THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Adaptive Stochastic Methods for Sampling Driven Systems

Citation for published version:
Jones, A & Leimkuhler, B 2011, 'Adaptive Stochastic Methods for Sampling Driven Systems', *The Journal of Chemical Physics*, vol. 135, 084125. https://doi.org/10.1063/1.3626941

Digital Object Identifier (DOI):
10.1063/1.3626941

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
The Journal of Chemical Physics

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 28. Apr. 2024

Jones, A & Leimkuhler, B 2011, 'Adaptive Stochastic Methods for Sampling Driven Systems' The Journal of Chemical Physics, vol. 135, 084 125

Adaptive Stochastic Methods for Sampling Driven Molecular Systems

Andrew Jones

School of Physics, University of Edinburgh, EH9 3JZ, UK

Ben Leimkuhler

The Maxwell Institute and School of Mathematics,

University of Edinburgh, EH9 3JZ, UK

Abstract

Thermostatting methods are discussed in the context of canonical sampling in the presence of driving stochastic forces. Generalisations of the Nosé-Hoover method and Langevin dynamics are introduced which are able to dissipate excess heat introduced by steady Brownian perturbation (without a priori knowledge of its strength) while preserving ergodicity. Implementation and parameter selection are considered. It is demonstrated using numerical experiments that the methods derived can adaptively control the target canonical ensemble in the presence of nonlinear driving perturbations.

I. INTRODUCTION

Ergodic sampling is an important requirement in most molecular simulation. Molecular dynamics simulations have finite, often modest, system sizes, and are based on Hamiltonian dynamics. Given a sufficiently strong chaotic mixing behavior, averages with respect to the NVE (constant particle number, volume and energy) ensemble can typically be calculated using (almost any) trajectory of the system; sampling the canonical or NVT ensemble, typically requires the use of an appropriate modification of the physical model, termed a thermostat. A popular scheme for this purpose is the Nosé-Hoover (NH) method, which supplements the physical system by an additional control variable and an auxiliary negative feedback loop to allow trajectories (projected to the original phase space) to sample the canonical ensemble. Because Hamiltonian dynamics, as well as many thermostats, are deterministic, it is possible for some types of models to be trapped in restricted subdomains of the thermodynamically accessible region of phase space, preventing an accurate sampling. As thermostats are artificial constructions, they can be designed to address the ergodicity problem, either by use of additional auxiliary variables (as in Nosé-Hoover Chains (NHC) [1]) or by adding a suitable stochastic perturbation in the equations for the physical or auxiliary variables (e.g. using Langevin dynamics [2] or the recently proposed Nosé-Hoover Langevin (NHL) method [3]). We describe all of these methods in the next section.

There are certain cases where one wishes to sample a canonical distribution, but this goal is frustrated by anomalous heating effects. A particular system of this type served as inspiration for the work presented here: the use of a Quantum-Drude model and Path Integral Molecular Dynamics [4]. In such systems, as in many other Car-Parinello type models [5], it is common to employ two temperatures, in this case hot Q-Drudes and cold atoms, resulting in heat flow between the two components. The artificial adiabatic separation of the two sets of degrees of freedom can allow a stable evolution; in numerical practice the electronic degrees of freedom are propagated using a smaller timestep than the nuclei of the atoms. The simulation of such non-equilibrium (open) systems is a wide and important area of research, which is currently much less well-understood than is the equilibrium case. Employing a Nosé-Hoover thermostat allows the excess heat to be removed from the physical system and deposited into the artificial reservoir (with a consequent steady increase in the artificial variable). Unfortunately when NHC, NHL or Langevin methods are employed in the

same manner, they do not reproduce the intended target temperature and the distributions obtained are modified, often in a complicated way which does not facilitate recovery of correct thermodynamic averages. Thus there is a fundamental challenge to combine the nonequilibrium sampling property of Nosé-Hoover with devices that provide ergodicity in those models that require it.

A number of other authors have encountered this problem [6–8] and some have proposed solutions. For example [6] introduces a complicated energy control into Langevin dynamics for two-temperature simulations. Brańka[7] discussed this problem in the case of Nosé-Hoover Chains and produced a particular solution that can be used only in cases where the heating is uniform and the rate can be precisely known beforehand. Here we propose new thermostats (Ad-NHC, Ad-NHL and Ad-Langevin) that are designed to adapt to cancel average cooling or heating effects, and so maintain a particular temperature in a fluctuating non-equilibrium environment, without a priori knowledge of the anomaly. These methods work by combining an equilibrium scheme with an additional Nosé-Hoover control whose purpose is to shift the ensemble, they are thus of a familiar form and straightforward to discretize. We also discuss the selection of parameters to ensure a rapid local convergence to statistical equilibrium. We compare the different thermostats in simple models. Our results suggest that all of the described methods can work well in the setting of a uniform heating of the degrees of freedom of the system, but only Ad-Langevin, with a large collision coefficient, can provide the strong internal equilibration property needed to control the distribution when the heating is rapid and non-uniform.

Our interest in this paper is in cases where it is specifically desired to force the dynamics to sample the Gibbs distribution but this is impeded due to an incorporated stochastic perturbation, coupling to another scale regime, limitations in the model formulation, or issues with the numerical methods used. In other applications it may be desired to allow the system to evolve far from equilibrium with e.g. only a control on temperature, in which case a different perspective should be taken than that put forward here.

II. THERMOSTATS FOR MOLECULAR SAMPLING

As a prelude to the results presented in this article, we describe below several thermostatting methods in common use. We adopt the following notation: $q_1, q_2, \ldots, q_{3N}$ represent the

position coordinates describing the system, $p_1, p_2, \ldots, p_{3N}$ are the corresponding momenta, $F_1, F_2, \ldots, F_{3N}$ the forces acting in each coordinate direction, k_B is Boltzmann's constant and T is the temperature. We write m_i for the mass coefficient associated to the ith coordinate, thus these are repeated in groups of 3. $K = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i}$ is the kinetic energy.

Nosé-Hoover Dynamics (NH)

$$\begin{aligned} \dot{q}_i &= \frac{p_i}{m_i}, \qquad i = 1, 2, \ldots, N \dot{p}_i = F_i - \xi p_i, \\ \dot{\xi} &= \left[2K - g k_B T\right]/Q. \end{aligned}$$

The system incorporates a coupling to an artificial variable ξ via a coupling parameter Q which is referred to as the thermal mass. g is the number of degrees of freedom in the system, typically $g = 3N$ if their are no conserved quantities besides the energy.

