling allows for a natural and highly effective implementation of sDFT on parallel architectures.

The source of errors in sDFT is statistical in nature and involves fluctuations, the magnitude of which can be controlled by statistical sampling theory and/or by variance-reducing techniques, such as the embedded-fragment method,[41−44] or the energy windowing approach.[45,46] In addition to statistical fluctuations, the sDFT estimates of the electron density and the forces exhibit bias errors resulting from the nonlinear nature of the SCF iterations.[44,47] The magnitude of the bias can be controlled by using the above-mentioned variance-reducing techniques.

Early implementations of sDFT were based on real-space grid representations of the electron density[40,42,43,47,48] and applied to relatively homogeneous systems: either to pure bulk silicon,[43,47] silicon with impurities,[46] H−He mixtures[49] or to finite-sized hydrogen-passivated silicon nanocrystals and water clusters.[42,50,51] We recently demonstrated that the noisy forces

Received: August 6, 2021
Published: January 31, 2022

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

produced by sDFT in the real-space grid representation can be used within a Langevin dynamics approach, to determine structural properties of such large systems.[42,52]

The real-space implementation of sDFT is especially useful as a starting point for postprocessing DFT-based methods, such as the stochastic GW for charge excitations,[53,54] the stochastic time-dependent DFT and Bethe-Salpeter equations for neutral excitations,[55−57] and for conductance calculations in warm dense matter.[49]

If one is only interested in the ground state atomistic structure, real-space grid representation could be quite expensive, and a more efficient representation may be beneficial. For this purpose, we recently developed an sDFT approach based on nonorthogonal atom-centered basis sets.[44] We found that the Hamiltonian within this compact basis has a much smaller energy range than in the real-space grid, allowing a significant speedup of sDFT calculations.

Despite the fact that sDFT with the nonorthogonal atom-centered basis set is designed to address the structural properties of large systems, up to now, we did not have the capability to estimate the forces on nuclei and therefore focused only on the electronic energy and density of states.[44] In this paper, we develop the necessary theory and computational tools for calculating the forces while maintaining the linear-scaling complexity of sDFT. In addition, we analyze the statistical fluctuations and the biases in the forces, using as a benchmark the heterogeneous system of the Tryptophan Zipper 2 (Trp-zip2) peptide solvated in water.

The manuscript is organized as follows: In Section 2, we introduce our formalism for the stochastic forces calculations. Then, in Section 3, we present the benchmark calculations on the Tryptophan Zipper 2 (Trp-zip2) peptide in solution. Finally, we summarize and discuss the results in Section 4.

2. FORCE CALCULATIONS IN STOCHASTIC DENSITY FUNCTIONAL THEORY

In this section, we describe the theory of the electronic forces on nuclei within the finite temperature KS-DFT formalism. We set the notations and describe the basis set representation we use for Kohn−Sham DFT in subsection 2.1 with the combined implementation using real space grids briefly described in subsection 2.2. Expressions for the forces are given in subsection 2.3 with a detailed derivation given in Appendix A. Finally, in subsections 2.4 and 2.5, we provide the details behind the stochastic evaluation of the electronic density and any other observables in sDFT (including the forces) and present the statistical errors involved.

2.1. Setting the Stage. The KS Hamiltonian is given by

$$\hat{h}_{\mathrm{KS}} = \hat{t}_s + \hat{v}_{pp}^{nl} + \hat{v}_{pp}^{loc} + v_{Hxc}[n](\boldsymbol{r}) \tag{1}$$

where $\hat{t}_s = -(1/2)\nabla^2$ (we use atomic units throughout the paper) is the electron kinetic energy operator, and $\hat{v}_{pp}^{nl} = \sum_{C\in nuclei}\hat{v}_{pp(C)}^{nl}$, and $\hat{v}_{pp}^{loc} = \sum_{C\in nuclei}v_{pp(C)}^{loc}(\hat{\boldsymbol{r}} - \boldsymbol{R}_C)$ are the nonlocal and local norm-conserving pseudopotential terms in the Kleinman-Bylander form,[58,59] for nucleus C, at position $\boldsymbol{R}_C$. The last potential term, $\hat{v}_{Hxc}$, is the Hartree and exchange correlation potential, depending on the electron density, $n(\boldsymbol{r})$

$$v_{Hxc}[n](\boldsymbol{r}) = \frac{\delta\mathcal{E}_{Hxc}[n]}{\delta n(\boldsymbol{r})} = \int\frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3r' + v_{xc}[n](\hat{\boldsymbol{r}}) \tag{2}$$

where $\mathcal{E}_{\mathrm{Hxc}}[n]$ is the Hartree and exchange-correlation energy functional.

We use a nonorthogonal atom-centered basis set, $\phi_\alpha(\boldsymbol{r})$, $\alpha = 1,...,K$, with an overlap matrix $S_{\alpha\gamma} = \langle\phi_\alpha|\phi_\gamma\rangle$, $\alpha, \gamma = 1,...,K$. Within such a basis set approach, the $K \times K$ DM is given as an operator involving a function of HS^{-1}

$$P = S^{-1}f(HS^{-1};\beta,\mu) \tag{3}$$

where $H_{\alpha\gamma} = \langle\phi_\alpha|\hat{h}_{\mathrm{KS}}|\phi_\gamma\rangle$ and

$$f(\varepsilon;\beta,\mu) \equiv \frac{1}{1+e^{\beta(\varepsilon-\mu)}} \tag{4}$$

is the Fermi−Dirac distribution function. The DM is used to calculate expected values of single-electron observables $\hat{o}$ as

$$\langle\hat{o}\rangle = 2\times\mathrm{Tr}[OP] \tag{5}$$

where O is the matrix representing $\hat{o}$ in the basis, with elements

$$O_{\alpha\gamma} = \langle\phi_\alpha|\hat{o}|\phi_\gamma\rangle \tag{6}$$

and the factor of 2 accounts for the electron's spin in a closed shell representation. For example, the expectation value of the density operator $\hat{n}(\boldsymbol{r})$ is the electron density, given by

$$n[P](\boldsymbol{r}) = \langle\delta(\boldsymbol{r}-\hat{\boldsymbol{r}})\rangle = 2\times\sum_{\alpha\gamma}P_{\alpha\gamma}\phi_\alpha(\boldsymbol{r})\phi_\gamma(\boldsymbol{r}) \tag{7}$$

The DM in eq 3 minimizes the total electronic free-energy:

$$\Omega[P] = \mathcal{E}[P] - \mu\mathcal{N}[P] - (k_B\beta)^{-1}\mathcal{S}_{ent}[P] \tag{8}$$

Here, $\mathcal{E}[P]$ is the electronic internal energy

$$\mathcal{E}[P] = 2\times\mathrm{Tr}[(T_s + V_{PP}^{nl} + V_{PP}^{loc})P] + \mathcal{E}_{Hxc}[n[P]]$$

and the number of electrons is given by

$$\mathcal{N}[P] = 2\times\mathrm{Tr}[SP]$$

The actual value we use for the chemical potential μ is tuned to enforce $\mathcal{N}[P]$ to be equal to the actual number of electrons in the system (see ref 44 for detail). Finally, $\mathcal{S}_{\mathrm{ent}}[P]$ is the entropy of the noninteracting electrons of the KS system, given by

$$\mathcal{S}_{ent}[P] = -2\times k_B\mathrm{Tr}[SP\ln SP + (1-SP)\ln(1-SP)]$$

Equations 1−7 must be solved together, and the resulting solution for the density $n(\boldsymbol{r})$ and the DM P is called the self-consistent field (SCF) solution to the KS equations. The procedure for reaching SCF solution is iterative: in each iteration, called an SCF cycle, P is calculated from H using eq 3, $n(\boldsymbol{r})$ is from P from which $v_{Hxc}[n](\boldsymbol{r})$ is calculated, and a new KS Hamiltonian matrix H is built.

