Transition to metallization in warm dense helium-hydrogen mixtures using stochastic density functional theory within the Kubo-Greenwood formalism

Yael Cytter

Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, Hebrew University of Jerusalem, Jerusalem 9190401, Israel

Eran Rabani[*]

Department of Chemistry, University of California, and Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
and Raymond and Beverly Sackler Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel

Daniel Neuhauser[†]

Department of Chemistry, University of California, Los Angeles, California 90095, USA

Martin Preising and Ronald Redmer[‡]

Institute of Physics, University of Rostock, A.-Einstein-Straße 23, 18059 Rostock, Germany

Roi Baer[§]

Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, Hebrew University of Jerusalem, Jerusalem 9190401, Israel

The Kubo-Greenwood (KG) formula is often used in conjunction with Kohn-Sham (KS) density functional theory (DFT) to compute the optical conductivity, particularly for warm dense matter. For applying the KG formula, all KS eigenstates and eigenvalues up to an energy cutoff are required and thus the approach becomes expensive, especially for high temperatures and large systems, scaling cubically with both system size and temperature. Here, we develop an approach to calculate the KS conductivity within the stochastic DFT framework, which requires knowledge only of the KS Hamiltonian but not its eigenstates and values. We show that the computational effort associated with the method scales linearly with system size and reduces in proportion to the temperature, unlike the cubic increase with traditional deterministic approaches. In addition, we find that the method allows an accurate description of the entire spectrum, including the high-frequency range, unlike the deterministic method which is compelled to introduce a high-frequency cutoff due to memory and computational time constraints. We apply the method to helium-hydrogen mixtures in the warm dense matter regime at temperatures of $\sim$60 kK and find that the system displays two conductivity phases, where a transition from nonmetal to metal occurs when hydrogen atoms constitute $\sim$0.3 of the total atoms in the system.

I. INTRODUCTION

The state of warm dense matter (WDM) is characterized by high atomic density, similar to conventional condensed matter systems, and elevated temperatures of several electron volts (1 eV $\approx 10^4$ K). This is an intermediate regime bridging plasma physics and condensed matter physics for which equations of state (EOSs) and other properties are of interest. One example appears in the study of hydrogen-helium mixtures under extreme conditions, where the EOS [1], phase separation, and physical properties such as conductivity [2] and miscibility [3] can be used to explain the luminosity and gravitational moments of planets such as Jupiter and other gas giants, as well as their formation and evolution characteristics [4–6]. Generally, EOSs and properties are calculated for various materials using first-principles methods, specifically the Kohn-Sham density functional theory (KS-DFT) at finite temperatures [7–10], often shows good agreement with experiments [10–12]. Within the KS-DFT framework, WDM conductivity is often obtained by using the Kubo-Greenwood (KG) formalism [13–16] with good results when compared to experiment. The KS-DFT and the KG electrical conductivity equation when applied to WDM requires large computational effort which increases dramatically with temperature and system size, because of the need to construct and propagate all the occupied KS eigenstates, as well as a sufficient number of unoccupied states, the number of which grows as T^3, where T is the temperature [17].

Recently, stochastic DFT (sDFT) approaches that circumvent the computational difficulties mentioned above have

---

[*]eran.rabani@berkeley.edu
[†]dxn@chem.ucla.edu
[‡]ronald.redmer@uni-rostock.de
[§]roi.baer@huji.ac.il

been developed [17–22] for ground/thermal state calculations. These have also served as a basis for developing time-dependent methodologies for description of materials properties [23–26]. It was shown that sDFT is especially useful for EOS calculations in the WDM regime since it involves a computational effort that scales as T^{-1} [17].

In this paper, we develop an approach for calculating the KG conductivity within the framework of sDFT. The main advantage of the approach is that it does not require any knowledge of the occupied or empty KS orbitals. We show and benchmark a stochastic method to sample the KG conductivity. We then use the method to study the conductivity in hydrogen-helium mixtures. Our approach is similar to previously developed stochastic conductivity approaches [27–29] but differs in essential implementation details and is unique in its combination with sDFT calculations.

In the paper, we present the development of the stochastic KG (sKG) method and provide important implementation details, as well as demonstrations of the method's validity and a discussion of the statistical errors and scaling in Sec. II. In Sec. III the sDFT-sKG method is applied to the study of the conductivity of mixtures of helium and hydrogen in the warm dense matter regime, targeting metallization and beyond-linear-mixing effects.

II. METHOD

A. Time-dependent linear response

The time-dependent expectation value of a many-body observable $\hat{B}$ *after* an impulsive perturbation is applied through the observable $\hat{A}$ to a system at time $t = 0$ (usually assumed in thermal equilibrium) is given, in the linear-response regime, as the following correlation function [13,30]: $C_{AB}(t) = i\theta(t)\text{Tr}\{\rho(\beta, \mu)[\hat{A}, \hat{B}(t)]\}$, where $\hat{B}(t) = e^{i\hat{H}t/\hbar}\hat{B}e^{-i\hat{H}t/\hbar}$, $\hat{H}$ is the unperturbed Hamiltonian, and $\theta(t)$ is the Heaviside function imposing causality. The expectation values are performed with respect to the many-body thermal density $\rho(\beta, \mu) = Z(\beta, \mu)^{-1}e^{-\beta(\hat{H}-\mu\hat{N})}$, where $Z(\beta, \mu)$ is the partition function at chemical potential μ and inverse temperature $\beta = \frac{1}{k_B T}$, k_B being the Boltzmann constant.