The main advantage of NH is that it is simple, deterministic and easy-to-implement. In the case of a system subject to a driving stochastic perturbation it is possible to show (see Section III) that Nosé-Hoover dynamics can still correctly regulate the system temperature, thus it is suited to non-equilibrium applications. Its significant disadvantage is that it can be shown to be non-ergodic for certain systems, including low dimensional models [9] and larger systems that are nearly harmonic (e.g. nearly an Einstein crystal [10]).

Nosé-Hoover Chains (NHC) of length r

$$\begin{aligned} \dot{q}_i &= \frac{p_i}{m_i}, \qquad \dot{p}_i = F_i - \xi_1 p_i, \qquad i = 1, 2, \ldots, 3N \\ \dot{\xi}_1 &= \left[2K - gkT\right]/Q_1 - \xi_2 \xi_1, \\ \dot{\xi}_2 &= \left[Q_1 \xi_1^2 - kT\right]/Q_2 - \xi_3 \xi_2, \qquad \ldots \quad , \\ \dot{\xi}_r &= \left[Q_{r-1} \xi_{r-1}^2 - kT\right]/Q_r. \end{aligned}$$

The equations incorporate parameters $Q_1, Q_2, \ldots, Q_r$ which represent 'thermal masses' associated to the r auxiliary variables $\xi_1, \xi_2, \ldots, \xi_r$. NHC dynamics [1] was created in an attempt to randomise the dynamics of NH through additional complexity. However, NHC has three major drawbacks. Firstly, it is somewhat complicated to implement, requiring very small timesteps and/or multi-stage integration schemes [11]. Secondly, NHC is not rigorously ergodic. As Hamiltonian dynamics, Nosé-Hoover, and Nosé-Hoover Chains are

all deterministic models, their phase space distributions evolve according to the Liouville equation, a hyperbolic PDE. As no diffusion is present, the effective ergodicity observed in some cases must arise from strong internal mixing that allows a dense coverage of the physical domain by an evolving set of initial conditions. Proving such a property in the absence of diffusion is exceedingly difficult. As we see below, for some choices of the parameters in NHC, a nonergodic behavior may be observed in simple models. Thirdly, and important particularly for this article, the ability of Nosé-Hoover to adapt to heating or cooling effects is lost when a chain is incorporated; it can be restored only if the heating/cooling rate is exactly known [7], but not in a general way, as we will see in the next section.

Langevin Dynamics (LD)

Langevin dynamics supplements the conservative vector field by noise and dissipation as follows:

$$\dot{q}_i \;=\; \frac{p_i}{m_i}\,, \qquad \dot{p}_i \;=\; F_i - \gamma p_i + \sigma \dot{w}_i,$$

where the $w_i(t)$ represent a Wiener independent processes with independent increments satisfying $w_i(0) = 0$, and, for $s < t$, $w_i(t) - w_i(s) \sim \mathcal{N}(0, t-s)$; ($\dot{w}_i$ can be viewed as Gaussian white noise). Relevant topics regarding stochastic differential equations may be found in [12] The coefficients γ and $\sigma = \sqrt{2kT\gamma m}$ are chosen by the fluctuation dissipation theorem in order that the canonical distribution is preserved. Modifications of the standard framework would allow different coefficients for each stochastic-dynamical perturbation.

Nosé-Hoover-Langevin Dynamics (NHL)

In Nosé-Hoover dynamics the thermostatting force is always in the direction of the momentum variable, a feature shared by Nosé-Hoover Chains; it has been argued (Bussi and Parinello [13]) that this feature accounts for its 'gentle' effect on dynamics, but it also may contribute to its lack of ergodicity. Those authors proposed a stochastic modification of kinetic energy which is intended to maintain the mild perturbation of dynamical quantities. An alternative, and in some ways simpler, approach was suggested in [14]. The idea

is just to modify the thermostat variable in Nosé-Hoover dynamics directly by a stochastic perturbation.

The NHL method [3, 14] is given by:

$$\begin{aligned} \dot{q}_i &= \frac{p_i}{m_i}\,, \qquad \dot{p}_i \;=\; F_i - \xi p_i, \\ \dot{\xi} &= [2K - gkT]\,/\mu \;+\; \sigma_\ell\,\dot{w} \;-\; \gamma_\ell\,\xi\,, \qquad \text{with} \quad \gamma_\ell \;=\; \tfrac{1}{2}\mu\sigma_\ell^2/kT, \end{aligned}$$

where $w(t)$ is a Wiener process. Due to the stochastic nature of this system, the parameter μ plays a different, although somewhat related, role to that of the parameter Q in Nosé-Hoover dynamics. The introduction of noise (the strength of which is measured by σ_ℓ) the distribution is in effect smeared out while the dissipative term balances it to maintain the correct temperature; this is analogous to introducing a Langevin thermostat applied to just the auxiliary degree of freedom. NHL dynamics was created in order to address the non-ergodicity of NH and NHC dynamics. NHL is only slightly more complex than NH dynamics, and simpler in structure than NHC. As shown in [3], NHL is ergodic for harmonic model systems, under a mild nonresonance assumption. This ergodicity property is weaker than that which can be shown for Langevin dynamics, in that for Langevin dynamics the ergodic nature does not depend at all on the dynamics of the system to which the Ornstein-Uhlenbeck process is coupled, whereas for NHL proofs are only available for specific systems. One typically observes ergodicity with this method, since internal processes will typically redistribute the fluctuations in energy to the internal degrees of freedom.

Despite its strong ergodic property, as we shall see in the next section, NHL has a similar problem to NHC when it comes to external heating or cooling. In Section IV we will consider the general and adaptive modification of NHL to correct the method in the nonequilibrium setting.