2.2. Combined Real-Space Grid and Basis Set Implementation. The theory described in the section above uses, in addition to the basis function $\phi_\alpha(\boldsymbol{r})$, also a Cartesian grid (with uniform grid-spacing h) which spans the space occupied by the electron density. The grid is used to evaluate the matrix elements of eq 6 of various observables $\hat{o}$, expressible as operators on the grid

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

$$O_{\alpha\gamma} = h^3 \sum_g \phi_\alpha(\boldsymbol{r}_g)[\hat{o}\phi_\gamma](\boldsymbol{r}_g) \tag{9}$$

where $\boldsymbol{r}_g$ are the grid points (g is a 3D index). Each matrix element of eq 9 can be evaluated efficiently, while we can also gain by parallel architecture, allowing different cores to independently compute different $\alpha\gamma$ pairs. [This requires a fast evaluation of basis functions $\phi_\alpha(\boldsymbol{r}_g)$ at the grid points. For this, we employ standard quantum-chemical Cartesian functions, expressible as sums of triple products, $\phi_\alpha(x, y, z) = \sum_p \xi_\alpha^p(x)\eta_\alpha^p(y)\zeta_\alpha^p(z)$ where the sums of $\xi_\alpha^p(x)$, $\eta_\alpha^p(y)$, and $\zeta_\alpha^p(z)$ are the primitive functions of the basis. At grid point $\boldsymbol{r}_g$, the basis function is a sum (over the primitive functions) of triple products formed from three 1D vectors: $\xi_\alpha^p(x_g)$, $\eta_\alpha^p(y_g)$, and $\zeta_\alpha^p(z_g)$ which are kept in memory. The same technique is used for the evaluation of the derivatives of the basis functions, which is relevant for the calculation of forces, see Supporting Information, Section S1.] In particular, the pseudopotentials $\hat{v}_{PP}^{nl/loc}$ are such grid operators. Evaluating the electron density of eq 7 at the grid points allows calculation of the density-dependent Hartree and XC potentials. For the former, we use fast Fourier transform techniques.[60]

2.3. Electronic Forces on the Nuclei. In this subsection, we give formal expressions for the electronic forces on the nuclei expressible as matrix trace operations, based on a finite temperature formalism presented in Appendix A. Our derivation and final results are similar yet differ in many ways with those of ref 61. We calculate the work done by the electrons as nucleus C is displaced by $\delta_C X$ in the x-coordinate. This work is the change in the free energy of eq 8, and therefore

$$-F_C\delta_C X = \delta_C\Omega \tag{10}$$

where F_C is the x-component of the force on the displaced nucleus. The atom displacement $\delta_C X$ has three types of effects: it causes an explicit change in its contribution to the pseudopotential $\hat{v}_{pp}^{nl/loc} \to \hat{v}_{pp}^{nl/loc} + \delta_C\hat{v}_{pp}^{nl/loc}$, it displaces the basis functions $\phi_\alpha \to \phi_\alpha + \delta_C\phi_\alpha$, and it induces a variation in the DM, $P \to P + \delta_C P$, since P is required to be the minimizer of the free energy. Note that due to this minimum principle $\delta_C\Omega$ is unaffected (to first order) by $\delta_C P$ so that the work done on the atom (see Appendix A)

$$-F_C\delta_C X = 2 \times \mathrm{Tr}[P(\delta_C H - (HS^{-1})\delta_C S)] \tag{11}$$

is given solely in terms of the variations in the Hamiltonian

$$(\delta_C H)_{\alpha\beta} = \langle\phi_\alpha|\delta_C(\hat{v}_{pp}^{nl} + \hat{v}_{pp}^{loc})|\phi_\beta\rangle + \langle\delta_C\phi_\alpha|\hat{h}_{KS}|\phi_\beta\rangle + \langle\phi_\alpha|\hat{h}_{KS}|\delta_C\phi_\beta\rangle \tag{12}$$

and the overlap

$$(\delta_C S)_{\alpha\beta} = \langle\delta_C\phi_\alpha|\phi_\beta\rangle + \langle\phi_\alpha|\delta_C\phi_\beta\rangle \tag{13}$$

matrices. The first term in eq 12 is the explicit change in the pseudopotential, giving the direct forces on the atom. The second and third terms in $\delta_C H$ (and similar terms in eq 13 for $\delta_C S$) are due to the variation in basis functions, and they lead to the so-called Pulay forces,[62] on the atom. More details concerning the calculation of $(\delta_C S)_{\alpha\beta}$ and $(\delta_C H)_{\alpha\beta}$ are given in Supporting Information, Section S1.

The estimation of the expectation value of a one-body observable $\hat{o}$, given by eq 5, requires the calculation of the trace of the matrix OP. By definition $\mathrm{Tr}[OP] = \sum_{k=1}^K (u^k)^T OP u^k$ where u^k is a set of K orthogonal unit vectors, and the numerical effort involves K applications of OP on a vector u, each of which scales quadratically, and thus the overall effort scales as $O(K^3)$.

One essential component in reducing the scaling of this step is to exploit the sparsity of the $S^{-1}H$ operation on a vector v, which is used within a Chebyshev expansion,[63] as a Fermi–Dirac function representing P (see eq 3). [The application of S^{-1} on a column vector involves repeated applications of S on the vector, within the preconditioned conjugate gradient method, implemented in the HSL-MA61 code. HSL is a collection of FORTRAN codes for large scale scientific computation (http://www.hsl.rl.ac.uk/).] This leads to the following method for applying the DM onto a vector v[8,64]

$$Pv = \sum_{n=0}^{N_C} a_n(\beta, \mu)v_n \tag{14}$$

where v_n, $n = 0,1,\ldots$ is obtained recursively

$$\begin{aligned} v_0 &= v \\ v_1 &= \left[\frac{HS^{-1} - \bar{E}}{\Delta E}\right]v_0 \\ v_{n+1} &= 2\left[\frac{HS^{-1} - \bar{E}}{\Delta E}\right]v_n - v_{n-1} \end{aligned} \tag{15}$$

Here, $\frac{HS^{-1} - \bar{E}}{\Delta E}$ is a shifted-scaled operator with eigenvalues in the interval $[-1,1]$ (So ΔE is equal to half the spectral range, and $\bar{E}$ is its center.). The expansion coefficients depend on β and μ characterizing the Fermi–Dirac function; they rapidly decay to zero once N_C exceeds a system size independent value determined by $\beta\Delta E$. With this technique, the step Pu^k involves a linear scaling effort, and since there are K such vectors, the complexity of the trace operation $\mathrm{Tr}[OP]$ is reduced from $O(K^3)$ to $O(K^2)$.[65]

2.4. Stochastic Estimation of Observables and Forces. In order to further reduce the numerical effort to linear scaling, we use a stochastic vector approach, where the trace is sampled using I stochastic vectors instead of *calculated* using a complete set of K orthonormal vectors. The calculation effort is reduced from $O(K^2)$ to $O(IK)$, and I is system independent. A full exposition of the method is given in ref 44. Here, we briefly mention the essential elements.

