Generalized Green's function molecular dynamics for canonical ensemble simulations

V. R. Coluci,[1] S. O. Dantas,[2] and V. K. Tewary[3]
[1]*School of Technology, University of Campinas–UNICAMP, Limeira, 13484-332 SP, Brazil*
[2]*Departamento de Física, ICE, Universidade Federal de Juiz de Fora, 36036-330 Juiz de Fora MG, Brazil*
[3]*Applied Chemical and Materials Division, National Institute of Standards and Technology, Boulder, Colorado 80305, USA*

(Received 19 December 2017; published 29 May 2018)

The need of small integration time steps (∼1 fs) in conventional molecular dynamics simulations is an important issue that inhibits the study of physical, chemical, and biological systems in real timescales. Additionally, to simulate those systems in contact with a thermal bath, thermostating techniques are usually applied. In this work, we generalize the Green's function molecular dynamics technique to allow simulations within the canonical ensemble. By applying this technique to one-dimensional systems, we were able to correctly describe important thermodynamic properties such as the temperature fluctuations, the temperature distribution, and the velocity autocorrelation function. We show that the proposed technique also allows the use of time steps one order of magnitude larger than those typically used in conventional molecular dynamics simulations. We expect that this technique can be used in long-timescale molecular dynamics simulations.

DOI: 10.1103/PhysRevE.97.053310

I. INTRODUCTION

Molecular dynamics (MD) simulations represent an important tool to study chemical, biological, and physical systems. In classical MD, Newton's equations of motion are numerically integrated for a system of interacting particles. To accurately describe the system evolution, typically small time steps (∼1 fs) on MD integration methods are necessary. Such small time steps pose a problem in simulating phenomena that occur at timescales of interest, which range from ns to ms, because the required number of integration steps is large ($\sim 10^{12}$), which may lead to a significant accumulation of numerical errors. Different strategies have been proposed to overcome the present timescale problem, each with limited success (for a review and other references, see Refs. [1–3]). These strategies are aimed at improving the temporal convergence of MD by using multiple time steps, including higher-order terms in time, computationally efficient integration algorithms, and/or by making simplifying assumptions applicable to specific systems being modeled.

A different scheme based upon the use of causal Green's functions has been developed by Tewary [4] to integrate the equations of motions in the microcanonical ensemble. In this scheme, named Green's function molecular dynamics (GFMD), the potential energy is expanded up to quadratic terms in atomistic displacements, in contrast to the conventional MD in which only the linear terms are retained. The advantage of the GFMD is that the equations of motions can still be solved exactly for arbitrary time steps. The GFMD has been shown to bridge the timescales from fs to μs in problems involving pulse propagation in one-dimensional lattices and ripples in graphene. However, despite the computational efficiency of the GFMD, it has not been applied to real problems. This may be because GFMD has no provision for including a thermostat. For practical MD applications, simulations within the canonical ensemble are necessary in order to model real systems that interact with thermal baths. This is precisely the need that we address in this paper.

In this work, we present a generalized formalism of the GFMD method based upon Langevin's equation. We illustrate our generalized GFMD method by applying it to simulate systems within the canonical ensemble. We show that our method can correctly describe their main thermodynamic properties with results comparable to established thermostating techniques. This capability of the GFMD, combined with its excellent temporal convergence, should make it very suitable for realistic simulations of discrete molecular systems.

Langevin's equation for Brownian motion [5] is used in the formulation of Langevin's thermostat [6]. In this thermostating technique, the degrees of freedom of the bath are omitted and a set of stochastic differential equations are used. Many integration algorithms have been proposed [7–14] to describe stochastic differential equations [15], and they have been compared among different aspects [16–21]. These algorithms include schemes from van Gunsteren and Berendsen [7], Brünger, Brooks, and Karplus [8], Skeel and Izaguirre ("Langevin impulse") [9], Mannella [10,11], and Euler, Heun, and Milstein [12,13]. Recently, Grønbech-Jensen and Farago proposed a Verlet-type algorithm that produces correct statistical-mechanics properties for large ensembles by exactly implementing the fluctuation-dissipation relationship in discrete time [14]. This method has been tested for simple [14] and complex [22–24] systems, and it has been incorporated in popular molecular dynamics simulation suites [23,24].

II. FORMALISM

Consider a system with N interacting particles of mass m_i $(i = 1, \ldots, N)$ in a thermal bath of temperature T. The instantaneous Newtonian force on the particle i is given by

$$m_i \frac{\partial^2 u_{i,\alpha}(t)}{\partial t^2} = -\frac{\partial V}{\partial u_{i,\alpha}} - \frac{m_i}{\tau_i}\frac{\partial u_{i,\alpha}(t)}{\partial t} + \eta_{i,\alpha}(t), \qquad (1)$$

where V is the potential energy, $u_{i,\alpha}$ is the α Cartesian component of the displacement of the particle i with respect

to its equilibrium position, and τ_i is the characteristic viscous damping time for the atom i (with friction coefficient $\xi_i \equiv m_i/\tau_i$). The α component of the random force on the particle i, $\eta_{i,\alpha}(t)$, is described by a white noise and through the fluctuation-dissipation theorem

$$\langle \eta_{i\alpha}(t)\rangle_\eta = 0,$$
$$\langle \eta_{i\alpha}(t)\eta_{j\beta}(t')\rangle_\eta = 2\xi_i k_b T \delta_{ij}\delta_{\alpha\beta}\delta(t-t'), \quad (2)$$

where k_b is Boltzmann's constant and $\langle\cdots\rangle_\eta$ is the average over many realizations of the random function $\eta(t)$. The nonanalytic behavior of $\eta(t)$ imposes difficulties to invent methods that are accurate for higher orders of the time step because additional Taylor terms, commonly used in integration methods for deterministic differential equations, produce no significant accuracy improvement [14–16].

We expand V as a Taylor series in atomic displacements as given below in the symbolic matrix notation:

$$V = V_0 - Fu + \tfrac{1}{2}u^T\phi u + \Delta V(u). \quad (3)$$

The constant term V_0 is taken to be zero without loss of generality. The first Taylor coefficient gives the force term $F_{i,\alpha}(t)$, the quadratic term is the force-constant term $\phi_{u_{i,\alpha},u_{j,\beta}} = \frac{\partial^2 V(u_{0i})}{\partial u_{i,\alpha}\partial u_{j,\beta}}$, $u_{0i,\alpha}$ is the displacement at $t=0$, and $\Delta V(u)$ includes $O(u^3)$ and higher-order nonlinear terms. In the GFMD method, we define an effective force term that includes the first force term and the nonlinear terms $O(u^3)$ as follows:

$$F_{\text{eff}\,i,\alpha}(t) \equiv F_{i,\alpha}(t) + \Delta F_{i,\alpha}(t), \quad (4)$$

where

$$\Delta F_{i,\alpha}(t) = -\frac{\partial \Delta V(u_{0i})}{\partial u_{i,\alpha}}. \quad (5)$$

The first-order term F in Eq. (3) is the first term in Taylor's expansion of the crystal potential energy in the Born–von Karman model. Physically, it represents the negative of the force on the atom i. This term is zero for a lattice in equilibrium, that is, at the potential minimum. If the lattice at equilibrium is subjected to a perturbation, such as by introducing a defect or by displacing an atom by applying mechanical or thermal stress, two things happen—a force is induced on each atom, and the second- and higher-order terms (the force constants) change. The effect of defects has been discussed in detail in Refs. [25,26]. In our model, based upon the Born–von Karman model, the inverse of the second term is the Green's function, which gives the response of the lattice. The force term and the higher-order terms give the lattice distortion or the strains in the perturbed lattice. If the force term is zero, there is no lattice distortion. This ideal case corresponds to an infinite perfect lattice, which is in stable equilibrium in the absence of external forces. The objective is to calculate the lattice distortion caused by the induced forces in the perturbed lattice.

The force term, which we treat as an effective force, includes linear and all nonlinear terms in the potential except the quadratic term. The effective force consists of the linear term $F_{i,\alpha}$ and $\Delta F_{i,\alpha}$, which formally contains the remaining nonlinear terms. In a pure harmonic model, $\Delta F_{i,\alpha}$ is zero and F is constant. In the general anharmonic case, $\Delta F_{i,\alpha}$ is a function of u and therefore t. The quadratic term is the force constant matrix, and its inverse gives the Green's function.

The effective force, which contains linear as well as nonlinear terms, is treated by iteration [4]. At each iteration, we calculate the effective force in terms of the displacements obtained in the previous iteration. This force is $F_{i,\alpha} + \Delta F_{i,\alpha}$. Effectively, it amounts to calculating $\Delta F_{i,\alpha}$ using the displacements obtained in the previous iteration and absorbing into $F_{\text{eff}\,i,\alpha}$. The effective force is then treated to be constant during the new iteration. The displacements in the new iteration are calculated by using the Green's function obtained from the inverse of the quadratic term. The iterations continue until the effective force term is zero, that is, less than a prescribed value. For a strictly harmonic nonthermal case, the assumption of constant force is exact and no iterations are necessary.