III. APPLICATION OF THERMOSTATTING METHODS TO DRIVEN SYSTEMS

We wish to formulate the schemes mentioned above in the setting of driven systems, by which we mean systems that include coupling to auxiliary, nonconservative, typically stochastic, forces which represent contact through boundaries or incompletely resolved in-

ternal degrees of freedom. Such systems may be written, without thermostat, in the form

$$\dot{q}_i = \frac{p_i}{m_i}\,, \qquad \dot{p}_i = F_i(q_1, q_2, \ldots, q_{3N}) + F_i^{\rm n}(q_1, q_2, \ldots, q_{3N}, t) \tag{1}$$

where F_i is the conservative force acting on the ith coordinate and $F_i^{\rm n}$ represents a generic non-reversible, non-conserving, possibly stochastic perturbation. We will often abbreviate these equations by writing q for the vector of positions, p for the vector of momenta, and F and $F^{\rm n}$ for the force vectors.

What could we logically require of a thermostat for such a model? At a minimum we would hope (i) that the thermostat would preserve (and be ergodic for) the canonical ensemble in case $F^n \equiv 0$. It also seems reasonable to require (ii) that the thermostat correctly adapt to a uniform thermal heating, as when F^n is a pure, constant amplitude white noise stochastic force that models a Brownian impulse. Among the methods given in the previous section, only Langevin and Nosé-Hoover-Langevin are rigorously ergodic for the canonical ensemble, i.e. they satisfy the first property.

On the other hand, although each of the thermostat methods described in the previous section can be naturally adapted to treat the nonequilibrium model (1), only the Nosé-Hoover method satisfies property (ii).

For Langevin dynamics, the excess Brownian heating, being of the same form as the heating introduced in the method itself, will obviously lead to an incorrect distribution (with a shifted temperature). However, with the schemes based on Nosé-Hoover, it is less obvious what happens to the excess heat. We consider each in turn below.

A. Nosé-Hoover Dynamics with Brownian Heating

Assume that Nosé-Hoover is coupled to a one degree of freedom system (energy $H = p^2/2m + U(q)$) subject to a steady heating by Brownian collisions. (The results contained below would naturally extend to systems of many degrees of freedom.) Nosé-Hoover acts to remove this excess heat. In this case it is possible to find a steady phase space distribution by studying the Fokker-Planck equation for the combined scheme.

The equations of motion are

$$\dot{q} \;=\; \frac{p}{m}\,, \qquad \dot{p} \;=\; F(q) - \xi p \;+\; \sigma_{\rm heat}\,\dot{w}_{\rm heat}\,, \qquad \dot{\xi} \;=\; \left[\frac{p^2}{m} - kT\right]/Q,$$

resulting in the Fokker-Planck Equation (subscript q, p, etc. represent partial differentiation with respect to to the indicated variables):

$$\begin{aligned}\rho_t &= -\left[\dot{q}\rho\right]_q - \left[(F(q)-\xi p)\rho\right]_p + \tfrac{1}{2}\sigma^2_{\text{heat}}\,\rho_{pp} - \left[\dot{\xi}\rho\right]_\xi \\ &= -\left[\frac{p}{m}\rho_q + F(q)\rho_p\right] + \tfrac{1}{2}\sigma^2_{\text{heat}}\,\rho_{pp} + \xi\left[p\rho_p + \rho\right] - \frac{1}{\beta Q}\left[\beta\frac{p^2}{m} - 1\right]\rho_\xi.\end{aligned}$$

We are interested in steady states of this equation, i.e. such that $\rho_t \equiv 0$. In the sequel we work with an unnormalized distribution for ease of exposition; normalization would not effect the result. If we assume that the distribution in q and p is separable from the thermostat, i.e. $\rho(q,p,\xi) = \rho^\beta(q,p) \times \hat{\rho}(\xi)$, with $\rho^\beta = e^{-\beta H}$, then the contents of the first brackets term cancel, and we simplify using $\rho_q = -\beta F(q)\rho$, $\rho_p = -\frac{\beta p}{m}\rho$ and $\rho_{pp} = \frac{\beta}{m}\left[\beta\frac{p^2}{m} - 1\right]\rho$ to arrive at the equation

$$\rho_t \equiv 0 = \left[\tfrac{1}{2}\tfrac{\beta}{m}\sigma^2_{\text{heat}}\,\rho - \xi\,\rho - \tfrac{1}{\beta Q}\,\rho_\xi\right]\left[\beta\tfrac{p^2}{m} - 1\right].$$

This tells us that the mean of thermostat variable ξ is shifted, and the distribution is otherwise unaffected:

$$\hat{\rho}(\xi) \;=\; \exp\left[-\tfrac{1}{2}\beta Q\left(\xi - \xi_{\text{heat}}\right)^2\right], \qquad \text{where} \quad \xi_{\text{heat}} \;\equiv\; \tfrac{1}{2}\tfrac{\beta}{m}\sigma^2_{\text{heat}} \;=\; \beta\dot{E}.$$

The implication of this is that Nosé-Hoover dynamics is capable of maintaining the correct temperature under noisy heating. We illustrate this property in Fig. 2 in Section VI, below.

In this model, the stochastic perturbation contacts each degree of freedom in the same way as in Langevin dynamics. Using an argument based on hypoellipticity [15, 16] it is possible to show that Nosé Hoover dynamics is actually ergodic in this case. Our interest, however, is in ultimately applying the schemes to more general dynamical models with complex, restricted, possibly deterministic, nonequilibrium forcings, thus it would be important to have schemes that provide ergodicity under more general types of perturbation.

B. Nosé-Hoover Chains with Brownian Heating

Considering, now, Nosé-Hoover Chains of length 2 (length $r > 2$ would be similar), we derive the Fokker-Planck equation under steady Brownian perturbation. The relevant stochastic differential equations are those given in Section II for NHC, with the force per-

turbed by a term of the form $\sigma_{\rm heat} w_{\rm heat}$. The distribution evolves according to:

$$\begin{aligned}
\rho_t &= -\left[\dot{q}\rho\right]_q - \left[(F(q)-\xi_1 p)\rho\right]_p + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} - \left[\dot{\xi}_1\rho\right]_{\xi_1} - \left[\dot{\xi}_2\rho\right]_{\xi_2} \\
&= -\left[\frac{p}{m}\rho_q + F(q)\rho_p\right] + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} + \xi_1\left[p\rho_p + \rho\right] - \frac{1}{\beta Q}\left[\beta\frac{p^2}{m} - 1\right]\rho_{\xi_1} \\
&\quad + \xi_2\left[\xi_1\rho_{\xi_1} + \rho\right] - \frac{1}{\beta Q_2}\left[\beta Q_1\xi_1^2 - 1\right]\rho_{\xi_2}.
\end{aligned}$$