Stochastic vectors, $\chi^T = (\chi^1,\ldots,\chi^K)$, have K random components, χ^k; each is a random variable taking the values ± 1 with equal probability. We refer the reader to Section S2 of the Supporting Information for definition and discussion of random variables (collectively denoted r), their expected values $\mathrm{E}[r]$, their variance $\mathrm{Var}[r]$, and the statistical methods for evaluating these quantities using finite samples. For each component of the stochastic vector, (1) $|\chi^k| = 1$ (2) $\mathrm{E}[\chi^k] = 0$, and therefore, $\mathrm{Var}[\chi^k] = 1$. Furthermore, the product $\chi^k\chi^j$ of any pair of components has a zero expected value ($\mathrm{E}[\chi^k\chi^j] = 0$, $k \neq j$), and hence, in matrix form

$$\mathrm{E}[\chi\chi^T] = \mathrm{Id} \tag{16}$$

where Id is the $K \times K$ identity matrix. We view eq 16 as the "stochastic resolution of the identity", and using it, we express the trace of the matrix OP as $\mathrm{Tr}[OP] = \mathrm{Tr}[OP\mathrm{E}[\chi\chi^T]] = \mathrm{E}[\mathrm{Tr}[OP\chi\chi^T]]$, which upon rearrangement gives the stochastic trace formula:[66]

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

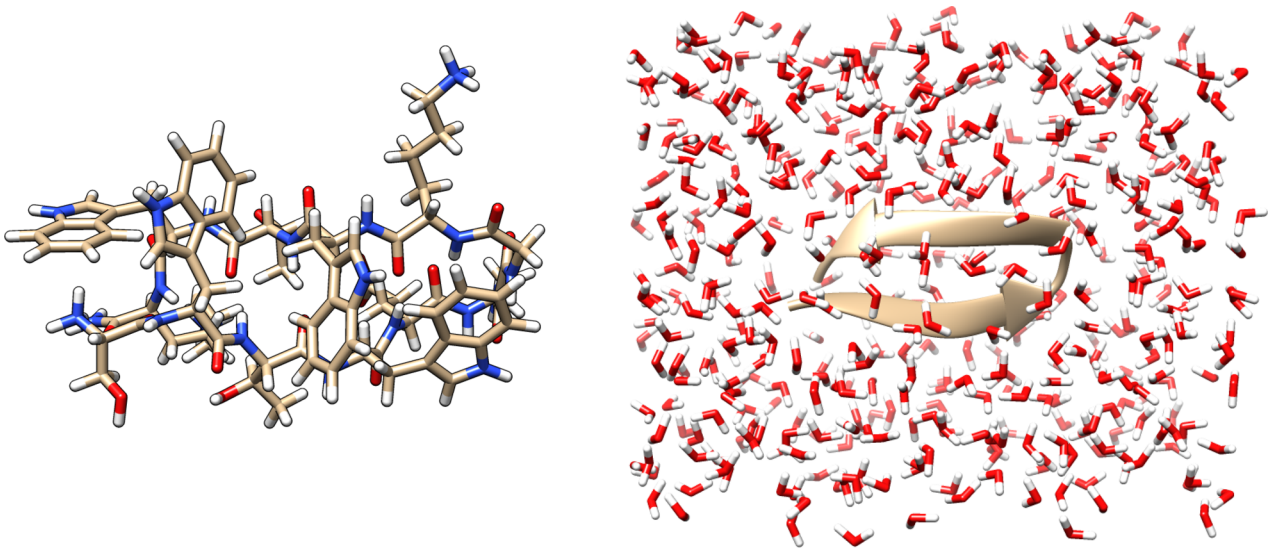


Figure 1. Left panel: Tryptophan Zipper 2 (Trp-zip2) peptide, composed of 220 atoms. Right panel: Trp-zip2 peptide (ribbon) solvated by 425 water molecules. The full system is composed of 1495 atoms and 4024 valence electrons, and 3118 basis functions are necessary to describe it using a minimal basis set.

$$\langle \hat{o} \rangle = 2 \times \mathrm{E}[\chi^T OP\chi] \tag{17}$$

The expected value $\mathrm{E}[\chi^T OP\chi]$ can be estimated using a sample of size I with

$$m_I = 2 \times \frac{1}{I} \sum_{i=1}^{I} \chi_i^T OP\chi_i \tag{18}$$

which establishes a 70% confidence interval $[m_I - \sigma_I, m_I + \sigma_I]$ for $\langle \hat{o} \rangle$ where

$$\sigma_I = \frac{s_I}{\sqrt{I}} \tag{19}$$

and $s_I = \sqrt{\frac{1}{I-1} \sum_{i=1}^{I} (\chi_i^T OP\chi_i - m_I)^2}$ is the standard deviation. We would like to highlight that since eq 18 is an average over i independent $\chi_i^T OP\chi_i$ terms, the computation is easily implemented to gain from parallel architecture.

We can use the stochastic trace to estimate the electron density at each grid point, based on eq 7. For this, we define stochastic orbitals which are stochastic linear combinations of the basis functions, defined on the grid as

$$\eta_i(\boldsymbol{r}_g) = \sum_{\alpha=1}^{K} \chi_i^\alpha \phi_\alpha(\boldsymbol{r}_g)$$

and *projected* stochastic orbitals

$$\xi_i(\boldsymbol{r}_g) = \sum_{\alpha=1}^{K} [P\chi_i]^\alpha \phi_\alpha(\boldsymbol{r}_g)$$

Using the above, we can now calculate the center of the confidence interval for the electron density at point $\boldsymbol{r}_g$ as the sample mean:

$$n_I(\boldsymbol{r}_g) = 2 \times \frac{1}{I} \sum_{i=1}^{I} \eta_i(\boldsymbol{r}_g) \xi_i(\boldsymbol{r}_g) \tag{20}$$

In ref 44 we have presented CPU times showing linear scaling in the calculation of sDFT observables.

The above technique can be used to evaluate the electronic forces on the nuclei as they too are formulated as matrix traces (see eq 11). The computational effort for evaluating the direct forces coming from $\hat{v}_{pp}^{nl}$ (the non-local pseudopotential) as well as all Pulay terms, for each degree of freedom, are independent of the system size since they are local (See Supporting Information Section S1.C. for detail). The computational effort for evaluating the force coming from $\hat{v}_{pp}^{loc}$ (the local pseudopotential), for each degree of freedom, will scale linearly unless specialized particle mesh methods (beyond the scope of this paper) are used.

