

## Equilibrium configurations of large nanostructures using the embedded saturated-fragments stochastic density functional theory

Eitam Arnon, Eran Rabani, Daniel Neuhauser, and Roi Baer

Citation: *The Journal of Chemical Physics* **146**, 224111 (2017); doi: 10.1063/1.4984931

View online: <http://dx.doi.org/10.1063/1.4984931>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/22>

Published by the [American Institute of Physics](#)

---

### Articles you may be interested in

[Strong correlation in incremental full configuration interaction](#)

*The Journal of Chemical Physics* **146**, 224104 (2017); 10.1063/1.4985566

[A Jeziorski-Monkhorst fully uncontracted multi-reference perturbative treatment. I. Principles, second-order versions, and tests on ground state potential energy curves](#)

*The Journal of Chemical Physics* **146**, 224108 (2017); 10.1063/1.4984616

[Low scaling random-phase approximation electron correlation method including exchange interactions using localised orbitals](#)

*The Journal of Chemical Physics* **146**, 174110 (2017); 10.1063/1.4981817

[Dealing with the exponential wall in electronic structure calculations](#)

*The Journal of Chemical Physics* **146**, 194107 (2017); 10.1063/1.4983207

[Laplace-transformed multi-reference second-order perturbation theories in the atomic and active molecular orbital basis](#)

*The Journal of Chemical Physics* **146**, 224101 (2017); 10.1063/1.4984591

---



**COMPLETELY  
REDESIGNED!**



**PHYSICS  
TODAY**

*Physics Today* Buyer's Guide  
Search with a purpose.

# Equilibrium configurations of large nanostructures using the embedded saturated-fragments stochastic density functional theory

Eitam Arnon,<sup>1</sup> Eran Rabani,<sup>2,3,a)</sup> Daniel Neuhauser,<sup>4,b)</sup> and Roi Baer<sup>1,c)</sup>

<sup>1</sup>*Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel*

<sup>2</sup>*Department of Chemistry, University of California and Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

<sup>3</sup>*The Raymond and Beverly Sackler Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel*

<sup>4</sup>*Department of Chemistry, University of California, Los Angeles, California 90095, USA*

(Received 9 February 2017; accepted 22 May 2017; published online 14 June 2017)

An *ab initio* Langevin dynamics approach is developed based on stochastic density functional theory (sDFT) within a new *embedded saturated fragment* formalism, applicable to covalently bonded systems. The forces on the nuclei generated by sDFT contain a random component natural to Langevin dynamics, and its standard deviation is used to estimate the friction term on each atom by satisfying the fluctuation–dissipation relation. The overall approach scales linearly with the system size even if the density matrix is not local and is thus applicable to ordered as well as disordered extended systems. We implement the approach for a series of silicon nanocrystals (NCs) of varying size with a diameter of up to 3 nm corresponding to  $N_e = 3000$  electrons and generate a set of configurations that are distributed canonically at a fixed temperature, ranging from cryogenic to room temperature. We also analyze the structure properties of the NCs and discuss the reconstruction of the surface geometry. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4984931>]

## I. INTRODUCTION

*Ab initio* molecular dynamics based on density functional theory (DFT) is becoming an important tool for studying the plethora of structural and dynamical processes in a broad range of systems in material science, chemistry, biology, and physics.<sup>1–11</sup> The application of this approach to very large systems is still limited by the computational scaling of the electronic structure portion of the calculation, regardless of whether one uses a Lagrangian-based or Born-Oppenheimer-based method. This is because of the cubic scaling involved in solving the Kohn-Sham equations coupled with the need to iterate to self-consistency or to propagate the Kohn-Sham (KS) orbitals, as both of these options further increase the computational times by an order of magnitude.

Significant advances in these respects have been made along two major directions. One primary direction is based on a Lagrangian formulation of density functional theory<sup>1,11</sup> and circumvents the need for SCF iterations by the propagation of the KS orbitals. This venue does not eliminate the cubic scaling and is therefore limited to relatively small systems. Another approach is based on linear-scaling techniques,<sup>12–15</sup> which reduces the algorithmic complexity by finding the density matrix directly, relying on its asymptotic sparseness in real-space. However, sparsity sets in only for very large systems,

limiting the applicability sparse-matrix methods, especially in 3D.

In a recent set of papers, we have introduced the stochastic DFT (sDFT) methods<sup>16–19</sup> which scale linearly (or even sub-linearly) with the system size and do not rely on the sparsity of the density matrix. sDFT is a general approach to the electronic structure based on a stochastic process and is applicable to extended ordered as well as disordered materials. Some of the techniques we use, based on the stochastic trace formula,<sup>20</sup> have been developed for the tight-binding electronic structure,<sup>21–23</sup> for molecular electronics<sup>24</sup> and for multi-exciton generation in nanocrystals.<sup>25</sup> The success of sDFT in reducing the scaling comes at a price of introducing a stochastic error in all its predictions, including forces, and that precludes application to *ab initio* molecular dynamics.

In this paper, we show that sDFT can be used to study equilibrium structural properties of large nanocrystals (NCs), despite the statistical fluctuations in the force estimates. For this, we invoke the Langevin equation following the work of Attaccalite and Sorella,<sup>26</sup> and generate a sequence of configurations distributed according to the canonical ensemble. These configurations can be used in a variety of applications for studying the structural, electronic and optical properties of NCs. Here we demonstrate their use for studying the structural properties of silicon nanocrystals (NCs) with a diameter of up to 3 nm, and  $N_e = 3000$  electrons.

The paper includes development of the embedded saturated fragments method which allows reducing the statistical errors in sDFT. This new method is inspired by, but more general than, the embedded fragments method developed in

<sup>a)</sup>Electronic mail: eran.rabani@berkeley.edu

<sup>b)</sup>Electronic mail: dxn@chem.ucla.edu

<sup>c)</sup>Electronic mail: roi.baer@huji.ac.il

Ref. 17. It uses small saturated fragments of the system and carves out the relevant part of the density to be embedded in the system. Hence, it is applicable not only to clusters of molecules, like Ref. 17, but also to covalently bonded systems, such as silicon NCs. The method is described in detail in Appendix A.

## II. METHODS

### A. Stochastic DFT

Kohn-Sham density functional theory<sup>27,28</sup> maps a system of  $N_e$  interacting electrons in an external electron-nucleus potential  $v_{eN}(\mathbf{r}) = -\frac{e}{4\pi\epsilon_0} \sum_N \frac{Z_N e}{|\mathbf{r} - \mathbf{R}_N|}$ , where  $\mathbf{R}_N$  ( $N = 1, 2, \dots$ ) are the nuclei positions and  $Z_N e$  are their charge ( $e$  is the electron charge), onto a system of *non-interacting* electrons (the KS system), having the same ground-state density  $n(\mathbf{r})$ . This mapping is performed by solving the KS equations<sup>27,28</sup>

$$\hat{h}_{KS} \phi_n(\mathbf{r}) = \varepsilon_n \phi_n(\mathbf{r}), \quad (1)$$

where the KS Hamiltonian is

$$\hat{h}_{KS} = -\frac{\hbar^2}{2m_e} \nabla^2 + v_{KS}(\mathbf{r}), \quad (2)$$

and the KS potential  $v_{KS}(\mathbf{r})$  is the sum of the external electron-nuclear potential  $v_{eN}(\mathbf{r})$ , the density-dependent Hartree potential  $v_H(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$ , and the exchange-correlation potential  $v_{xc}(\mathbf{r})$ ,

$$v_{KS}(\mathbf{r}) = v_{eN}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \quad (3)$$

In the KS system, the density is expressed in terms of the normalized single electron KS eigenstates  $\phi_n(\mathbf{r})$  and eigenvalues  $\varepsilon_n$ ,

$$n(\mathbf{r}) = 2 \sum_n \theta(\mu - \varepsilon_n) |\phi_n(\mathbf{r})|^2, \quad (4)$$

where  $\theta(x)$  is the Heaviside function and  $\mu$  is the chemical potential chosen so that  $2 \sum_n \theta(\mu - \varepsilon_n) = N_e$ . Equations (1)–(4) must be solved self-consistently since  $\hat{h}_{KS}$  depends on the density. While the entire scheme is a significant simplification over the original many-electron problem, it remains a challenge for large systems since the computational effort scales as  $O(N_e^3)$ .