We again assume a steady product distribution, incorporating a unnormalized Gibbs density in the physical variables, writing

$$\rho(q,p,\xi_1,\xi_2) = \rho^\beta\hat{\rho}(\xi_1,\xi_2), \qquad \rho^\beta = e^{-\beta H}, \qquad \hat{\rho}(\xi_1,\xi_2) = e^{-\beta G(\xi_1,\xi_2)},$$

then introducing these into the right hand side of the Fokker-Planck equation, eliminating the derivatives with respect to q and p, and setting the right hand side to zero yields

$$\begin{aligned}
0 &= \tfrac{1}{2}\sigma_{\rm heat}^2\,(\beta/m)\left[\frac{\beta p^2}{m} - 1\right] \\
&\quad + \xi_1(-\beta p^2/m + 1) - \frac{1}{Q_1}\left[\beta\frac{p^2}{m} - 1\right]G_{\xi_1} \\
&\quad + \xi_2\left[-\beta\xi_1 G_{\xi_1} + 1\right] - \frac{1}{Q_2}\left[\beta Q_1\xi_1^2 - 1\right]G_{\xi_2}.
\end{aligned}$$

Combining all the terms involving p (and noting that G is independent of p) we must have

$$\tfrac{1}{2}\sigma_{\rm heat}^2\,\beta^2 p^2/m^2 - \xi_1(\beta p^2/m) - \frac{1}{Q_1}\left(\beta\frac{p^2}{m}G_{\xi_1}\right) \equiv 0.$$

Thus

$$G(\xi_1,\xi_2) = -\tfrac{1}{2}Q_1\left(\xi_1 - \xi_{\rm heat}\right)^2 + R,$$

where R is some smooth function of ξ_2 only. Reinserting this formula into stationarity equation, we find that this reduces to

$$-\xi_2(-\beta\xi_1 Q_1(\xi_1 - \xi_{\rm heat}) + 1) - \frac{1}{Q_2}(\beta Q_1\xi_1^2 - 1)R'(\xi_2) = 0.$$

We see that it is only possible to satisfy this equation for R independent of ξ_1 if $\xi_{\rm heat} = 0$. The implication of this is that Nosé-Hoover Chains are not capable of adapting to maintain the canonical ensemble, in the case of Brownian heating. The second thermostat in the chain frustrates the action of the first.

Brańka [7] discovered this problem previously, and introduced a simple fix to the dynamics of the second thermostat to take account of the shifted distribution of the first, by adding

the $\xi_{\rm heat}$ term explicitly (B in that paper), but unfortunately it depends on knowing $\xi_{\rm heat}$ a priori.

Nosé-Hoover dynamics on the other hand does not need explicit knowledge of the heating term a priori; it can effectively measure it instead, and, as we shall see later, can even adapt should it change.

C. Nosé-Hoover Langevin with Brownian Heating

As we have just done for NH and NHC, we also consider the NHL scheme under Brownian heating (see Fig. 4). The Fokker-Planck Equation for this system, which has noise in both the physical and artificial variables, becomes:

$$\begin{aligned}\rho_t &= -\left[\dot{q}\rho\right]_q - \left[(F(q)-\xi p)\rho\right]_p + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} - \left[\left((2K-kT)/\mu - \gamma_\ell\xi\right)\rho\right]_\xi + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} \\ &= -\left[\frac{p}{m}\rho_q + F(q)\rho_p\right] + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} + \xi\left[p\rho_p + \rho\right] - \frac{1}{\beta\mu}\left[\beta\frac{p^2}{m} - 1\right]\rho_\xi \\ &\quad + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} + \gamma\left[\xi\rho_\xi + \rho\right] = \mathcal{L}_{\rm NHL}\rho.\end{aligned}$$

Assuming $\rho = \rho^\beta\hat{\rho}(\xi)$ and using the simplifying expressions for ρ_p and ρ_{pp}, we find

$$\mathcal{L}_{\rm NHL}\rho = \left[\tfrac{1}{2}\tfrac{\beta}{m}\sigma_{\rm heat}^2\,\rho - \xi\,\rho - \tfrac{1}{\beta\mu}\,\rho_\xi\right]\left[\beta\tfrac{p^2}{m} - 1\right] + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} + \gamma_\ell\left[\xi\rho_\xi + \rho\right].$$

Once again, the first term tells us

$$\log\hat{\rho}(\xi) = -\tfrac{1}{2}\beta\mu\left(\xi - \xi_{\rm heat}\right)^2,$$

Substituting this into the original equation, we find that the distribution is stationary if

$$\frac{1}{2}\sigma_\ell^2\beta\mu[\beta\mu(\xi-\xi_{\rm heat})^2/2 - 1] + \gamma_\ell[1 - \xi\beta\mu(\xi-\xi_{\rm heat})] \equiv 0.$$

For this quadratic to vanish identically, each of its coefficients must vanish. The constant term yields

$$\gamma_\ell = \frac{1}{2}\sigma_\ell^2\beta\mu.$$

Using this, the remaining conditions can only be satisfied if $\xi_{\rm heat} = 0$.

Therefore Nosé-Hoover Langevin dynamics is also incapable of adapting to maintain the desired distribution under heating, although it fails in a somewhat different way, as is also

evident in Fig. 4 vs Fig. 3. In the case of NHL, the damping term ($\gamma_\ell \xi$) fails to take account of the shift in the distribution of ξ. Analogous to Brańka's fix for NHC[7] we could correct for the problem by a shift of the damping coefficient: $(\gamma_\ell \xi) \to (\gamma_\ell \xi - \xi_{\rm heat})$, but it again requires knowledge of $\xi_{\rm heat}$ a priori, and thus would only be relevant in the simple model setting of steady Brownian heating.

IV. THERMOSTATS ADAPTED FOR DRIVEN SYSTEMS

As we have seen, Nosé-Hoover (NH) dynamics can adapt automatically to heating effects, but Nosé-Hoover-Chains (NHC) or Nosé-Hoover-Langevin (NHL) dynamics fail to do so. As our goal is to provide a generic method that can address nonergodic systems, for example those dominated by harmonic components (which includes QD-PIMD), we would like to simultaneously correct the ergodicity problem while adapting to heating effects.