The SCF cycle of KS theory in sDFT involves using our best estimate for the density, i.e., $n_I(\boldsymbol{r})$, to build the Hamiltonian. Since $n_I(\boldsymbol{r})$ includes an uncertainty (a fluctuation), the resulting Hamiltonian matrix H also has a fluctuation. Then, plugging H into the Chebyshev expansion from which a new $n_I(\boldsymbol{r})$ is calculated converts the fluctuation into a bias, as discussed in Section S2.C of the Supporting Information. Thus, after the SCF converges, all expectation values have both an uncertainty σ_I and a bias error, which we define as

$$\Delta\rho_I = |\mathrm{E}[m_I] - \langle \hat{o} \rangle^{dDFT}|$$

The estimation of the uncertainty σ_I can be done using eq 19, but the estimation of $\Delta\rho_I$ is more complicated since we need to determine $\mathrm{E}[m_I]$. We discuss this issue when we determine the bias error in the force (see Section 3).

2.5. Embedded Fragments Approach. In order to mitigate the fluctuation and bias errors, we developed a basis set version of the embedded-fragment (EF) approach,[41−44] which can be described in a general way as introducing a correction term to the sDFT calculation. We first split all the atoms in the system into F fragments, such that each atom, and all basis functions centered on it, belong to one and only one fragment. If the fragments are chosen such that their size is independent of the total system size, with sublinear scaling and minimal increase in computation time we can calculate the electron density in each fragment, using 1.) deterministic DFT $n_{\mathrm{dDFT}}^f(\boldsymbol{r})$ ($f = 1,...,F$) and 2.) stochastic DFT $n_I^f(\boldsymbol{r})$. We then use the difference

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

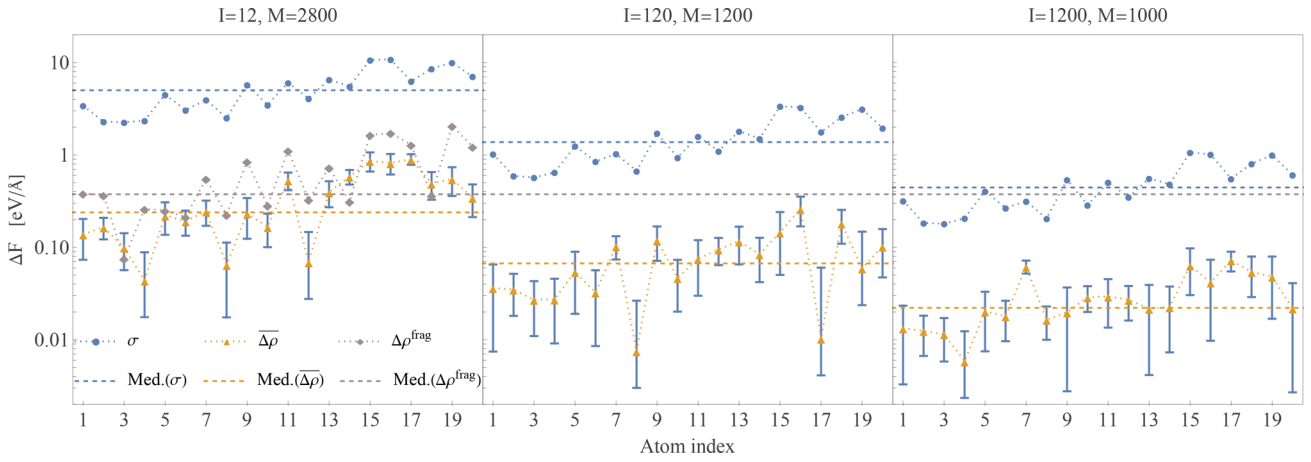


Figure 2. Statistical errors in the sDFT forces acting on the 20 nitrogen atoms in the solvated-TrpZip2 system calculated using I = 12, 120, 1200 stochastic vectors (see left, center, and right panels). For each nitrogen atom, we show the uncertainty σ_C (blue dots) and the estimate in the bias $\Delta\rho_C$ (orange triangles), see eqs 24 and 25 in the text, with error bars calculated as $\pm\sigma_C/\sqrt{M}$. In the I = 12 column, we also plot $\Delta\rho_C^{frag} = \|\boldsymbol{F}_C^{frag} - \boldsymbol{F}_C^{dDFT}\|$ (gray diamonds), where $\boldsymbol{F}_C^{frag}$ is the dDFT force vector on the nitrogen atom C from the peptide-only fragment calculation. The dotted lines connecting the markers are presented as a guide for the eye, while the dashed horizontal lines are medians over all atoms of σ_C and $\Delta\rho_C$. For simplification of the image, in the I = 120, 1200 columns, we only present the median of $\Delta\rho_C^{frag}$ (gray dashed line) taken over all 20 nitrogen atoms.

$$\Delta n^f(\boldsymbol{r}) = n^f_{\mathrm{dDFT}}(\boldsymbol{r}) - n_I^f(\boldsymbol{r}) \tag{21}$$

as a correction to the sDFT calculation of the density $n_I(\boldsymbol{r})$ on the entire system:

$$n_I^{EF}(\boldsymbol{r}) = n_I(\boldsymbol{r}) + \sum_{f=1}^{F} \Delta n^f(\boldsymbol{r}) \tag{22}$$

We note that the correct result, $n_I^{EF}(\boldsymbol{r}) = n_{dDFT}(\boldsymbol{r})$, is obtained in two limits: 1) when F = 1 (i.e., the entire system is a fragment) and 2) when $I \to \infty$, so $n_I^f(\boldsymbol{r}) \to n^f_{\mathrm{dDFT}}(\boldsymbol{r})$, etc. Similarly, the expectation value of any operator of interest, $\hat{o}$

$$\langle\hat{o}\rangle_I^{EF} = \langle\hat{o}\rangle_I + \sum_f \langle\Delta\hat{o}^f\rangle_I \tag{23}$$

where $\langle\Delta\hat{o}^f\rangle_I = \langle\hat{o}^f\rangle_{\mathrm{dDFT}} - \langle\hat{o}^f\rangle_I$. The EF approach is applicable to the forces calculation, by choosing $\hat{o}$ to be the relevant operators from eq 11. For further details on the implementation of the embedded fragments method in our program, see Supporting Information, Section S3.

3. STATISTICAL ANALYSIS OF SDFT FORCES IN THE TRYPTOPHAN ZIPPER 2 PEPTIDE

Our test system is a Tryptophan Zipper 2 (Trp-zip2) peptide (pdb 1le1), composed of 220 atoms (left panel of Figure 1), solvated with 425 water molecules, and built using a universal force field (UFF) in ArgusLab[67,68] (right panel of Figure 1). For benchmark calculations, we focused on the 20 nitrogen atoms of the peptide (indexed by C) and calculated the forces acting on each Cartesian degree of freedom. In these calculations, the embedded-fragment method was used, for which we chose to consider the peptide as a single fragment and then divided the 425 water molecules into 27 fragments, with an average size of 16 molecules.