One of the important applications of linear-response theory is the prediction of the frequency-dependent conductivity

$$\sigma(\omega) = \frac{2\pi e^2}{\Omega m_e^2 \hbar} \frac{\text{Im}[\tilde{C}_{PP}(\omega)]}{\omega}, \tag{1}$$

where Ω is the volume of the simulation cell and $\tilde{C}_{PP}(\omega)$ is the Fourier transform of the momentum-momentum correlation function,

$$\tilde{C}_{PP}(\omega) = \int_0^\infty C_{PP}(t)e^{-i\omega t}e^{-\frac{1}{2}\eta^2 t^2}dt, \tag{2}$$

and η is a small real parameter. In the limit $\omega \to 0$ L'Hopital's rule can be used to assess the DC conductivity:

$$\sigma(0) = \frac{2\pi e^2}{\Omega m_e^2 \hbar} \lim_{\omega \to 0} \frac{\partial \text{Im}\tilde{C}_{PP}(\omega)}{\partial \omega}. \tag{3}$$

For noninteracting particles, with a single-particle Hamiltonian $\hat{h}$, having eigenvalues ε_n and eigenstates $|n\rangle$, $n = 1, 2, \ldots$, the correlation function reduces to the following

expression:

$$C_{AB}(t) = -2\theta(t)\text{ImTr}[f_{\text{FD}}(\hat{h})\hat{a}(1 - f_{\text{FD}}(\hat{h}))\hat{b}(t)], \tag{4}$$

where $\hat{a}$, $\hat{b}$ are the single-particle perturbing and observed operators, respectively, $\hat{b}(t) = e^{i\hat{h}t}\hat{b}e^{-i\hat{h}t}$, and

$$f_{\text{FD}}(\hat{h}) \equiv \frac{1}{1 + e^{\beta(\hat{h}-\mu)}} \tag{5}$$

is the Fermi-Dirac distribution. Combining Eqs. (4) and (2) and taking the formal limit $\eta \to 0$ gives the Kubo-Greenwood (KG) conductivity [13,31]:

$$\sigma(\omega) = \frac{2\pi e^2}{\Omega m_e^2 \hbar\omega} \sum_{m,n}^{N_g} f_{mn}|p_{mn}|^2 \delta(\omega - \varepsilon_{nm}/\hbar), \tag{6}$$

where N_g is the number of grid points, $f_{mn} \equiv f_{\text{FD}}(\varepsilon_m) - f_{\text{FD}}(\varepsilon_n)$, $\varepsilon_{nm} = \varepsilon_n - \varepsilon_m$, and $p_{nm} = \langle n|\hat{p}|m\rangle$. For practical reasons, the summation over the occupied and unoccupied states is determined according to an energy cutoff and as a result the conductivity spectrum can be calculated only up to a corresponding frequency cutoff.

B. Stochastic calculation of the response function

To calculate the KG conductivity in a stochastic manner the stochastic trace formula [32] can be used to estimate the trace in Eq. (4). However, we found that a smaller statistical noise can be obtained if the stochastic trace is applied to the following equivalent but more symmetrical expression:

$$C_{PP}(t) = -2\theta(t)\text{ImTr}[\sqrt{f_{\text{FD}}}\hat{p}(1 - f_{\text{FD}})\hat{p}(t)\sqrt{f_{\text{FD}}}]. \tag{7}$$

To apply the stochastic trace formula, we define a set of stochastic orbitals χ, represented on the grid such that $\langle r_g|\chi_i\rangle = (\delta x)^{-3/2}e^{i\theta_g^i}$, where $\theta_g \in [0, 2\pi]$ is a random phase and δx is the grid spacing. The stochastic expression for $C_{PP}(t)$ is given by

$$C_{PP}(t) = -2\theta(t)\text{E}\{\text{Im}\langle\xi|\hat{p}(1 - f_{\text{FD}}(\hat{h}))e^{i\hat{h}t}\hat{p}e^{-i\hat{h}t}|\xi\rangle\}, \tag{8}$$

where $|\xi\rangle = \sqrt{f_{\text{FD}}(\hat{h})}|\chi\rangle$ and E{...} designates an expectation value.

The procedure consists of the following schematic steps:

(1) Set $n = 0$, $|\eta_j\rangle = |\xi_j\rangle$, $|\zeta_j\rangle = [1 - f_{\text{FD}}(\hat{h})]\hat{p}|\xi_j\rangle$, and the time step $\Delta t = \frac{\pi}{\Delta E}$, where $\Delta E = E_{\text{max}} - E_{\text{min}}$ and E_{max} (E_{min}) is the maximal (minimal) eigenvalue of $\hat{h}$ (the condition is required to avoid aliasing). The time step determines the cutoff frequency of the spectrum and $N_{\text{ts}} = \frac{\Delta E}{\Delta\omega} = \frac{\pi}{\Delta\omega\Delta t}$ is the total number of time steps for achieving a spectral resolution of $\Delta\omega$.

(2) Calculate $C_{PP}^j(n\Delta t) = -2\text{Im}\langle\zeta_j|\hat{p}|\eta_j\rangle$.

(3) Set $n = n + 1$, $|\eta_j\rangle = e^{-i\hat{h}\Delta t}|\eta_j\rangle$, $|\zeta_j\rangle = e^{-i\hat{h}\Delta t}|\zeta_j\rangle$.

(4) Go to 2 and repeat until $n = N_{\text{ts}}$.

(5) The response function is then averaged over I_σ (the number of stochastic orbitals), yielding $C_{PP}(n\Delta t) \approx \frac{1}{I_\sigma}\sum_{j=1}^{I_\sigma} C_{PP}^j(n\Delta t)$. This response function is then discrete-Fourier transformed and used to obtain the frequency-dependent conductivity [Eq. (1)].

The process is easily parallelized, since each element $C_{PP}^j(n\Delta t)$ is calculated independently before averaging in the

final step. In our calculations we do not use the above procedure directly because using Chebyshev expansions for the evolution operator, we found a way to expedite the calculation as described in Sec. II E.

The stochastic-KG (sKG) procedure forms a postprocessing step after a sDFT calculation [17,18] which provides the self-consistent KS Hamiltonian $\hat{h}$. The stochastic calculation requires two sets of stochastic orbitals. One set is used to perform the sDFT calculation with which $\hat{h}$ is determined; this set will be denoted "sDFT-o's." A second set, used in the sKG calculation to determine the conductivity, is denoted "sKG-o's." The KS wave functions are expanded using plane waves although the nonlocal part of the pseudopotentials are implemented using a real-space grid, to acheive high efficiency. For all the stochastic calculations in this paper, we used the local density approximation (LDA) [33] and Troullier-Martins norm-conserving pseudopotentials [34] within the Kleinman-Bylander representation [35].

C. Validation of the method

To validate the method we compare conductivity estimates with that of the well-established Quantum Espresso (QE) package [36], for carrying out both the DFT and the KG (using the KGEC module [37]) calculations. The results for H_{256} (at 4000 K) and a single He atom (at 3200 K) are shown in the panels of Fig. 1 and the density of states (DOS) $\rho(\varepsilon)$ is shown in the insets. The H_{256} nuclear configuration was obtained using an *ab initio* molecular dynamics (AIMD) simulation using VASP. It can be seen that for both systems the stochastic conductivity spectra and the DOS curves are in close agreement with the corresponding deterministic estimates of QE throughout the entire frequency/energy range.