Here we show that there is a simple way to combine the adaptive property of NH with the ergodic property of NHL, creating a thermostat that is barely more complex to implement, and both flexible and robust. Specifically we propose to augment the system by a second thermostatting variable which is separately adjusted using a Nosé-Hoover-like deterministic dynamics. Our Adaptive-NHL scheme thus takes the form

$$\begin{aligned}
\dot{q} &= \frac{p}{m}\,, \qquad \dot{p} \;=\; F(q) - (\xi+\chi)p \;+\; F^{\rm n}(q,t)\,,\\
\dot{\xi} &= \left[2K - kT\right]/\mu \;+\; \sigma_\ell\,\dot{w}_\ell \;-\gamma_\ell \xi\,,\\
\dot{\chi} &= \left[2K - kT\right]/Q.
\end{aligned}$$

In the case of steady Brownian heating, the Fokker-Planck equation now becomes

$$\begin{aligned}
\rho_t &= -\Big[\dot{q}\rho\Big]_q - \Big[(F(q)-\xi p)\rho\Big]_p + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} - \left[\left(\left[2K-kT\right]/\mu - \gamma_\ell\xi\right)\rho\right]_\xi + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} - \left[\dot{\chi}\rho\right]_\chi\\
&= -\left[\frac{p}{m}\rho_q + F(q)\rho_p\right] + \tfrac{1}{2}\sigma_{\rm heat}^2\,\rho_{pp} + \xi\left[p\rho_p + \rho\right] + \chi\left[p\rho_p + \rho\right]\\
&\quad -\frac{1}{\beta\mu}\left[\beta\frac{p^2}{m} - 1\right]\rho_\xi - \frac{1}{\beta Q}\left[\beta\frac{p^2}{m} - 1\right]\rho_\chi + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} + \gamma\left[\xi\rho_\xi + \rho\right].
\end{aligned}$$

We will show that this system has a steady-state solution which is a product distribution of the form $\rho = \rho^\beta \hat{\rho}(\xi)\tilde{\rho}(\chi)$. With this assumption, the right hand side may be written

$$\mathcal{L}_{\rm Ad-NHL}\rho \;=\; \left[\tfrac{1}{2}\sigma_{\rm heat}^2\,\rho - \xi\,\rho - \chi\,\rho - \frac{1}{\beta\mu}\rho_\xi - \frac{1}{\beta Q}\rho_\chi\right]\left[\beta\frac{p^2}{m} - 1\right] + \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} + \gamma_\ell\left[\xi\rho_\xi + \rho\right],$$

we can rearrange the expression and show it is separable:

$$\begin{aligned}\mathcal{L}_{\text{Ad-NHL}}\rho \;=\; & \left[-\xi\rho - \frac{1}{\beta\mu}\rho_\xi\right]\left[\beta\frac{p^2}{m} - 1\right] \;+\; \tfrac{1}{2}\sigma_\ell^2\,\rho_{\xi\xi} + \gamma_\ell\left[\xi\rho_\xi + \rho\right] \\ & + \left[-\chi\rho + \chi_{\text{heat}}\rho - \frac{1}{\beta Q}\rho_\chi\right]\left[\beta\frac{p^2}{m} - 1\right].\end{aligned}$$

With the proposed simple form, we must have

$$\begin{aligned}\hat{\rho}(\xi) \;&=\; \exp\left[-\tfrac{1}{2}\beta\mu\,\xi^2\right], \\ \tilde{\rho}(\chi) \;&=\; \exp\left[-\tfrac{1}{2}\beta Q\,(\chi - \chi_{\text{heat}})^2\right], \qquad \text{where} \quad \chi_{\text{heat}} \;\equiv\; \tfrac{1}{2}\tfrac{\beta}{m}\sigma_{\text{heat}}^2.\end{aligned}$$

Essentially, we have divided the original auxiliary variable in NHL into two parts: ξ, which behaves as a NHL thermostat and promotes ergodicity, and χ, which behaves as an NH thermostat and allows an adaptive adjustment in the presence of stochastic heating.

It is not essential to restrict to the stochastic case. An Ad-NHC method can be constructed in an entirely analogous way, i.e. by combining Nosé-Hoover Chain and Nosé-Hoover thermostats in a single united framework. In this case, it is possible to demonstrate a steady state consisting of the NHC steady state multiplied by $\tilde{\rho}(\chi)$.

Numerical discretization of the Ad-NHL and Ad-NHC methods is based on the corresponding implementations for the original NHL and NHC methods, as the computations of ξ and χ are independent of each other. Note that the extension of either method to a system with an arbitrary number of degrees of freedom is straightforward, but the theoretical property demonstrated above will only hold in case the heating is uniform in all the degrees of freedom, so that a single uniform translation is obtained for the Gaussian distribution of χ. Alternatively, if the system is subject to a nonuniform heating, then the system would need to rapidly equilibrate this perturbation in order for the Nosé-Hoover variable to correctly compensate for it. We examine this behavior later in the section on numerical experiments. Numerical methods for Ad-NHL and Ad-NHC are given in Appendix A.

It is clear that the Langevin method, in which the parameters are frozen a priori, will be unable to dissipate excess heat adaptively. When the noise is introduced in a uniform way, we have already seen in the previous section that the Nosé-Hoover method effectively adapts the Langevin damping coefficient to preserve the canonical distribution. However, in many cases the heat may be introduced in a subset of the degrees of freedom, and may

not be present at all times. The idea of our adaptive Langevin method is just to combine Langevin and Nosé-Hoover thermostats to adaptively handle many separate situations in a single formulation.

The method is as follows:

$$\begin{aligned} \dot{q} &= \frac{p}{m}, \qquad \dot{p} = F(q) - (\gamma + \chi)p + F^{\mathrm{n}}(q,t) + \sigma \dot{w}, \\ \dot{\chi} &= [2K - kT]/Q, \end{aligned}$$

where $w(t)$ is a Wiener process and $\sigma = \sqrt{2kT\gamma m}$. An ambient stochastic bath is combined with an adaptive temperature regulator to allow thermostatting of a wide variety of systems. The proof that this method correctly compensates for uniform Brownian heating is already found in the previous section (where we considered the Nosé-Hoover method with steady heating). A numerical method for Ad-Langevin is given in Appendix A.