To study the statistical errors, we performed the sDFT calculations using an increasing number of stochastic vectors, I = 12, 120, 1200, according to eq 17. To estimate the magnitudes of the bias and the uncertainty, we repeated the calculations M times (using independent random number generator seeds) from which we calculated a sample average force vector

$$\bar{\boldsymbol{F}}_C = \frac{1}{M}\sum_{m=1}^{M} \boldsymbol{F}_C^m$$

and a 3 × 3 force covariance matrix

$$\Sigma^2 = \frac{1}{M}\sum_{m=1}^{M} (\boldsymbol{F}_C^m - \bar{\boldsymbol{F}}_C)(\boldsymbol{F}_C^m - \bar{\boldsymbol{F}}_C)^T$$

as an estimate for the covariance of the sDFT calculation. As the forces acting on each atom are represented as 3-dimensional vectors (over the Cartesian coordinates), we would like to obtain scalar values, irrespective of the way the Cartesian axes are defined, in order to estimate the uncertainty and bias of the sDFT forces. [In addition to the analysis given here, we also present the distribution of the errors $F_C^m - F_C^{dDFT}$ in the Supporting Information, Section S2.D.] For a canonical estimate of the uncertainty, we use an average over the eigenstates of Σ^2

$$\sigma_C = \sqrt{\frac{1}{3}\mathrm{Tr}\Sigma^2} \tag{24}$$

where $F_C^{dDFT} = \|\boldsymbol{F}_C^{dDFT}\|$ is the magnitude of the dDFT electronic force on atom C. For a canonical estimate of the bias in the force, we use the L_2-Norm of the error in the average force vector:

$$\Delta\rho_C = \|\bar{\boldsymbol{F}}_C - \boldsymbol{F}_C^{dDFT}\| \tag{25}$$

In Figure 2, we present data for the statistical errors in the forces of the 20 nitrogen atoms, ordered by an atom index according to their distance from the center of the peptide (1 closest, 20 furthest). The estimates for the uncertainty in the forces, σ_C of eq 24, are plotted in blue circles, while the estimates of the bias $\Delta\rho_C$ of eq 25, with an error bar calculated as $\pm\sigma_C/\sqrt{M}$, are plotted in orange triangles with blue error bars. The medians over all nitrogen atoms are plotted as dashed lines. The used number of stochastic vectors, I, as well as the number of repetitions, M, is shown above each panel.

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

We found that stable estimates of σ_C are obtained even when using a small number of $M \approx 50$ repetitions and observe that they obey the expected $1/\sqrt{I}$ behavior in accordance with the central limit theorem. Since the variance is given by the matrix elements of the system, (see Supporting Information, Section S2.C, eq S3), the pattern seen for σ_C as a function of atom index is almost unchanged for different values of I. To estimate the bias, we need a good estimate of $\mathrm{E}[m_I]$ (the expected value of the forces when calculated using I stochastic vectors in eq 18). As σ_C is much larger than $\Delta\rho_C$, a very large number of repetitions, M, was required in order to achieve a good enough estimate of $\mathrm{E}[m_I]$ such that $\Delta\rho_C$ values are useful estimates of the bias. It is clear from the error bars that for almost all nitrogen atoms we have good estimates of the bias.

In the $I = 12$ column, for an added perspective, we plot, in gray diamonds, the error $\Delta\rho_C^{frag} = \|\boldsymbol{F}_C^{frag} - \boldsymbol{F}_C^{dDFT}\|$, where $\boldsymbol{F}_C^{frag}$ is the force vector on the nitrogen atom C from a dDFT calculation on its peptide only (gas-phase) fragment. The median is given again, in a dashed line. We observe that the values of $\Delta\rho_C^{frag}$ for the atoms closer to the center of the fragment are mostly smaller than those further away, causing a similar pattern in the sDFT errors. When comparing the median of $\Delta\rho_C^{frag}$ (plotted for all panels in a gray dashed line) with those of the stochastic results, we see they are higher even for the $I = 12$ stochastic vectors case, whereas for the cases of $I = 120$, 1200 we observe a reduction in the errors, showing that overall sDFT significantly improves the force estimation in comparison to the deterministic fragment calculation. [We base this conclusion on the medians of $\Delta\rho_C$. The same conclusions are valid also when considering the largest error, $\max_C\{\Delta\rho_C\}$.]

Additional sDFT calculations on a smaller system, composed of the Trp-zip2 peptide and only 195 solvating water molecules, show that for a given number of stochastic orbitals ($I = 12$) the uncertainty and bias are very similar to the case of the original solvated system (see Supporting Information, Section S4). This suggests the statistical errors are roughly independent of system size.

4. SUMMARY AND CONCLUSIONS

We have presented a method for force calculations within finite temperature sDFT in nonorthogonal atom-centered basis sets. The forces are random variables evaluated using the stochastic trace formula applied to various operators derived from the free energy and are therefore, like all sDFT observables, characterized by statistical errors, a fluctuation, and a bias. The calculation of the forces is adapted to benefit from the embedded-fragment methodology. These calculations are dominated by the SCF sDFT convergence step, and therefore the times for force calculations are similar to those reported in ref 44.

In Section 3, we presented benchmarking calculations, focusing on the statistical errors in the force estimates for the 20 nitrogen atoms of a solvated Tryptophan Zipper 2 peptide system. The results are given as a function of I, the number of stochastic vectors used in the calculation according to eq 17. The uncertainty in the sDFT forces follows the expected $1/\sqrt{I}$ behavior in accordance with the central limit theorem. Using a very large number of repetitions, we were also able to uncover the bias and determine that it is at least an order of magnitude smaller than the uncertainty. The magnitude of the force bias is of the order of 0.065 eV/Å ($\sim 10^{-3} E_h/a_0$) when 120 stochastic orbitals are used, independently of system size. A back-of-the-envelope calculation shows that this magnitude of bias is sufficiently small to ensure that the bond lengths estimated by stochastic DFT (within a Langevin molecular dynamics simulation) will deviate by less than 1% from those predicted by a deterministic calculation. [Assuming the minimum of the Born−Oppenheimer potential is harmonic with a local force constant k, the bond length deviation δR due to a force perturbation δF obeys $|k\delta R| = |\delta F|$. In typical solids and molecules, k is on the order of 5 to 100 eV Å^{-2},[69,70] so for δF of the order of 0.065 eV/Å, we find $\delta R \lesssim 0.01$ Å, 1% or less for most bond lengths of interest.] Indeed, this fact was demonstrated using a Langevin Dynamics simulation on silicon nanocrystals,[42] within a real-space representation sDFT. Our present results indicate that sDFT based on nonorthogonal atom-centered basis sets can be also used successfully in this way.