To further test the method we also looked at He_{128} systems in the temperature range 27–57 kK and density range 0.71–0.83 $\frac{\text{g}}{\text{cm}^3}$. To obtain a set of nuclear configurations, an AIMD trajectory was run using the PBE [38] exchange-correlation (XC) functional, employing the plane-wave code VASP [39,40]. Snapshots of the nuclear configurations were then taken from the equilibrated part of the simulation, as described in Ref. [12]. For each snapshot we performed a sDFT calculation using 160 sDFT-o's to obtain the Hamiltonian. The standard deviation was estimated by using five different sets of 160 sKG-o's, each different from the set used for the Hamiltonian calculation, to avoid bias. For a given nuclear snapshot, the stochastic calculation produces a conductivity spectrum with a certain stochastic error. The stochastic fluctuations in our case turned out to be larger than the fluctuations arising from averaging over the different nuclear configurations. We therefore present here results obtained from one snapshot only. To calculate the discretized momentum-momentum correlation function $C_{PP}(\Delta t \times n)$ we used $n = 600$ time steps with $\Delta t = 0.25 \, \hbar E_h^{-1}$. The conductivity spectrum is then obtained from Eq. (8) using a Gaussian-broadening parameter of $\eta = 0.036$ eV.

The full spectrum and the standard deviation involved in the calculation as described above are shown in Fig. 2. The advantage of the stochastic method is apparent when looking at frequencies higher than ~70 eV, where the deterministic calculation of Ref. [12] gives no contributions above this
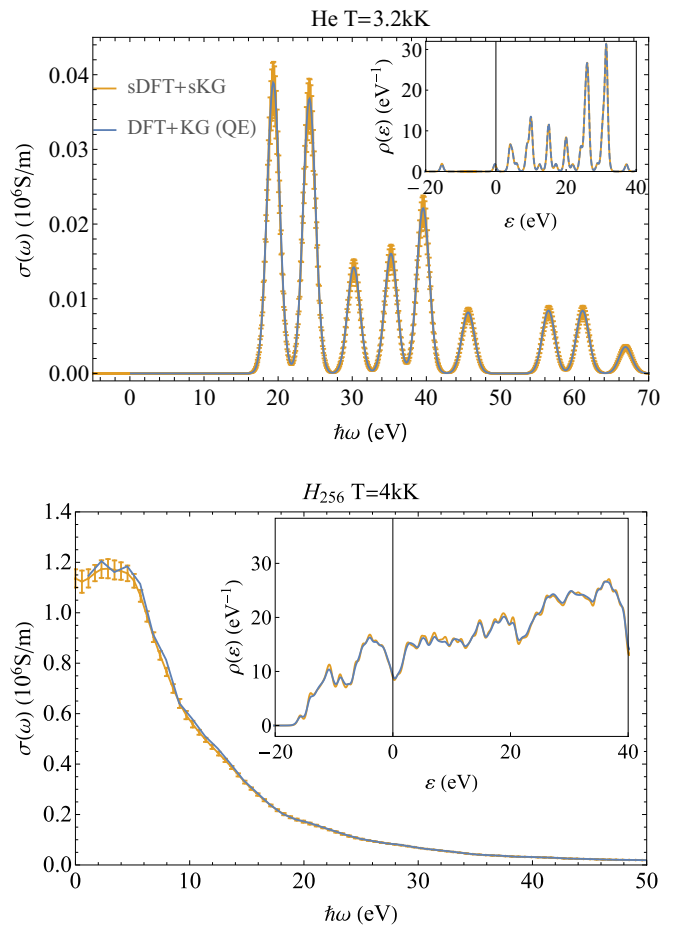

FIG. 1. The conductivity $\sigma(\omega)$ and DOS $\rho(\varepsilon)$ (in the insets) using stochastic versus deterministic (Quantum Espresso) calculations. We show two examples, each calculated in a periodic simulation cell of length L at the Γ point: an insulator He atom at 3200 K with $L = 5.3$ Å (upper panel) and a metallic system, H_{256} at 4000 K, with $L = 8$ Å (lower panel). The deterministic QE results used 200 KS eigenstates for the first system and 1700 for the second. For the stochastic DFT calculation we used 960 stochastic orbitals for the insulator and 480 orbitals for the conductor. The conductivity of both systems was calculated using 120 stochastic orbitals. A kinetic energy cutoff of 762 eV for He and 525 eV for H_{256} was used. Each peak was Gaussian broadened, deploying a width parameter equal to $\eta = 1.2$ eV [in sKG this parameter affects Eq. (2)].

cutoff energy which has to be introduced in plane-wave DFT codes like VASP. The deterministic frequency range could have increased in principle by including more KS states, but this would require an excessive computational effort. Careful analysis with respect to the cutoff energy shows that equation-of-state data and the low-frequency conductivity can be converged properly (see, e.g., Refs. [11,12]). The sKG calculation on the other hand samples states from the entire energy spectrum and therefore exhibits the physically correct asymptotic decay of $\omega^{-5/2}$, as expected for the free-electron gas. The correct high-frequency asymptotic behavior enables calculation of the Thomas-Reiche-Kuhn sum rule [41,42], which states that the total oscillator strength per electron
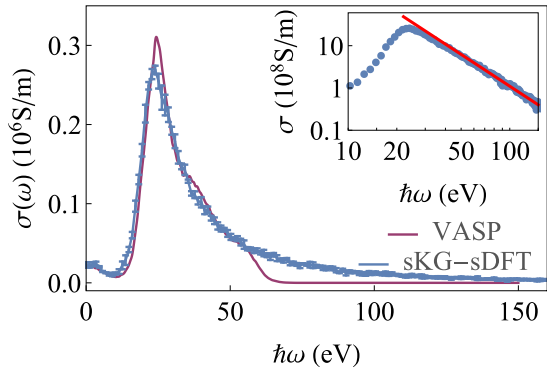
FIG. 2. The conductivity spectrum for a He_{128} system at $T = 27$ kK and density of $0.83 \frac{g}{cm^3}$. The sKG-sDFT LDA conductivity with error bars (with $I_\sigma = I_H = 160$) is compared to the deterministic results by VASP based on PBE, as described in Ref. [12]. The deterministic calculations were done with an energy cutoff of 800 eV using 570 KS states. Inset: The spectrum decay. The red line is proportional to $\omega^{-5/2}$.

f_{osc}/N_e, where

$$f_{osc} = \frac{m_e \Omega}{\pi e^2} \int_{-\infty}^{\infty} \sigma(\omega) d\omega, \qquad (9)$$

and $\sigma(\omega)$ is the conductivity defined in Eq. (1), should be equal to 1. The actual calculated values of f_{osc}/N_e are shown in Table I for three He_{128} systems (one of which we considered in Fig. 2 and two others, of different temperature and density, are given for further demonstration) and are indeed very close to the theoretical value of 1, signifying that the calculations are converged with respect to the number of states and the total time of propagation.