An important step towards reducing the computational scaling of KS-DFT was recently proposed by Baer, Neuhauser, and Rabani (BNR),<sup>16</sup> where the density of Eq. (4) was expressed as a trace over the projected density operator,<sup>16</sup>

$$n(\mathbf{r}) = 2\text{Tr} \left[ \theta(\mu - \hat{h}_{KS}) \delta(\mathbf{r} - \hat{\mathbf{r}}) \right]. \quad (5)$$

The problem now shifts into calculating self-consistently the trace in Eq. (5) [since  $\hat{h}_{KS}$  depends on  $n(\mathbf{r})$ ] rather than solving the KS equations by brute-force diagonalization. When the trace is performed using the KS eigenstates, the computational cost remains  $O(N_e^3)$  similar to the traditional approach. However, since the trace is invariant to the basis, alternative schemes that potentially lead to improved scaling can be used. One such scheme is based on the concept of a stochastic trace formula, which reduces the scaling of the trace operation by introducing a controlled statistical error.<sup>20</sup>

Using the stochastic trace formula, the density can be estimated as a symmetrized stochastic trace formula, given by<sup>16</sup>

$$n^I(\mathbf{r}) = \left\langle \langle \chi | \sqrt{\theta_\beta(\mu - \hat{h}_{KS})} 2\delta(\mathbf{r} - \hat{\mathbf{r}}) \times \sqrt{\theta_\beta(\mu - \hat{h}_{KS})} | \chi \rangle \right\rangle_\chi, \quad (6)$$

where  $\langle \dots \rangle_\chi$  denotes an average over  $I$  stochastic orbitals  $|\chi\rangle$ , defined as

$$\langle \mathbf{r} | \chi \rangle = h^{-3/2} e^{i\varphi_r} \quad (7)$$

for each grid point  $\mathbf{r}$ , the parameter  $h$  (not to be confused with the KS Hamiltonian  $\hat{h}_{KS}$  operator) is the grid spacing, and  $\varphi_r$  are statistically independent random variables in the range  $[0, 2\pi]$  ( $\langle e^{i\varphi_r} e^{-i\varphi_{r'}} \rangle_\varphi = \delta_{rr'}$ ). The density  $n(\mathbf{r})$  is, strictly speaking, given by the limit  $n(\mathbf{r}) = \lim_{I \rightarrow \infty} n^I(\mathbf{r})$  and we approximate it with a finite  $I$ . The Heaviside function in Eq. (6) is smoothed by the function  $\theta_\beta(\varepsilon) \equiv \frac{1}{2} \text{erfc}[\beta\varepsilon]$ , where  $\beta$  is a large constant satisfying  $\beta E_g \gg 1$ , where  $E_g$  is the KS-DFT fundamental gap. Throughout this paper we set the value of  $\beta$  to  $100E_h^{-1}$ . The action of  $\sqrt{\theta_\beta(\mu - \hat{h}_{KS})}$  on  $|\chi\rangle$  is evaluated by a Chebyshev expansion in powers of the sparse KS Hamiltonian,  $\hat{h}_{KS}$ .<sup>29</sup> The length of the Chebyshev series is determined by the value of  $(\mu - E_{min})/\Delta E$  and  $\beta\Delta E$ , where  $\Delta E = (E_{max} - E_{min})/2$  and  $E_{min}/max$  are the minimal and maximal eigenvalues of  $\hat{h}_{KS}$ . Under the conditions of the present systems, the length of the series is  $\sim 3000$  terms.

The stochastic trace evaluation [Eq. (6)] reduces the computational scaling of KS-DFT to  $O(N_e)$  and for certain properties even to a sub-linear scaling.<sup>16</sup> Linear-scaling complexity is achieved due to the following facts: (1) the application of a Hamiltonian to a stochastic orbital,  $\hat{h}_{KS} |\chi\rangle$ , requires a linear scaling effort (irrespective of the structure of the orbital); (2) the length of the Chebyshev series is independent (or at most weakly dependent) of the system size; and (3) only a system-size independent number of stochastic orbitals are required. This type of assumptions is different from the linear-scaling approaches depending on density matrix sparsity,<sup>12,30</sup> which assume that locality of orbitals is not completely destroyed by the repeated operation of the Hamiltonian.

A converged self-consistent solution of Eq. (6) provides an estimate of the electron density and in addition can be used to generate other quantities, such as the density of states (DOS), the total energy per electron, and the forces acting on the nuclei. All estimates contain a statistical error that can be controlled by increasing the number of stochastic orbitals ( $I$ ) used to evaluate the trace in Eq. (6). Of particular relevance to this work are the Cartesian forces exerted by the electrons on  $N$  nuclei ( $\alpha = 1, \dots, N$ ), which can be evaluated through the Hellmann-Feynman theorem,<sup>31,32</sup>

$$f_\alpha = - \int \frac{\partial v_{eN}(\mathbf{r})}{\partial \mathbf{R}_\alpha} n^I(\mathbf{r}) d^3r. \quad (8)$$

It should be stressed that for finite sampling, these forces are only approximately commensurate with the stochastic estimate of the energy (which is not used in the sampling procedure at all), as discussed in Appendix B. The stochastic estimate of

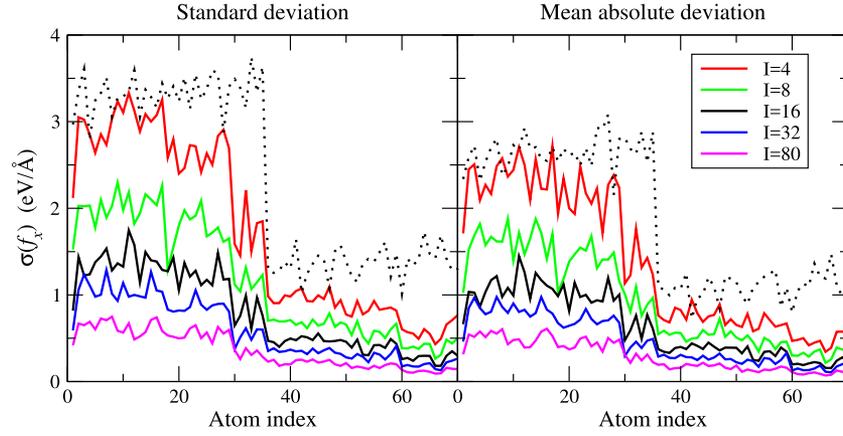


FIG. 1. The x-component of the atomic force statistics for the 71 atoms of  $\text{Si}_{35}\text{H}_{36}$  calculated by sDFT (black dashed line for  $I = 16$ ) and efsDFT (solid lines, using passivated fragments of size smaller or equal to  $\text{Si}_5$ , depending on the way surface atoms are treated). For each atom  $\alpha = 1, \dots, N$  and number of stochastic orbitals  $I$ , we present the standard deviation (STD)  $\sigma(f_x) = \sqrt{\langle (f_\alpha^x - \langle f_\alpha^x \rangle)^2 \rangle_I}$  (left) and the mean-absolute-deviation (MAD) from the deterministic DFT (dDFT) value,  $\langle |f_\alpha^x - (f_\alpha^x)^{\text{det}}| \rangle_I$ . These statistics were calculated using 60 independent efsDFT/sDFT runs. Atoms are ordered by their distance from the origin, the first 35 atoms are Si atoms followed by 36 H atoms.

the Hellmann-Feynman forces is an excellent estimator of the deterministic forces, as can be seen in Fig. 1, where the mean absolute deviation is dominated by the fluctuations and not by additional bias terms.

