JCTC Journal of Chemical Theory and Computation

Colored-Noise Thermostats à la Carte

Michele Ceriotti,*,† Giovanni Bussi,‡ and Michele Parrinello†

Computational Science, Department of Chemistry and Applied Biosciences, ETH Zürich, USI Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland and Università di Modena e Reggio Emilia and INFM-CNR-S3, via Campi 213/A, 41100 Modena, Italy

Received October 23, 2009

Abstract: Recently, we have shown how a colored-noise Langevin equation can be used in the context of molecular dynamics as a tool to obtain dynamical trajectories whose properties are tailored to display desired sampling features. In the present paper, after having reviewed some analytical results for the stochastic differential equations forming the basis of our approach, we describe in detail the implementation of the generalized Langevin equation thermostat and the fitting procedure used to obtain optimal parameters. We also discuss the simulation of nuclear quantum effects and demonstrate that by carefully choosing parameters one can successfully model strongly anharmonic solids such as neon. For the reader's convenience, a library of thermostat parameters and some demonstrative code can be downloaded from an online repository.

1. Introduction

Stochastic differential equations (SDE) have been used to model the time evolution of processes characterized by random behavior in fields as diverse as physics and economics. In particular, the Langevin equation (LE) has been regularly applied in the study of Brownian motion and used extensively in molecular dynamics (MD) computer simulations as a convenient and efficient tool to obtain trajectories which sample the constant-temperature, canonical ensemble.[1,2]

In its original form, the Langevin equation is based on the assumption of instantaneous system−bath interactions, which corresponds to the values of the random force being uncorrelated at different times. A non-Markovian, generalized version of the LE arises in the context of Mori−Zwanzig theory.[3,4] In this theory, if one considers a harmonic system coupled with a harmonic bath, it is possible to integrate out the degrees of freedom of the bath. This leaves one with a linear stochastic equation where both the friction and the noise have a finite memory. The conventional LE is recovered in the limit of a clear separation between the characteristic time scale of the system's dynamics and that of the system−bath interaction.

This class of non-Markovian SDEs has been extensively used to model the dynamics of open systems interacting with a physically relevant bath (see, e.g., refs 5–7). Instead, our recent works[8,9] have used colored(correlated)-noise SDEs as a device to sample efficiently statistical distributions in molecular-dynamics (MD) simulations. These works aimed to show how a stochastic thermostat suitable for Car−Parrinello-like dynamics[8] could be constructed, and how nuclear quantum effects can be included in a large class of problems at a fraction of the cost of path-integrals calculations.[9] In these applications the real dynamics is lost, and one focuses only on the efficient calculation of static ensemble averages.

In this paper we discuss the practical implementation of the generalized Langevin equation (GLE) thermostat that we used in the two cases mentioned above. We also provide the reader with the analytical and numerical tools needed to construct SDEs tailored to their own sampling needs. Throughout we take advantage of the dimensional reduction scheme, which allows one to exploit the equivalence between non-Markovian dynamics and Markovian dynamics in higher dimensionality. In doing this, we supplement the physical coordinates with additional degrees of freedom,[4] whose equations of motion are taken as linear, so as to simplify the formalism and analytical derivations.

In the Appendices we recall some of the properties of multidimensional stochastic processes,[5,10–12] which are useful

* Corresponding author e-mail: michele.ceriotti@phys.chem.ethz.ch.

† ETH Zürich.

‡ Università di Modena e Reggio Emilia and INFM-CNR-S3.

10.1021/ct900563s
Published on Web 03/01/2010

to our discussion, and present a short comparison of the GLE thermostat and the widely used massive Nosé−Hoover chains.[13−16] A simple FORTRAN90 code implementing our method to the dynamics of a harmonic oscillator and a library of optimized thermostat parameters can be downloaded from an online repository.[17]

2. Generalized Langevin Thermostat

2.1. Markovian and Non-Markovian Formulations. The Langevin equation for a particle with position q and momentum p, subject to a potential $V(q)$, can be written as

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -V'(q) - a_{pp}p + b_{pp}\xi(t) \end{aligned} \tag{1}$$

where $\xi(t)$ represents an uncorrelated, Gaussian-distributed random force with unitary variance and zero mean [$\langle\xi\rangle = 0$, $\langle\xi(t)\xi(0)\rangle = \delta(t)$]. Here and in what follows we use mass-scaled coordinates. Furthermore, for consistency, the friction coefficient (usually denoted by γ) is here given the symbol a_{pp}, while b_{pp} is the intensity of the random force. In this notation, the fluctuation−dissipation theorem (FDT) reads $b_{pp}^2 = 2a_{pp}k_BT$. If this relation holds, the dynamics generated by eq 1 will sample the canonical ensemble at temperature T.[4,18]

As explained in the Introduction, in order to bypass the complexity of dealing with a non-Markovian formulation directly, we supplement the system with n additional degrees of freedom $\mathbf{s} = \{s_i\}$, which are linearly coupled to the physical momentum and between themselves. The resulting SDE can be cast into the compact form

$$\begin{aligned} \dot{q} &= p \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} &= \begin{pmatrix} -V'(q) \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_p^T \\ \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_p^T \\ \bar{\mathbf{b}}_p & \mathbf{B} \end{pmatrix} \begin{pmatrix} \xi \\ \boldsymbol{\xi} \end{pmatrix} \end{aligned} \tag{2}$$

Here, $\boldsymbol{\xi}$ is a vector of $n + 1$ uncorrelated Gaussian random numbers with $\langle\xi_i(t)\xi_j(0)\rangle = \delta_{ij}\delta(t)$. Clearly, eq 1 is recovered when $n = 0$. For a harmonic potential $V(q) = \omega^2q^2/2$, eqs 2 are linear and an Ornstein−Uhlenbeck (OU) process is recovered whose time propagation can be evaluated analytically. In the nonlinear case one can use the Trotter decomposition to split the dynamics into a linear part, which evolves the $(p, \mathbf{s})$ momenta, and a nonlinear part, which evolves Hamilton's equations.[19] This is facilitated by the fact that the dynamics of $(p, \mathbf{s})$ alone is linear, and its exact finite-time propagator can be analytically evaluated (see section 2.5).

Here and in the rest of the paper, we adopt the same notation introduced in ref 9 to distinguish between matrices acting on the full state vector $\mathbf{x} = (q, p, \mathbf{s})^T$ or on parts of it as illustrated below:

$$\begin{array}{c|ccc} & q & p & \mathbf{s} \\ q & m_{qq} & m_{qp} & \mathbf{m}_q^T \\ p & \bar{m}_{qp} & m_{pp} & \mathbf{m}_p^T \\ \mathbf{s} & \bar{\mathbf{m}}_q & \bar{\mathbf{m}}_p & \mathbf{M} \end{array} \quad \begin{matrix} \} \mathbf{M}_p \\ \} \mathbf{M}_{qp} \end{matrix} \tag{3}$$

The Markovian dynamical eqs 2 are equivalent to a non-Markovian process for the physical variables only. This is best seen by first considering the evolution of the $(p, \mathbf{s})$ variables in the free-particle analogue of eqs 2. The additional degrees of freedom $\mathbf{s}$ can be integrated away, and one is left with (see ref 4 and Appendix A)

$$\dot{p} = -\int_{-\infty}^{t} K(t - s)p(s)\mathrm{d}s + \zeta(t) \tag{4}$$

where the memory kernel $K(t)$ is related to the elements of $\mathbf{A}_p$ by

$$K(t) = 2a_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}}\bar{\mathbf{a}}_p \tag{5}$$

On the basis of the fact that the free-particle dynamics of $(p, \mathbf{s})$ is an OU process, one also finds than the relationship between the static covariance matrix $\mathbf{C}_p = \langle(p, \mathbf{s})^T(p, \mathbf{s})\rangle$, the drift matrix $\mathbf{A}_p$, and the diffusion matrix $\mathbf{B}_p$ is given by

$$\mathbf{A}_p\mathbf{C}_p + \mathbf{C}_p\mathbf{A}_p^T = \mathbf{B}_p\mathbf{B}_p^T \tag{6}$$

Note the remarkable formal analogy between eq 6 and the equations for the orthogonality constraints in Car−Parrinello dynamics, see, e.g., ref 20. In Appendix A we show that setting $\mathbf{C}_p = k_BT$ is sufficient to satisfy the FDT. In this case, eq 6 fixes $\mathbf{B}_p$ once $\mathbf{A}_p$ is given. FDT also implies that the colored-noise autocorrelation function $H(t) = \langle\zeta(t)\zeta(0)\rangle$ is equal to $k_BTK(t)$, whereas the more complex relation between $K(t)$ and $H(t)$, valid in the general case, is reported in eq 27.