This iteration technique is similar to the conventional MD in the sense that in any iteration the energy is calculated in terms of the displacements calculated in the previous iteration. The main difference between the conventional MD and GFMD is that in the GFMD we retain terms up to the quadratic term in the potential at each iteration, whereas in the conventional MD only the first-order term is retained. The inclusion of the second-order term improves the convergence of the GFMD by several orders of magnitude [4].

Using the transformation $u_{i,\alpha}(t) \to \sqrt{m_i}u_{i,\alpha}(t)$, Eq. (1) can be written in the following matrix form (resembling the system's set of stochastic differential equations):

$$\left(\mathbb{I}\frac{\partial^2}{\partial t^2} + \mathbb{X}\frac{\partial}{\partial t} + \mathbb{D}\right)\mathbb{U} = \mathbb{F}(t) + \Delta\mathbb{F}(t) + \mathbb{F}^R(t), \quad (6)$$

where $\mathbb{U}$, $\mathbb{F}$, $\Delta\mathbb{F}$, and $\mathbb{F}^R$ are vectors whose components are $\sqrt{m_i}u_{i,\alpha}$, $(1/\sqrt{m_i})F_{i,\alpha}$, $(1/\sqrt{m_i})\Delta F_{i,\alpha}$, and $(1/\sqrt{m_i})\eta_{i,\alpha}$, respectively. $\mathbb{X}$ is a diagonal matrix with the elements $1/\tau_i$, $\mathbb{D}$ is a square matrix with the elements $(1/\sqrt{m_i m_j})\phi_{i,\alpha,j,\beta}$ (the so-called system's dynamical matrix), and $\mathbb{I}$ is the unit matrix.

The formal solution of Eq. (6) is

$$\mathbb{U}(t) = \left(\mathbb{I}\frac{\partial^2}{\partial t^2} + \mathbb{X}\frac{\partial}{\partial t} + \mathbb{D}\right)^{-1}\mathbb{F}_{\text{eff}}(t), \quad (7)$$

where $\mathbb{F}_{\text{eff}}(t) \equiv \mathbb{F}(t) + \Delta\mathbb{F}(t) + \mathbb{F}^R(t)$.

The inverse operator in Eq. (7) is the causal Green's function $\mathbb{G}(t-t')$, defined as a solution for $t > t'$, of

$$\left(\mathbb{I}\frac{\partial^2}{\partial t^2} + \mathbb{X}\frac{\partial}{\partial t} + \mathbb{D}\right)\mathbb{G}(t-t') = \mathbb{I}\delta(t-t'). \quad (8)$$

The Laplace transform $\mathcal{L}$ of Eq. (8) is

$$[s^2\mathbb{I} + s\mathbb{X} + \mathbb{D}]\mathcal{L}[\mathbb{G}] - (s\mathbb{I} + \mathbb{X})\mathbb{G}(0) - \mathbb{G}'(0)$$
$$= \mathbb{I}\mathcal{L}[\delta(t-t')], \quad (9)$$

where s is the Laplace variable conjugate to t. For the initial conditions $\mathbb{G}(0) = \mathbb{G}'(0) = 0$, the Laplace transform of the Green's function takes the form

$$\mathcal{L}[\mathbb{G}] = \frac{\mathbb{I}}{s^2\mathbb{I} + s\mathbb{X} + \mathbb{D}}. \quad (10)$$

Therefore, the Laplace transform of Eq. (6) is

$$[s^2\mathbb{I} + s\mathbb{X} + \mathbb{D}]\mathcal{L}[\mathbb{U}] - (s\mathbb{I} + \mathbb{X})\mathbb{U}(0) - \mathbb{C}(0) = \mathcal{L}[\mathbb{F}_{\text{eff}}]$$

or

$$\mathcal{L}[\mathbb{U}] = \mathcal{L}[\mathbb{G}]\mathcal{L}[\mathbb{F}_{\text{eff}}] + \mathcal{L}[\mathbb{H}]\mathbb{U}(0) + \mathcal{L}[\mathbb{G}]\mathbb{C}(0), \quad (11)$$

where

$$\mathcal{L}[\mathbb{H}] \equiv \frac{s\mathbb{I}+\mathbb{X}}{s^2\mathbb{I}+s\mathbb{X}+\mathbb{D}}. \tag{12}$$

Because $\mathbb{D}$ is real and symmetric, we can write $\mathbb{V}^T\mathbb{D}\mathbb{V} = \mathbb{E}^2$, where $\mathbb{V}$ is an orthogonal matrix with the eigenvectors of $\mathbb{D}$, and $\mathbb{E}^2$ is the diagonal matrix with the eigenvalues E_i^2 of $\mathbb{D}$. Multiplying Eq. (11) by $\mathbb{V}^T$ and using $\mathbb{I} = \mathbb{V}\mathbb{V}^T$, Eq. (11) is written as

$$\mathbb{U}^{L,*} = \mathbb{K}_G\,\mathbb{F}_{\rm eff}^{L,*} + \mathbb{K}_H\,\mathbb{U}^{L,*}(0) + \mathbb{K}_G\,\mathbb{C}^{L,*}(0), \tag{13}$$

where $\mathbb{K}_G \equiv \mathbb{V}^T\mathcal{L}[\mathbb{G}]\mathbb{V}$, $\mathbb{K}_H \equiv \mathbb{V}^T\mathcal{L}[\mathbb{H}]\mathbb{V}$, $\mathbb{U}^{L,*} \equiv \mathbb{V}^T\mathcal{L}[\mathbb{U}]$, $\mathbb{F}_{\rm eff}^{L,*} \equiv \mathbb{V}^T\mathcal{L}[\mathbb{F}_{\rm eff}]$, $\mathbb{U}^{L,*}(0) \equiv \mathbb{V}^T\mathbb{U}(0)$, and $\mathbb{C}^{L,*}(0) \equiv \mathbb{V}^T\mathbb{C}(0)$, with the superscript L standing for Laplace transform, and $*$ standing for the multiplication by $\mathbb{V}^T$.

Finally, the inverse Laplace transform of Eq. (13) is

$$\mathbb{U}^* = \mathcal{L}^{-1}[\mathbb{K}_G\,\mathbb{F}_{\rm eff}^{L,*}] + \mathcal{L}^{-1}[\mathbb{K}_H\,\mathbb{U}^{L,*}(0)] + \mathcal{L}^{-1}[\mathbb{K}_G\,\mathbb{C}^{L,*}(0)]. \tag{14}$$

If we assume the same viscous damping time τ for all atoms, then $\mathbb{X} = \tau^{-1}\mathbb{I}$, and Eq. (14) can be simplified to

$$\mathbb{U}^* = \mathcal{L}^{-1}[\mathbb{K}_G\,\mathbb{F}_{\rm eff}^{L,*}] + \mathcal{L}^{-1}[(s+\tau^{-1})\mathbb{K}_G\,\mathbb{U}^{L,*}(0)] + \mathcal{L}^{-1}[\mathbb{K}_G\,\mathbb{C}^{L,*}(0)], \tag{15}$$

where $\mathbb{K}_G$ is now a diagonal matrix with elements $K_i = (s^2 + s/\tau + E_i^2)^{-1}$. The elements of Eq. (15) are written as (see Appendix A for the derivations)

$$U_i^*(t) = \frac{\left(F_i^* + F_i^{R*}\right)}{E_i^2}\left[1 - \frac{e^{-t/(2\tau)}}{2\omega_i\tau}(2\tau\dot{S}(t) + S(t))\right] + \frac{C_i^*(0)S(t)}{\omega_i}e^{-t/(2\tau)} + \frac{U_i^*(0)}{2\omega_i\tau}(2\tau\dot{S}(t) + S(t))e^{-t/(2\tau)}, \tag{16}$$

where $S(t) \equiv \sin(\omega_i t)$ with $\omega_i \equiv (2\tau)^{-1}\sqrt{4E_i^2\tau^2 - 1}$ for the case in which $4E_i^2\tau^2 > 1$, and $S(t) \equiv \sinh(\omega_i t)$ with $\omega_i \equiv (2\tau)^{-1}\sqrt{1 - 4E_i^2\tau^2}/2\tau$ for $4E_i^2\tau^2 < 1$, and

$$U_i^*(t) = \left(F_i^* + F_i^{R*}\right)\tau^2\left[4 - 2e^{-t/(2\tau)}\left(2 + \frac{t}{\tau}\right)\right] + C_i^*(0)te^{-t/(2\tau)} + U_i^*(0)e^{-t/(2\tau)}\left(1 + \frac{t}{2\tau}\right) \tag{17}$$

for the case in which $4E_i^2\tau^2 = 1$. The velocities of the particles are obtained from $U_i^*(t)$ by $C_i^*(t) = \frac{dU_i^*(t)}{dt}$. If time is discretized in time steps Δt as $t_n = n\Delta t$, then the displacement (and velocity) $U_i(t+\Delta t)$ depends only on the displacement at the previous time $U_i(t)$. Equations (16) and (17) are the extended GFMD (e-GFMD) equations for canonical ensemble simulations. The original e-GFMD equations for the microcanonical ensemble [4] can be obtained from Eq. (16) for the limit $\tau \to \infty$.