It is instructive to discuss the efficiency and accuracy of the basis set[44] vs real-space grid[40,42] representations of sDFT calculations. For this, we used the $Si_{35}H_{36}$ system, comparing the 6-31G basis set calculations with those of a real-space grid having 64^3 points and grid spacing of $\delta x = 0.5a_0$ (For more information about this comparison, see the Supporting Information, Section S5.). We find that the time for application of the density matrix to a random vector in the 6-31G basis is a factor of 30 faster than in the grid representation. On the other hand, surprisingly, the standard deviation of fluctuations in a typical Si force component is about 5 times larger in the basis set calculation than in the grid. Therefore, we need a factor of $5^2 = 25$ more stochastic vectors (because their number is proportional to the square of the standard deviation) in the basis set calculation for achieving the same fluctuation error. If we had only a single processor, the two representations would thus require a similar numerical effort for achieving a given fluctuation goal: the grid is 30 times slower but requires a factor of 25 less samplings. Due to the highly parallelizable nature of sDFT, the necessary extra sampling required by the basis-set-based calculation does not automatically lead to increased wall-times, if additional CPUs can be offered. We conclude that the basis-set-based calculations can achieve smaller wall-times than real-space grids, given additional CPUs.

APPENDIX A: DERIVATION OF THE CHANGES IN FREE ENERGY

Here, we derive the force expression of eq 11. The force is given by the change in free energy

$$\Omega[P] = \mathcal{E}[P] - \mu\mathcal{N}[P] - (k_B\beta)^{-1}\mathcal{S}_{ent}[P]$$

due to displacement of the nuclei. When nuclei are displaced, the DM also changes, and we will show that under any change in the density matrix $P \to P + \delta_0 P$, while keeping the nuclei fixed, the free energy of eq 8 does not change when P is given by eq 3. This will be done by examining each term in the above equation separately and summing over all of them. Then, we will consider the direct change in free energy due to a displacement of the nuclei (while P is held constant). It is only this latter change which affects the free energy.

A.1. The Variation in $\mathcal{N}[P]$

Starting from

$$\mathcal{N}[P] = \int n[P](\boldsymbol{r})\, d\boldsymbol{r}$$

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

and

$$n[P](\boldsymbol{r}) = 2 \times \sum_{\alpha\gamma} P_{\alpha\gamma}\phi_\alpha(\boldsymbol{r})\phi_\gamma(\boldsymbol{r})$$

Combining these we see

$$\mathcal{N}[P] = 2 \times \mathrm{Tr}[SP]$$

We consider two types of variations: δ_0 which changes the DM but not the atoms and δ_C which changes the position of atom C (and thus affects the basis functions associated with that atom) but not P.

1. $P \to P + \delta_0 P$ (assuming nuclei are constant): Here

$$\delta_0 n(\boldsymbol{r}) = 2 \times \sum_{\alpha\gamma} \delta_0 P_{\alpha\gamma}\phi_\alpha(\boldsymbol{r})\phi_\gamma(\boldsymbol{r}) \tag{26}$$

so

$$\delta_0 \mathcal{N}[P] = 2 \times \mathrm{Tr}[S\delta_0 P] \tag{27}$$

2. Nucleus C moves by $\delta_C X$ (and $\phi_\alpha \to \phi_\alpha + \delta_C\phi_\alpha$) (constraining P to be constant): the change in the density is

$$\delta_C n(\boldsymbol{r}) = 2 \times \sum_{\alpha\gamma} P_{\alpha\gamma}[\delta_C\phi_\alpha(\boldsymbol{r})\phi_\gamma(\boldsymbol{r}) + \phi_\alpha(\boldsymbol{r})\delta_C\phi_\gamma(\boldsymbol{r})]$$

so

$$\delta_C \mathcal{N}[P] = 2 \times \mathrm{Tr}[P\delta_C S] \tag{28}$$

using the change in the overlap matrix

$$(\delta_C S)_{\alpha\beta} = \langle \delta_C\phi_\alpha|\phi_\beta\rangle + \langle \phi_\alpha|\delta_C\phi_\beta\rangle \tag{29}$$

A.2. The Variation in $\mathcal{E}[P]$

Starting from

$$\mathcal{E}[P] = 2 \times \mathrm{Tr}[(T_s + V_{PP}^{nl} + V_{PP}^{loc})P] + \mathcal{E}_{Hxc}[n[P]]$$

we have two types of variations: δ_0 which changes the DM but not the atoms and δ_C which changes the position of atom C (and thus affects the basis functions associated with that atom) but not P.

1. $P \to P + \delta_0 P$ (freezing the nuclei). We have that

$$\delta_0 \mathcal{E}_{Hxc}[n[P]] = \int v_{Hxc}(n[P](\boldsymbol{r}))\delta_0 n(\boldsymbol{r})\, d\boldsymbol{r}$$

so using eq 26

$$\delta_0 \mathcal{E}[P] = 2 \times \mathrm{Tr}[H\delta_0 P] \tag{30}$$

2. Nucleus C moves by $\delta_C X$ (and $\phi_\alpha \to \phi_\alpha + \delta_C\phi_\alpha$) (constraining P to be constant): we find

$$\delta_C \mathcal{E}[P] = 2 \times \mathrm{Tr}[P\delta_C H] \tag{31}$$

where

$$(\delta_C H)_{\alpha\beta} = \langle \phi_\alpha|\delta_C(\hat{v}_{pp}^{nl} + \hat{v}_{pp}^{loc})|\phi_\beta\rangle + \langle \delta_C\phi_\alpha|\hat{h}_{KS}|\phi_\beta\rangle + \langle \phi_\alpha|\hat{h}_{KS}|\delta_C\phi_\beta\rangle \tag{32}$$

A.3. The Variation in $\mathcal{S}_{\mathrm{ent}}[P]$

Starting from $\mathcal{S}_{ent}[P] = -2 \times k_B \mathrm{Tr}[SP \ln(SP) + (1 - SP) \ln(1 - SP)]$,

1. $P \to P + \delta_0 P$ (freezing the nuclei). We have by derivation that

$$\delta_0 \mathcal{S}_{ent}[n[P]] = -2 \times k_B \mathrm{Tr}\left[\ln\left(\frac{SP}{1 - SP}\right) S\delta_0 P\right] \tag{33}$$

2. Nucleus C moves by $\delta_C X$ (and $\phi_\alpha \to \phi_\alpha + \delta_C\phi_\alpha$) (constraining P to be constant): we find

$$\delta_C \mathcal{S}_{ent}[n[P]] = -2 \times k_B \mathrm{Tr}\left[\ln\left(\frac{SP}{1 - SP}\right) P\delta_C S\right] \tag{34}$$

A.4. The Variation in $\Omega[P]$

Here, we combine the above results, while using the relationship

$$\left(\mu - \beta^{-1} \ln \frac{SP}{1 - SP}\right) = HS^{-1}$$

which we find by substituting in eq 3 for P.

1. $P \to P + \delta_0 P$ (freezing the nuclei). Using eqs 27, 30, and 33, we have

$$\delta_0 \Omega = 2 \times \mathrm{Tr}\left[\left\{H - \left(\mu - \beta^{-1} \ln \frac{SP}{1 - SP}\right) S\right\}\delta_0 P\right]$$

leading to

$$\delta_0 \Omega = 0$$

This reflects the fact that P of eq 3 minimizes Ω $[P]$.