Since there is no explicit coupling between the position q and the additional momenta $\mathbf{s}$, one can check that exactly the same dimensional reduction can be performed in the case of an arbitrary potential coupling p and q and that eqs 2 correspond to the non-Markovian process

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\frac{\partial V}{\partial q} - \int_{-\infty}^{t} K(t - s)p(s)\mathrm{d}s + \zeta(t) \end{aligned} \tag{7}$$

In the memory kernel eq 5, $\mathbf{A}$ can be chosen to be a general real matrix and can have complex eigenvalues, provided they have a positive real part. This results in a $K(t)$ that is a linear combination of exponentially damped oscillations. Therefore, a vast class of non-Markovian dynamics can be represented by Markovian equations such as eqs 2.

2.2. Exact Solution in the Harmonic Limit. The thermostats typically used in MD simulations have a few parameters that are chosen by trial and error. A thermostat based on eqs 2 depends on a much larger number of parameters, and hence the fitting procedure is more complex. It is therefore important to find ways to compute a priori analytical estimates so as to guide the tuning of the thermostat.

To this end, we examine the harmonic oscillator, which is commonly used to model physical and chemical systems. By choosing $V(q) = \omega^2q^2/2$ the force term in eqs 2 becomes linear and the dynamics of $\mathbf{x} = (q, p, \mathbf{s})^T$ is the OU process $\dot{\mathbf{x}} = -\mathbf{A}_{qp}\mathbf{x} + \mathbf{B}_{qp}\boldsymbol{\xi}$. In eqs 2 the $\mathbf{s}$ degrees of freedom are coupled to the momentum only. Therefore, most of the additional entries in $\mathbf{A}_{qp}$ and $\mathbf{B}_{qp}$ are zero, and the equations for $\mathbf{x}$ read

$$\begin{pmatrix} \dot{q} \\ \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} = -\begin{pmatrix} 0 & -1 & \mathbf{0} \\ \omega^2 & a_{pp} & \mathbf{a}_p^T \\ \mathbf{0} & \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} q \\ p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} 0 & 0 & \mathbf{0} \\ 0 & & \mathbf{B}_p \\ \mathbf{0} & & \end{pmatrix} \begin{pmatrix} 0 \\ \xi \\ \end{pmatrix} \quad (8)$$

The exact finite-time propagator for eq 8 can be computed, and so it is possible to obtain any ensemble average or time-correlation function analytically. Of course, one is most interested in the expectation values of the physical variables q and p. In particular, one can obtain the fluctuations $\langle q^2 \rangle$ and $\langle p^2 \rangle$ and correlation functions of the form $\langle q^2(t)q^2(0) \rangle$, which can be used to measure the coupling between the thermostat and the system. The resulting expressions are simple to evaluate but lengthy, and we refer the reader to Appendix B for their explicit form.

One can envisage, using the estimates computed for an oscillator of frequency ω, to predict and hence optimize the response of a normal mode of a similar frequency in the system being studied. Furthermore, thanks to the properties of eq 8, one does not need to perform a normal-modes analysis to turn this idea into a practical method. Consider indeed a perfect harmonic crystal and apply an independent instance of the GLE thermostat to the three Cartesian coordinates of each atom. It is easy to see that since eq 8 is linear and contains Gaussian noise, the thermostatted equations of motion are invariant under any orthogonal transformation of the coordinates. Therefore, the resulting dynamics can be described on the basis of the normal modes just as in ordinary Hamiltonian lattice dynamics. As a consequence, each phonon will respond independently as a 1-D oscillator with its own characteristic frequency. Thus, to tune the GLE thermostat, one only needs the analytical results in the one-dimensional case, evaluated as a function of ω. The parameters can then be optimized for a number of different purposes, based solely on minimal information on the vibrational spectrum of the system under investigation, without any knowledge of the eigenmodes of the phonons.

The invariance properties of the GLE thermostat lead to additional advantages. For instance, we can contrast its behavior with that of Nosé−Hoover (NH) chains, based on equations which are quadratic in p (see Appendix C). As a consequence of the nonlinearity, the efficiency of an NH chains thermostat for a multidimensional oscillator depends on the orientation of the eigenmodes relative to the Cartesian axes, an artifact which is absent in our case.

Having set the background, we now turn to the description of the various applications of eqs 2.

2.3. Efficient Canonical Sampling. We first discuss the design of a GLE which can optimally sample phase space. In this case, the target stationary distribution is the canonical ensemble, so the equations of motion need to satisfy the detailed-balance condition. Still, there is a great deal of freedom available in the choice of the autocorrelation kernel or, equivalently, in the choice of the $\mathbf{A}_p$ matrix. These free parameters can be used to optimize the sampling efficiency. To this end, we must first define an appropriate merit function. Standard choices are the autocorrelation times of the potential and total energy (V and H, respectively):

$$\tau_V = \frac{1}{\langle V^2 \rangle - \langle V \rangle^2} \int_0^\infty \langle (V(t) - \langle V \rangle)(V(0) - \langle V \rangle) \rangle \mathrm{d}t$$
$$\tau_H = \frac{1}{\langle H^2 \rangle - \langle H \rangle^2} \int_0^\infty \langle (H(t) - \langle H \rangle)(H(0) - \langle H \rangle) \rangle \mathrm{d}t \quad (9)$$

In the harmonic case, these can be readily computed in terms of correlation times of q^2 and p^2 (see Appendix B) and will depend on $\mathbf{A}_p$ and the oscillator's frequency ω. For example, one easily finds that in the white-noise limit with no additional degrees of freedom as in eq 1

$$\tau_H(\omega) = \frac{1}{a_{pp}} + \frac{a_{pp}}{4\omega^2}, \qquad \tau_V(\omega) = \frac{1}{2a_{pp}} + \frac{a_{pp}}{2\omega^2} \quad (10)$$

Both response times are constant in the high-frequency limit and increase quadratically in the low-frequency extreme of the spectrum. For a given frequency one can choose a_{pp} so as to minimize the correlation time, thus enhancing sampling. It should be noted that eqs 10 contain a "trivial" dependence on ω, as one expects that sampling a normal mode would require at least a time on the order of its vibrational period. One can thus define a renormalized $\kappa(\omega) = [\tau(\omega)\omega]^{-1}$ as a measure of the efficiency of the coupling. In the white-noise case, $\kappa = 1$ for the optimally coupled frequency ($\omega_H = a_{pp}/2$ and $\omega_V = a_{pp}$, respectively) and decreases linearly for lower and higher values of ω.

While this result in itself provides a guide to choose a good value of the friction coefficient in conventional (white-noise) Langevin dynamics, we can enhance the value of $\kappa(\omega)$ over a broader frequency range by using a colored-noise SDE. If we want to obtain canonical sampling, the FDT has to hold, so that $\mathbf{C}_p = k_\mathrm{B}T$. We therefore consider the entries of $\mathbf{A}_p$ as the only independent parameters, since $\mathbf{B}_p$ is then determined by eq 6.

In practice, we set up a fitting procedure in which we choose a set of frequencies ω_i distributed over a broad range ($\omega_{\min}$, $\omega_{\max}$). For an initial guess for the thermostat matrix $\mathbf{A}_p$ we compute $\kappa(\omega)$ for each of these frequencies. We then vary $\mathbf{A}_p$ so as to optimize $\min_i \kappa(\omega_i)$ and aim at a sampling efficiency on the range ($\omega_{\min}$, $\omega_{\max}$) which is as high and frequency independent as possible. We will discuss this fitting procedure in more detail in section 3.

In Figure 1 we compare the optimized $\kappa(\omega)$ for different frequency ranges and number of additional degrees of freedom. We find empirically that $\kappa(\omega) = 1$ is the best result which can be attained and that nearly optimal efficiency can be reached over a very broad range of frequencies. This constant efficiency decreases slightly as the fitted range is extended, regardless of the number n of s_i employed. For a given frequency range, however, increasing n has the effect of making the response flatter.