Note that in deriving Eqs. (16) and (17), we have assumed that E_i^2 is real and positive. This is, of course, the stability condition of the solid, which requires the dynamical matrix to be positive-definite. Hence, the treatment given here is applicable to only stable solids that can sustain phonons, such that the atoms vibrate about their positions of equilibrium but do not move away. If an eigenvalue is negative, it would indicate unstable phonons. This would indicate displacement of atoms away from their position of equilibrium such as in diffusion. Use of GFMD for modeling such processes will be the subject of a later paper.

III. RESULTS

To validate the e-GFMD, we applied it to study the one-dimensional atomic chain comprised of N atoms of mass $m = 12$ amu initially separated by a distance $l_0 = 1$ Å and located at positions $x_i(t) = x_{0i} + u_i(t)$ ($x_{0i} = [-(N-1)/2 + (i-1)]l_0$ for $i = 1, \ldots, N$). The extremities of the chain are kept fixed, and only the interaction between neighboring atoms was considered.

Although the case of equal masses is not general, Eq. (15) can be applied to a large range of real solids including one-, two-, and three-dimensional monatomic systems. Examples include conventional metals such as Na, K, and Cu; semiconductors like silicon, germanium, and diamond; as well as modern advanced 1D materials like carbon and boron nanotubes and 2D ones like graphene, silicene, germanene, and phosphorene [3]. In addition, 1D systems are also of topical interest because of their relevance to soft matter applications. For example, many polymers and DNA molecules can be modeled as chain structures (see, for example, Refs. [27–32]). Of course, in these systems the molecules have additional degrees of freedom, but their underlying structure is essentially 1D, to which the GFMD method should be applicable. Further, it is quite straightforward to extend the present 1D method to 2D and 3D systems.

The interaction between the atoms separated by x was described by Morse's potential $V(x) = V_0\{\exp[-2\gamma(x-l_0)] - 2\exp[-\gamma(x-l_0)]\}$ with $V_0 = 1.724\,35$ eV and $\gamma = 1.0$ Å^{-1}. For this set of parameters, the nearest-neighbor harmonic force constant κ is 55.25 N/m.

We first tested e-GFMD for $\tau \to \infty$ to reproduce the behavior within the microcanonical ensemble. For this test, we set all the initial velocities to zero and the initial displacements to zero except the one of the central atom, which was chosen to be $d_0 = 10^{-4}$ Å. Such a configuration allows us to compare the system behavior with the exact one available for the atom displacements for a linear chain within the harmonic approximation [4]. For comparison, we also calculated the displacement evolution with the velocity-Verlet method [33]. Both velocity-Verlet and e-GFMD are able to reproduce the exact behavior of the system for a time step of 1 fs (not shown), but only e-GFMD reproduces the correct behavior for a larger time step of 10 fs (Fig. 1). This is expected since GFMD provides exact results in the harmonic approximation. Also, e-GFMD exhibits excellent energy conservation with relative energy variations 100 times smaller than the ones obtained from the velocity-Verlet method (inset of Fig. 1).

We then applied e-GFMD to the atomic chain for the canonical ensemble (finite τ). For the calculations on the canonical ensemble, we used two reference methods for comparison. The first one was the frequently used Brünger-Brooks-Karplus (BBK) method [8] and the second one was the recently developed Grønbech-Jensen-Farago (G-JF) thermostat [14].

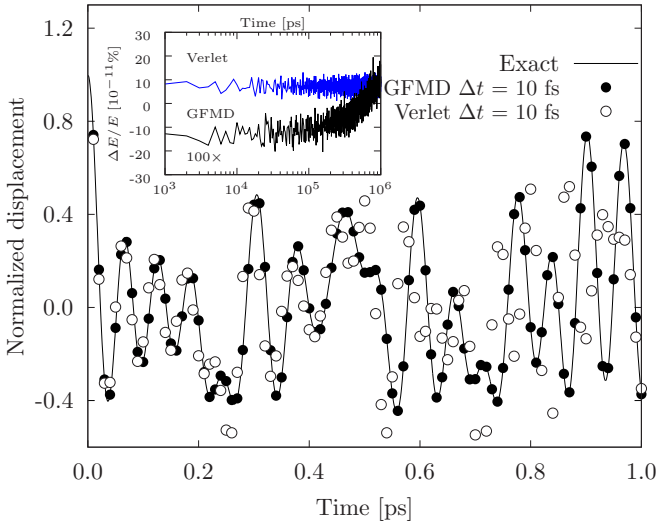

FIG. 1. Evolution of the normalized displacement of the central atom $[u_0(t)/d_0]$ obtained from the e-GFMD and velocity-Verlet methods within the microcanonical ensemble. The exact solution for a linear chain with 23 atoms is also shown for comparison [4]. The inset shows the total energy variation, normalized by the initial total energy, for a time step of 1 fs. Data from e-GFMD are scale up 100 times to a better visualization.

As in Langevin's thermostating techniques, the damping time (τ) plays a crucial role in temperature control and it has to be carefully chosen since each simulated system has specific behaviors. Unappropriated choices of τ can result in average temperatures that may significantly differ from the target one. To establish the appropriate range of Δt and τ that allows temperature control, we determine the dependence of the kinetic temperature $T_k = \sum m_i v_i^2/(Nk_b)$ on damping time for e-GFMD (Fig. 2) and for BBK and G-JF (inset of Fig. 2)

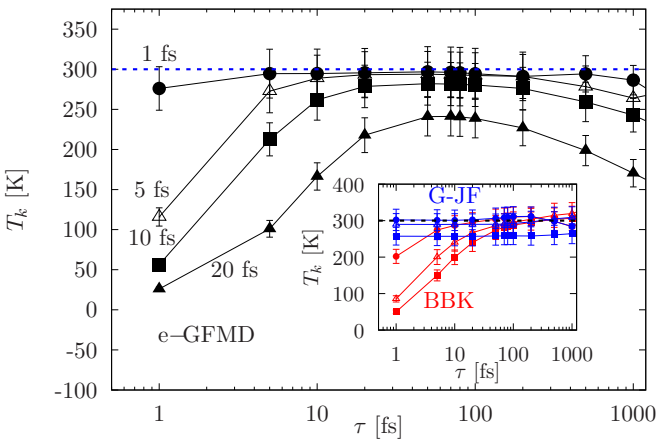

FIG. 2. Average kinetic temperature (symbols) and temperature standard deviation (error bars) as a function of the damping time for different time steps ($N = 203$) for e-GFMD. The equilibration time was 2000 steps and the average as well as the standard deviation were calculated for the following 1000 steps. The blue horizontal line indicates the target temperature. The inset shows the behavior for the BBK (red) and G-JF (blue) methods. The time steps are indicated with the same symbols used for e-GFMD.

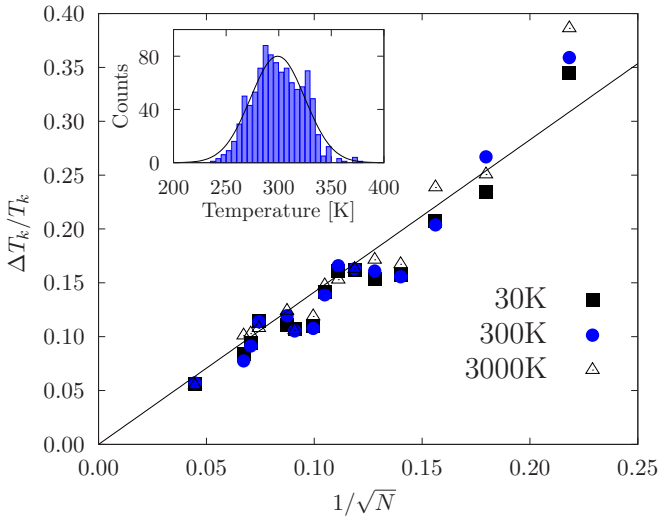

FIG. 3. Relative kinetic temperature fluctuation $\Delta T_k/T_k$ as a function of the size of the linear chain for different target temperatures T. The straight line corresponds to the expected behavior for the canonical ensemble ($\Delta T_k/T_k = \sqrt{2/N}$). For each chain, the system was equilibrated for 1 ns and the average was calculated for the following 1 ps ($\tau = 100$ fs, $\Delta t = 1$ fs). The inset shows the kinetic temperature distribution for the final 1 ps for a chain of 223 atoms compared with the Boltzmann distribution (black curve).

techniques. Whereas G-JF shows no dependence of T_k on the damping time, e-GFMD and BBK show similar behaviors. For our testing system, we found that e-GFMD is able to control the temperature for time steps up to 10 fs for a damping time of $\simeq$100 fs.