2. Nucleus C moves by $\delta_C X$ (and $\phi_\alpha \to \phi_\alpha + \delta_C\phi_\alpha$) (since a variation in P does not affect the value of Ω, we can take it as a constant): using eqs 28, 31, and 34, we find

$$\delta_C \Omega = 2 \times \mathrm{Tr}[P(\delta_C H - HS^{-1}\delta_C S)] \tag{35}$$

The change in free energy is composed of two terms: a term due to the energy, $\delta_C H$ (which includes a direct change and a Pulay term, see eq 32), and a change due to entropy, which depends purely on Pulay changes in the overlap matrix, $\delta_C S$ (see (29)).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00794.

Evaluation of force matrix elements (Section S1), basic concepts in statistics (Section S2), embedded fragment method (Section S3), system size dependency of statistical errors (Section S4), and efficiency of representations (Section S5) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Eran Rabani − *Department of Chemistry, University of California, Berkeley, California 94720, United States; Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; The Raymond and Beverly Sackler Center of Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel;* orcid.org/0000-0003-2031-3525; Email: eran.rabani@berkeley.edu

Roi Baer − *Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel;* orcid.org/0000-0001-8432-1925; Email: roi.baer@huji.ac.il

Authors

Ben Shpiro − *Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel*

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466

Marcel David Fabian − *Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00794

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

R.B. and E.R. gratefully thank the Binational Science Foundation grant No. 2018368. E.R. acknowledges support from the Center for Computational Study of Excited State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05CH11231 as part of the Computational Materials Sciences Program.

■ REFERENCES

(1) Marx, D.; Hutter, J. *Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: 2009; Google-Books-ID: VRZUw8Wk4CIC.

(2) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*; Cambridge University Press: 2004; Google-Books-ID: iqDJ2hjqBMEC, DOI: 10.1017/CBO9780511816581.

(3) Graziani, F.; Desjarlais, M. P.; Redmer, R.; Trickey, S. B. *Frontiers and Challenges in Warm Dense Matter*; Springer Science & Business: 2014; Vol. 96, DOI: 10.1007/978-3-319-04912-0.

(4) Huggins, D. J.; Biggin, P. C.; Dämgen, M. A.; Essex, J. W.; Harris, S. A.; Henchman, R. H.; Khalid, S.; Kuzmanic, A.; Laughton, C. A.; Michel, J.; Mulholland, A. J.; Rosta, E.; Sansom, M. S. P.; van der Kamp, M. W. Biomolecular simulations: From dynamics and mechanisms to computational assays of biological activity. *WIREs Comput. Mol. Sci.* **2019**, *9*, e1393.

(5) Yang, W. T. Direct Calculation of Electron-Density in Density-Functional Theory. *Phys. Rev. Lett.* **1991**, *66*, 1438−1441.

(6) Li, X.; Nunes, W.; Vanderbilt, D. Density-Matrix Electronic-Structure Method with Linear System- Size Scaling. *Phys. Rev. B* **1993**, *47*, 10891−10894.

(7) Ordejon, P.; Drabold, D. A.; Grumbach, M. P.; Martin, R. M. Unconstrained Minimization Approach for Electronic Computations That Scales Linearly with System Size. *Physical Review B-Condensed Matter* **1993**, *48*, 14646−14649.

(8) Goedecker, S.; Colombo, L. Efficient Linear Scaling Algorithm for Tight-Binding Molecular Dynamics. *Phys. Rev. Lett.* **1994**, *73*, 122−125.

(9) Nunes, R. W.; Vanderbilt, D. Generalization of the Density-Matrix Method to a Nonorthogonal Basis. *Physical Review B-Condensed Matter* **1994**, *50*, 17611−17614.

(10) Wang, Y.; Stocks, G. M.; Shelton, W. A.; Nicholson, D. M. C.; Szotek, Z.; Temmerman, W. M. Order-N Multiple Scattering Approach to Electronic Structure Calculations. *Phys. Rev. Lett.* **1995**, *75*, 2867−2870.

(11) Hernandez, E.; Gillan, M. J. Self-Consistent First-Principles Technique with Linear Scaling. *Physical Review B-Condensed Matter* **1995**, *51*, 10157−10160.

(12) Goedecker, S. Low Complexity Algorithms for Electronic Structure Calculations. *J. Comput. Phys.* **1995**, *118*, 261−268.

(13) Ordejon, P.; Artacho, E.; Soler, J. M. Self-consistent order-N density-functional calculations for very large systems. *Physical Review B-Condensed Matter* **1996**, *53*, R10441−R10444.

(14) Bowler, D. R.; Aoki, M.; Goringe, C. M.; Horsfield, A. P.; Pettifor, D. G. A comparison of linear scaling tight-binding methods. *Modell. Simul. Mater. Sci. Eng.* **1997**, *5*, 199−222.

(15) Baer, R.; Head-Gordon, M. Chebyshev expansion methods for electronic structure calculations on large molecular systems. *J. Chem. Phys.* **1997**, *107*, 10003−10013.

(16) Palser, A. H. R.; Manolopoulos, D. E. Canonical purification of the density matrix in electronic- structure theory. *Physical Review B-Condensed Matter* **1998**, *58*, 12704−12711.

(17) Goedecker, S. Linear scaling electronic structure methods. *Rev. Mod. Phys.* **1999**, *71*, 1085−1123.

(18) Scuseria, G. E. Linear scaling density functional calculations with Gaussian orbitals. *J. Phys. Chem. A* **1999**, *103*, 4782−4790.

(19) Galli, G. Large-scale electronic structure calculations using linear scaling methods. *Physica Status Solidi B-Basic Research* **2000**, *217*, 231−249.

(20) Adhikari, S.; Baer, R. Augmented Lagrangian method for order-N electronic structure. *J. Chem. Phys.* **2001**, *115*, 11−14.

(21) Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D. The SIESTA method for ab initio order-N materials simulation. *J. Phys. C* **2002**, *14*, 2745−2779.

(22) Skylaris, C. K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Using ONETEP for accurate and efficient O(N) density functional calculations. *J. Phys. C* **2005**, *17*, 5757−5769.

(23) Gillan, M. J.; Bowler, D. R.; Torralba, A. S.; Miyazaki, T. Order-N first-principles calculations with the CONQUEST code. *Comput. Phys. Commun.* **2007**, *177*, 14−18.

(24) Ochsenfeld, C.; Kussmann, J.; Lambrecht, D. S. *Reviews in Computational Chemistry*; Wiley-Blackwell: 2007; pp 1−82, DOI: 10.1002/9780470116449.ch1.

(25) Havu, V.; Blum, V.; Havu, P.; Scheffler, M. Efficient integration for all-electron electronic structure calculation using numeric basis functions. *J. Comput. Phys.* **2009**, *228*, 8367−8379.

(26) Lin, L.; Lu, J.; Ying, L.; Weinan, E. Pole-based approximation of the Fermi−Dirac function. *Chinese Annals of Mathematics, Series B* **2009**, *30*, 729.

(27) Ozaki, T. Efficient low-order scaling method for large-scale electronic structure calculations with localized basis functions. *Phys. Rev. B* **2010**, *82*, 075131.