These sDFT forces can be expressed as

$$\mathbf{f}_\alpha = \mathbf{f}_\alpha^{\text{det}} + \mathbf{f}_\alpha^{\text{fluc}} + \mathbf{f}_\alpha^{\text{bias}}, \quad (9)$$

where  $\mathbf{f}_\alpha^{\text{det}}$  is the deterministic (generally unknown) force,  $\mathbf{f}_\alpha^{\text{fluc}}$  is the pure fluctuating term, and  $\mathbf{f}_\alpha^{\text{bias}}$  is the bias expected to be proportional to  $\frac{1}{I}$  in leading order. The choice of  $I$  should be large enough to reduce  $\mathbf{f}_\alpha^{\text{bias}}$  to negligible values, and the only source of error in the procedure is then the statistical fluctuations proportional to  $\frac{1}{\sqrt{I}}$  with vanishing mean ( $\langle \mathbf{f}_\alpha^{\text{fluc}} \rangle = 0$ ).

## B. Embedded saturated fragments sDFT

The reduction of the scaling in sDFT is achieved by replacing the deterministic, numerically exact, trace evaluation with a stochastic sampling of the density. In return, this leads to statistical errors in the computed observables. To reduce the size of the statistical fluctuations, an embedded saturated fragments method is introduced inspired by (but different from) the method of Ref. 17. The latter approach was suitable mainly for systems composed of proximate but chemically separated molecules (like clusters of water molecules, for examples). The present method is applicable for fragmenting covalently bonded systems, like silicon NCs.

In this approach, the system is divided into  $F$  small fragments that are possibly overlapping. The division to fragments is flexible, and any desired physically motivated fragmentation can be used. The density is then a sum of the fragment density and a small correction term,

$$n(\mathbf{r}) = n_F(\mathbf{r}) + \Delta n(\mathbf{r}), \quad (10)$$

where  $n_F(\mathbf{r}) = \sum_{f=1}^F n_f(\mathbf{r})$  is the density generated by the individual fragments obtained from a deterministic KS-DFT (dDFT) calculation for each fragment and  $\Delta n(\mathbf{r}) = (n^l(\mathbf{r}) - n_F^l(\mathbf{r}))$  is a correction term evaluated using stochas-

tic orbitals. Here,  $n^l(\mathbf{r})$  is given by Eq. (6) and  $n_F^l(\mathbf{r}) = \sum_{f=1}^F n_f^l(\mathbf{r})$  is a sum over a stochastic estimate of the fragments density. In the limit  $I \rightarrow \infty$ , Eqs. (6) and (10) are identical and equal to the deterministic density. For finite values of  $I$ , the size of the statistical fluctuations of the two approaches are quite different. Since the deterministic fragmented density,  $n_f(\mathbf{r})$ , provides a reasonable approximation for the full density  $n(\mathbf{r})$ , the correction term,  $\Delta n(\mathbf{r})$ , which is evaluated stochastically, is rather small, leading to a reduced variance in the relevant observables (forces, DOS, total energy per electron, etc.) compared to the direct stochastic approach of Eq. (6). An equivalent viewpoint is that the fragmentation is a device for reducing the variance in the stochastic evaluation of the density. This is evident by rewriting Eq. (10) in the following form:

$$n(\mathbf{r}) = n^l(\mathbf{r}) + \sum_{f=1}^F (n_f(\mathbf{r}) - n_f^l(\mathbf{r})), \quad (11)$$

and the implementation of this form is described in Appendix A.

To assess the accuracy of the embedded saturated fragmented sDFT (efsDFT), we calculated the standard deviations (STDs) and mean absolute deviations (MADs) with respect the deterministic DFT of the atomic forces in an  $\text{Si}_{35}\text{H}_{36}$  NC using hydrogen passivated  $\text{Si}_5$  fragments. The results are shown in Fig. 1.<sup>33,43–45</sup> The STDs and MADs decrease as  $1/\sqrt{I}$ , indicating that the bias in the force estimation is negligible. The standard deviations in the sDFT forces are larger by a factor of  $\approx 3$  compared to those of efsDFT. This implies that the required number of stochastic orbitals in efsDFT is nearly an order of magnitude smaller than in sDFT for similar STDs. The STDs can be further reduced by using larger fragments as discussed below (cf. Fig. 7).

## C. Langevin dynamics based on efsDFT

The standard approach to generate canonically distributed configurations using *ab initio* techniques is based on molecular dynamics, which requires as input accurate force estimates

for each atomic degree of freedom. Since the forces generated by efsDFT contain a stochastic component, we use Langevin dynamics (LD) instead of molecular dynamics to sample configurations according to the Boltzmann distribution. A LD trajectory<sup>34–37</sup> is a sequence of configurations  $(\mathbf{p}, \mathbf{q})^m = (\mathbf{p}(t_m), \mathbf{q}(t_m))$  at discrete “times”  $t_m = m\Delta t$ , where  $\Delta t$  is the time step, and  $\mathbf{q} \equiv (\mathbf{q}_1, \dots, \mathbf{q}_N)$  and  $\mathbf{p} \equiv (\mathbf{p}_1, \dots, \mathbf{p}_N)$  are the Cartesian coordinates and conjugate momenta, respectively, for the  $N$  atoms. The trajectory is a solution of the Langevin equation (LE) of motion,<sup>38</sup>

$$\mu_\alpha \ddot{\mathbf{q}}_\alpha = \mathbf{f}_\alpha(\mathbf{q}) - \gamma_\alpha \mathbf{p}_\alpha + \boldsymbol{\eta}_\alpha, \quad (12)$$

where  $\mu_\alpha$  is the mass of the atom  $\alpha$ ,  $\gamma_\alpha$  is its friction constant, and  $\mathbf{f}_\alpha = \mathbf{f}_\alpha^{\text{det}} + \mathbf{f}_\alpha^{\text{fluc}}$  is the total efsDFT force acting on it, including deterministic and fluctuating parts [see Eq. (9)]. The bias is assumed negligible, so that  $\langle \mathbf{f}_\alpha \rangle = \mathbf{f}_\alpha^{\text{det}}$ . In Eq. (12)  $\boldsymbol{\eta}_\alpha$  is an additional uncorrelated white-noise force introduced so as to satisfy the fluctuation-dissipation (FD) relation. We require that the total random fluctuation on each atom obeys

$$\langle \boldsymbol{\eta}_\alpha(t) \rangle = \langle \mathbf{f}_\alpha^{\text{fluc}}(t) \rangle = 0$$

and

$$\begin{aligned} & \langle (\boldsymbol{\eta}_\alpha(t) + \mathbf{f}_\alpha^{\text{fluc}}(t)) \otimes (\boldsymbol{\eta}_{\alpha'}(t') + \mathbf{f}_{\alpha'}^{\text{fluc}}(t')) \rangle \\ &= \langle \boldsymbol{\eta}_\alpha(t) \otimes \boldsymbol{\eta}_{\alpha'}(t') \rangle + \langle \mathbf{f}_\alpha^{\text{fluc}}(t) \otimes \mathbf{f}_{\alpha'}^{\text{fluc}}(t') \rangle \\ &= \mathbf{I}_{3 \times 3} \sigma_\alpha^2 \delta_{\alpha\alpha'} \delta(t - t'), \end{aligned} \quad (13)$$

where  $\alpha, \alpha' = 1, \dots, N$  are atom indices,  $\langle \dots \rangle$  designates average over the atomic force distribution,  $\mathbf{I}_{3 \times 3}$  is the  $3 \times 3$  unit matrix, and  $\sigma_\alpha$  is the atomic force STD of atom  $\alpha$ , which is taken to satisfy the fluctuation-dissipation relation,

$$\sigma_\alpha^2 = 2\mu_\alpha \gamma_\alpha k_B T. \quad (14)$$

We use the Verlet-like algorithm<sup>39</sup> for numerically integrating the LE of motion at a fixed temperature  $T$  and a predefined time step  $\Delta t$ . The positions and momenta in time step  $m+1$  depend on the positions and momenta in time step  $m$  as well as on the forces in time step  $m$  and the additional white noise  $\boldsymbol{\eta}_\alpha^m$  is sampled from a Gaussian distribution such that the discretized version of Eq. (13) holds:  $\langle (\boldsymbol{\eta}_\alpha^m + \mathbf{f}_\alpha^{\text{fluc}}) \otimes (\boldsymbol{\eta}_{\alpha'}^n + \mathbf{f}_{\alpha'}^{\text{fluc}}) \rangle \Delta t = \mathbf{I} \sigma_\alpha^2 \delta_{\alpha\alpha'} \delta_{mn}$ ,