V. PARAMETER SELECTION FOR AD-NHL

In general, it is difficult to fully analyze the statistical convergence behavior in a complex molecular model. However, as described in [17] in the case of NHL, if one assumes that the system of interest is near to thermal equilibrium, it is possible to make assumptions which allow for computation of the convergence rate, specifically the rate of convergence of the kinetic energy. We perform such an analysis here for the Ad-NHL scheme which is slightly more complicated than NHL and reveals the different roles of the parameters μ and Q in the method, and the interplay between them. We perform this analysis under the assumption of null nonequilibrium forcing, $F^n \equiv 0$. Our purpose here is to obtain values for the parameters which ensure a good convergence rate of averages, and also to shed some light on the way in which equilibrium is achieved.

It should be emphasized that parameter selection in the case of Ad-NHL relies on statistical assumptions which are likely to hold in the vicinity of equilibrium. These assumptions allow us to give precise ranges for parameters for which a certain rate of convergence would be expected. Importantly, due to the presence of the stochastic terms, the parameters may be chosen independently of the need in deterministic thermostats such as Nosé-Hoover to achieve a resonance with the underlying dynamics of the physical system under study.

At any time t, the phase variables and auxiliary variables are distributed according to the

distribution $\rho(q,p,\xi,\chi,t) = \rho(\cdot,t)$. For a suitable function $f = f(\cdot)$ of the state variables, we will denote by $\mathbb{E}^t f$ the expectation with respect to the unnormalized density $\rho(\cdot,t)$, i.e. a spatial averaging defined by

$$\mathbb{E}^t f := \langle f \rangle_{\rho(\cdot,t)} = \frac{\int_\Omega f(\cdot)\rho(\cdot,t)d(\cdot)}{\int_\Omega \rho(\cdot,t)d(\cdot)},$$

where Ω represents the product of phase space and R^2. Thus $\mathbb{E}^t f$ can be viewed as a time-dependent quantity. Define

$$x = x(t) = \mathbb{E}^t(2K - gkT) = \mathbb{E}^t \sum_{i=1}^N \frac{p_i^2}{m} - gkT.$$

In order to write simplified equations for the evolution of expectation, we need to make a near-equilibrium assumption. For a system with energy in the standard form $H(q,p) = K(p) + U(q)$, positions and momenta are statistically independent, thus $\mathbb{E}^t f(q)g(p) = \mathbb{E}^t f(q)\mathbb{E}^t g(p)$, but this independence is not guaranteed away from equilibrium conditions. We suppose that for a given system sufficiently near equilibrium, the distribution can be viewed as canonical with an incorrect temperature $T(t) \approx T$, with, for Ad-NHL, the unnormalized density $\rho = e^{-\beta(t)H}\bar{\rho}(\xi,\eta,t)$, $\beta(t) = (kT(t))^{-1}$. This is a very strong assumption. However, the simulations of [17] on a Lennard-Jones system confirm the analysis presented here is applicable in that case. It is of course possible to imagine slow converging relaxation processes for which the assumption would be invalid, but this assumption allows us to formulate an elementary analysis of the convergence rate for temperature, isolating the regulating effect of the auxiliary device.

With the indicated assumption, we have that $\mathbb{E}^t f(q)g(p) = \mathbb{E}^t f(q)\mathbb{E}^t g(p)$, and, moreover $\mathbb{E}^t pF(q) = \mathbb{E}^t p\mathbb{E}^t F(q) = 0$, as H is even in p.

From the Ad-NHL equations of motion, with no driving terms, we have

$$\dot{K} = \sum_{i=1}^N p_i\dot{p}_i/m_i = \sum_{i=1}^N p_i[F_i - (\xi+\chi)p_i]/m_i.$$

Using the near-equilibrium assumption,

$$\mathbb{E}^t\left(p_i F_i(q)\right) = \mathbb{E}^t p_i \cdot \mathbb{E}^t F_i, \qquad \mathbb{E}^t\left(\xi K\right) = \mathbb{E}^t \xi \cdot \mathbb{E}^t K,$$

hence,

$$\dot{x} = 2\mathbb{E}^t\dot{K} = -2(\mathbb{E}^t\xi + \mathbb{E}^t\chi)2\mathbb{E}^t K.$$

Define $y = \mathbb{E}^t\xi$ and $z = \mathbb{E}^t\chi$, then as the expectation of a Brownian increment is zero, we arrive at the system of equations

$$\begin{aligned}\dot{x} &= -2(y+z)(x+gkT),\\ \dot{y} &= x/\mu - \gamma_\ell y,\\ \dot{z} &= x/Q.\end{aligned}$$

The equilibrium state is characterized by $x = y = z = 0$. Therefore, we linearize the nonlinear equations at this point to arrive at:

$$\frac{d}{dt}\begin{bmatrix}\delta x\\ \delta y\\ \delta z\end{bmatrix} = \begin{bmatrix}0 & -2gkT & -2gkT\\ \mu^{-1} & -\gamma_\ell & 0\\ Q^{-1} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta x\\ \delta y\\ \delta z\end{bmatrix}.$$

The eigenvalues are solutions of the characteristic equation

$$p(\lambda) = \lambda^3 + \gamma_\ell\lambda^2 + 2(\hat{\mu}^{-1} + \hat{Q}^{-1})\lambda + \hat{Q}^{-1}\gamma_\ell = 0.$$

where $\hat{\mu} = \mu/2gkT$, and $\hat{Q} = Q/2gkT$. The eigenvalues are all in the left half plane for all values of the parameters, but for some choices of γ_ℓ, μ, Q, there may be weak damping. Indeed, unlike in the case of Langevin dynamics, a very large value of γ_ℓ will be counterproductive with some eigenvalues close to the imaginary axis. In Appendix B we have carefully analyzed the eigenvalues of the matrix and obtained conditions for certain damping behavior to be realized.

Our investigation shows that there is an optimal choice of the parameters to achieve the most rapid convergence ('critical damping') similar to the case of NHL dynamics, but in the Ad-NHL method (and unlike for NHL), there is not guaranteed to be an interval of the damping parameter which places all eigenvalues on the negative real axis, unless the ratio μ/Q is sufficiently small.