Clearly this scheme will work optimally in harmonic or quasi-harmonic systems, and anharmonicity will introduce deviations from the predicted behavior. In the extreme case of diffusive systems such as liquids, one has to ask the question of how much diffusion will be affected by the thermostat, especially since in an overdamped LE equation the diffusive modes are considerably slowed down (see, e.g., ref 21). To estimate the impact of the thermostat on the

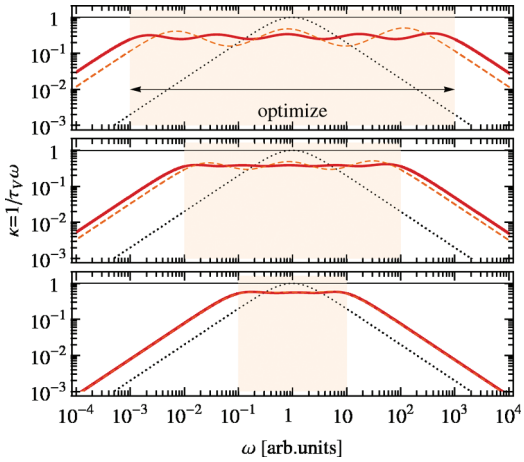

Figure 1. Sampling efficiency as estimated from eq 9 for a harmonic oscillator, plotted as a function of the frequency ω. The $\kappa(\omega)$ curve for a white-noise Langevin thermostat optimized for $\omega = 1$ (black, dotted lines, eq 10) is contrasted with those for a set of optimized GLE thermostats. The panels, from bottom to top, contain the results fitted, respectively, over a frequency range spanning 2, 4, and 6 orders of magnitudes around $\omega = 1$. Dark, continuous lines correspond to matrices with $n = 4$, and dashed, lighter lines correspond to $n = 2$. The GLE curves correspond to the sets of parameters kv_4−2, kv_2−2, kv_4−4, kv_2−2, kv_4−6, kv_2−6, which can be downloaded from an online repository.[17]

diffusion, we define the free-particle diffusion coefficient D^* as that calculated switching off the physical forces. Its value when a GLE thermostat is used is

$$\frac{mD^*}{k_B T} = \frac{1}{\langle p^2 \rangle} \int_0^\infty \langle p(t)p(0) \rangle \, \mathrm{d}t = [\mathbf{A}_p^{-1}]_{pp} = (a_{pp} - \mathbf{a}_p^T \mathbf{A}^{-1} \tilde{\mathbf{a}}_p)^{-1} \tag{11}$$

where we assumed the FDT to hold. In practical cases, if an estimate of the unthermostated (intrinsic) diffusion coefficient D is available, one should choose the matrix $\mathbf{A}_p$ in such a way that $D^* \gg D$, so that the thermostat will not behave as an additional bottleneck for diffusion. Equation 11 has the interesting consequence that D^* can be enhanced either by reducing the overall strength of the noise, as in white-noise LE, or by carefully balancing the terms in the denominator of eq 11.

We found empirically that for an $\mathbf{A}_p$ matrix fitted to harmonic modes over the frequency range $(\omega_{\min}, \omega_{\max})$ the diffusion coefficient computed by eq 11 is $D^* \approx k_B T/m\omega_{\min}$. This latter expression gives a useful recipe for choosing the minimal frequency to be considered when fitting a GLE thermostat for a system whose diffusion coefficient can be roughly estimated.

2.4. Frequency-Dependent Thermostatting. The ability to control the strength of the thermostat−system coupling as a function of the frequency, demonstrated above, points quite naturally at more sophisticated applications. For instance, one can apply two thermostats with distinct target temperatures and different efficiencies $\kappa(\omega)$ (see Figure 2). Obviously, such a simulation is not an equilibrium one, since energy is systematically injected in some modes and removed from others, but leads to a steady state that has useful

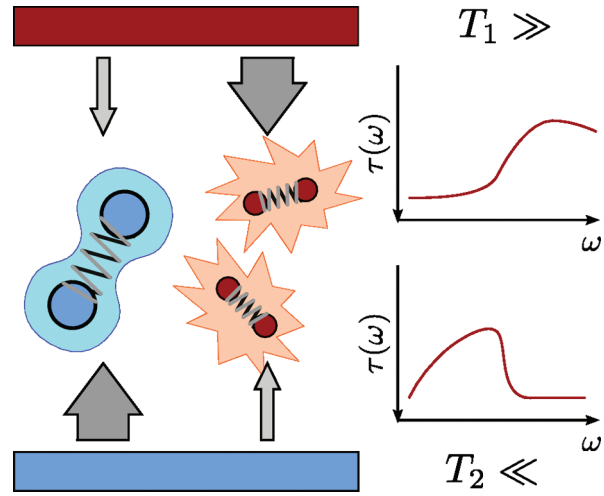

Figure 2. Cartoon representing a two-thermostat setup, which we take as the simplest example of a stochastic process violating the fluctuation−dissipation theorem. If the relaxation time versus frequency curves for the two thermostats are different, a steady state will be reached in which normal modes corresponding to different frequencies will equilibrate at different effective temperatures.

properties. Indeed, the normal modes will couple differently to the two thermostat, so that the effective temperature of each mode can be controlled as a function of ω. This two-thermostats example is just an instance of a broader class of stochastic processes, for which the FDT is violated. In general, we can relax the assumption that $\mathbf{C}_p = k_B T$, and for a given drift matrix we can choose a $\mathbf{B}_p$, which is suitable to our purpose.

Returning to the harmonic oscillator case, one can solve exactly the dynamics for a given choice of $\mathbf{A}_p$, $\mathbf{B}_p$, and frequency ω. The resulting dynamics is performed in the $(n+2)$-dimensional space defined by the variables $(q, p, \mathbf{s})$ according to eq 8. For a compact notation, we used the full matrices $\mathbf{A}_{qp}$ and $\mathbf{B}_{qp}$. The full $\mathbf{C}_{qp}(\omega)$, which defines the stationary distribution in the steady state, can be computed solving an equation analogous to eq 6

$$\mathbf{A}_{qp}\mathbf{C}_{qp} + \mathbf{C}_{qp}\mathbf{A}_{qp}^T = \mathbf{B}_{qp}\mathbf{B}_{qp}^T \tag{12}$$

One can tune the free parameters ($\mathbf{A}_p$ and $\mathbf{B}_p$) so as to make the $c_{qq}(\omega)$ and $c_{pp}(\omega)$ elements of the extended covariance matrix as close as possible to the desired target functions $\langle q^2 \rangle(\omega)$ and $\langle p^2 \rangle(\omega)$.

In a previous paper[9] we applied this method to obtain $\langle q^2 \rangle(\omega)$ and $\langle p^2 \rangle(\omega)$ in agreement with the values appropriate for a quantum harmonic oscillator and obtained a good approximation to the quantum-corrected structural properties in quasi-harmonic systems. Many other applications can be envisaged, which take advantage of frequency-dependent thermostatting. For instance, one could use this technique in accelerated sampling methods,[22–24] which work by artificially heating the low-frequency modes while keeping the other modes at the correct temperature.

2.5. Implementation. The implementation of a GLE thermostat in molecular-dynamics simulations is straightforward. Here, we consider the case of a velocity−Verlet

integrator, which updates positions and momenta by a time step Δt, according to the scheme:

$$\begin{aligned} p &\leftarrow p - V'(q)\Delta t/2 \\ q &\leftarrow q + p\Delta t \\ p &\leftarrow p - V'(q)\Delta t/2 \end{aligned} \tag{13}$$

Equations 13 can be obtained using Trotter splitting in a Liouville operator formalism.[25] In the same spirit one can introduce our GLE thermostat by performing two free-particle steps by $\Delta t/2$ on the $(p, \mathbf{s})$ variables:[19]

$$\begin{aligned} (p, \mathbf{s}) &\leftarrow \mathscr{P}[(p, \mathbf{s}), \Delta t/2] \\ p &\leftarrow p - V'(q)\Delta t/2 \\ q &\leftarrow q + p\Delta t \\ p &\leftarrow p - V'(q)\Delta t/2 \\ (p, \mathbf{s}) &\leftarrow \mathscr{P}[(p, \mathbf{s}), \Delta t/2] \end{aligned} \tag{14}$$