$$\begin{aligned} \mathbf{q}_\alpha^{m+1} &= \mathbf{q}_\alpha^m + b_\alpha \Delta t \mu_\alpha^{-1} \mathbf{p}_\alpha^m + \frac{1}{2} b_\alpha \Delta t^2 \mu_\alpha^{-1} (\mathbf{f}_\alpha^m + \boldsymbol{\eta}_\alpha^{m+1}), \\ \mathbf{p}_\alpha^{m+1} &= a_\alpha \mathbf{p}_\alpha^m + \frac{1}{2} \Delta t (a_\alpha \mathbf{f}_\alpha^m + \mathbf{f}_\alpha^{m+1} + 2b_\alpha \boldsymbol{\eta}_\alpha^{m+1}), \end{aligned} \quad (15)$$

where  $a_\alpha = b_\alpha (1 - \frac{1}{2} \gamma_\alpha \Delta t)$  and  $b_\alpha^{-1} = 1 + \frac{1}{2} \gamma_\alpha \Delta t$ . The algorithm allows for stable and accurate solutions of the LE with time step comparable to that used in molecular dynamics simulations for similar systems. It treats the additional white noise component  $\boldsymbol{\eta}_\alpha$  of the force differently from the force  $\mathbf{f}_\alpha = \mathbf{f}_\alpha^{\text{det}} + \mathbf{f}_\alpha^{\text{fluc}}$  that result from the efsDFT calculation containing deterministic and fluctuating components that cannot be separated.

In Fig. 2, we plot for  $\text{Si}_{35}\text{H}_{36}$  the running average of the transient temperature,  $T^m$ , calculated from the kinetic energy

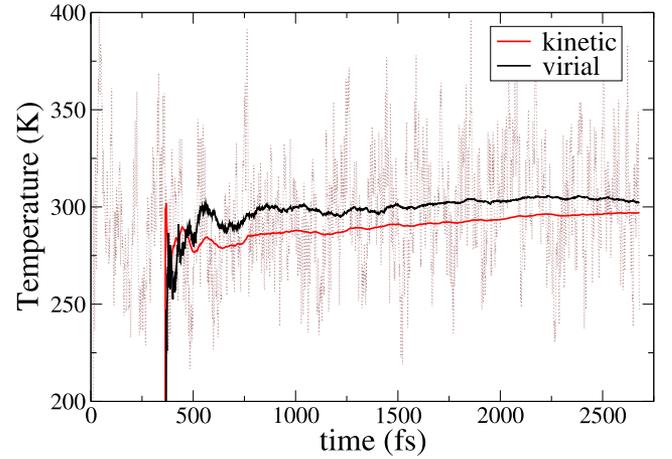


FIG. 2. The Verlet-like<sup>39</sup> temperature in  $\text{Si}_{35}\text{H}_{36}$  evaluated with efsDFT using transient kinetic energy (dotted line), running average kinetic energy for Si (red), and the running average virial (black) using  $I = 30$  stochastic orbitals, a time step of  $\Delta t = 1.2$  fs, and the friction coefficients  $\gamma_{\text{Si}} = \gamma_{\text{H}} = 0.04 \text{ fs}^{-1}$ . In the canonical distribution, the average kinetic energy  $T$  is equal to  $\sqrt{\frac{3N}{2}} \delta T$ , where  $\delta T$  is the fluctuation in the kinetic energy.<sup>40</sup> The standard deviation of the transient shown as a dotted line is  $\delta T = 29$  K and multiplied by  $\sqrt{3N/2}$ , where  $N = 71$  is the number of atoms in the system, gives 299.3 K, which is close to the designated temperature. The average kinetic energy of Si and H is 315 K and 285 K, respectively.

$$T_K^m = \frac{2}{3Nk_B} \sum_{\alpha} (\mathbf{p}_\alpha^m)^2 / 2\mu_\alpha \quad (16)$$

and from the virial estimator,

$$T_V^m = -\frac{1}{3Nk_B} \sum_{\alpha} (\mathbf{f}_\alpha^m + \boldsymbol{\eta}_\alpha^m) \cdot (\mathbf{q}_\alpha^m - \langle \mathbf{q}_\alpha \rangle). \quad (17)$$

In the above,  $\langle \mathbf{q}_\alpha \rangle$  is the time average of the coordinate of atom  $\alpha$ . The initial positions of the Si atoms were taken from the bulk values. All surface Si atoms with more than two dangling bonds were removed, and the remaining surface Si atoms were passivated using one or two H atoms placed in a tetrahedral geometry at the Si–H distance of 1.47 Å. The momenta were sampled from a Boltzmann distribution at  $T = 300$  K. This non-equilibrium initial configuration relaxes towards equilibrium.

The agreement in Fig. 2 between the two temperature estimators is consistent with a proper sampling of the canonical distribution of both positions and velocities. The small discrepancies at the longest averaging time are due to the large fluctuations of the transient temperature, particularly when using the virial estimator. We have also calculated the fluctuations in the kinetic energy and found good agreement with the corresponding analytical value (see caption of the figure). The two atomic species have a slightly  $\pm 5\%$  deviations in the temperatures. These deviations are slightly inconsistent with equipartition theorem and may result from several factors, such as the finite time step of the Langevin propagator,<sup>39</sup> incomplete SCF convergence,<sup>40</sup> and insufficiently accurate estimate of the amount of white noise  $\boldsymbol{\eta}_\alpha^m$  in Eq. (13) required to fulfill the fluctuation-dissipation relation.

#### D. Determining the optimal friction

The effect of  $\gamma_{\text{Si}}$  on the configurational relaxation and on the velocity autocorrelation decay is illustrated in Fig. 3 for

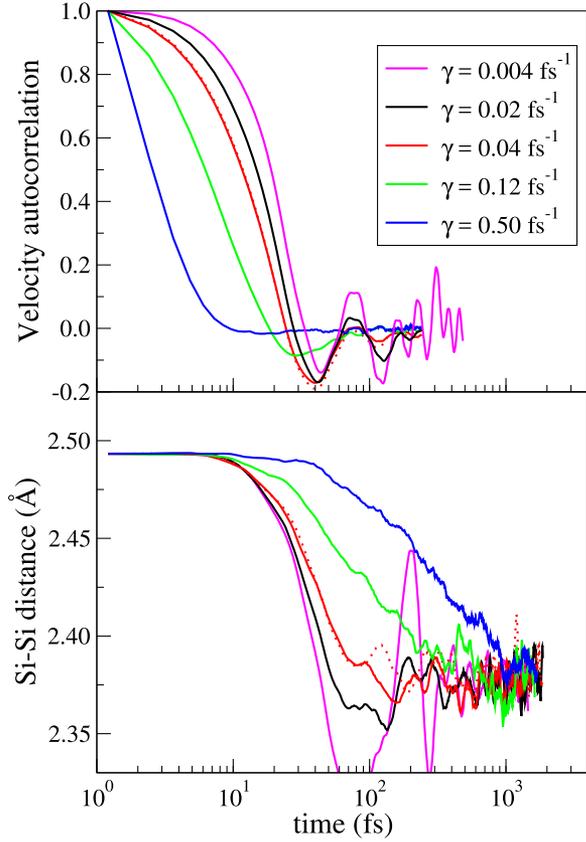


FIG. 3. The normalized velocity autocorrelation function (top panel) and the mean nearest-neighbor Si-Si distance (bottom panel) in  $\text{Si}_{35}\text{H}_{36}$  as a function of time for a LD trajectory at  $T=300$  K with time step  $\Delta t=1.2$  fs calculated using a dDFT based LD for different values of  $\gamma = \gamma_{\text{Si}} = \gamma_{\text{H}}$ . The dashed curve corresponds to efsDFT based LD calculation with  $\gamma = 0.04 \text{ fs}^{-1}$  and  $I=30$  stochastic orbitals. The simulation started intentionally from an inflated configuration in order to measure the relaxation time.