Besides the ability to control the temperature around the reservoir temperature T (target temperature), a thermostating technique should be able to reproduce the properties of the canonical ensemble. For example, temperature fluctuations $[(\Delta T_k)^2 = \langle T_k^2\rangle - \langle T_k\rangle^2]$ are expected to decrease with the increase of the system size as $\Delta T_k/T_k = \sqrt{2/N}$ for the canonical ensemble [6]. We found that e-GFMD correctly describes such behavior (Fig. 3) with a Gaussian distribution of the instantaneous kinetic temperature (inset Fig. 3). Another property is the distribution of the velocities that should follow the Maxwell-Boltzmann distribution. Figure 4 shows the resulting distributions of the velocities obtained by e-GFMD. Gaussian fittings show standard deviations of 4.6 ± 0.3 Å/ps (for $\Delta t = 1$ fs) and 4.4 ± 0.4 Å/ps (for $\Delta t = 10$ fs) that are in agreement with the expected value of 4.56 Å/ps for 300 K.

To further test e-GFMD, we calculate the normalized velocity autocorrelation function (VACF). The VACF is an important thermodynamic quantity because it can reveal the dynamical processes acting on the system and it can be used to calculate properties such as the phonon dispersion and the diffusion coefficient that can be directly compared with experiments. The VACF $C_{v_0}(\zeta)$ was calculated as [34] $C_{v_0}(\zeta) = \lim_{t\to\infty} \langle v_0(t+\zeta)v_0(t)\rangle/\langle v_0^2(t)\rangle$, where the $\langle\ \rangle$ is the average over different time intervals ζ, and r_0 and v_0 are the position and velocity of the central atom, respectively (see the Appendix B for the derivations and the final expression for VACF). To validate e-GFMD VACF results, we compared them

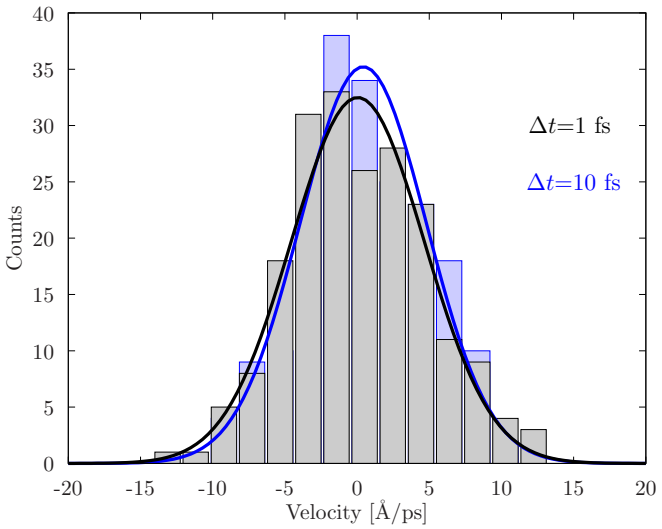

FIG. 4. Distributions of the velocities for a linear chain ($N = 203$, $T = 300$ K, $\tau = 100$ fs) for time steps of 1 and 10 fs after 10^5 time steps. The curves represent the Gaussian fittings.

with the exact VACF calculated for a linear chain of equal masses in contact with a thermal bath within the harmonic approximation (see Appendix B). The three methods e-GFMD, BBK, and G-JF were able to reproduce the exact VACF for time steps of 1 and 5 fs. However, when the time step is increased to 10 fs, the correct VACF behavior was only reproduced by e-GFMD (Fig. 5).

For estimating numerical convergence and accuracy in real solids for nonlinear terms in the potential energy, the limitation on the time step size in GFMD arises from the "space" part of the MD equations. As far as the time part is concerned, the time step can be arbitrarily large because the solution is exact in time. However, the time step should not be too large because the displacements during the time step in that iteration must be small enough so that the cubic and higher-order terms in displacements in the potential energy are negligible. This constraint, in general, is less severe than in the conventional MD [4].

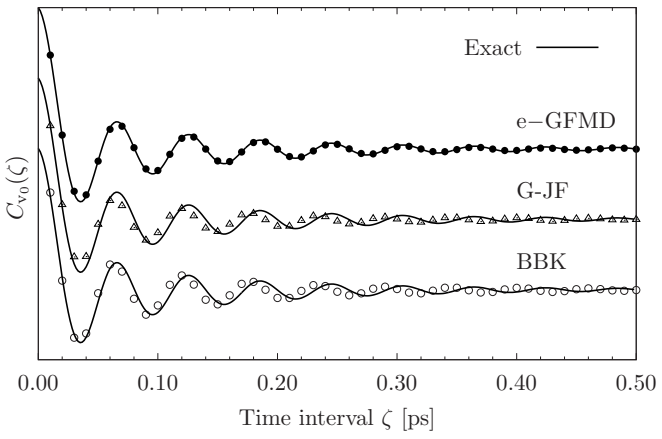

FIG. 5. Normalized velocity autocorrelation function obtained from the e-GFMD, BBK, and G-JF methods for $\Delta t = 10$ fs ($N = 203$, $\tau = 100$ fs, $T = 30$ K, and total simulation time of 10 ns).

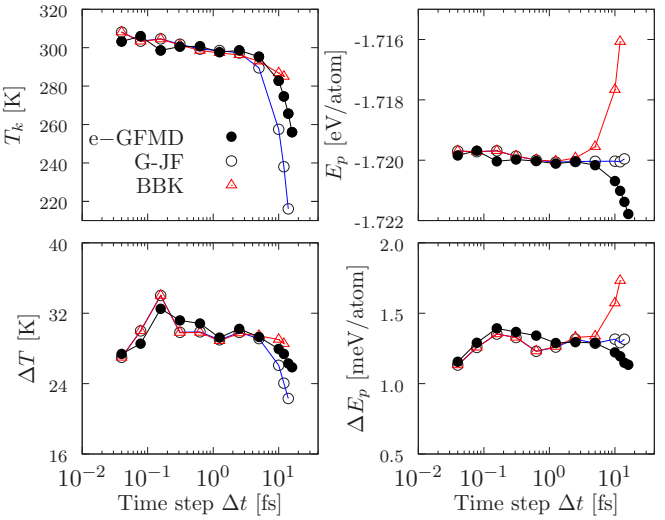

FIG. 6. Convergence of the kinetic temperature (T_k) and the potential energy (E_p) (and their respective standard deviations) with decreasing time step for e-GFMD, BBK, and G-JF methods. The properties were averaged for 25 000 time steps after an equilibration of 5×10^5 time steps for e-GFMD and 10^6 time steps for BBK, and G-JF methods ($N = 203$, $\tau = 100$ fs, and $T = 300$ K).

The convergence properties of e-GFMD with decreasing time step were analyzed by calculating the average values of the kinetic temperature and potential energy as a function of time step (Fig. 6). e-GFMD shows convergence of both kinetic temperature and potential energy with decreasing time step similar to the BBK and G-JF methods. Significant differences occur for time steps larger than 5 fs. The observed decreasing of T_k with decreasing time step is due to the fact that position and velocity coordinates are not exactly conjugated in discrete-time dynamics [24]. For the potential energy, the correct behavior is only reproduced by the G-JF method. This happens when the integrator shows the correct values of measured configurational thermodynamics quantities with increasing time step. For this case, the potential energy shows no dependence on the time step [23,24]. Therefore, in terms of accuracy, the overall convergence results for the kinetic temperature and potential energy indicate that e-GFMD lies between the G-JF and BBK methods.

A crude estimate of the computational advantage of the GFMD relative to conventional MD can be obtained as follows. Suppose u_c is the calculated value of u after one iteration and $\mathcal{N}_{\mathrm{MD}}$ iterations are needed to get the correct value u_a. Assuming linear convergence, $u_c \sim u_a/\mathcal{N}_{\mathrm{MD}}$. In MD the leading neglected term is $O(u^2)$. Hence the error in MD is $O(u^2) \sim O[1/(\mathcal{N}_{\mathrm{MD}})^2]$. Similarly the error in GFMD is $O(u^3) \sim O[1/(\mathcal{N}_{\mathrm{GFMD}})^3]$. Thus, to have the same error in MD and GFMD, $O[1/(\mathcal{N}_{\mathrm{MD}})^2] \sim O[1/(\mathcal{N}_{\mathrm{GFMD}})^3]$. Hence the estimated convergence improvement factor in GFMD, defined as $\mathcal{N}_{\mathrm{MD}}/\mathcal{N}_{\mathrm{GFMD}}$, is $(\mathcal{N}_{\mathrm{GFMD}})^{1/2} \sim 10^{3-4}$ for $\mathcal{N} \sim 10^{6-8}$. This obviously is a very attractive feature of the GFMD.