At variance with thermostats based on second-order equations of motion such as Nosé−Hoover, where a multiple time-step approach is required to obtain accurate trajectories,[26,27] this free-particle step can be performed without introducing additional sampling errors. The exact finite-time propagator for $(p, \mathbf{s})$ reads

$$\mathscr{P}[(p, \mathbf{s}), \Delta t]^T = \mathbf{T}(\Delta t)(p, \mathbf{s})^T + \mathbf{S}(\Delta t)\boldsymbol{\xi}^T \tag{15}$$

where $\boldsymbol{\xi}$ is a vector of $n + 1$ uncorrelated Gaussian numbers and the matrices $\mathbf{T}$ and $\mathbf{S}$ can be computed once at the beginning of the simulation and for all degrees of freedom.[10,28] The relations between $\mathbf{T}$, $\mathbf{S}$, $\mathbf{A}_p$, $\mathbf{C}_p$, and Δt read

$$\mathbf{T} = e^{-\Delta t \mathbf{A}_p}, \qquad \mathbf{S}\mathbf{S}^T = \mathbf{C}_p - e^{-\Delta t \mathbf{A}_p}\mathbf{C}_p e^{-\Delta t \mathbf{A}_p^T}$$

It is worth pointing out that when FDT holds, the canonical distribution is invariant under the action of eq 15, whatever the size of the time step. A useful consequence of this property is that in the rare cases where applying eq 15 introduces a significant overhead over the force calculation, the thermostat can be applied every m steps of dynamics using a stride of $m\Delta t$. This will change the trajectory but does not affect the accuracy of sampling.

The velocity−Verlet algorithm (eq 13) introduces finite-Δt errors, whose effect needs to be monitored. In microcanonical simulations, this is routinely done by checking conservation of the total energy H. Following the work of Bussi et al.[29] we introduce a conserved quantity $\tilde{H}$, which can be used for the same purpose

$$\tilde{H} = H - \sum_i \Delta K_i \tag{16}$$

where ΔK_i is the change in kinetic energy due to the action of the thermostat at the ith time step and the sum is extended over the past trajectory. In cases where the FDT holds, such as that described in section 2.3, the drift of the effective energy quantitatively measures the violation of detailed balance induced by the velocity−Verlet step, similarly to refs 19 and 29. In the cases where the FDT does not hold, such as the frequency-dependent thermostatting described in section 2.4, the conservation of this quantity just measures the accuracy of the integration, similarly to refs 30 and 31.

3. Fitting of Colored-Noise Parameters

A key feature of our approach resides in the possibility to optimize the performance of the thermostat based on analytical estimates, making the method effectively parameterless. Such optimization, however, is not trivial, even if computationally inexpensive. The relationship between $\mathbf{A}_p$, $\mathbf{B}_p$, and the correlation properties of the resulting trajectory is highly nonlinear. Furthermore, we found empirically that many local minima exist which greatly hinder the optimization process. With these difficulties in mind, we provide a downloadable library of fitted parameters[17] which can be adapted to most of the foreseeable applications, according to the prescriptions given in section 3.4. Details about the fitting procedure are given in the following three sections.

3.1. Parameterization of GLE Matrices. A number of constraints must be enforced on the drift and diffusion matrices in order to guarantee that the resulting SDE is well behaved. It is therefore important to find a representation of the matrices such that during fitting these conditions are automatically enforced and that the parameters space is efficiently explored. A first condition, required to yield a memory kernel with exponential decay, is that all the eigenvalues of $\mathbf{A}_p$ must have positive real part. A second requirement is that the kernel $K(\omega)$ is positive for all real ω. This ensures that the stochastic process will be consistent with the second law of thermodynamics.[32]

Finding the general conditions for $\mathbf{A}_p$ to satisfy this second constraint is not simple. However, we can state that a sufficient condition for $K(\omega) > 0$ is that $\mathbf{A}_p + \mathbf{A}_p^T$ is positive definite. For simplicity we shall assume such a positivity condition to hold, since we found empirically that this modest loss of generality does not significantly affect the accuracy or the flexibility of the fit. Moreover, in the case of canonical sampling, $\mathbf{A}_p + \mathbf{A}_p^T > 0$ is also required in order to obtain a real diffusion matrix, since $\mathbf{B}_p\mathbf{B}_p^T = k_B T(\mathbf{A}_p + \mathbf{A}_p^T)$ according to eq 6.

One would like to find a convenient parametrization, which automatically enforces these constraints. This is best done by writing $\mathbf{A}_p = \mathbf{A}_p^{(S)} + \mathbf{A}_p^{(A)}$, the sum of a symmetric and antisymmetric part. Since any orthogonal transform of the $\mathbf{s}$ degrees of freedom would not change the dynamics (see Appendix A), one can assume without loss of generality that the $\mathbf{A}^{(S)}$ block in $\mathbf{A}_p^{(S)}$ is diagonal (see eq 3 for the naming convention). Since in the general case the antisymmetric $\mathbf{A}_p^{(A)}$ does not commute with $\mathbf{A}_p^{(S)}$, we will assume it to be full, while $\mathbf{A}_p^{(S)}$ can be written in the form

$$\mathbf{A}_p^{(S)} = \begin{pmatrix} a & a_1 & a_2 & \dots & a_n \\ a_1 & \alpha_1 & 0 & \dots & 0 \\ a_2 & 0 & \alpha_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_n & 0 & 0 & \dots & \alpha_n \end{pmatrix} \tag{17}$$

In order to enforce the positive definiteness, one uses an analytical Cholesky decomposition $\mathbf{A}_p^{(S)} = \mathbf{Q}_p\mathbf{Q}_p^T$ with

$$\mathbf{Q}_p = \begin{pmatrix} q & q_1 & q_2 & \dots & q_n \\ 0 & d_1 & 0 & \dots & 0 \\ 0 & 0 & d_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix} \tag{18}$$

and $\alpha_i = d_i^2$, $a_i = d_i q_i$, and $a = q^2 + \Sigma_i q_i^2$. Such a parametrization guarantees that $\mathbf{A}_p$ will generate a dynamics with a stationary probability distribution and requires $2n + 1$ parameters for the symmetric part (the elements of $\mathbf{Q}_p$, eq 18) and $n(n + 1)/2$ for the antisymmetric part $\mathbf{A}_p^{(A)}$. If we want the equilibrium distribution to be the canonical, we must enforce the FDT and $\mathbf{B}_p\mathbf{B}_p^T$ is uniquely determined.

If we aim at a generalized formulation, which allows for frequency-dependent thermalization, there are no constraints on the choice of $\mathbf{B}_p$ other than the fact that both $\mathbf{B}_p\mathbf{B}_p^T$ and the covariance $\mathbf{C}_p$ must be positive definite. Clearly, a real, lower triangular $\mathbf{B}_p$ is the most general parametrization of a positive-definite $\mathbf{B}_p\mathbf{B}_p^T$ and amounts to introducing $(n + 1)(n + 2)/2$ extra parameters. Together with the assumption that $\mathbf{A}_p^{(S)} > 0$, the condition $\mathbf{B}_p\mathbf{B}_p^T > 0$ is sufficient to ensure that the unique symmetric $\mathbf{C}_p$ which satisfies eq 6 is also positive definite.

3.2. Fitting for Canonical Sampling. Armed with such a robust and fairly general parametrization, one only needs to define a merit function to be optimized. Again, we first consider the simpler case of canonical sampling. Here, we want to obtain a flat response over a wide, physically relevant frequency range ($\omega_{\min}$, $\omega_{\max}$). We have chosen the form

$$\chi_1 = \left[\sum_i |\log \kappa(\omega_i)|^m \right]^{1/m} \tag{19}$$

where ω_is are equally spaced on a logarithmic scale over the fitted range. If a large value of m is chosen, the ω_i which yields the lowest efficiency is weighted more and a flat response curve is obtained. We found empirically that values of m larger than 10 lead to a proliferation of local minima and hinder efficient optimization. To resolve this, one can use the optimal parameters for $m = 2$ as input for further refinement at larger m until convergence is achieved.