(28) Bowler, D.; Miyazaki, T. Methods in electronic structure calculations. *Rep. Prog. Phys.* **2012**, *75*, 036503.

(29) Moussa, J. E. Minimax rational approximation of the Fermi−Dirac distribution. *J. Chem. Phys.* **2016**, *145*, 164108.

(30) Ratcliff, L. E.; Mohr, S.; Huhs, G.; Deutsch, T.; Masella, M.; Genovese, L. Challenges in large scale quantum mechanical calculations: Challenges in large scale quantum mechanical calculations. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1290.

(31) 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.

(32) 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, Á. 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.

(33) 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.

(34) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A.; Ehlert, S.; Elstner, M.; van der Heide, T.; Hermann, J.; Irle, S.; Kranz, J. J.; Köhler, C.; Kowalczyk, T.; Kubar, T.; Lee, I. S.; Lutsker, V.; Maurer, R. J.; Min, S. K.; Mitchell, I.; Negre, C.; Niehaus, T. A.; Niklasson, A. M. N.; Page, A. J.; Pecchia, A.; Penazzi, G.; Persson, M. P.; Řezáč, J.; Sánchez, C. G.; Sternberg, M.; Stöhr, M.; Stuckenberg, F.; Tkatchenko, A.; Yu, V. W.-z.; Frauenheim, T. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101.
(35) Aradi, B.; Hourahine, B.; Frauenheim, T. DFTB+, a sparse matrix-based implementation of the DFTB method. *J. Phys. Chem. A* **2007**, *111*, 5678−5684.
(36) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment. *J. Chem. Phys.* **2001**, *114*, 5149−5155.
(37) Witt, W. C.; Beatriz, G.; Dieterich, J. M.; Carter, E. A. Orbital-free density functional theory for materials research. *J. Mater. Res.* **2018**, *33*, 777−795.
(38) Karasiev, V. V.; Sjostrom, T.; Trickey, S. B. Finite-temperature orbital-free DFT molecular dynamics: Coupling Profess and Quantum Espresso. *Comput. Phys. Commun.* **2014**, *185*, 3240−3249.
(39) Kohn, W. Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms. *Phys. Rev. Lett.* **1996**, *76*, 3168−3171.
(40) Baer, R.; Neuhauser, D.; Rabani, E. Self-Averaging Stochastic Kohn-Sham Density-Functional Theory. *Phys. Rev. Lett.* **2013**, *111*, 106402.
(41) Neuhauser, D.; Baer, R.; Rabani, E. Communication: Embedded fragment stochastic density functional theory. *J. Chem. Phys.* **2014**, *141*, 041102.
(42) 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.
(43) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Overlapped embedded fragment stochastic density functional theory for covalently-bonded materials. *J. Chem. Phys.* **2019**, *150*, 034106.
(44) Fabian, M. D.; Shpiro, B.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic density functional theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, *9*, e1412.
(45) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Energy window stochastic density functional theory. *J. Chem. Phys.* **2019**, *151*, 114116.
(46) 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.
(47) Cytter, Y.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic Density Functional Theory at Finite Temperatures. *Phys. Rev. B* **2018**, *97*, 115207.
(48) Baer, R.; Rabani, E. Communication: Biexciton generation rates in CdSe nanorods are length independent. *J. Chem. Phys.* **2013**, *138*, 051102.
(49) Cytter, Y.; Rabani, E.; Neuhauser, D.; Preising, M.; Redmer, R.; Baer, R. Transition to metallization in warm dense helium-hydrogen mixtures using stochastic density functional theory within the Kubo-Greenwood formalism. *Phys. Rev. B* **2019**, *100*, 195101.
(50) Neuhauser, D.; Rabani, E.; Cytter, Y.; Baer, R. Stochastic Optimally Tuned Range-Separated Hybrid Density Functional Theory. *J. Phys. Chem. A* **2016**, *120*, 3071−3078.
(51) Lee, A. J.; Chen, M.; Li, W.; Neuhauser, D.; Baer, R.; Rabani, E. Dopant levels in large nanocrystals using stochastic optimally tuned range-separated hybrid density functional theory. *Phys. Rev. B* **2020**, *102*, 035112.
(52) Arnon, E.; Rabani, E.; Neuhauser, D.; Baer, R. Efficient Langevin dynamics for "noisy" forces. *J. Chem. Phys.* **2020**, *152*, 161103.
(53) Neuhauser, D.; Gao, Y.; Arntsen, C.; Karshenas, C.; Rabani, E.; Baer, R. Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic GW Approach. *Phys. Rev. Lett.* **2014**, *113*, 076402.
(54) Vlček, V.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic GW Calculations for Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 4997−5003.
(55) Rabani, E.; Baer, R.; Neuhauser, D. Time-dependent stochastic Bethe-Salpeter approach. *Phys. Rev. B* **2015**, *91*, 235302.
(56) Gao, Y.; Neuhauser, D.; Baer, R.; Rabani, E. Sublinear scaling for time-dependent stochastic density functional theory. *J. Chem. Phys.* **2015**, *142*, 034106.
(57) Vlček, V.; Baer, R.; Neuhauser, D. Stochastic time-dependent DFT with optimally tuned range-separated hybrids: Application to excitonic effects in large phosphorene sheets. *J. Chem. Phys.* **2019**, *150*, 184118.
(58) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993−2006.
(59) Kleinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425−1428.
(60) 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.
(61) Niklasson, A. M. A note on the Pulay force at finite electronic temperatures. *J. Chem. Phys.* **2008**, *129*, 244107.
(62) Pulay, P. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. *Mol. Phys.* **1969**, *17*, 197−204.
(63) Tal-Ezer, H.; Kosloff, R. An Accurate and Efficient Scheme for Propagating the Time-Dependent Schrodinger-Equation. *J. Chem. Phys.* **1984**, *81*, 3967−3971.
(64) 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.
(65) Fabian, M.; Shpiro, B.; Baer, R. Linear scalability of density functional theory calculations without imposing electron localization. 2021. https://arxiv.org/abs/2108.13478 (accessed 2022-01-21).
(66) Hutchinson, M. F. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Commun. Stat Simul Comput.* **1990**, *19*, 433−450.
(67) Thompson, M. A. *ArgusLab 4.0.1*; Planaria Software LLC: Seattle, WA, 2004.
(68) Thompson, M. Molecular docking using ArgusLab, an efficient shape-based search algorithm and the AScore scoring function. *ACS Meeting, Philadelphia, PA*, 2004; CINF 42.
(69) Gonze, X. *Advances in Quantum Chemistry*; Elsevier: 1998; Vol. *33*, pp 225−239, DOI: 10.1016/S0065-3276(08)60438-4.
(70) Zou, W.; Tao, Y.; Freindorf, M.; Cremer, D.; Kraka, E. Local vibrational force constants − From the assessment of empirical force constants to the description of bonding in large systems. *Chem. Phys. Lett.* **2020**, *748*, 137337.

https://doi.org/10.1021/acs.jctc.1c00794
J. Chem. Theory Comput. 2022, 18, 1458−1466