$\text{Si}_{36}\text{H}_{35}$ . In order to decrease the number of unknown parameters, we set the values of  $\gamma_{\alpha}$  and  $\sigma_{\alpha}$  to be identical for all atoms of the same type (i.e., Si or H in the systems studied here). To achieve such a uniform value of  $\sigma$ , we introduced white noise  $\eta_{\alpha}^i = \sqrt{(\sigma_{\alpha}^i)^2 - \langle (f_{\alpha}^{i,\text{fluc}})^2 \rangle}$  for each degree of freedom [see Eq. (13)], where  $\langle (f_{\alpha}^{i,\text{fluc}})^2 \rangle$  is estimated by a separate set of runs on the initial NC configuration using several independent sets of stochastic orbitals. Note that we have tested that

TABLE I. Value of various parameters for the LD based on efsDFT calculations: The friction coefficients  $\gamma$ , number of stochastic orbitals  $I$ , time-step  $\Delta t$ , and the wall time per single SCF iteration  $t_{\text{iter}}$ .

NC	T(K)	$\gamma$ (fs <sup>-1</sup> )		$I$	$\Delta t$ (fs <sup>-1</sup> )	$t_{\text{iter}}$ (min)
		H	Si			
$\text{Si}_{35}\text{H}_{36}$	30	0.12	0.04	120	1.2	1
	300	0.04	0.04	30	1.2	1
$\text{Si}_{147}\text{H}_{100}$	30	0.12	0.04	120	1.2	2
	300	0.12	0.04	30	1.2	2
$\text{Si}_{705}\text{H}_{300}$	30	0.12	0.04	120	1.2	10
	300	0.12	0.04	92	1.2	10

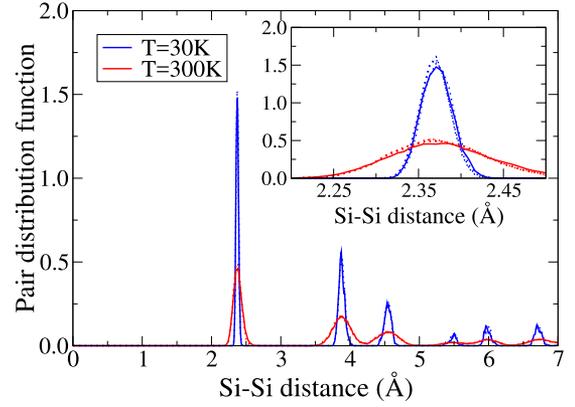


FIG. 4. The Si-Si pair distribution function  $g(r)$  for  $\text{Si}_{35}\text{H}_{36}$  calculated using dDFT (dotted lines) and efsDFT based LD (solid lines, see Table I for parameters) at  $T=30$  K (blue curves) and 300 K (red curves). Inset: Details of the first (nearest neighbor) peak. The dotted lines are for different friction coefficients  $\gamma = \gamma_{\text{Si}} = \gamma_{\text{H}}$  in the range 0.02–0.5 fs<sup>-1</sup>.

the magnitude of the sDFT force fluctuation  $\langle (f^{\text{fluc}})^2 \rangle$  is not sensitive to the particular configuration used.

As expected, the configurational relaxation time increases with increasing values of  $\gamma_{\text{Si}}$  with the opposite trend for the decay time of the velocity autocorrelation function. Based on the results of  $\text{Si}_{36}\text{H}_{35}$  presented in Fig. 3, we conclude that a friction coefficient of  $0.04 \text{ fs}^{-1}$  is sufficiently small for this system, with respect to minimizing both velocity and pair distance autocorrelation times. Although lower values of the friction coefficients could decrease the correlation time further, they would require reducing the statistical noise, which would be expensive to achieve using sDFT. Thus we chose the friction constants,  $\gamma_{\text{Si}} = 0.04 \text{ fs}^{-1}$  for Silicon and  $\gamma_{\text{H}} = 0.12 \text{ fs}^{-1}$  for the lighter H atoms. These values were used for the larger systems described in Sec. III (see Table I). Note that the results shown in Fig. 3, which were generated using LD under dDFT, could have been equally well generated under efsDFT. This is shown explicitly for  $\gamma_{\text{Si}} = 0.04 \text{ fs}^{-1}$  (dotted red line) proving that the

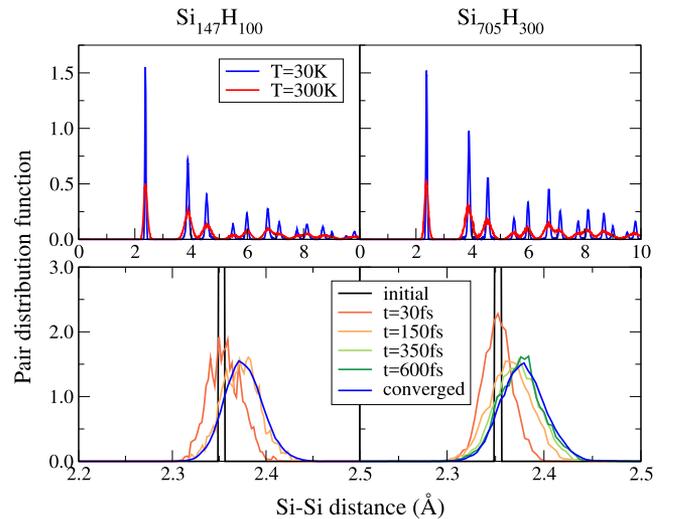


FIG. 5. The Si-Si pair distribution function  $g(r)$  for  $\text{Si}_{147}\text{H}_{100}$  (left) and  $\text{Si}_{705}\text{H}_{300}$  (right) calculated using efsDFT based LD. Upper panels:  $g(r)$  for  $T=30$  K (blue curves) and  $T=300$  K (red curves). Lower panels: The first peak of  $g(r)$  at 30 K shown for several times. The calculation parameters are given in Table I.

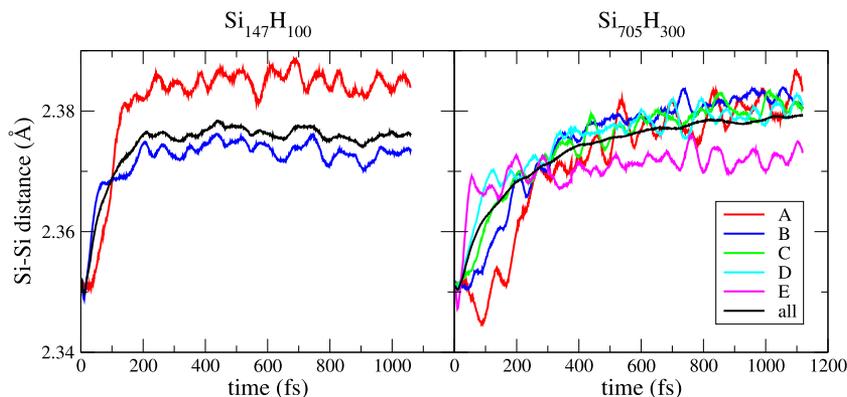


FIG. 6. Si-Si nearest neighbor distance averaged over atoms in shells A-E (see Table II for definition), at 30 K for  $\text{Si}_{147}\text{H}_{100}$  (left) and  $\text{Si}_{705}\text{H}_{300}$  (right) as a function of time. The calculation parameters are given in Table I.

relaxation times are similar to those of the dDFT based LD calculation with the same value of  $\gamma_{\text{Si}}$ .

### E. Validation of LD within efsDFT

Validation of the structure obtained using efsDFT based LD is demonstrated using the pair distribution function  $g(r)$ .<sup>34</sup> For finite size NCs, the average number of neighbors at a distance  $r$  is expected to be smaller than the bulk value due to surface atoms with a smaller number of neighbors. Figure 4 shows a close agreement between the dDFT and efsDFT based LD estimates of  $g(r)$  of the  $\text{Si}_{35}\text{H}_{36}$  NC at two temperatures. The inset focuses on the first peak in  $g(r)$ , comparing the efsDFT to dDFT at  $T = 30$  and 300 K.