This procedure can be modified so as to provide an efficient thermostat which can be used in Car−Parrinello-like dynamics. In this case, the GLE has to act as a low-pass filter in which only the low ionic frequencies are affected and fast electronic modes are not perturbed. To obtain this effect, we compute eq 19 only for the ω_i's which are smaller than a cutoff frequency ω_{CP} and introduce an additional term

$$\chi_2 = \left[\sum_{\omega_i > \omega_{CP}} \max\left[\log \kappa(\omega_i) - k \log \frac{\omega_{CP}}{\omega_i}, 0 \right]^m \right]^{1/m} \tag{20}$$

χ_2 enforces a steep decrease of $\kappa(\omega)$ above ω_{CP} with a slope k on a logarithmic scale. Values of k as large as 9 can be used, which guarantee an abrupt drop in thermalization efficiency for the fast modes (see Figure 3).

3.3. Nonthermal Noise and Quantum Thermostat. We now discuss the case in which the thermostat is permitted to violate FDT in order to achieve frequency-dependent equilibration. For these applications, one must also fit the fluctuations $c_{pp}(\omega)$ and $c_{qq}(\omega)$ to some target function $\tilde{c}_{pp}$ and $\tilde{c}_{qq}$. We shall not treat the general case but rather investigate the example of the quantum thermostat (ref 9). The procedure followed provides a clear guide for future extensions to different applications.

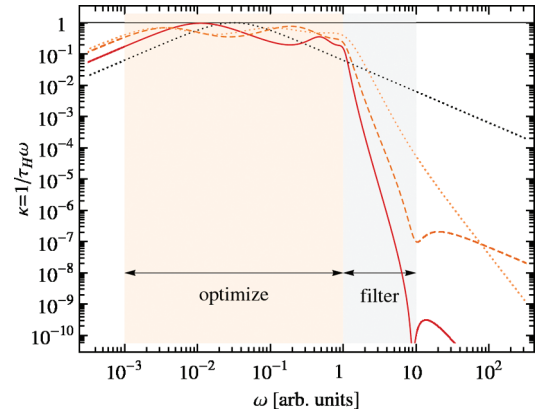

Figure 3. Thermostatting efficiency, as estimated from eq 9, for a colored-noise thermostat optimized for Car−Parrinello dynamics. Sampling efficiency is optimized for $\omega \in (10^{-3}, 1)$, and an abrupt drop in efficiency is enforced for $\omega \in (1, 10)$ using the penalty function eq 20in the fitting. The continuous (dark red) curve corresponds to $k = 9$, the dashed (orange) curve to $k = 6$, and the dotted (light orange) curve to $k = 3$. The $\kappa(\omega)$ curve for a white-noise thermostat centered on the optimized range is also reported for reference (dotted black curve). The three curves correspond to the parameters set, cp-9_4−3, cp-6_4−3 and cp-3_4−3.[17]

In order to reproduce quantum ions effects, one must selectively heat high-frequency phonons for which zero-point energy effects are important without affecting the low-frequency modes which behave classically. The required frequency dependence of the variance for this case is that of a quantum oscillator, i.e., $\tilde{c}_{pp}(\omega) = \omega^2\tilde{c}_{qq}(\omega) = \hbar\omega/2 \coth \hbar\omega/2k_BT$. The $\omega \to 0$ classical limit can be proved to correspond to two conditions on the elements of the free-particle covariance matrix $\mathbf{C}_p$, namely, $c_{pp} = k_BT$ and $\mathbf{a}_p^T\mathbf{A}^{-1}\mathbf{c}_p = 0$. One could enforce such constraints exactly by considering the entries of $\mathbf{C}_p$ as independent fitting parameters and obtaining the diffusion matrix from eq 6. We found however that this choice makes it difficult to obtain a positive-definite $\mathbf{B}_p\mathbf{B}_p^T$ and that the fitting becomes more complex and inefficient.

As an alternative, we decided to enforce the low-frequency limit with an appropriate penalty function

$$\chi_3 = (c_{pp}/k_BT - 1)^2 + (\mathbf{a}_p^T\mathbf{A}^{-1}\mathbf{c}_p/k_BT)^2 \tag{21}$$

to be optimized together with the sampling efficiency (eq 19) and a term which measures how well the finite-frequency fluctuations were fitted

$$\chi_4 = \left[\sum_i \left| \log \frac{c_{qq}(\omega_i)}{\tilde{c}_{qq}(\omega_i)} \right|^m + \left| \log \frac{c_{pp}(\omega_i)}{\tilde{c}_{pp}(\omega_i)} \right|^m \right]^{1/m} \tag{22}$$

Since the low-frequency limit is already enforced by eq 21, we compute eq 22 on a set of points equally spaced between

the maximum frequency ω_{max} and one-half of the onset frequency for quantum effects $\omega_q = k_B T/\hbar$.

3.4. Transferability of Fitted Parameters. The scheme described in the previous sections allowed us to obtain matrices suitable for all the applications discussed in previous works. Furthermore, it provides a starting point for obtaining matrices which one might deem useful for novel applications. However, the reader is advised that the fitting is still far from being a black-box procedure. It is thus necessary to experiment with a combination of different initial parameters and minimization schemes. We found the downhill simplex method[33] to be particularly effective but resorted to simulated annealing when the optimization got stuck in a local minimum. There is a great deal of arbitrariness in the choice of the terms in eqs 19−22 and in their weighted combination $\chi = \Sigma w_i \chi_i$. To make the procedure even more delicate, we observe that in high-n cases the parameters tend to collapse into "degenerate" minima, where the full dimensionality of the search space is not exploited. This phenomenon can be successfully circumvented by enforcing an even spacing of the eigenvalues of $\mathbf{A}$ over the frequency range of interest and slowly releasing this restraint during the later stages of optimization.

However, the problems mentioned above have no major practical consequences, as the computation of analytical estimates is inexpensive and one can afford a great deal of trial and error during the optimization. Moreover, fitted parameters can be reused, since the optimized parameters can be easily transferred to similar problems because of the scaling properties of the dynamics (eq 8).

In fact, one can see that if the drift and covariance matrices ($\mathbf{A}_p$, $\mathbf{C}_p$) lead to the efficiency curves $\kappa(\omega)$ and fluctuations $c_{pp}(\omega)$, the scaled matrices ($\alpha\mathbf{A}_p$, $\beta\mathbf{C}_p$) will yield $\kappa(\alpha^{-1}\omega)$ and the fluctuations $\beta c_{pp}(\alpha^{-1}\omega)$. This means that if $\mathbf{A}_p$ is optimized for sampling over the range (ω_{min}, ω_{max}), $\alpha\mathbf{A}_p$ will be optimal over ($\alpha\omega_{min}$, $\alpha\omega_{max}$). We also remark that if ($\mathbf{A}_p$, $\mathbf{C}_p$) are fitted to the quantum harmonic oscillator fluctuations at temperature T, ($\alpha\mathbf{A}_p$, $\alpha\mathbf{C}_p$) will be suitable for temperature αT. Care must be taken in this case to ensure that the scaled frequency range still encompasses the whole vibrational spectrum of the system being studied.

4. Understanding the Quantum Thermostat

As discussed in ref 9, one must pay a great deal of attention when using a "quantum thermostat" because energy is transferred between modes of different frequency as a consequence of the anharmonic coupling. This is reminiscent of zero-point energy (ZPE) leakage which plagues semiclassical approaches to the computation of nuclear quantum effects.[34,35] In the cases we explored so far, empirical evidence suggests that quasi-harmonic solids can be treated with good accuracy down to temperatures as low as 10% of the Debye temperature Θ_D. Clearly, the ultimate test to assess the accuracy of the method is a comparison with path-integral calculations to be performed on a similar but computationally cheaper model, such as a smaller size box or a simpler force field.