Viewing the eigenvalues as functions of the parameter γ_ℓ they exhibit one of the two behaviors shown below in Figure 1.

Based on Figure 1 and the analysis of the appendix, we propose two representative choices for parameterization, as detailed in Table I. We stress that these are only examples of suitable parameters.

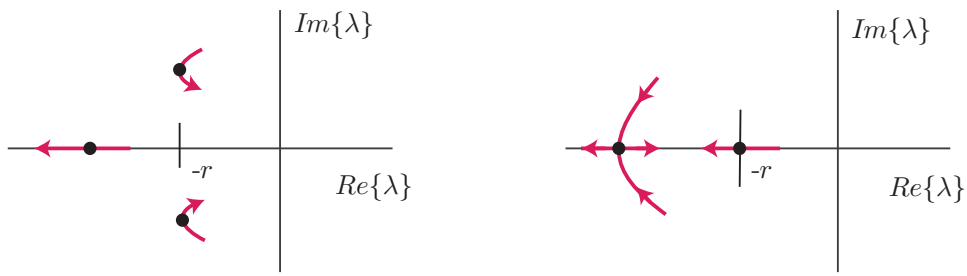

FIG. 1: Eigenvalues of the linearized system of ordinary differential equations obtained by averaging the Nosé-Hoover-Langevin method over the evolving distribution. The arrows indicate the change in the eigenvalues as γ_ℓ is increased. At left, the eigenvalue branches remain separated; hence the optimal choice corresponds to the point of maximal dissipation. At right, the complex eigenvalue pair meets on the negative real axis.

convergence behavior	Q	μ	γ_ℓ
damped oscillation w./ amplitude e^{-rt}	$\frac{gkT}{2r^2}$	$0.5Q$	$4r$
stable node converging as e^{-rt}	$\frac{2gkT}{r^2}$	$.103Q$	$5.8r$

TABLE I: Table of parameter choices for Ad-NHL.

The calculation described above can be extended to the case of Brownian heating, without any difficulty, and even to the more general case, provided we may assume

$$\mathbb{E}^t p_i F_i^n = 0.$$

We emphasize that optimal choice of parameters might well be to allow a small damped oscillation, since in this case the damping rate is improved for larger values of the thermal masses. Moreover, the choice should be informed by consideration of the numerical methods used for simulating the system. (We have already seen that small thermal masses may lead to less stable numerical simulations.)

This type of analysis cannot be used to choose parameters in Nosé-Hoover or NHC, since, in those cases, even under an assumption of statistical independence (an odd concept in the setting of a deterministic method!) one finds that the system for the expectations does not have an asymptotically stable equilibrium point consistent with thermal equilibrium, for any choice of the coefficients.

Using a similar methodology as described above, we derive a 2-dimensional system to understand the convergence of kinetic energy in Langevin dynamics in the near equilibrium regime:

$$\begin{aligned} \dot{x} &= -2(y+\gamma)(x+gkT), \\ \dot{y} &= x/Q, \end{aligned}$$

where x is as before and $y = \mathbb{E}^t\chi$. Linearizing, we have the system $\frac{d}{dt}[\delta x, \delta y]^T = A[\delta x, \delta y]^T$, where

$$A = \begin{bmatrix} -2\gamma & -2gkT \\ Q^{-1} & 0 \end{bmatrix}.$$

The eigenvalues are

$$\lambda = \frac{-2\gamma \pm \sqrt{4\gamma^2 - \hat{Q}^{-1}}}{2}, \qquad \hat{Q} = \frac{Q}{2gkT}.$$

For large γ these are real, but one tends to the origin. Critical damping occurs when

$$\gamma = \frac{1}{2}\sqrt{\frac{2gkT}{Q}},$$

which gives a double eigenvalue at $-\gamma/2$. This gives $\gamma = 2r$, $Q = \frac{gkT}{2r^2}$ for convergence with exponential rate r. Of course it is possible that γ, which is the collision parameter of the Langevin dynamics, has been provided from physical considerations. In this case, our formula gives a value of Q for the kinetic energy device.

In practice (see Section VI), we found that in cases where the system is subject to nonuniform heating, the key issue was whether externally supplied heat is sufficiently rapidly equilibrated to the various degrees of freedom, thus it is the internal relaxation rate that is crucial. For Ad-Langevin we can estimate this by ignoring the Nosé-Hoover kinetic device and focussing just on the Langevin dynamics; the rate for internal temperature convergence is proportional to γ (and is independent of Q) in the case of Ad-Langevin. This rate is very difficult to predict in the case of Ad-NHL, since it requires knowledge of the spectral gap associated to the Fokker-Planck equation which is problem dependent.

VI. RESULTS FROM SIMULATION

In this section we examine the techniques discussed within this article for a harmonic oscillator with stochastic driving force showing the relative ability of each method to dissipate

external heating. We also consider two examples that are meant to test the ability of the schemes to adapt to more challenging (and probably more realistic) perturbations: a model subject to time-dependent external heating, and a nonlinear model problem testing the internal equilibration in response to a perturbation. These are small dimensional problems, however the properties being tested here would be generic in a wide range of models.

A. Steady uniform heating

To begin, we verified the theoretical results presented in Sections III and IV by applying the various methods discussed to thermostat a harmonic oscillator subject to steady Brownian heating. Parameters $Q, Q_1, Q_2, \mu, \gamma_\ell$ were all taken to be one; as the system is naturally ergodic under uniform heating, their exact choice was not very critical to obtain the results shown. Figures 2, 3, and 4 show the performance of Nosé-Hoover, Nosé-Hoover Chains (of length 2 or 5), and Nosé-Hoover Langevin, demonstrating that only the first is able to correctly recover the Gaussian momentum distribution at the prescribed temperature.