## III. RESULTS

In Secs. I and II, we presented the methods and assessed the accuracy and validity of the efsDFT based LD. Here we apply the method to study the structural properties of larger NCs exceeding  $N_e = 3000$  electrons. The Si-Si pair distribution functions  $g(r)$  at two temperatures  $T$  (30 and 300 K) are displayed in the upper panel of Fig. 5 for  $\text{Si}_{147}\text{H}_{100}$  and  $\text{Si}_{705}\text{H}_{300}$ . Temperature broadens the peaks by a factor of 2–3 without significantly changing the peak position.

In the lower panel of Fig. 5, we plot the transient and relaxed  $g(r)$  at 30 K for the two systems, focusing on the first, nearest neighbor peak. As described also for  $\text{Si}_{35}\text{H}_{36}$ , the initial positions of the Si atoms for both systems were taken from the experimental bulk values, and all surface Si atoms with more than two dangling bonds were removed. The remaining surface Si atoms was then passivated using one or two H atoms placed in a tetrahedral position at the Si-H distance of 1.47 Å. The initially sharp peak broadens and shifts to longer Si-Si bond lengths as the system relaxes towards thermal equilibrium. For

30 K, the relaxation times are 180 and 650 fs for  $\text{Si}_{147}\text{H}_{100}$  and  $\text{Si}_{705}\text{H}_{300}$ , respectively. For 300 K, they are 180 and 250 fs, respectively.

The relaxation transient is studied in greater detail in Fig. 6, where the average nearest-neighbor bond lengths are shown for  $\text{Si}_{147}\text{H}_{100}$  (spherical shells A-B) and  $\text{Si}_{705}\text{H}_{300}$  (shells A-E); see Table II for the definition and properties of the shells. In  $\text{Si}_{705}\text{H}_{300}$  the deep layer shells (A-D) relax slower than those near the surface showing that relaxation progresses from the surface inwards. The difference between the relaxation times of the two systems is correlated with the smaller frequency,  $\omega$ , of the breathing mode of the larger NC. In the limit of an overdamped motion (as is the case here since  $\gamma^2 \gg \omega^2$ ), the relaxation is dominated by two time scales proportional to  $\gamma^{-1}$  and  $(\omega^2/\gamma)^{-1}$ . The former leads to a fast relaxation, while the latter is slower and depends on the value of  $\omega^{-2}$ . The ratio of the breathing mode frequency for the two particles is  $\frac{\omega_L^2}{\omega_S^2} \approx 2.8$  (L/S for large/small) assuming that the breathing mode frequency scales linearly with the NC diameter.<sup>41</sup> This is similar to the ratio of the relaxation times ( $650/180 = 3.6$ ) for the lower temperature. At the higher temperature, one needs to consider anharmonic effects which are more pronounced in the large NC with lower acoustic phonons. Another noticeable feature in Fig. 6 is that the Si-Si bonds seem slightly shorter in  $\text{Si}_{147}\text{H}_{100}$  than in  $\text{Si}_{705}\text{H}_{300}$ . This results from the difference in the bond distance of atoms in the outer shell, while the inner shell atoms have similar bond distances.

## IV. CONCLUSIONS

In this paper, we developed an *ab initio* Langevin dynamics approach based on a new embedded saturated fragment stochastic DFT method. We showed how the noisy forces resulting from the efsDFT calculation are used to generate a set of configurations that are distributed canonically at cryogenic and room temperatures. Choosing the friction coefficients and the number of stochastic orbitals according to the criteria presented in Sec. II D, thermalization is reached within  $\approx 100$  time steps for these materials, since the method is trivially parallelizable, larger computer resources would allow to easily reduce the friction coefficients thus greatly improving the sampling efficiency. While the methods presented here have already allowed impressive achievements, such as deter-

TABLE II. The shells of the silicon NCs used for analyzing the bond length relaxation in Fig. 6: Their inner and outer radii (in Å), the number of Si atoms  $N_{\text{Si}}$ , and the number of nearest neighbor (NN) Si-Si pairs  $N_{\text{NN}}$ .

Shell	$R_{\text{in}}$	$R_{\text{out}}$	$N_{\text{Si}}$	$N_{\text{NN}}$
A	0	5.5	35	52
B	5.5	8.5	113	158
C	9.0	11.6	153	163
D	11.6	13.6	200	189
E	13.6	15.1	205	168

mining the structural properties of silicon NCs of 3 nm diameter containing more than 3000 electrons, larger systems still, of unprecedented size, are now coming within our grasp due to the linear-scaling highly parallelizable features of sDFT.

## ACKNOWLEDGMENTS

E.R. acknowledges support from the Physical Chemistry of Inorganic Nanostructures Program, KC3103, Office of Basic Energy Sciences of the United States Department of Energy under Contract No. DE-AC02-05CH11232. D.N. acknowledges support from the NSF Grant No. DMR/BSF-1611382. R.B. acknowledges support from the US-Israel Binational Science foundation under the BSF-NSF program, Grant No. 2015687.

## APPENDIX A: THE EMBEDDED SATURATED FRAGMENTS APPROACH

Here, we provide the technical details for the embedded fragment method described in Sec. II. The method corrects the stochastic estimate  $\langle \hat{A} \rangle^I$  for the expectation value  $\langle \hat{A} \rangle$  of a one-body operator  $\hat{A}$ , using calculations performed on  $F$  separate fragments [see Eq. (11)],

$$\langle \hat{A} \rangle = \langle \hat{A} \rangle^I + \sum_{f=1}^F \Delta A_f^I, \quad (\text{A1})$$

where the stochastic correction due to fragment  $f$  is

$$\Delta A_f^I = \langle \hat{A}_f \rangle - \langle \hat{A}_f \rangle^I, \quad (\text{A2})$$

and the deterministic and the stochastic estimates,  $\langle \hat{A}_f \rangle$  and  $\langle \hat{A}_f \rangle^I$ , are calculated directly on the fragment itself.

Previous implementations of embedded fragment sDFT were applied to systems of many weakly interacting molecules where the selection of fragments or clusters of such molecules was natural.<sup>17</sup> We now describe a new method for defining and carrying calculations with fragments which can break up covalently bonded systems, such as silicon NCs. The large system is divided into  $F$  small fragments composed of one or more bonded atoms each. The surface dangling bonds of the fragment are passivated using an H atom placed in 1.46 Å from the Si atom, in the direction of the neighboring atom which is not included in the fragment. This forms a saturated fragment. For a saturated fragment  $f$ , the deterministic KS-DFT method is applied to determine the KS eigenvalues  $\varepsilon_n^f$  and eigenfunctions  $\psi_n^f(\mathbf{r})$ . Further, occupation numbers  $(p_n^f)^2 = \frac{1}{2} \text{erfc}(\beta(\varepsilon_n^f - \mu_f))$  are introduced for determining the saturated fragment density  $n_{sf}(\mathbf{r}) = \sum_n (p_n^f)^2 \psi_n^f(\mathbf{r})^2$ . The fragment density  $n_f(\mathbf{r}) = c_f(\mathbf{r})^2 n_{sf}(\mathbf{r})$  is ‘‘carved out’’ of  $n_{sf}(\mathbf{r})$  using a carving function  $c_f(\mathbf{r})^2$ . Thus

$$n_f(\mathbf{r}) = c_f(\mathbf{r})^2 \sum_n (p_n^f)^2 \psi_n^f(\mathbf{r})^2, \quad (\text{A3})$$

where, inspired by Hirshfeld partitioning,<sup>42</sup> the carving function is defined as

$$c_f(\mathbf{r}) = \sqrt{\frac{\sum_{a \in f} n_a^{(0)}(\mathbf{r})}{\sum_{a \in sf} n_a^{(0)}(\mathbf{r})}},$$

where  $n_a^{(0)}(\mathbf{r})$  is the spherical density of neutral atom  $a$ . The temperature parameter  $\beta$  in the definition of the population  $p_n^f$  is chosen to be the same value as that of the sDFT calculation, while the chemical potential  $\mu_f$  of each fragment is determined by the condition of neutrality of the fragment,

$$\int n_f(\mathbf{r}) d\mathbf{r} = \int \sum_{a \in f} n_a^{(0)}(\mathbf{r}) d\mathbf{r}. \quad (\text{A4})$$