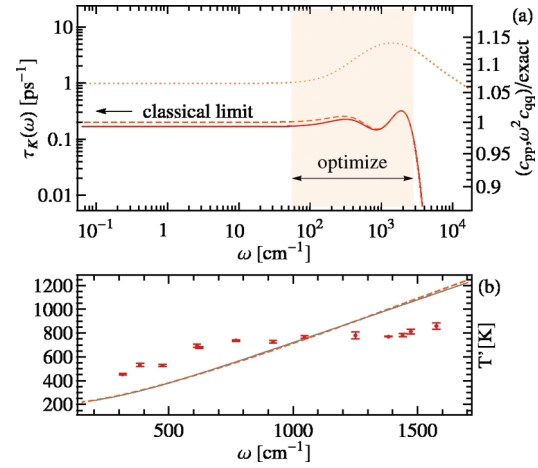

Figure 4. (a) ω dependence of the kinetic energy correlation time $\tau_K(\omega)$ (light, dotted line) and the ratio of the fitted fluctuations $c_{pp}(\omega)$ (dashed line) and $\omega^2 c_{qq}(\omega)$ (full line) with the exact, quantum-mechanical target function. (b) Normal-mode-projected kinetic temperature for a few selected phonons in diamond. The dashed line is the value expected from the fitted $c_{pp}(\omega)$, while the full line is the exact, quantum-mechanical expectation value for a harmonic oscillator. Calculations have been performed with the parameters qt-20_6_BAD.[17]

One would like however to obtain some qualitative measure of the quality of the fit and gauge the transferability of a given set of parameters. To this end, we first state a couple of empirical rules and then validate them on two fairly different real systems. A first observation is that it is useless to push the fitting of the fluctuations $c_{pp}(\omega)$ and $c_{qq}(\omega)$ to very high accuracy if this comes at the expense of the coupling efficiency. In fact, we would be trading a small, controlled fitting error with a possibly larger, uncontrollable, and system-dependent error stemming from anharmonicity. Second, we observed that in order to contrast more effectively the flow of energy between different phonons, one should try to reduce the correlation time of the kinetic energy τ_K, rather than focus solely on the terms in eq 9, which are better suited to measure sampling efficiency. In fact, a low $\tau_K(\omega)$ corresponds to a slightly overdamped regime, where sampling efficiency is suboptimal but ZPE is enforced more tightly.

To demonstrate these concepts in a real system, we performed some calculations with a Tersoff model of diamond at a temperature $T = 200$ K. At this low temperature, slightly below $0.1\Theta_D$, quantum effects are very strong and we therefore expect to have problems maintaining the large difference in temperature between the stiff and soft phonons. Using a very harmonic system such as diamond is particularly useful, since one can monitor directly the efficiency of the thermostat by projecting the atomic velocities on a selection of normal modes. Hence, a projected kinetic temperature $T'(\omega)$ can be computed and its value checked against the predictions in the harmonic limit in the same spirit as in ref 9. In Figure 4 we report the results with a matrix fitted taking into account only the terms in eqs 21 and 22. Even in a harmonic system such as diamond there are major errors due to ZPE leakage from the high-frequency

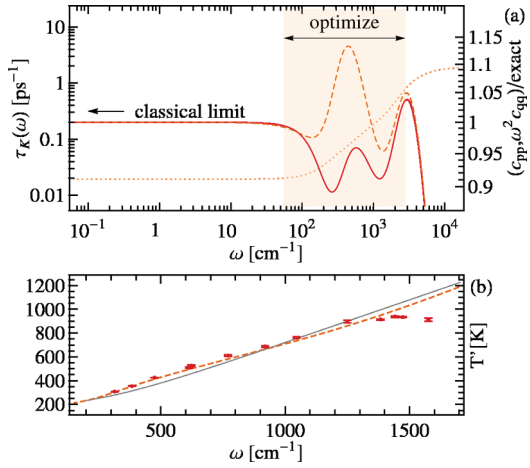

Figure 5. (a) ω dependence of the kinetic energy correlation time $\tau_K(\omega)$ (light, dotted line) and the ratio of the fitted fluctuations $c_{pp}(\omega)$ (dashed line) and $\omega^2 c_{qq}(\omega)$ (full line) with the exact, quantum-mechanical target function. (b) Normal-mode-projected kinetic temperature for a few selected phonons in diamond. The dashed line is the value expected from the fitted $c_{pp}(\omega)$, while the full line is the exact, quantum-mechanical expectation value for a harmonic oscillator. Calculations have been performed with the parameters qt-20_6.[17]

to the low-frequency modes, which the thermostat compensates only partially. These poor results should be compared with those of Figure 5. Here, we also introduced in the fit a term analogous to 19 to reduce the value of $\tau_K(\omega)$. The projected kinetic temperature now agrees almost perfectly with the analytical predictions $c_{pp}(\omega)$ for most of the modes. The only ones displaying significant deviations are the faster ones, for which the value of $\tau_K(\omega)$ is slightly larger. The $c_{pp}(\omega)$ curve deviates by nearly 10% from the exact, quantum-mechanical expectation value. However, thanks to the more efficient coupling, the errors due to anharmonicities are better compensated, and in actuality, the overall error is much smaller than for the parameters presented in Figure 4.

To test whether these prescriptions work for less harmonic problems, we now turn to a completely different system; namely, the structural properties of solid neon at 20 K. At variance with diamond, quantum-ions effects are less pronounced, but the system is close to its melting temperature and is significantly anharmonic. As shown in Figure 6, the agreement between our results and those of accurate path-integral calculations[36] is almost perfect if the parameters of Figure 5 are used. As expected, large errors are present if qt-20_6_BAD is used. Further improvements on the fitting strategy and the application to strongly anharmonic systems is currently being investigated and will be the subject of further work.

5. Conclusions

In this paper we discussed in detail the use of colored-noise dynamics based on Ornstein−Uhlenbeck processes as a tool for performing molecular dynamics. Applications range from enhanced sampling, which we demonstrate in the harmonic limit and will be applied to real systems in forthcoming

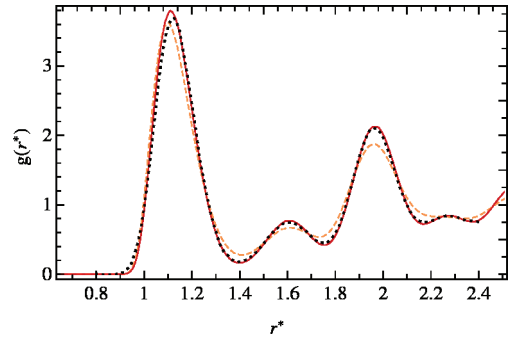

Figure 6. Radial distribution function as computed from fully converged path-integral calculations[36] (black, dotted line) and a quantum-thermostat MD trajectory for a Lennard−Jones model of solid neon at $T = 20$ K. Distances are in reduced units. Full line corresponds to the parameters set qt-20_6 (cf. Figure 5) and lighter, dashed line to the set qt-20_6_BAD (cf. Figure 4).

publications, to thermostats for adiabatically separated problems and frequency-dependent thermalization.

Our idea exploits the linear nature of the OU stochastic differential equations, which allows one to use the one-dimensional harmonic oscillator as a simple but physically motivated reference model. On the basis of the analytical prediction obtained in that case, we describe a recipe for fitting the thermostat parameters so as to obtain the desired response properties in real systems. The procedure is not simple, and we are considering different approaches to make it more robust and effective. Fortunately, however, fitted matrices can be easily transferred from one system to another. With this in mind we provided an extensive library of optimized parameters,[17] which makes fitting unnecessary for most applications.

We also comment on practical issues concerning the implementation of the generalized-Langevin thermostat in a molecular-dynamics program and its use in applications. In particular, we discuss in detail how one can use colored noise to model nuclear quantum effects.[9] We provide some empirical rules to guide the fitting in this difficult case, and we demonstrate that a normal-mode analysis in a quasi-harmonic system is a valuable tool for assessing the quality of a set of parameters. We believe that further investigation will find many other applications for colored noise in molecular dynamics and in computer simulations of molecular systems in general. As an example, we are currently investigating the use of a zero-temperature, optimal-sampling GLE thermostat in order to perform structural optimization. On similar lines and taking inspiration from "quantum annealing",[37,38] one can envisage using frequency-dependent thermalization to improve the performance of simulated annealing.