At intermediate frequencies, we find (Fig. 2) a close overall agreement between the deterministic and stochastic estimates of the conductivity spectra, despite the fact that both methods make use of different XC functionals. The most conspicuous feature of the spectrum in this range is its peak at $\hbar\omega_{peak} \approx 25$ eV, featuring the maximal deviation between the two spectra which is nonetheless small, with a 10% difference in height and 0.3 eV difference in the value of $\hbar\omega_{peak}$.

The DC conductivities for three different systems are displayed in the third and fourth columns of Table I and

TABLE I. The total oscillator strength per electron f_{osc}/N_e [see Eq. (9)] and the DC conductivity calculated using VASP based on PBE [12] and the sDFT-sKG based on LDA employing $I_\sigma = I_H = 160$ stochastic orbitals. The statistics for the stochastic calculation is obtained from 5 different sKG runs and that of the deterministic calculation was taken from 5 points in the proximity of the DC conductivity to evaluate the $\omega \to 0$ limit.

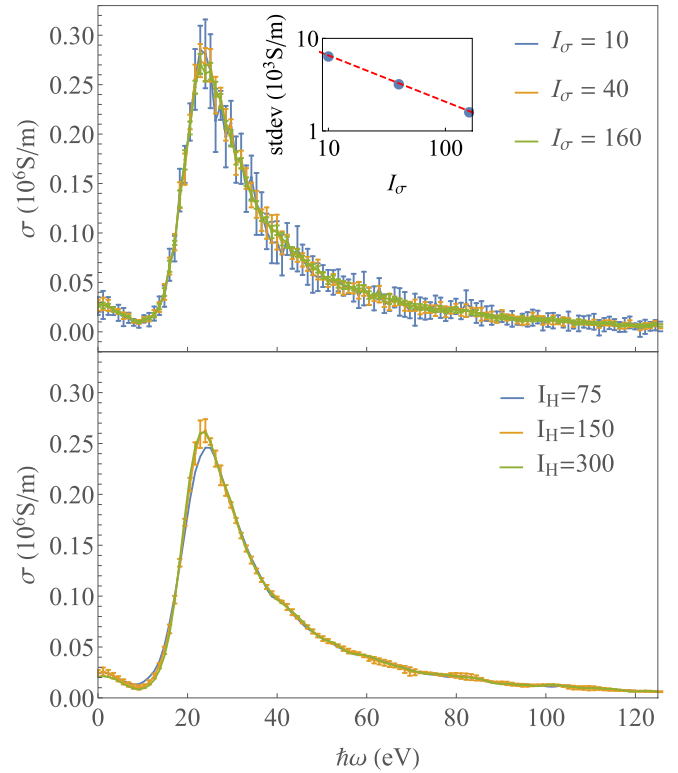| System			σ_{DC} (10^6 S/m)	
$\rho/(g/cm^3)$	T/kK	f_{osc}/N_e	VASP/PBE	sKG-sDFT/LDA
0.71	29	1.010	0.021 ± 0.001	0.026 ± 0.002
0.83	27	0.998	0.018 ± 0.001	0.02 ± 0.003
0.75	57	1.014	0.110 ± 0.002	0.10 ± 0.020

FIG. 3. The conductivity spectrum of He_{128} at 27 kK and density of 0.83 g/cm^3. Top panel: The conductivity based on one sDFT Hamiltonian (using $I_H = 150$) calculated with an increasing number I_σ of sKG-o's. Inset: The standard deviation (stdev) of the conductivity, averaged over all frequencies as a function of I_σ. The dashed line is proportional to $I_\sigma^{-1/2}$. Bottom panel: The conductivity based on three sDFT Hamiltonians, each obtained using I_H sDFT-o's. In order not to clutter the plot, error bars are given only for the $I_H = 150$ sDFT-o's calculation. The sKG calculations were all done using $I_\sigma = 150$ sKG-o's.

the agreement between the deterministic and stochastic zero-frequency limit is shown.

D. Analysis of the statistical errors

There are three sources of statistical errors in the calculation. The sDFT, which produces the Hamiltonian with which the conductivity is calculated with Eqs. (2) and (3), contributes two of the errors. One is the fluctuation which is measured by the standard deviation of the results, and is proportional to $I_H^{-1/2}$, where I_H is the number sDFT-o's. The second is the bias, related to the deviance of the average from the exact value, discussed in previous works [17,21], that is proportional to I_H^{-1}. In addition to the errors in the sDFT stage, the stochastic evaluation of the momentum-momentum correlation function also contributes an additional fluctuation. The effect of the two errors arising from the sDFT calculation on the conductivity spectrum is displayed at the bottom panel of Fig. 3. We show three spectra, each based on a distinct sDFT Hamiltonian, calculated using different values of I_H. For the case of $I_H = 150$ ten different sDFT-o sets were used in order to assess the fluctuation stemming from the stochastic procedure. For all three conductivity calculations, we used the

same set of $I_\sigma = 150$ sKG-o's, thereby leading to a similar fluctuation, so that we can focus on the errors resulting from the sDFT process. The spectra based on $I_H = 300$ are within the error bars of the $I_H = 150$ for all frequencies considered, while the spectrum that is based on $I_H = 75$ exhibits a deviation from the other two, especially near the $\omega \sim 25$ eV peak. Since the fluctuation is small, we deduce that this difference can be attributed to the bias component of the statistical error, and that it is small at $I_H = 75$ and much smaller than the fluctuation when $I_H \geqslant 150$.

Having discussed the two errors connected with the stochastic nature of the Hamiltonian, we now address the random fluctuations that arise from the sKG calculation. For this purpose, we take one of the sDFT Hamiltonians above (that was calculated using $I_H = 150$ sDFT-o's) and perform three conductivity spectra calculations on it using different values I_σ of sKG-o's. The resulting spectra are shown in the upper panel of Fig. 3. The inset shows that the standard deviation, averaged over all frequencies, decreases according to the central-limit theorem as expected. Since the sKG-o's are used to directly sample the trace, the statistical error should be a "pure" fluctuation, with no bias. Therefore, while the peak in this example exhibits a decrease as I_σ increases, we attribute that behavior to a fluctuation.