In addition to its computational advantage, GFMD has a very major physics advantage. The second-order terms in displacements define the phononic or elastic characteristics of the system [3]. These are neglected in the conventional

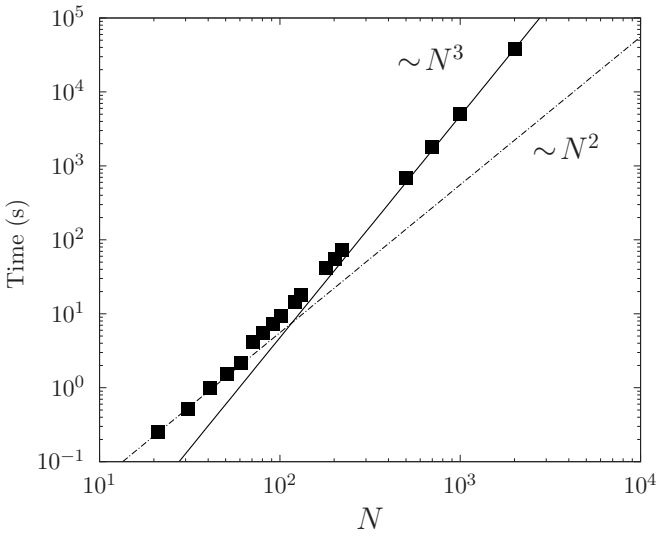

FIG. 7. Computation time of the e-GFMD method for 100 time steps ($\Delta t = 1$ fs, $\tau = 100$ fs, and $T = 300$ K). The simulations were carried out on a 3.3 GHz-IBM Power 755 processor. The curves for costs proportional to N^3 and N^2 are also indicated for comparison.

MD, which includes only the first-order term in displacements. Phonon frequencies and statistics are defined by the second-order term in the potential energy near a potential minimum at which the first-order term is exactly zero. This makes molecular dynamics a rather inefficient representation for phonons. The limitation of the conventional MD becomes even more severe considering that one has to represent the potential energy in a space of higher dimensions in which the energy may have local minima and saddle points, where the first-order terms will be vanishingly small. In our model, we include the second-order term in the potential energy that exactly determines the phonon frequency and the elastic response of the solid.

The use of e-GFMD may be computationally more expensive than the BBK and G-JF techniques in the conventional MD because e-GFMD involves the diagonalization of the dynamical matrix $\mathbb{D}$ for the whole lattice, which is a $O(N^3)$ process (Fig. 7). The computational cost should be partly offset because of the need for much fewer iterations. For example, if the potential energy can be expressed exactly in a quadratic form, only a single iteration is needed to obtain the displacement at time t from the displacement at time t_0. On the other hand, it is also possible to diagonalize the dynamical matrix for the whole lattice by using a single-particle approximation as in conventional MD. This would be an $O(N)$ process, but many more iterations will be needed even for harmonic crystals. Moreover, preliminary results [35] indicate that high-performance computing coupled with simple parallelism techniques can reduce the computational cost of GFMD.

Finally, we foresee different applications of e-GFMD beyond the classical MD. For example, it is shown that the convergence total energy calculations can be improved by the use of analytical integration techniques [36]. Thus, coupling e-GFMD with the Car-Parrinello method may improve total energy calculations for electronic structure calculations.

IV. CONCLUSIONS

In summary, we show that Green's function molecular dynamics can be generalized to allow thermostating and to describe systems within the canonical ensemble. The e-GFMD also exhibit high accuracy for dynamical properties even for larger time steps, making it a reliable alternative or, at least, complementary technique for long-timescale molecular dynamics simulations.

ACKNOWLEDGMENTS

We acknowledge the financial support from CNPq, UFJF, FAPEMIG, and FAPESP (Grants No. 2010/50646-6 and No. 2016/01736-9).

APPENDIX A: DERIVATION OF e-GFMD EQUATIONS

The inverse transform of each term on the right-hand side of Eq. (15),

$$\mathbb{U}^* = \mathcal{L}^{-1}\left[\mathbb{K}_G\, \mathbb{F}^{L,*}_{\text{eff}}\right] + \mathcal{L}^{-1}[(s+\tau^{-1})\mathbb{K}_G\, \mathbb{U}^{L,*}(0)] + \mathcal{L}^{-1}[\mathbb{K}_G\, \mathbb{C}^{L,*}(0)], \tag{A1}$$

is calculated as follows. For the first term, we have

$$\begin{aligned}\mathcal{L}^{-1}[K_i \mathcal{L}\,[F_{\text{eff}\,i}]] &= \mathcal{L}^{-1}\left[\left(\frac{1}{s^2+s/\tau+E_i^2}\right)F^*_{\text{eff}\,i}(s)\right] \\ &= \mathcal{L}^{-1}[F_1(s)F_2(s)] \\ &= \int_0^t f_1(\epsilon) f_2(t-\epsilon) d\epsilon. \end{aligned} \tag{A2}$$

For the case in which $F_2(s) = F^*_{\text{eff}\,i}(s) = F_i^*/s + F_i^{R*}/s$, where F_i^* is a constant, then $f_2(t)$ is the step function, i.e.,

$$f^*_{\text{eff}\,i}(t) = \begin{cases} F_i^* + F_i^{R*} & \text{if } t \geqslant 0, \\ 0 & \text{if } t < 0. \end{cases} \tag{A3}$$

Therefore,

$$\mathcal{L}^{-1}[K_i \mathcal{L}\,[F_{\text{eff}\,i}]] = \left(F_i^* + F_i^{R*}\right)\int_0^t \mathcal{L}^{-1}\left[\frac{1}{s^2+s/\tau+E_i^2}\right](\epsilon) d\epsilon. \tag{A4}$$

If we define $\omega_i \equiv \sqrt{\frac{1}{4\tau^2} - E_i^2}$, then

$$\mathcal{L}^{-1}\left[\frac{1}{s^2+s/\tau+E_i^2}\right] = e^{-t/(2\tau)}\left(\frac{e^{\omega_i t} - e^{-\omega_i t}}{2\omega_i}\right). \tag{A5}$$

We have three possibilities for (A5), depending on the value of τ and E_i. If $\frac{1}{4\tau^2} > E_i^2$, then

$$\mathcal{L}^{-1}\left[\frac{1}{s^2+s/\tau+E_i^2}\right] = \frac{e^{-t/(2\tau)}}{\omega}\sinh \omega_i t. \tag{A6}$$

If $\frac{1}{4\tau^2} < E_i^2$, then $\omega_i' \equiv j\omega_i$ ($j = \sqrt{-1}$) and

$$\mathcal{L}^{-1}\left[\frac{1}{s^2+s/\tau+E_i^2}\right] = \frac{e^{-t/(2\tau)}}{\omega_i'}\sin \omega_i' t. \tag{A7}$$

And finally, if $\frac{1}{4\tau^2} = E_i^2$, then

$$\mathcal{L}^{-1}\left[\frac{1}{s^2+s/\tau+E_i^2}\right] = t e^{-t/(2\tau)}. \tag{A8}$$

Therefore, the first term $\mathcal{L}^{-1}[K_i\mathcal{L}\,[F_{\text{eff}\,i}]]$ has the following values:

$$\frac{\left(F_i^*+F_i^{R*}\right)}{E_i^2}\left[1-e^{-t/(2\tau)}\left(\cosh(\omega_i t)+\frac{\sinh(\omega_i t)}{2\omega_i\tau}\right)\right] \quad \text{(A9)}$$

if $\frac{1}{4\tau^2} > E_i^2$,

$$\frac{\left(F_i^*+F_i^{R*}\right)}{E_i^2}\left[1-e^{-t/(2\tau)}\left(\cos(\omega_i' t)+\frac{\sin(\omega_i' t)}{2\omega_i'\tau}\right)\right] \quad \text{(A10)}$$

if $\frac{1}{4\tau^2} < E_i^2$, and

$$\left(F_i^*+F_i^{R*}\right)\tau^2[4-2e^{-t/(2\tau)}(2+t/\tau)] \quad \text{(A11)}$$

if $\frac{1}{4\tau^2} = E_i^2$.

Using the same analysis, the second term on the right-hand side of Eq. (A1) is calculated as

$$\begin{aligned}
&\mathcal{L}^{-1}[(s+\tau^{-1})\,K_i U_i^*(0)] \\
&\quad=\mathcal{L}^{-1}\left[\frac{s+\tau^{-1}}{s^2+s/\tau+E_i^2}U_i^*(0)\right] \\
&\quad=\begin{cases} U_i^*(0)e^{-t/(2\tau)}\left[\frac{\sinh(\omega_i t)}{2\omega_i\tau}+\cosh(\omega_i t)\right] & \text{if } \frac{1}{4\tau^2} > E_i^2, \\ U_i^*(0)e^{-t/(2\tau)}\left[\frac{\sin(\omega_i' t)}{2\omega_i'\tau}+\cos(\omega_i' t)\right] & \text{if } \frac{1}{4\tau^2} < E_i^2, \\ U_i^*(0)e^{-t/(2\tau)}\left[1+\frac{t}{2\tau}\right] & \text{if } \frac{1}{4\tau^2} = E_i^2, \end{cases}
\end{aligned} \quad \text{(A12)}$$

and the third term as

$$\begin{aligned}
\mathcal{L}^{-1}[K_i C_i^*(0)] &= \mathcal{L}^{-1}\left[\frac{C_i^*(0)}{s^2+s/\tau+E_i^2}\right] \\
&= C_i^*(0)\frac{e^{-t/(2\tau)}}{\omega_i}\sinh(\omega_i t) \text{ if } \frac{1}{4\tau^2} > E_i^2, \\
&= C_i^*(0)\frac{e^{-t/(2\tau)}}{\omega_i'}\sin(\omega_i' t) \text{ if } \frac{1}{4\tau^2} < E_i^2, \\
&= C_i^*(0)te^{-t/(2\tau)} \text{ if } \frac{1}{4\tau^2} = E_i^2.
\end{aligned} \quad \text{(A13)}$$