Defining non-orthogonal functions  $\tilde{\psi}_n^f(\mathbf{r}) = c_f(\mathbf{r}) p_n^f \psi_n^f(\mathbf{r})$ , the fragment density of Eq. (A3) becomes  $n_f(\mathbf{r}) = 2 \sum_n \tilde{\psi}_n^f(\mathbf{r})^2$ , so the chemical potential is determined from the condition

$$2 \sum_n \langle \tilde{\psi}_n^f | \tilde{\psi}_n^f \rangle = \int \sum_{a \in f} n_a^{(0)}(\mathbf{r}) d\mathbf{r}. \quad (\text{A5})$$

After determining  $\mu_f$  and in order to construct the reduced density matrix (RDM), we orthogonalize the functions  $\tilde{\psi}_n^f(\mathbf{r})$  by diagonalizing the overlap matrix  $S_{mm'}^f = \langle \tilde{\psi}_m^f | \tilde{\psi}_{m'}^f \rangle$ , obtaining the unitary matrix  $U_f$  of eigenvectors and the eigenvalues  $s_m^f > 0$  (so that  $U_f^T S_f U_f = \text{diag}[s_1^f, s_2^f, \dots]$ ). The orthogonal wavefunctions are  $\phi_m^f(\mathbf{r}) = \sum_n \tilde{\psi}_n^f(\mathbf{r}) U_{nm}^f$  and the norm is  $\langle \phi_m^f | \phi_m^f \rangle = s_m^f$ . Using the new wave functions, the unsaturated fragment density is given by

$$n_f(\mathbf{r}) = 2 \sum_m \phi_m^f(\mathbf{r})^2$$

and the RDM by

$$\hat{\theta}_f = 2 \sum_m |\phi_m^f\rangle \langle \phi_m^f|.$$

Using the RDM, we express the unsaturated fragment expectation value appearing in Eq. (A2) as

$$\langle \hat{A}_f \rangle \equiv \text{tr} [\hat{\theta}_f \hat{A}] = \text{tr} \left[ \sqrt{\hat{\theta}_f} \hat{A} \sqrt{\hat{\theta}_f} \right],$$

where

$$\sqrt{\hat{\theta}_f} = \sqrt{2} \sum_m (s_m^f)^{-1/2} |\phi_m^f\rangle \langle \phi_m^f|. \quad (\text{A6})$$

By choosing the fragment grid-points to be a subset of the full system grid, each stochastic orbital  $\chi_i$  ( $i = 1, \dots, I$ ) of the full system appears as a stochastic orbital on the fragment grid and can be used to perform the stochastic estimate appearing in Eq. (A2) as

$$\langle \hat{A}_f \rangle^I = \frac{1}{I} \sum_i \left\langle \chi_i \left| \sqrt{\hat{\theta}_f} \hat{A} \sqrt{\hat{\theta}_f} \right| \chi_i \right\rangle_f,$$

where the subscript  $f$  on the left denotes integration over the fragment grid. The difference  $\Delta A_f^I = \langle \hat{A}_f \rangle - \langle \hat{A}_f \rangle^I$  in Eq. (A2) can now be written in a unified form as

$$\Delta A_f^I = 2 \sum_{mm'} \Delta_{mm'}^{fI} \langle \phi_m^f | \hat{A} | \phi_{m'}^f \rangle_f, \quad (\text{A7})$$

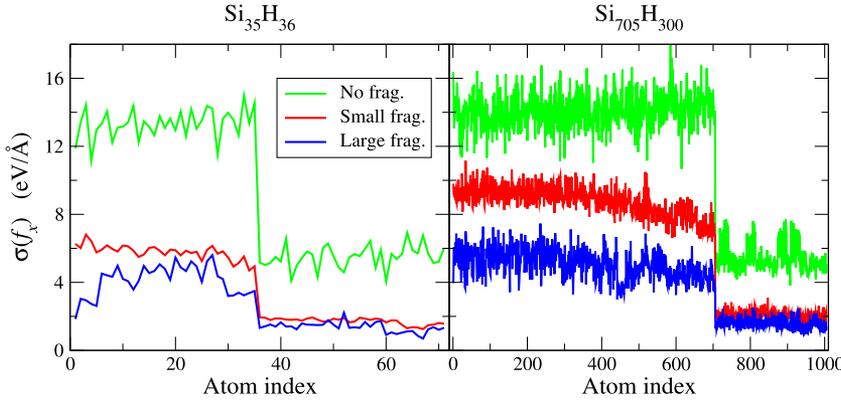


FIG. 7. Efficacy of fragments on the inherent sDFT STD  $\sigma_1(f_x)$  of the  $x$ -component of the force on each atom of  $\text{Si}_{35}\text{H}_{36}$  (left) and  $\text{Si}_{705}\text{H}_{300}$  (right) NCs. The inherent STD  $\sigma_1$  is the actual STD  $\sigma$  times  $\sqrt{I}$ . Si atoms are shown first followed by H atoms, where atoms are ordered by distance from the NC center. Calculations are done on the configuration cut out from the bulk silicon, where H atoms were placed near the surface for saturating the dangling bonds.

where

$$\Delta_{mm'}^{fI} \equiv \delta_{mm'} - \frac{1}{I} \sum_i \frac{\langle \chi_i | \phi_m^f \rangle_f \langle \phi_{m'}^f | \chi_i \rangle_f}{\sqrt{s_m^f s_{m'}^f}}. \quad (\text{A8})$$

Hence, by calculating the matrix  $\Delta_{mm'}^{fI}$ , all types of expectation value corrections can be obtained from Eq. (A7).

The efficacy of embedded fragments in sDFT force calculations is achieved through a reduction of the STD  $\sigma(f_x)$  of a force component. The STD  $\sigma(f_x)$  is proportional to  $1/\sqrt{I}$ , where  $I$  is the number of stochastic orbitals and the proportionality constant, denoted by  $\sigma_1(f_x) = \sqrt{I}\sigma(f_x)$ , is called the *inherent* STD. This quantity depends on the NC characteristics but not on the number of stochastic orbitals. In Fig. 7 we plot the inherent STD on each atom for  $\text{Si}_{35}\text{H}_{36}$  and  $\text{Si}_{705}\text{H}_{300}$  as a function of fragment size. Even the use of the smallest fragments reduces the inherent force STD by a significant factor, 1.6 (for  $\text{Si}_{705}\text{H}_{300}$ ) to 2.3 (for  $\text{Si}_{35}\text{H}_{36}$ ). Using larger fragments reduces the STD by an additional factor of  $\approx 1.5$ , with increasing effect for larger systems since the electron density in the larger fragments is similar to that of the full system. It is interesting to see that for the forces there is no noticeable sublinear scaling: the inherent STD for both systems is similar, with the larger system having slightly higher ( $\approx 5\%$ ) STD.

In summary, the embedded fragment sDFT method serves as a way to expedite the sDFT calculation by a judicious choice of fragment size and composition. As the fragment size grows, the numerical effort invested in sDFT decreases (due to reduction of STD) while in dDFT it increases. For example, consider Fig. 7 where we showed that increasing the fragment size by a factor of 10–20 reduces the STD by a factor of 2 and therefore the sDFT CPU time by a factor of  $\approx 2^2 = 4$ . On the other hand since the fragments are ten-fold larger, the amount of dDFT work on them increases (cubically) by a factor of more than  $\sim 10^3$ . Clearly then, the optimal fragment size is system dependent. Embedded fragments have the additional benefit of providing an initial density for the SCF calculation, significantly reducing the number of SCF cycles.