E. Algorithmic implementation and scaling of the algorithm

The computational time of the sKG algorithm, as described in Sec. II B and Eq. (7), is determined mainly by application of $f_{\mathrm{FD}}(\hat{h})$, $\sqrt{f_{\mathrm{FD}}(\hat{h})}$, and the time evolution operator $e^{-i\hat{h}\Delta t}$, all functions of the Hamiltonian $\hat{h}$ on given wave functions. Each of these Hamiltonian functions can be applied by using Chebyshev expansions [43–46], where the Hamiltonian is applied to the wave function repeatedly N_C times. The length of the expansion N_C is proportional to $\Delta E = E_{\max} - E_{\min}$ where $E_{\max}$ and $E_{\min}$ are upper and lower bounds on the maximal and minimal eigenvalues of $\hat{h}$, respectively. For the Fermi-Dirac functions $f_{\mathrm{FD}}(\hat{h})$ and $\sqrt{f_{\mathrm{FD}}(\hat{h})}$ the Chebyshev expansion length N_C is proportional to $\beta \Delta E$.

Propagating the wave function φ to different times can be performed with several Chebyshev expansions:

$$\varphi_n = e^{-i\hat{h}(n\Delta t)}\varphi = \sum_{m=0}^{N_C(n\Delta t)} a_m(n\Delta t)\phi_m, \qquad (10)$$

where ϕ_m are the Chebyshev recursion wave functions [47]. Note that the φ_n's for the different values of n are different linear combinations of the same recursion wave functions ϕ_m, but summed with different expansion coefficient $a_m(n\Delta t)$. We can therefore generate one set of $\phi_1, \ldots, \phi_{N_C}$ for generating the required set of φ_n's. The Chebyshev expansion length N_C is determined as the smallest integer for which $|a_m(n\Delta t)| < 10^{-9}$ for all $m \geqslant N_C$. Clearly, N_C depends on $n\Delta t$; hence the notation $N_C(n\Delta t)$. The expansions in Eq. (10) are highly beneficial since most of the computational effort goes to applying the Hamiltonian on the different wave functions, that is, calculating the set of ϕ's. Thus, to find the optimized number of simultaneously calculated time steps n, in Fig. 4 we look at the number of Chebyshev terms required *per time step*, $N_C(n\Delta t)/n$, alongside the wall time for every choice
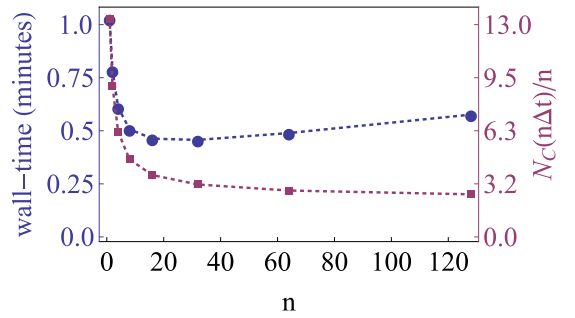

FIG. 4. The CPU time (blue circled markers) and $N_C(n\Delta t)/n$ (purple square markers) as a function of the number n of time step propagation operators used in the calculation. The calculation was done for He_{128} using $N_g = 60^3$ grid points, at 57 kK on 8 processors, where $N_{\mathrm{ts}} = 128$ and $\Delta t = 0.25$ a.u., for one dipole direction.

of n. It can be seen that $N_C(n\Delta t)/n$ is highest at $n = 1$ and as n increases, its value drops steeply toward an asymptotic plateau value smaller by a factor of ~ 4. It is seen that using this approach CPU times indeed decrease but due to an additional overhead of the calculation only a factor of 2 is obtained.

The computational effort for the sKG procedure has a near-linear scaling with system size N as does the sDFT, and this is due to the following two reasons: (1) the Hamiltonian $\hat{h}$ action on a wave function involves an $O(N_g \ln N_g)$ numerical complexity (this is the operation count of the fast Fourier transform involved in the kinetic energy operation), where $N_g \propto N$ is the number of grid points; and (2) the number of such Hamiltonian operations is $N_C \times I_\sigma$, where N_C (the Chebyshev expansion length) and I_σ (the number of sKG-o's) are both system-size independent. For the same reasons the sDFT calculation also scales linearly with N (as shown before in Ref. [18]). Furthermore, the computational effort in sDFT decreases as the temperature increases in proportion to $1/T$ [17], due to the fact that the FD Chebyshev expansion length N_C is proportional to $\Delta E/k_B T$ [48] and the energy range ΔE is system-size independent. The $O(N/T)$ scaling with system size and temperature we report here should be compared to the $O(N^3 T^3)$ scaling of the deterministic calculation based on Eq. (6), which requires calculation of all the occupied (and many unoccupied) states, the number of which is proportional to T^3 (based on the electron gas density of states). The system-size scaling can be seen in actual calculation, as shown in Fig 5, where the wall time for the DFT + KG calculation is shown as a function of the number of atoms in the system (keeping the density and temperature fixed as the number of atoms increases) for the deterministic case (using QE) and sDFT + sKG. For small system sizes the QE calculation is considerably faster than the stochastic approach. However, as the system size increases, due to liner scaling, the stochastic approach becomes competitive. At $N = 128$ we find a crossover and already for $N = 432$ the stochastic calculation is 10 times faster than the deterministic one.

III. MIXED He/H WDM SYSTEMS

As an application of the method, we study the conductivity and DOS of various systems with different hydrogen-helium mixtures at a temperature of 57 kK and constant volume.
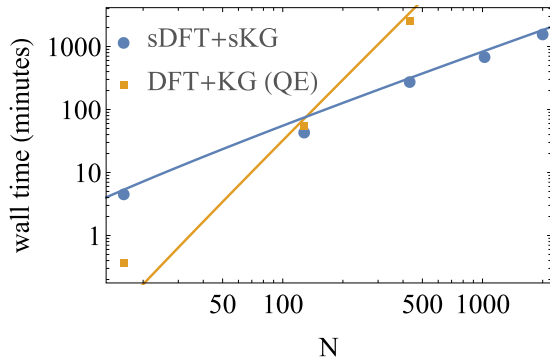
FIG. 5. Wall time for stochastic DFT + KG calculations with 600 time steps with $\Delta t = 0.25\,\hbar E_h^{-1}$ as well as a deterministic calculation done using Quantum Espresso (QE), as a function of the number of He atoms, at 0.75 $\frac{g}{cm^3}$ and 57 kK. The orange curve is proportional to N^3, and the blue curve is linear with N, the number of atoms.