With those terms, Eq. (A1) is written as

$$\begin{aligned}
U_i^*(t) &= \frac{\left(F_i^*+F_i^{R*}\right)}{E_i^2}\left[1-e^{-t/(2\tau)}\left(\cos(\omega_i' t)+\frac{\sin(\omega_i' t)}{2\omega_i'\tau}\right)\right] \\
&\quad + C_i^*(0)\frac{\sin(\omega_i' t)}{\omega_i^{\pm}}e^{-t/(2\tau)} \\
&\quad + U_i^*(0)\left(\cos(\omega_i' t)+\frac{\sin(\omega_i' t)}{2\omega_i'\tau}\right)e^{-t/(2\tau)}
\end{aligned} \quad \text{(A14)}$$

if $\frac{1}{4\tau^2} < E_i^2$,

$$\begin{aligned}
U_i^*(t) &= \frac{\left(F_i^*+F_i^{R*}\right)}{E_i^2}\left[1-e^{-t/(2\tau)}\left(\cosh(\omega_i t)+\frac{\sinh(\omega_i t)}{2\omega_i\tau}\right)\right] \\
&\quad + C_i^*(0)\frac{\sinh(\omega_i t)}{\omega_i^{\pm}}e^{-t/(2\tau)} \\
&\quad + U_i^*(0)\left(\cosh(\omega_i t)+\frac{\sinh(\omega_i t)}{2\omega_i\tau}\right)e^{-t/(2\tau)}
\end{aligned} \quad \text{(A15)}$$

if $\frac{1}{4\tau^2} > E_i^2$, and

$$\begin{aligned}
U_i^*(t) &= \left(F_i^*+F_i^{R*}\right)\tau^2\left[4-2e^{-t/(2\tau)}\left(2+\frac{t}{\tau}\right)\right] \\
&\quad + C_i^*(0)te^{-t/(2\tau)} + U_i^*(0)e^{-t/(2\tau)}\left(1+\frac{t}{2\tau}\right)
\end{aligned} \quad \text{(A16)}$$

if $\frac{1}{4\tau^2} = E_i^2$.

The velocities $C_i^*(t) = \frac{dU_i^*(t)}{dt}$ are written as

$$\begin{aligned}
C_i^*(t) &= \left(\frac{F_i^*+F_i^{R*}}{E_i^2}-U_i^*(0)\right)\left[1+\frac{1}{4\omega_i'^2\tau^2}\right]\omega_i'\sin(\omega_i' t) \\
&\quad \times e^{-t/(2\tau)} + C_i^*(0)\left[\cos(\omega_i' t)-\frac{\sin(\omega_i' t)}{2\omega_i'\tau}\right]e^{-t/(2\tau)}
\end{aligned} \quad \text{(A17)}$$

if $\frac{1}{4\tau^2} < E_i^2$,

$$\begin{aligned}
C_i^*(t) &= \left(\frac{F_i^*+F_i^{R*}}{E_i^2}-U_i^*(0)\right)\left[-1+\frac{1}{4\omega_i^2\tau^2}\right]\omega_i\sinh(\omega_i t) \\
&\quad \times e^{-t/(2\tau)} + C_i^*(0)\left[\cosh(\omega_i t)-\frac{\sinh(\omega_i t)}{2\omega_i\tau}\right]e^{-t/(2\tau)}
\end{aligned} \quad \text{(A18)}$$

if $\frac{1}{4\tau^2} > E_i^2$, and finally

$$\begin{aligned}
C_i^*(t) &= \left(F_i^*+F_i^{R*}\right)te^{-t/(2\tau)} + C_i^*(0)e^{-t/(2\tau)}\left(1-\frac{t}{2\tau}\right) \\
&\quad - U_i^*(0)\frac{te^{-t/(2\tau)}}{4\tau^2}
\end{aligned} \quad \text{(A19)}$$

if $\frac{1}{4\tau^2} = E_i^2$.

APPENDIX B: VELOCITY AUTOCORRELATION FUNCTION

Here we derive the exact velocity autocorrelation function for a linear chain of N (N odd) particles of the same mass with harmonic nearest-neighbor interactions in a thermal bath of temperature T. The particles are located at la, where $l=-(N-1)/2,\ldots,(N-1)/2$ and a is the initial separation between neighboring atoms. The velocity autocorrelation function for the central atom ($l=0$) of the chain C_{v_0} is given by

$$C_{v_0}(\zeta) = \lim_{t\to\infty}\frac{\langle v_0(t+\zeta)v_0(t)\rangle}{\langle v_0^2(t)\rangle}, \quad \text{(B1)}$$

where ζ is the delay time, $\langle v_0(t+\zeta)v_0(t)\rangle = (1/t)\int_0^t v_0(t'+\zeta)\,v_0(t')dt'$. The equation of motion for a harmonic linear chain is

$$\frac{d^2u_l(t)}{dt^2} = \frac{\kappa}{m}[u_{l+1}(t)-2u_l+u_{l-1}(t)] - \frac{1}{\tau}\frac{du_l(t)}{dt} + \frac{1}{m}\eta_l(t). \quad \text{(B2)}$$

For a linear chain with the atoms at the extremities [$l=-(N-1)/2$ and $l=(N-1)/2$] fixed, we write

$$u_l(t) = \frac{1}{N}\sum_{k_i}\cos[k_i la]u_{k_i}(t), \quad \text{(B3)}$$

$$\eta_l(t) = \frac{1}{N}\sum_{k_i}\cos[k_i la]\eta_{k_i}(t), \quad \text{(B4)}$$

where $k_i = (2i+1)\pi/(N-1)$, $i = -(N-1)/2, \ldots, (N-1)/2$. The values of k_i were obtained by imposing that the displacements for the atoms at the extremities are zero for all times, i.e., $\cos[(N-1)ka/2] = 0$.

Equation (B2) is written as

$$\frac{dv_{k_i}(t)}{dt} = -\omega_0^2(k_i)u_{k_i}(t) - \frac{1}{\tau}v_{k_i}(t) + \frac{1}{m}\eta_{k_i}(t) \tag{B5}$$

and

$$\frac{du_{k_i}(t)}{dt} = v_{k_i}(t), \tag{B6}$$

where $\omega_0^2(k_i) = (2\kappa/m)[1-\cos(k_i)]$. The solutions of Eqs. (B5) and (B6) are

$$u_{k_i}(t) = \left[\cosh\left(\omega_{k_i}t\right) + \frac{1}{2\omega_{k_i}\tau}\sinh\left(\omega_{k_i}t\right)\right]e^{-t/2\tau}\ u_{k_i}(0) + \frac{1}{\omega_{k_i}}\sinh\left(\omega_{k_i}t\right)\ e^{-t/2\tau}\ v_{k_i}(0) + \frac{1}{m}\int_0^t \frac{1}{\omega_{k_i}}\sinh\left[\omega_{k_i}(t-t')\right]e^{-(t-t')/2\tau}\ \eta_{k_i}(t')dt' \tag{B7}$$

and

$$v_{k_i}(t) = \left(1 - \frac{1}{4w_{k_i}^2\tau^2}\right)\omega_{k_i}\sinh\left(\omega_{k_i}t\right)e^{-t/2\tau}\ u_{k_i}(0) + \left[\cosh\left(\omega_{k_i}t\right) - \frac{1}{2\omega_{k_i}\tau}\sinh\left(\omega_{k_i}t\right)\right]e^{-t/2\tau}\ v_{k_i}(0) + \frac{1}{m}\int_o^t \left[\cosh\left[\omega_{k_i}(t-t')\right] - \frac{1}{2\omega_{k_i}\tau}\sinh\left[w_{k_i}(t-t')\right]\right]e^{-(t-t')/2\tau}\ \eta_{k_i}(t')dt' \tag{B8}$$

if $2\omega_o(k_i)\tau < 1$ [$\omega_{k_i}^2 \equiv \omega_o^2(k_i) - \frac{1}{4\tau^2}$], and

$$u_{k_i}(t) = \left[\cos\left(\omega_{k_i}t\right) + \frac{1}{2\omega_{k_i}\tau}\sin\left(\omega_{k_i}t\right)\right]e^{-t/2\tau}\ u_{k_i}(0) + \frac{1}{\omega_{k_i}}\sin\left(\omega_{k_i}t\right)\ e^{-t/2\tau}\ v_{k_i}(0) + \frac{1}{m}\int_0^t \frac{1}{\omega_{k_i}}\sin\left[\omega_{k_i}(t-t')\right]e^{-(t-t')/2\tau}\ \eta_{k_i}(t')dt' \tag{B9}$$

and

$$v_{k_i}(t) = -\left(1 + \frac{1}{4w_{k_i}^2\tau^2}\right)\omega_{k_i}\sin\left(\omega_{k_i}t\right)e^{-t/2\tau}u_{k_i}(0) + \left[\cos\left(\omega_{k_i}t\right) - \frac{1}{2\omega_{k_i}\tau}\sin\left(\omega_{k_i}t\right)\right]e^{-t/2\tau}v_{k_i}(0) + \frac{1}{m}\int_0^t \left[\cos\left[\omega_{k_i}(t-t')\right] - \frac{\sin\left[w_{k_i}(t-t')\right]}{2\omega_{k_i}\tau}\right]e^{-(t-t')/2\tau}\eta_{k_i}(t')dt' \tag{B10}$$

if $2\omega_o(k_i)\tau > 1$ [$\omega_{k_i}^2 \equiv \frac{1}{4\tau^2} - \omega_o^2(k_i)$].