Appendix A: Memory Kernels for the Non-Markovian Formulation

The connection between the Markovian (eq 2) and non-Markovian (eq 7) formulations of the colored-noise Langevin equation can be understood using techniques similar to those adopted in Mori−Zwanzig theory.[4,11] Let us first consider a very general, multidimensional OU process, where we single

out some degrees of freedom (**y**) that we wish to integrate away, leaving only the variables marked as **x**.

$$\begin{pmatrix}\dot{\mathbf{x}}\\ \dot{\mathbf{y}}\end{pmatrix} = -\left(\begin{array}{c|c}\mathbf{A}_{xx} & \mathbf{A}_{xy}\\ \hline \mathbf{A}_{yx} & \mathbf{A}_{yy}\end{array}\right)\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix} + \begin{pmatrix}\mathbf{B}_{x\xi}\\ \mathbf{B}_{y\xi}\end{pmatrix}\boldsymbol{\xi} \tag{23}$$

Assuming that the dynamics has finite memory, one can safely take $\mathbf{y}(-\infty) = 0$ and the ansatz

$$\mathbf{y}(t) = \int_{-\infty}^{t} e^{-(t-t')\mathbf{A}_{yy}}[-\mathbf{A}_{yx}\mathbf{x}(t') + \mathbf{B}_{y\xi}\boldsymbol{\xi}(t')]\mathrm{d}t' \tag{24}$$

Substituting into eq 23, one sees that **y** can be eliminated from the dynamics of **x** and arrives at

$$\begin{aligned}\dot{\mathbf{x}}(t) &= -\int_{-\infty}^{t}\mathbf{K}(t-t')\mathbf{x}(t')\mathrm{d}t' + \boldsymbol{\zeta}(t)\\ \mathbf{K}(t) &= 2\mathbf{A}_{xx}\delta(t) - \mathbf{A}_{xy}e^{-t\mathbf{A}_{yy}}\mathbf{A}_{yx} \quad (t \geq 0)\\ \boldsymbol{\zeta}(t) &= \mathbf{B}_{x\xi}\boldsymbol{\xi}(t) - \int_{-\infty}^{t}\mathbf{A}_{xy}e^{-(t-t')\mathbf{A}_{yy}}\mathbf{B}_{y\xi}\boldsymbol{\xi}(t')\mathrm{d}t'\end{aligned} \tag{25}$$

One can see that eqs 25 are invariant under any orthogonal transformation of the **y** dynamical variables, meaning that such a transformation leaves the dynamics of the **x**'s unchanged.

The colored noise is better described in terms of its time-correlation function, $\mathbf{H}(t) = \langle\boldsymbol{\zeta}(t)\boldsymbol{\zeta}(0)^T\rangle$. Let us first introduce the symmetric matrix $\mathbf{D} = \mathbf{B}\mathbf{B}^T$, whose parts we shall label using the same scheme used for **A** in eq 23. We shall also need $\mathbf{Z}_{yy} = \int_0^\infty e^{-\mathbf{A}_{yy}t}\mathbf{D}_{yy}e^{-\mathbf{A}_{yy}^T t}\,\mathrm{d}t$. With these definitions in mind, one finds

$$\mathbf{H}(t) = \delta(t)\mathbf{D}_{xx} + \mathbf{A}_{xy}e^{-t\mathbf{A}_{yy}}[\mathbf{Z}_{yy}\mathbf{A}_{xy}^T - \mathbf{D}_{yx}] \quad (t \geq 0) \tag{26}$$

Note that the value of $\mathbf{H}(t)$ for $t < 0$ is determined by the constraint $\mathbf{H}(-t) = \mathbf{H}(t)^T$; the value of $\mathbf{K}(t)$ instead is irrelevant for negative times: we will assume $\mathbf{K}(-t) = \mathbf{K}(t)^T$ to hold, since this will simplify some algebra below.

Let us now switch to the case of the free-particle counterpart of eqs 2, which is relevant to the memory functions entering eqs 7. Here, we want to integrate away all the **s** degrees of freedom, retaining only the momentum p. Hence, we can transform eqs 25 and 26 to the less cumbersome form

$$\begin{aligned}K(t) &= 2a_{pp}\delta(t) - \mathbf{a}_p^T e^{-|t|\mathbf{A}}\bar{\mathbf{a}}_p\\ H(t) &= d_{pp}\delta(t) + \mathbf{a}_p^T e^{-|t|\mathbf{A}}[\mathbf{C}\mathbf{a}_p - \mathbf{d}_p]\end{aligned} \tag{27}$$

This compact notation hides certain relevant property of the memory kernels, which are more apparent when the kernels are written in their Fourier representation. If $\mathbf{D}_p = \mathbf{B}_p\mathbf{B}_p^T$ is transformed according to eq 6. $K(\omega)$ and $H(\omega)$ read

$$\begin{aligned}K(\omega) &= 2a_{pp} - 2\mathbf{a}_p^T\frac{\mathbf{A}}{\mathbf{A}^2+\omega^2}\bar{\mathbf{a}}_p\\ H(\omega) &= K(\omega)\left(c_{pp} - \mathbf{a}_p^T\frac{\mathbf{A}}{\mathbf{A}^2+\omega^2}\mathbf{c}_p\right) +\\ &\quad 2\omega^2\left(\mathbf{a}_p^T\frac{1}{\mathbf{A}^2+\omega^2}\mathbf{c}_p\right)\left(1 + \mathbf{a}_p^T\frac{1}{\mathbf{A}^2+\omega^2}\bar{\mathbf{a}}_p\right)\end{aligned} \tag{28}$$

It is seen that the memory functions (hence the dynamical trajectory) are independent of the value of **C**, the covariance of the fictitious degrees of freedom. Moreover, a sufficient condition for the FDT to hold is readily found. By setting $c_{pp} = k_\mathrm{B}T$ and $\mathbf{c}_p = 0$, one obtains $H(\omega) = k_\mathrm{B}TK(\omega)$, which is precisely the FDT for a non-Markovian Langevin equation. Since the value of **C** is irrelevant we can take $\mathbf{C}_p = k_\mathrm{B}T$, which simplifies the algebra and leads to numerically stable trajectories.

Appendix B: Covariance Matrix and Correlation Times for the Harmonic Oscillator

Given **A** and **C** matrices (the drift term and static covariance for a generic OU process), one can find the diffusion matrix **B** by an expression analogous to eq 6. The same relation can be used to obtain the elements of **C** given the drift and diffusion matrices by solving the linear system. However, the covariance matrix can be computed more efficiently by finding the eigendecomposition of $\mathbf{A} = \mathbf{O}\,\mathrm{diag}(\alpha_i)\,\mathbf{O}^{-1}$ and computing

$$C_{ij} = \sum_{kl}\frac{O_{ik}[\mathbf{O}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{O}^{-1T}]_{kl}O_{jl}}{\alpha_k + \alpha_l} \tag{29}$$

Now, let **x** be the vector describing the trajectory of the OU process. In order to compute τ_H or τ_V (eq 9) one needs time-correlation functions of the form $\langle x_i(t)x_j(t)x_k(0)x_l(0)\rangle$. The corresponding, non-normalized integrals

$$\tau_{ijkl} = \int_0^\infty[\langle x_i(t)x_j(t)x_k(0)x_l(0)\rangle - \langle x_ix_j\rangle\langle x_kx_l\rangle]\mathrm{d}t \tag{30}$$

can be computed in terms of the tensorial quantity

$$X_{ijkl} = \sum_{mn}\frac{O_{im}[\mathbf{O}^{-1}\mathbf{C}]_{ml}O_{jn}[\mathbf{O}^{-1}\mathbf{C}]_{nk}}{\alpha_m + \alpha_n} \tag{31}$$

as $\tau_{ijkl} = (1/4)\,(X_{ijkl} + X_{ijlk} + X_{klij} + X_{lkij})$. For example, if we consider the full OU process in the harmonic case, one computes

$$\tau_H = \frac{\omega^4\tau_{qqqq} + 2\omega^2\tau_{qqpp} + \tau_{pppp}}{\omega^4c_{qq}^2 + 2\omega^2c_{qp}^2 + c_{pp}^2}, \qquad \tau_V = \frac{\tau_{qqqq}}{c_{qq}^2} \tag{32}$$

where we use an obvious notation for the indices in τ_{ijkl}.

Appendix C: Comparison with Nosé−Hoover Chains

The most widespread techniques for canonical sampling in MD are probably white-noise Langevin and Nosé−Hoover chains (NHC). The white-noise Langevin can be considered as a limiting case of the thermostatting method we describe in this work, but NHC is based on a radically different philosophy. It is therefore worth performing a brief comparison between the latter and the GLE thermostat.