For each system, we obtained a set of thermally distributed nuclear configurations using the electron force field (eFF) [49] dynamics as implemented in LAMMPS [50], which has been shown to give a good description of the pair correlation and equations of state of first row materials under extreme conditions [51,52]. For He$_{128}$ at 57 kK we generate a set of Boltzmann-distributed configurations using both an empirical force field and an *ab initio* approach taken from Ref. [12]. The configurations were then used to average the results over the thermal fluctuations of the nuclei. In Fig. 6 we compare the two sets of results and show that while giving two visibly different spectra they share similar trends with peaks/troughs located at nearly identical frequencies. Comparing the two DOSs we see small differences in the occupied state energies while they are nearly identical at the unoccupied state energies. Comparing the correlation functions $g(r)$, we find that AIMD gives significant weight to He pairs approaching as close as 0.5 Å while the eFF does not. Both functions show a peak at 1.1 Å, but it is more significant in AIMD. It is perhaps surprising that despite the rather large differences in
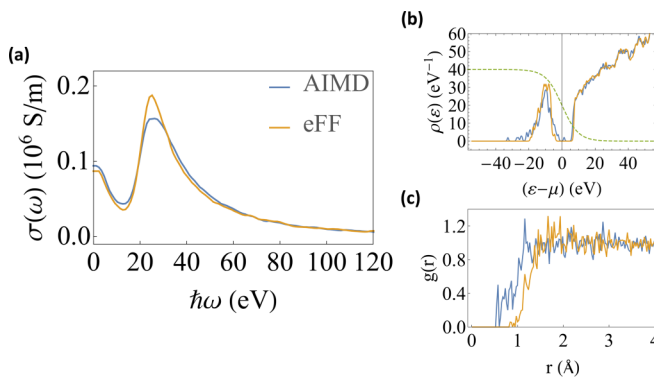

FIG. 6. Comparison of the calculated conductivity $\sigma(\omega)$ (a), the DOS $\rho(\varepsilon)$ (b), and the radial pair correlation $g(r)$ (c) for He$_{128}$ at 57 kK and density of 0.75 $\frac{g}{cm^3}$ based on configurations generated by AIMD [12] versus eFF dynamics. The dashed line in panel (b) shows the Fermi-Dirac level occupation, transiting from a value of 1 at low energies ε to 0 at high energies.

the pair correlation between the two methods, the electronic properties, as mentioned above, are not very different.

We characterize the mixture by the hydrogen fraction in the system

$$X_{\mathrm{H}} = \frac{N_{\mathrm{H}}}{N_{\mathrm{H}} + N_{\mathrm{He}}}, \tag{11}$$

where N_{H} and N_{He} are the number of hydrogen and helium atoms, respectively. For practical purposes, this ratio is achieved by holding the total number of atoms $N_{\mathrm{H}} + N_{\mathrm{He}}$ in the simulation cell constant and equal to 1024.

We ran five molecular dynamics trajectories at fixed volume (minimum image periodic boundary conditions for $L = 39.4a_0$) and temperature ($T = 57$ kK) with interactions between He and H described by the eFF force field with a cutoff of $6.45a_0$. Each trajectory started with the same ordered configuration, and a different velocity allocation, equilibrated, and then ran for a total of 3 ps with time step of 10^{-3} fs, needed because of consideration of both electronic and nuclear timescales. The duration of the trajectories corresponded to the correlation time of 3 ps estimated using the same data. The final nuclear configuration for each trajectory represented a set of five uncorrelated H-He mixtures. For each structure, a sDFT calculation determined the Hamiltonian $\hat{h}$ which was used for the sKG calculation of the conductivity spectrum. For both sDFT and sKG an identical simulation box and grid of $N_g = 120^3$ points were used which correspond to a grid spacing of $\delta x = 0.33a_0$. The sDFT calculation was based on $I_{\mathrm{H}} = 120$ sDFT-o's and the sKG calculation used a distinct set of $I_\sigma = 120$ sKG-o's.

In Fig. 7 the conductivity spectra and the DOS for three different mixtures are displayed. These two characteristics are closely related and will therefore be discussed together. The statistical fluctuation in the DOS (lower panel), denoted as error bars, was determined by running sDFT calculations on the five distinct configuration snapshots as described above, each using a different set of sDFT-o's. These five Hamiltonians were then used for evaluating the conductivity (upper panel) employing a different set of sKG-o's to avoid additional bias. The resulting five conductance spectra and DOSs were used for estimating the thermally averaged curves and their associated statistical errors. It is seen in the upper panel of Fig. 7 that the statistical fluctuations are small compared to the difference between the curves and they do not seem to increase as a function of the hydrogen atomic fraction X_{H} and therefore, only one snapshot was used in all other calculations.

When a relatively small fraction of hydrogen atoms is present in the system, it gives rise to a small peak at 3 eV inside the helium energy gap in the DOS (see the lower panel of Fig. 7). As the hydrogen concentration increases the He gap fills with states until it is no longer visible and at the same time the DOS of the valence band (seen in the figure at around −10 eV) decreases steadily. Both effects show a gradual transition to metallization as the hydrogen ratio grows. At high energies the DOS of all mixtures converges to the free-electron limit.