Thus,

$$v_l(t+\zeta)v_l(t) = \frac{1}{N^2}\sum_{k_i'}\cos[k_i' la]v_{k_i'}(t+\zeta)\sum_{k_i}\cos[k_i la]v_{k_i}(t). \tag{B11}$$

For the particular case of the central atom ($l = 0$), we have

$$v_0(t+\zeta)v_0(t) = \frac{1}{N^2}\sum_{k_i'}v_{k_i'}(t+\zeta)\sum_{k_i}v_{k_i}(t). \tag{B12}$$

If we define

$$A_{k_i}(t) \equiv -\left(1 + \frac{1}{4w_{k_i}^2\tau^2}\right)w_{k_i}\sin\left(w_{k_i}t\right)e^{-t/2\tau}u_{k_i}(0),$$

$$B_{k_i}(t) \equiv \left[\cos\left(w_{k_i}t\right) - \frac{\sin\left(w_{k_i}t\right)}{2w_{k_i}\tau}\right]e^{-t/2\tau}\ v_{k_i}(0),$$

$$I_{k_i}(t) \equiv \frac{1}{m}\int_0^t \left\{\cos\left[w_{k_i}(t-t')\right] - \frac{\sin\left[w_{k_i}(t-t')\right]}{2w_{k_i}\tau}\right\}e^{-(t-t')/2\tau}\eta_{k_i}(t')dt'$$

if $2\omega_o(k_i)\tau > 1$, and

$$A'_{k_i}(t) \equiv \left(1 - \frac{1}{4w_{k_i}^2\tau^2}\right) w_{k_i} \sinh\left(w_{k_i}t\right) e^{-t/2\tau}\, u_{k_i}(0),$$

$$B'_{k_i}(t) \equiv \left[\cosh\left(w_{k_i}t\right) - \frac{\sinh\left(w_{k_i}t\right)}{2w_{k_i}\tau}\right] e^{-t/2\tau} v_{k_i}(0),$$

$$I'_{k_i}(t) \equiv \frac{1}{m}\int_0^t \left\{\cosh\left[w_{k_i}(t-t')\right] - \frac{\sinh\left[w_{k_i}(t-t')\right]}{2w_{k_i}\tau}\right\} e^{-(t-t')/2\tau}\eta_{k_i}(t')dt'$$

if $2\omega_o(k_i)\tau < 1$, then Eq. (B12) is written as

$$v_0(t+\zeta)v_0(t) = \frac{1}{N^2}\sum_{k'_i,k_i}[A_{k'_i}(t+\zeta) + B_{k'_i}(t+\zeta) + I_{k'_i}(t+\zeta)] \times \left[A_{k_i}(t) + B_{k_i}(t) + I_{k_i}(t)\right] \tag{B13}$$

if $2\omega_o(k_i)\tau < 1$ and a similar expression for $2\omega_o(k_i)\tau > 1$. The time average is then written as

$$\lim_{t\to\infty}\langle v_0(t+\zeta)v_0(t)\rangle = \frac{1}{N^2}\sum_{k'_i,k_i}\sum_{p=1}^{9}\lim_{t\to\infty} T_p, \tag{B14}$$

where

$$T_1 = \left\langle A_{k'_i}(t+\zeta)A_{k_i}(t)\right\rangle,$$
$$T_2 = \left\langle A_{k'_i}(t+\zeta)B_{k_i}(t)\right\rangle,$$
$$T_3 = \left\langle A_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle,$$
$$T_4 = \left\langle B_{k'_i}(t+\zeta)A_{k_i}(t)\right\rangle,$$
$$T_5 = \left\langle B_{k'_i}(t+\zeta)B_{k_i}(t)\right\rangle,$$
$$T_6 = \left\langle B_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle,$$
$$T_7 = \left\langle I_{k'_i}(t+\zeta)A_{k_i}(t)\right\rangle,$$
$$T_8 = \left\langle I_{k'_i}(t+\zeta)B_{k_i}(t)\right\rangle,$$
$$T_9 = \left\langle I_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle.$$

When $2\omega_o(k_i)\tau > 1$ (and consequently $2\omega_i\tau > 1$), the terms T_1, T_2, T_4, and T_5 are zero because these terms involve integrals $I = (2\tau/t)\int_0^{t/2\tau} f(s)e^{-4s}ds$, where $f(s)$ are products involving $\sin(2\omega_i\tau s)$ and $\cos(2\omega'_i\tau s)$ [e.g., $\sin(2\omega_i\tau s)\cos(2\omega'_i\tau s)$], which vanish for $t\to\infty$. When $2\omega_o(k_i)\tau < 1$ (and $2\omega_i\tau < 1$), $f(s)$ involves combinations of hyperbolic sine and cosine functions [e.g., $\sinh(2\omega_i\tau s)\cosh(2\omega'_i\tau s)$] and I is also zero when $t\to\infty$.

By applying the ergodic theorem to replace the time average $\langle\cdots\rangle$ by the average over many realizations of the random function $\eta(t)$ (i.e., $\langle\cdots\rangle_\eta$), the terms T_3, T_6, T_7, and T_8 can also be shown to be zero. For example,

$$\begin{aligned} T_3 &= \left\langle A_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle = \left\langle A_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle_\eta \\ &= A_{k'_i}(t+\zeta)\left\langle I_{k_i}(t)\right\rangle_\eta \\ &= \frac{A_{k'_i}(t+\zeta)}{m}\int_0^t\left[\cos\left[\omega_{k_i}(t-t')\right] - \frac{\sin\left[\omega_{k_i}(t-t')\right]}{2\omega_{k_i}\tau}\right] \\ &\quad\times e^{-(t-t')/2\tau}\left\langle\eta_{k_i}(t')\right\rangle_\eta dt', \end{aligned} \tag{B15}$$

which is zero because $\langle\eta_l(t)\rangle_\eta = 0$. Therefore, the only term that is not zero is T_9:

$$\begin{aligned} T_9 &= \left\langle I_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle = \left\langle I_{k'_i}(t+\zeta)I_{k_i}(t)\right\rangle_\eta \\ &= \frac{1}{m^2}\int_0^{t+\zeta}\int_0^t\left[\cos\left[\omega_{k_i}(t-t'_2)\right] - \frac{\sin\left[\omega_{k_i}(t-t'_2)\right]}{2\omega_{k_i}\tau}\right]\left[\cos\left[\omega_{k'_i}(t+\zeta-t'_1)\right] - \frac{\sin\left[\omega_{k'_i}(t+\zeta-t'_1)\right]}{2\omega_{k'_i}\tau}\right] \\ &\quad\times e^{-(t+\zeta-t'_1)/2\tau}e^{-(t-t'_2)/2\tau}\left\langle\eta_{k'_i}(t'_1)\eta_{k_i}(t'_2)\right\rangle_\eta dt'_2dt'_1 \end{aligned} \tag{B16}$$

if $2\omega_o(k_i)\tau < 1$, and

$$\begin{aligned} T_9 &= \left\langle I'_{k'_i}(t+\zeta)I'_{k_i}(t)\right\rangle = \left\langle I'_{k'_i}(t+\zeta)I'_{k_i}(t)\right\rangle_\eta \\ &= \frac{1}{m^2}\int_0^{t+\zeta}\int_0^t\left[\cosh\left[\omega_{k_i}(t-t'_2)\right] - \frac{\sinh\left[\omega_{k_i}(t-t'_2)\right]}{2\omega_{k_i}\tau}\right]\left[\cosh\left[\omega_{k'_i}(t+\zeta-t'_1)\right] - \frac{\sinh\left[\omega_{k'_i}(t+\zeta-t'_1)\right]}{2\omega_{k'_i}\tau}\right] \\ &\quad\times e^{-(t+\zeta-t'_1)/2\tau}e^{-(t-t'_2)/2\tau}\left\langle\eta_{k'_i}(t'_1)\eta_{k_i}(t'_2)\right\rangle_\eta dt'_2dt'_1 \end{aligned} \tag{B17}$$

if $2\omega_o(k_i)\tau > 1$.