In the "massive" version of the NH thermostat,[13,14] each component of the physical momentum is coupled to an additional degree of freedom with a fictitious mass Q by means of a second-order equation of motion. The resulting dynamics ensures that the physically relevant degrees of

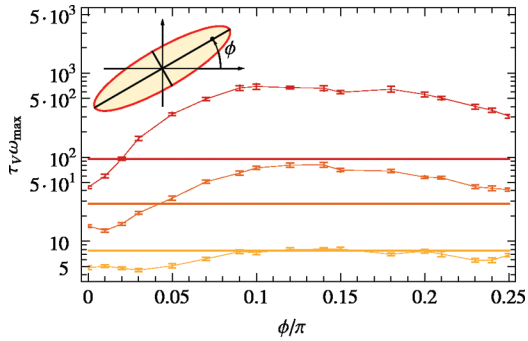

Figure 7. Correlation time for the potential energy of a 2-D harmonic oscillator as a function of the angle between the eigenmodes and the Cartesian axes. τ_V is computed for different values of the condition number $\omega_{max}/\omega_{min}$, from bottom to top 10, 31.6, and 100. Thin lines serve as an aid for the eye, connecting the results obtained in the three cases using a massive NH chains thermostat with four additional degrees of freedom and $Q = k_BT/\omega_{max}^2$. Error bars are also shown for individual data points. Thick lines correspond to the (constant) result predicted for a GLE thermostat using respectively, the thermostat parameters kv_2−1, centered on $0.32\omega_{max}$, kv_4−2, centered on $0.18\omega_{max}$, and kv_4−2 centered on $0.1\omega_{max}$. The values obtained in actual GLE simulations agree with the predictions within the statistical error bar and are not reported.

freedom will sample the correct, constant-temperature ensemble with the advantage of having deterministic equations of motion and a well-defined conserved quantity. However, in the harmonic case, trajectories are poorly ergodic. This problem can be addressed by coupling the fictitious momentum to a second bath variable with a similar equation of motion. By repeating this process further a "Nosé−Hoover chain" can be formed, which ensures that the dynamics is sufficiently chaotic to achieve efficient sampling.[15,41] The drawback of this approach is that the thermostat equations are quadratic in momenta. It is therefore difficult to obtain analytical predictions for the properties of the dynamics, and the integration of the additional degrees of freedom must be performed with a multiple time-step approach, which makes the thermostat more expensive.

To examine the performances of NHC and GLE, one could envisage comparing the sampling efficiency as defined by the correlation times (eq 9). Obtaining such estimates is not straightforward, not only because the harmonic case cannot be treated analytically but also because in the multidimensional case the properties of the trajectory will not be invariant under an orthogonal transformation of coordinates, as discussed in section 2. The simplest model we can conceive for comparing NHC and GLE is therefore a two-dimensional harmonic oscillator with different vibrational frequencies on the two normal modes and adjustable relative orientations of the eigenvectors with respect to the thermostatted coordinates.

The resulting τ_V is reported in Figure 7: in the highly anisotropic cases, the efficiency of the NH chains depends dramatically on the orientation of the axes, while for well-conditioned problems is almost constant. The linear stochastic thermostat, on the other hand, has a predictable response, which is completely independent of orthogonal transforms of the coordinates. In the one-dimensional case, or when eigenvectors are perfectly aligned with the axes, NH chains are very efficient for all modes with frequency $\omega < (k_BT/Q)^{1/2}$. One should however consider that in the absence of an exact propagator choosing a small Q implies that integration of the trajectory for the chains will become more expensive.

Obviously, such a simple toy model does not give quantitative information on the behavior in real-life cases, where modes of different frequencies coexist with anharmonicity and diffusive behavior. However, it demonstrates that the colored-noise Langevin thermostat performs almost as well as the axis-aligned NH chains. Furthermore, unlike the NHC, there are no unpredictable failures for anisotropic potentials.

Acknowledgment. We thank David Manolopoulos for important suggestions and fruitful discussion, Marcella Iannuzzi for having implemented the GLE thermostat in CP2K,[39] Alessandro Curioni for implementation in CPMD,[40] and Grigorios Pavliotis and Michela Ottobre for discussion and references on stochastic processes. We also acknowledge Stefano Angioletti-Uberti, Paolo Elvati, Hagai Eshet, Kuntal Hazra, Rustam Khaliullin, and Tom Markland for discussion and having preliminarily tested the thermostat in real applications, providing valuable feedback. We are especially in debt to Gareth Tribello, who contibuted to testing and greatly helped us improve the manuscript.

Note Added after ASAP Publication. This article was published ASAP on March 1, 2010. Equation 20 has been modified. The correct version was published on March 4, 2010.

References

(1) Schneider, T.; Stoll, E. *Phys. Rev. B* **1978**, *17*, 1302–1322.

(2) Adelman, S. A.; Brooks, C. L. *J. Phys. Chem.* **1982**, *86*, 1511.

(3) Zwanzig, R. *Phys. Rev.* **1961**, *124*, 983–992.

(4) Zwanzig, R. *Nonequilibrium statistical mechanics*; Oxford University Press: New York, 2001.

(5) Martens, C. C. *J. Chem. Phys.* **2002**, *116*, 2516–2528.

(6) Wang, J.-S. *Phys. Rev. Lett.* **2007**, *99*, 160601.

(7) Kantorovich, L. *Phys. Rev. B* **2008**, *78*, 094304.

(8) Ceriotti, M.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2009**, *102*, 020601.

(9) Ceriotti, M.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2009**, *103*, 030603.

(10) Gardiner, C. W. *Handbook of Stochastic Methods*, 3rd ed.; Springer: Berlin, 2003.

(11) Łuczka, J. *Chaos* **2005**, *15*, 026107.

(12) Marchesoni, F.; Grigolini, P. *J. Chem. Phys.* **1983**, *78*, 6287.

(13) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.

(14) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(15) Martyna, G. J.; Tuckerman, M. E.; Klein, M. L. *J. Chem. Phys.* **1992**, *97*, 2635.

(16) Tobias, D. J.; Martyna, G. J.; Klein, M. L. *J. Phys. Chem.* **1993**, *97*, 12959–12966.

(17) GLE4MD; http://gle4md.berlios.de (accessed Jan 12, 2010).

(18) Kubo, R. *Rep. Prog. Phys.* **1966**, *29*, 255–284.

(19) Bussi, G.; Parrinello, M. *Phys. Rev. E* **2007**, *75*, 056707.

(20) Marx, D.; Hutter, J. Ab initio molecular dynamics: Theory and Implementation. In *Modern Methods and Algorithms of Quantum Chemistry Proceedings*, 1st ed.; Grotendorst, J., Ed.; NIC Series: Julich, Germany, 2000; Vol. 1, pp 301−449.

(21) Bussi, G.; Parrinello, M. *Comput. Phys. Commun.* **2008**, *179*, 26.

(22) Rosso, L.; Mináry, P.; Zhu, Z.; Tuckerman, M. E. *J. Chem. Phys.* **2002**, *116*, 4389.

(23) VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 203–208.

(24) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2006**, *426*, 168–175.

(25) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.

(26) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *J. Chem. Phys.* **1996**, *104*, 5579–5588.

(27) Jang, S.; Voth, G. A. *J. Chem. Phys.* **1997**, *107*, 9514–9526.

(28) Fox, R. F.; Gatland, I. R.; Roy, R.; Vemuri, G. *Phys. Rev. A* **1988**, *38*, 5938–5940.

(29) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

(30) Bruneval, F.; Donadio, D.; Parrinello, M. *J. Phys. Chem. B* **2007**, *111*, 12219.

(31) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. *J. Chem. Theory Comput.* **2007**, *3*, 1100–1105.

(32) Ford, G. W.; Lewis, J. T.; O' Connell, R. F. *Phys. Rev. A* **1988**, *37*, 4419–4428.

(33) Nelder, J. A.; Mead, R. *Comp. J.* **1965**, *7*, 308–313.

(34) Alimi, R.; García-Vela, A.; Gerber, R. B. *J. Chem. Phys.* **1992**, *96*, 2034.

(35) Habershon, S.; Manolopoulos, D. E. *J. Chem. Phys.* **2009**, *131*, 244518.

(36) Singer, K.; Smith, W. *Mol. Phys.* **1988**, *64*, 1215–1231.

(37) Lee, Y.-H.; Berne, B. J. *J. Phys. Chem. A* **2000**, *104*, 86.

(38) Santoro, G. E.; Tosatti, E. *J. Phys. A* **2006**, *39*, R393.

(39) The CP2K developers group. CP2K; http://cp2k.berlios.de (accessed Jan 12, 2010).

(40) CPMD; Copyright IBM Corp 1990−2006, Copyright MPI für Festkörperforschung Stuttgart 1997−2001.

(41) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J.; Klein, M. L. *J. Chem. Phys.* **1993**, *99*, 2796–2808.

CT900563S