The sKG conductivity follows the changes seen in the DOS. Consider first the DC conductivity, shown in the lower panel of Fig. 8, which remains relatively constant as the hydrogen fraction grows until $X_{\mathrm{H}}^{\mathrm{crit}} \sim 0.3$. Beyond this value
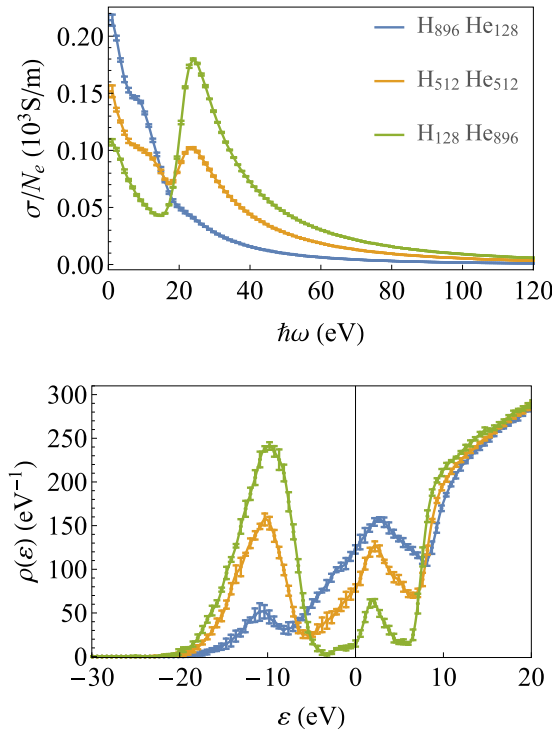
FIG. 7. The conductivity (upper panel) and the DOS (lower panel) of different systems containing 1024 atoms with different hydrogen percentages at $T = 57$ kK and an average atomic volume of $60a_0^3$ per atom. The conductivity was normalized according to the number of electrons in the system. The DOS is shifted so that the chemical potential is zero.

of the hydrogen fraction the DC conductivity increases nearly linearly with χ_H as a result of the energy gap filling in the DOS, allowing more transitions at low energies. Due to the finite temperature and therefore partial occupation there exist zero-frequency transitions even at helium-dominated systems causing the DC conductivity to change only by a factor of 2.5 when moving from pure-helium to pure-hydrogen systems. The peak in the He-dominated spectrum, as seen in the upper panel of Fig. 7, appears at around 25 eV and corresponds to the transition from the highest density in the occupied band to the nonoccupied band threshold levels (as seen in the DOS at 10 eV). Furthermore, at higher He concentrations due to the energy gap, transitions in 15 eV become less probable, resulting in a local minimum in the conductivity at this frequency.

Next, we consider the frequency ω_{max} for which the conductivity is maximal, plotted as a function of X_H in the top panel of Fig. 8. This frequency displays an abrupt shift of ω_{max} from ~ 25 eV to 0 (DC) as X_H passes through the critical value of $X_H^{crit} \sim 0.3$. This critical value indicates an abrupt nonmetal-to-metal transition in the H-He system as reported in Ref. [2] for considerably lower temperatures. This critical hydrogen concentration is well within the range of the Mott criterion for metallization in pure hydrogen, as seen in Ref. [53], which shows that it occurs at $n_H^{1/3} a_0 \approx 0.25$ for temperatures up to 15 kK. In the present system, we find the metallization density at $n_H^{1/3} a_0 \approx 0.18$, which seems
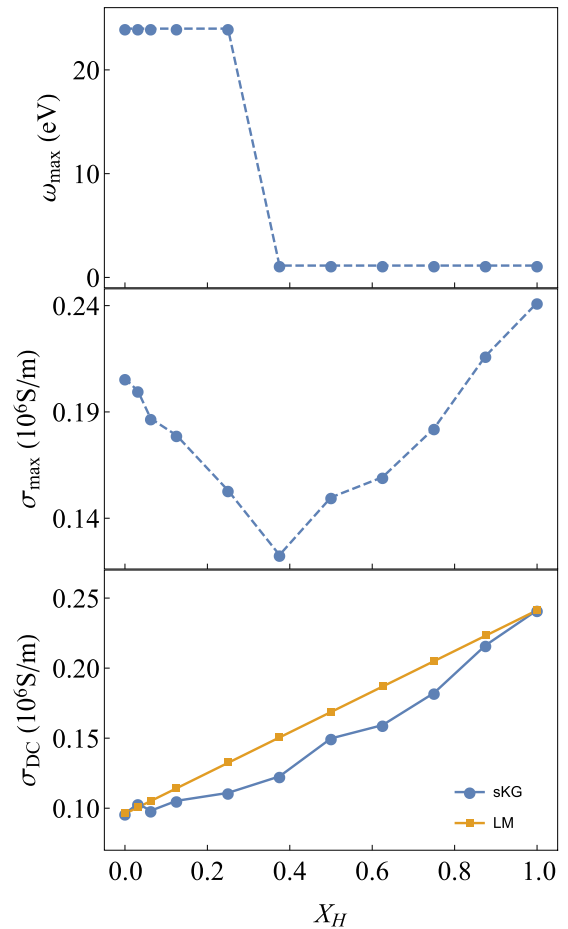

FIG. 8. The maximal conductivity frequency ω_{max} (top panel), maximal conductivity σ_{max} (middle panel), and the DC conductivity of the actual (round blue markers) and the linear-mixing model σ_{LM} (square orange markers), as a function of the hydrogen ratio X_H in He-H mixtures.

reasonable considering the fact that we are looking at a substantially higher temperature in which thermal effects promote the conductivity onset.

The finite ω_{max} is a result of the energy gap in what is generally an insulating system (He dominated) and the zero ω_{max} signifies its disappearance, allowing many of the energy transitions to occur at infinitesimal energy values. In the middle panel of Fig. 8 the transition through X_H^{crit} is seen as a qualitative change in the behavior of the maximal conductivity σ_{max}, which initially decreases as X_H approaches X_H^{crit}, and then increases as X_H grows further.

Finally, we compare the spectra in the different concentrations to a model of linear averaging of pure helium and pure hydrogen spectra, defined as

$$\sigma_{LM}(X_H; \omega) = X_H \sigma(1; \omega) + (1 - X_H)\sigma(0; \omega). \quad (12)$$

It can be seen in the lower panel of Fig. 8 that the DC conductivity $\sigma_{LM}(X_H; 0)$ based on the linear-averaging model is typically greater than the corresponding value calculated using sKG. This is due to the fact that in the actual system the environment each atom experiences includes, on average,

a mixture of He and H atoms, while in the linear-averaging model each atom is surrounded by atoms of its own kind.

IV. SUMMARY AND CONCLUSIONS

In this work we presented a stochastic approach, sKG, to calculate the conductivity using the Kubo-Greenwood formalism on top of a sDFT calculation. We showed that sKG conductivity can approach the values of the deterministic KS conductivity determined by the KG method when the number of sDFT-o's and sKG-o's are increased systematically. Moreover, the uniform sampling of all states of the system by sKG allows it to describe equally well the low-end, middle, and high-end ranges of the spectrum, while the deterministic method is limited to lower energies due to memory and CPU constraints. The computational effort of the method scales linearly with system size and is inversely proportional to the temperature, similarly to the sDFT calculations [17], while the deterministic approach has cubic scaling in both system size and temperature.