Using that

$$\begin{aligned}\left\langle \eta_{k_i'}(t_1')\eta_{k_i}(t_2')\right\rangle_\eta &= \sum_{l,l'} \cos[k_i' l' a]\cos[k_i l a]\langle \eta_{l'}(t_1')\eta_l(t_2')\rangle_\eta \\ &= \sum_{l,l'} \cos[k_i' l' a]\cos[k_i l a]\frac{2mk_bT}{\tau}\delta_{l,l'}\delta(t_1'-t_2') \\ &= \frac{2mk_bT}{\tau}\sum_{l} \cos[k_i' l a]\cos[k_i l a]\delta(t_1'-t_2') \\ &= \frac{2mk_bT}{\tau}\sum_{l}\{\cos[(k_i'-k_i)la] - \sin[k_i' l a]\sin[k_i l a]\}\delta(t_1'-t_2') \\ &= \frac{2Nmk_bT}{\tau}\delta(t_1'-t_2')\{\delta_{k_i'-k_i,0} - \sum_{l}\sin[k_i' l a]\sin[k_i l a]\}\end{aligned}$$

and

$$\sum_{l=-(N-1)/2}^{(N-1)/2} \sin[k_i' l a]\sin[k_i l a] = 0,$$

then

$$\left\langle \eta_{k_i'}(t_1')\eta_{k_i}(t_2')\right\rangle_\eta = \frac{2Nmk_bT}{\tau}\delta(t_1'-t_2')\delta_{k_i'-k_i,0}. \tag{B18}$$

Finally, applying the δ functions to $\sum_{k_i,k_i'}\lim_{t\to\infty} T_9$, this term reduces to

$$\begin{aligned}\sum_{k_i}\lim_{t\to\infty} T_9 = \frac{2Nk_bT}{m\tau}\sum_{k_i}\lim_{t\to\infty}\int_0^{\min[t+\zeta,t]} & e^{-\zeta/2\tau}e^{-(t-t')/\tau}\left\{\cos\left[\omega_{k_i}(t-t')\right] - \frac{\sin\left[\omega_{k_i}(t-t')\right]}{2\omega_{k_i}\tau}\right\} \\ &\times\left\{\left[\cos\left[\omega_{k_i}(\zeta+t-t')\right] - \frac{\sin\left[\omega_{k_i}(\zeta+t-t')\right]}{2\omega_{k_i}\tau}\right]\right\}dt' \\ = \frac{Nk_bT}{m}e^{-\zeta/2\tau}\sum_{k_i}&\left\{\cos\left[\omega_{k_i}\zeta\right] - \frac{\sin\left[\omega_{k_i}\zeta\right]}{2\omega_{k_i}\tau}\right\}\end{aligned} \tag{B19}$$

if $2\omega_o(k_i)\tau < 1$, and

$$\begin{aligned}\lim_{t\to\infty} T_9 = \frac{2Nk_bT}{m\tau}\lim_{t\to\infty}\int_0^{\min[t+\zeta,t]} & e^{-\zeta/2\tau}e^{-(t-t')/\tau}\left\{\cosh\left[\omega_{k_i}(t-t')\right] - \frac{\sinh\left[\omega_{k_i}(t-t')\right]}{2\omega_{k_i}\tau}\right\} \\ &\times\left\{\left[\cosh\left[\omega_{k_i}(\zeta+t-t')\right] - \frac{\sinh\left[\omega_{k_i}(\zeta+t-t')\right]}{2\omega_{k_i}\tau}\right]\right\}dt' \\ = \frac{Nk_bT}{m}e^{-\zeta/2\tau}\sum_{k_i}&\left\{\cosh\left[\omega_{k_i}\zeta\right] - \frac{\sinh\left[\omega_{k_i}\zeta\right]}{2\omega_{k_i}\tau}\right\}\end{aligned} \tag{B20}$$

if $2\omega_o(k_i)\tau > 1$. Therefore,

$$\lim_{t\to\infty}\langle v_0(t+\zeta)v_0(t)\rangle = \frac{1}{N^2}\sum_{k_i}\lim_{t\to\infty} T_9 = \frac{k_bT}{Nm}e^{-\zeta/2\tau}\sum_{k_i} f\left(\omega_{k_i},\zeta,\tau\right), \tag{B21}$$

where

$$f(\omega_{k_i},\zeta,\tau) = \cos\left[\omega_{k_i}\zeta\right] - \frac{\sin\left[\omega_{k_i}\zeta\right]}{2\omega_{k_i}\tau} \tag{B22}$$

if $2\omega_o(k_i)\tau < 1$, and

$$f(\omega_{k_i},\zeta,\tau) = \cosh\left[\omega_{k_i}\zeta\right] - \frac{\sinh\left[\omega_{k_i}\zeta\right]}{2\omega_{k_i}\tau} \tag{B23}$$

if $2\omega_o(k_i)\tau > 1$. Because

$$\lim_{t\to\infty}\langle v_0(t)^2\rangle = \frac{k_bT}{Nm}\sum_{k_i} f\left(\omega_{k_i},0,\tau\right) = \frac{k_bT}{Nm}\sum_{k_i} 1 = \frac{k_bT}{m}, \tag{B24}$$

we have

$$C_{v_0}(\zeta) = \frac{e^{-\zeta/2\tau}}{N}\sum_{k_i} f\left(\omega_{k_i},\zeta,\tau\right). \tag{B25}$$

For the general case of the atom l, Eq. (B25) becomes

$$C_{v_l}(\zeta) = \frac{e^{-\zeta/2\tau}}{\sum_{k_i}\cos^2[k_i la]}\sum_{k_i} f\left(\omega_{k_i},\zeta,\tau\right)\cos^2[k_i la]. \tag{B26}$$

[1] A. F. Voter, F. Montalenti, and T. C. Germann, Annu. Rev. Mater. Res. **32**, 321 (2002).
[2] D. C. Rapaport, *The Art of Molecular Dynamics Simulation* (Cambridge University Press, New York, 2004).
[3] *Modeling, Characterization and Production of Nanomaterials*, edited by V. K. Tewary and Y. Zhang (Elsevier, Amsterdam, 2015).
[4] V. K. Tewary, Phys. Rev. B **80**, 161409 (2009).
[5] P. Langevin, C. R. Acad. Sci. (Paris) **146**, 530 (1908).
[6] P. H. Hünenberger, Adv. Polym. Sci. **173**, 105 (2005).
[7] W. F. van Gunsteren and H. J. C. Berendsen, Mol. Phys. **45**, 637 (1982).
[8] A. Brünger, C. L. Brooks III, and M. Karplus, Chem. Phys. Lett. **105**, 495 (1984).
[9] R. D. Skeel and J. A. Izaguirre, Mol. Phys. **100**, 3885 (2002).
[10] R. Mannella, Phys. Rev. E **69**, 041107 (2004).
[11] R. Mannella, SIAM J. Sci. Comput. **27**, 2121 (2006).
[12] G. N. Milstein, Theory Probab. Appl. **19**, 557 (1975).
[13] G. N. Milstein, Theory Probab. Appl. **23**, 396 (1978).
[14] N. Grønbech-Jensen and O. Farago, Mol. Phys. **111**, 983 (2013).
[15] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1999).
[16] A. Greiner, W. Strittmatter, and J. Honerkamp, J. Stat. Phys. **51**, 95 (1988).
[17] B. Mishra and T. Schlick, J. Chem. Phys. **105**, 299 (1996).
[18] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel, J. Chem. Phys. **114**, 2090 (2001).
[19] G. Zhang and T. Schlick, Mol. Phys. **84**, 1077 (2006).
[20] K. Burrage, I. Lenane, and G. Lythe, SIAM J. Sci. Comput. **29**, 245 (2007).
[21] W. Wang and R. D. Skeel, Mol. Phys. **101**, 2149 (2009).
[22] N. Grønbech-Jensen and O. Farago, J. Chem. Phys. **141**, 194108 (2014).
[23] N. Grønbech-Jensen, N. R. Hayre, and O. Farago, Comput. Phys. Commun. **185**, 524 (2014).
[24] E. Arad, O. Farago, and N. Grønbech-Jensen, Isr. J. Chem. **56**, 629 (2016).
[25] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, in *Theory of Lattice Dynamics in the Harmonic Approximation*, 2nd ed., Solid State Physics Suppl. 3, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1971).
[26] V. K. Tewary, Adv. Phys. **22**, 757 (1973).
[27] H. Yamakawa and T. Yoshizaki, *Models for Polymer Chains*, in *Helical Wormlike Chains in Polymer Solutions* (Springer, Berlin, 2016).
[28] H. Vandebroek and C. Vanderzande, Phys. Rev. E **92**, 060601 (2015).
[29] T. X. Hoang, H. L. Trinh, A. Giacometti, R. Podgornik, J. R. Banavar, and A. Maritan, Phys. Rev. E **92**, 060701 (2015).
[30] L. Piseri and G. Zerbi, J. Chem. Phys. **48**, 3561 (1968).
[31] M. Bohdaneck, Macromolecules **16**, 1483 (1983).
[32] J. D. Hoffman, Polymer **24**, 3 (1983).
[33] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[34] M. A. Desposito and A. D. Vinales, Phys. Rev. E **80**, 021111 (2009).
[35] F. Andrijauskas and V. R. Coluci (unpublished).
[36] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulus, Rev. Mod. Phys. **64**, 1045 (1992).