## APPENDIX B: STOCHASTIC ESTIMATES OF THE FORCES AND ENERGY PERTURBATIONS

In the stochastic method, the electronic density is [see Eq. (6)]:

$$n(r) = 2 \langle \chi | \hat{P} \delta(r - \hat{r}) \hat{P} | \chi \rangle, \quad (\text{B1})$$

where  $\hat{P} = \sqrt{\theta_\mu}$  is the Chebyshev expansion of the projection operator, depending on  $\beta$  and  $\mu$ , on the occupied space of  $\hat{h}_{KS}$ , and the energy is

$$\begin{aligned} E &= 2 \langle \chi | \hat{P} \hat{T} \hat{P} | \chi \rangle + \int v_{eN}(r; R) n(r) dr + E_{HXC}[n], \\ &= 2 \langle \chi | \hat{P} [\hat{T} + v_{eN}(r; R)] \hat{P} | \chi \rangle + E_{HXC}[n], \end{aligned} \quad (\text{B2})$$

where  $E_{HXC}[n]$  is the Hartree-exchange-correlation energy functional, depending only on the electronic density  $n(r)$ . Under variation in position of nuclei  $R$ ,

$$\begin{aligned} \delta E &= 2 \langle \chi | \hat{P} [\hat{T} + v(\hat{r}, R)] \delta \hat{P} | \chi \rangle \\ &\quad + 2 \langle \chi | \delta \hat{P} [\hat{T} + v(\hat{r}, R)] \hat{P} | \chi \rangle \\ &\quad + \langle \chi | \hat{P} \delta v(\hat{r}, R) \hat{P} | \chi \rangle \\ &\quad + \int v_{HXC}(r) \delta n(r) dr, \end{aligned} \quad (\text{B3})$$

which using

$$\begin{aligned} \delta n(r) &= 2 \langle \chi | \delta \hat{P} \delta(r - \hat{r}) \hat{P} | \chi \rangle \\ &\quad + 2 \langle \chi | \hat{P} \delta(r - \hat{r}) \delta \hat{P} | \chi \rangle \end{aligned} \quad (\text{B4})$$

can be written as

$$\begin{aligned} \delta E &= 2 \langle \chi | \hat{h}_{KS} \hat{P} \delta \hat{P} + \delta \hat{P} \hat{h}_{KS} | \chi \rangle \\ &\quad + \langle \chi | \hat{P} \delta v(\hat{r}, R) \hat{P} | \chi \rangle, \\ &= 2 \langle \chi | \hat{h}_{KS} \hat{P} \delta \hat{P} + \delta \hat{P} \hat{h}_{KS} | \chi \rangle \\ &\quad + \int n(r) \delta v(r; R) d^3r \end{aligned} \quad (\text{B5})$$

The average of the second term on the right leads to the Hellmann-Feynman force Eq. (8), while the first term can be shown to vanish when a full sampling is made on  $\chi$  and when  $\beta \rightarrow \infty$ , owing to the fact that within these limits  $P\delta P P = 0$ .

<sup>1</sup>R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).

<sup>2</sup>R. N. Barnett, U. Landman, A. Nitzan, and G. Rajagopal, *J. Chem. Phys.* **94**, 608 (1991).

<sup>3</sup>M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).

<sup>4</sup>G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).

<sup>5</sup>M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello, *J. Chem. Phys.* **103**, 150 (1995).

<sup>6</sup>D. Marx and J. Hutter, “*Ab initio* molecular dynamics: Theory and implementation,” in *Modern Methods and Algorithms of Quantum Chemistry, Proceedings, NIC Series Vol. 3*, edited by J. Grotendorst (John von Neumann Institute for Computing, Julich, 2000), p. 329.

- <sup>7</sup>H. B. Schlegel, J. M. Millam, S. S. Iyengar, G. A. Voth, A. D. Daniels, G. E. Scuseria, and M. J. Frisch, *J. Chem. Phys.* **114**, 9758 (2001).
- <sup>8</sup>J. M. Herbert and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **7**, 3269 (2005).
- <sup>9</sup>J.-Y. Raty, F. Gygi, and G. Galli, *Phys. Rev. Lett.* **95**, 096103 (2005).
- <sup>10</sup>M. Bockstedte, A. Kley, J. Neugebauer, and M. Scheffler, *Comput. Phys. Commun.* **107**, 187 (1997).
- <sup>11</sup>M. Cawkwell, A. M. Niklasson, and D. M. Dattelbaum, *J. Chem. Phys.* **142**, 064512 (2015).
- <sup>12</sup>S. Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999).
- <sup>13</sup>M. J. Gillan, D. R. Bowler, A. S. Torralba, and T. Miyazaki, *Comput. Phys. Commun.* **177**, 14 (2007).
- <sup>14</sup>R. Baer and M. Head-Gordon, *Phys. Rev. Lett.* **79**, 3962 (1997).
- <sup>15</sup>J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejón, and D. Sanchez-Portal, *J. Phys.: Condens. Matter* **14**, 2745 (2002).
- <sup>16</sup>R. Baer, D. Neuhauser, and E. Rabani, *Phys. Rev. Lett.* **111**, 106402 (2013).
- <sup>17</sup>D. Neuhauser, R. Baer, and E. Rabani, *J. Chem. Phys.* **141**, 041102 (2014).
- <sup>18</sup>Y. Gao, D. Neuhauser, R. Baer, and E. Rabani, *J. Chem. Phys.* **142**, 034106 (2015).
- <sup>19</sup>D. Neuhauser, E. Rabani, Y. Cytter, and R. Baer, *J. Phys. Chem. A* **120**, 3071 (2015).
- <sup>20</sup>M. F. Hutchinson, *Commun. Stat. Simul. Comput.* **19**, 433 (1990).
- <sup>21</sup>D. A. Drabold and O. F. Sankey, *Phys. Rev. Lett.* **70**, 3631 (1993).
- <sup>22</sup>L. W. Wang, *Phys. Rev. B* **49**, 10154 (1994).
- <sup>23</sup>H. Röder, R. Silver, D. Drabold, and J. J. Dong, *Phys. Rev. B* **55**, 15382 (1997).
- <sup>24</sup>R. Baer, T. Seideman, S. Ilani, and D. Neuhauser, *J. Chem. Phys.* **120**, 3387 (2004).
- <sup>25</sup>R. Baer and E. Rabani, *Nano Lett.* **12**, 2123 (2012).
- <sup>26</sup>C. Attaccalite and S. Sorella, *Phys. Rev. Lett.* **100**, 114501 (2008).
- <sup>27</sup>P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- <sup>28</sup>W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- <sup>29</sup>H. Tal-Ezer and R. Kosloff, *J. Chem. Phys.* **81**, 3967 (1984).
- <sup>30</sup>R. Baer and M. Head-Gordon, *J. Chem. Phys.* **107**, 10003 (1997).
- <sup>31</sup>J. Hellman, *Einführung in die Quantenchemie* (Springer, Deuticke, Leipzig, 1937).
- <sup>32</sup>R. P. Feynman, *Phys. Rev.* **56**, 340 (1939).
- <sup>33</sup>All calculations in this work use real-space grids of spacing  $\Delta x = 0.5a_0$  and Troullier-Martins norm-conserving pseudopotentials<sup>43</sup> within the Kleinman-Bylander approximation.<sup>44</sup> Fast Fourier transforms were used for applying the kinetic energy operator and for determining the Hartree potentials, and the method of Ref. 45 was used for treating the long range Coulomb interactions in a finite simulation cell with periodic boundary conditions. DFT calculations were performed under the local density approximation (LDA).
- <sup>34</sup>M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987), p. xix, 385.
- <sup>35</sup>N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, 1992), Vol. 1.
- <sup>36</sup>R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, USA, 2001).
- <sup>37</sup>D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, 2002), p. xxii, 638.
- <sup>38</sup>P. Langevin, *C. R. Acad. Sci. Paris* **146**, 530 (1908).
- <sup>39</sup>N. Grønbech-Jensen and O. Farago, *Mol. Phys.* **111**, 983 (2013).
- <sup>40</sup>E. Martínez, M. J. Cawkwell, A. F. Voter, and A. M. Niklasson, *J. Chem. Phys.* **142**, 154120 (2015).
- <sup>41</sup>E. Ghavanloo, S. Fazelzadeh, T. Murmu, and S. Adhikari, *Phys. E (Amsterdam, Neth.)* **66**, 228 (2015).
- <sup>42</sup>F. Hirshfeld, *Theor. Chim. Acta* **44**, 129 (1977).
- <sup>43</sup>N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- <sup>44</sup>L. Kleinman and D. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- <sup>45</sup>G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.* **110**, 2810 (1999).