As an application of the method, we studied the conductivity and DOS for mixed hydrogen and helium systems at a constant volume and temperature ($T = 57$ kK) ensemble. We found that the system displays two conductivity phases,

where a transition from insulator to metal occurs at a hydrogen atomic fraction of $X_H \approx 0.3$.

The method enlarges the scope of sDFT to study properties of warm dense matter for very large systems at high temperatures. This could be significant when large inhomogeneous systems are studied or in systems where the mixing occurs on the nanoscale.

[1] B. Militzer, Phys. Rev. B **87**, 014202 (2013).

[2] W. Lorenzen, B. Holst, and R. Redmer, Phys. Rev. B **84**, 235109 (2011).

[3] M. Schöttler and R. Redmer, Phys. Rev. Lett. **120**, 115703 (2018).

[4] D. J. Stevenson, Phys. Rev. B **12**, 3999 (1975).

[5] N. Nettelmann, B. Holst, A. Kietzmann, M. French, R. Redmer, and D. Blaschke, Astrophys. J. **683**, 1217 (2008).

[6] T. Guillot, Science **286**, 72 (1999).

[7] P. L. Silvestrelli, A. Alavi, M. Parrinello, and D. Frenkel, Phys. Rev. Lett. **77**, 3149 (1996).

[8] T. R. Mattsson and G. Wahnstrom, Phys. Rev. B **56**, 14944 (1997).

[9] M. Pozzo, C. Davies, D. Gubbins, and D. Alfè, Nature (London) **485**, 355 (2012).

[10] B. B. L. Witte, P. Sperling, M. French, V. Recoules, S. H. Glenzer, and R. Redmer, Phys. Plasmas **25**, 056901 (2018).

[11] B. Holst, R. Redmer, and M. P. Desjarlais, Phys. Rev. B **77**, 184201 (2008).

[12] M. Preising, W. Lorenzen, A. Becker, R. Redmer, M. D. Knudson, and M. P. Desjarlais, Phys. Plasmas **25**, 012706 (2018).

[13] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

[14] S. Mazevet, M. Torrent, V. Recoules, and F. Jollet, High Energy Density Phys. **6**, 84 (2010).

[15] B. Holst, M. French, and R. Redmer, Phys. Rev. B **83**, 235120 (2011).

[16] M. P. Desjarlais, J. D. Kress, and L. A. Collins, Phys. Rev. E **66**, 025401(R) (2002).

[17] Y. Cytter, E. Rabani, D. Neuhauser, and R. Baer, Phys. Rev. B **97**, 115207 (2018).

[18] R. Baer, D. Neuhauser, and E. Rabani, Phys. Rev. Lett. **111**, 106402 (2013).

[19] D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **141**, 041102 (2014).

[20] Y. Cytter, D. Neuhauser, and R. Baer, J. Chem. Theory Comput. **10**, 4317 (2014).

[21] M. D. Fabian, B. Shpiro, E. Rabani, D. Neuhauser, and R. Baer, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **9**, e1412 (2019).

[22] M. Chen, R. Baer, D. Neuhauser, and E. Rabani, J. Chem. Phys. **150**, 034106 (2019).

[23] Y. Gao, D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Phys. **142**, 034106 (2015).

[24] D. Neuhauser, R. Baer, and D. Zgid, J. Chem. Theory Comput. **13**, 5396 (2017).

[25] S. Hernandez, Y. Xia, V. Vlček, R. Boutelle, R. Baer, E. Rabani, and D. Neuhauser, Mol. Phys. **116**, 2506 (2018).

[26] T. Y. Takeshita, W. A. de Jong, D. Neuhauser, R. Baer, and E. Rabani, J. Chem. Theory Comput. **13**, 4605 (2017).

[27] L.-W. Wang, Phys. Rev. B **49**, 10154 (1994).

[28] R. Baer, T. Seideman, S. Ilani, and D. Neuhauser, J. Chem. Phys. **120**, 3387 (2004).

[29] T. Iitaka, S. Nomura, H. Hirayama, X. W. Zhao, Y. Aoyagi, and T. Sugano, Phys. Rev. E **56**, 1222 (1997).

[30] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).

[31] D. A. Greenwood, Proc. Phys. Soc. **71**, 585 (1958).

[32] M. F. Hutchinson, Commun. Stat. Simul. Comput. **19**, 433 (1990).

[33] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[34] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[35] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[36] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).

[37] L. Calderín, V. V. Karasiev, and S. B. Trickey, Comput. Phys. Commun. **221**, 118 (2017).

[38] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[39] J. Hafner, Comput. Phys. Commun. **177**, 6 (2007).

[40] G. Kresse and J. Furthmuller, Phys. Rev. B **54**, 11169 (1996).

[41] W. Kuhn, Z. Phys. **33**, 408 (1925).

[42] F. Reiche and W. Thomas, Z. Phys. B **34**, 510 (1925).

[43] R. Kosloff, J. Phys. Chem. **92**, 2087 (1988).

[44] S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).

[45] Y. H. Huang, D. J. Kouri, and D. K. Hoffman, Chem. Phys. Lett. **243**, 367 (1995).

[46] R. Baer and M. Head-Gordon, J. Chem. Phys. **107**, 10003 (1997).

[47] The Chebyshev recursion is $\phi_{m+1} = 2\hat{h}_N \phi_m - \phi_{m-1}$, with $\phi_0 = \varphi$ and $\phi_1 = \hat{h}_N \phi_0$, where $\hat{h}_N = \frac{\hat{h} - \bar{E}}{\Delta E}$, $\bar{E} = \frac{1}{2}(E_{max} + E_{min})$ and $\Delta E = \frac{1}{2}(E_{max} - E_{min})$.

[48] R. Baer and M. Head-Gordon, Phys. Rev. Lett. **79**, 3962 (1997).

[49] A. Jaramillo-Botero, J. Su, A. Qi, and W. A. Goddard, J. Comput. Chem. **32**, 497 (2010).

[50] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[51] H. Kim, J. T. Su, and W. A. Goddard, Proc. Natl. Acad. Sci. USA **108**, 15101 (2011).

[52] J. T. Su and W. A. Goddard, Phys. Rev. Lett. **99**, 185003 (2007).

[53] W. Lorenzen, B. Holst, and R. Redmer, Phys. Rev. Lett. **102**, 115701 (2009).