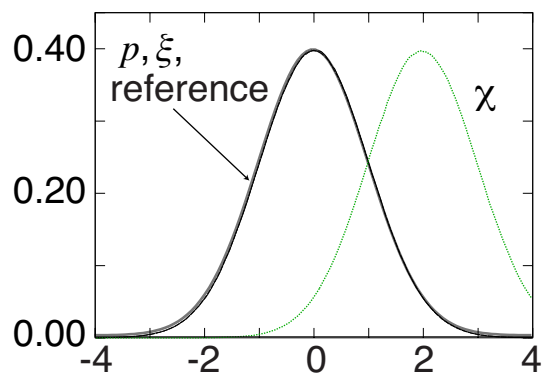

FIG. 2: Nosé-Hoover dynamics has been applied to the harmonic oscillator with a brownian stochastic perturbation. The observed distribution of the thermostat variable ξ (dotted line) is correctly shifted in such a way as to remove excess heat, giving the correct temperature (seen here as the variance of p). The average rate of heat-damping can be seen in the position of the peak of ξ (in units of kT).

Figure 5 shows that the corresponding Ad-NHL and Ad-NHC methods both recover the correct distributions for this model. We saw little difference in our simulations between these methods, up to the issues of reliability and robustness for certain parameter choices for equilibrium cases discussed in Section II.

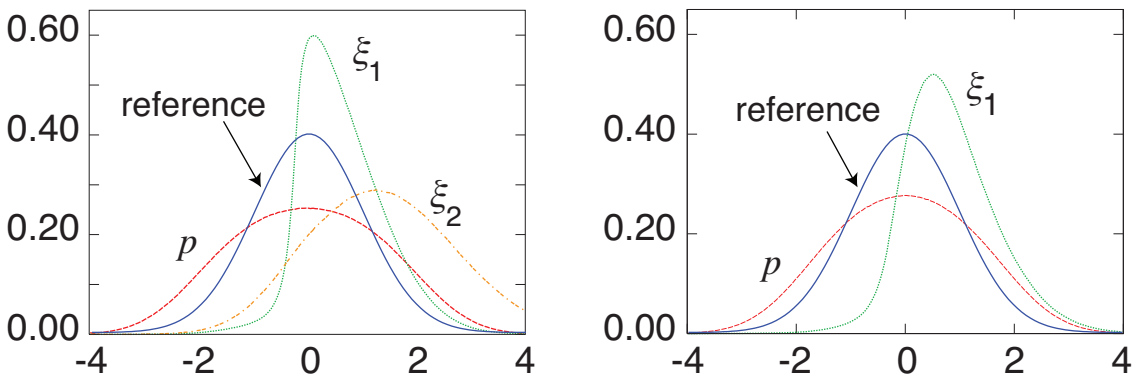

FIG. 3: Nosé-Hoover Chains of length 2 and 5 were applied to the driven harmonic oscillator and can be seen to fail to respond correctly to heating. While the first thermostat variable ξ_1 has a tendency to damp heat, it is frustrated by the action of the second thermostat ξ_2. As a result too much heat remains in the system and this is reflected in the widening of the momentum (p) distribution.

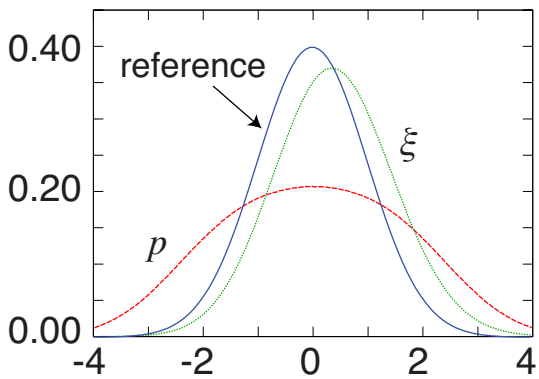

FIG. 4: NHL does not respond correctly to heating. As for NHC, the thermostat, which is designed to give ergodic sampling at the target temperature specified in the control law, is unable to adapt to the presence of a driving stochastic perturbation. Note that the observed distributions of ξ in NHL (dotted line) and ξ_1 in NHC (Fig. 3) are different, indicating that the way the two methods fail is subtly different.

B. Time dependent driving force

To test of the ability of the method to respond to a variable rate of heating, we drove the simple oscillator with a time-dependent term of the form $\sigma_{\text{heat}}(t)\dot{w}_{\text{heat}}$, i.e. the amplitude of the noise process depends on time. When this test was performed with a smooth function such as $\sigma_{\text{heat}}(t) = 2(1 + \sin \omega t)$, the open variant methods gave reasonable results, although it lost some accuracy in the distibution of the position variable ("twice removed" from the stochastic perturbation), and in both variables when the driving frequency was increased. Graphs are shown in Fig. 6 for Ad-NHL. (Ad-NHC and Ad-Langevin results were similar.)

Problems became apparent when the smooth noise amplitude was replaced by a square

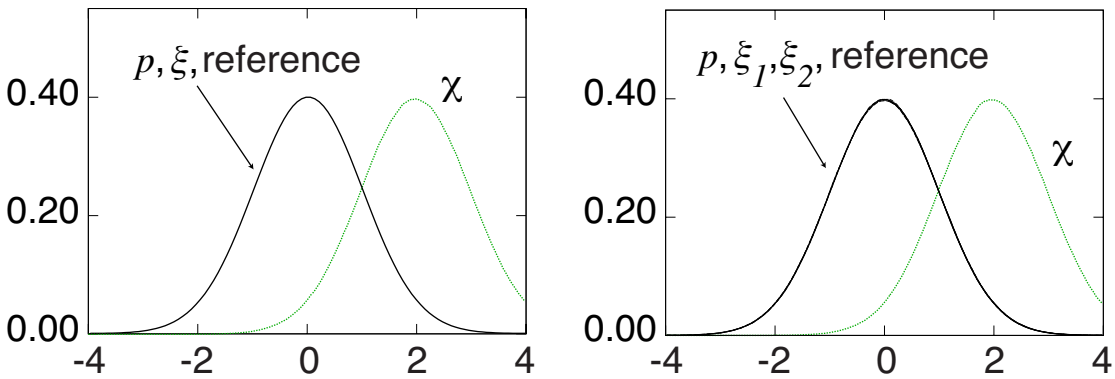

FIG. 5: The adaptive NHL (Ad-NHL) at left responds correctly to Brownian heating, as intended. The distribution of ξ is not distorted because heat is dissipated via the additional thermostat variable χ. The heating rate can be read off the peak of the distribution of χ (in units of kT). In a similar way, Ad-NHC methods with chains of length 2 (shown) or 5 respond correctly to steady heating.

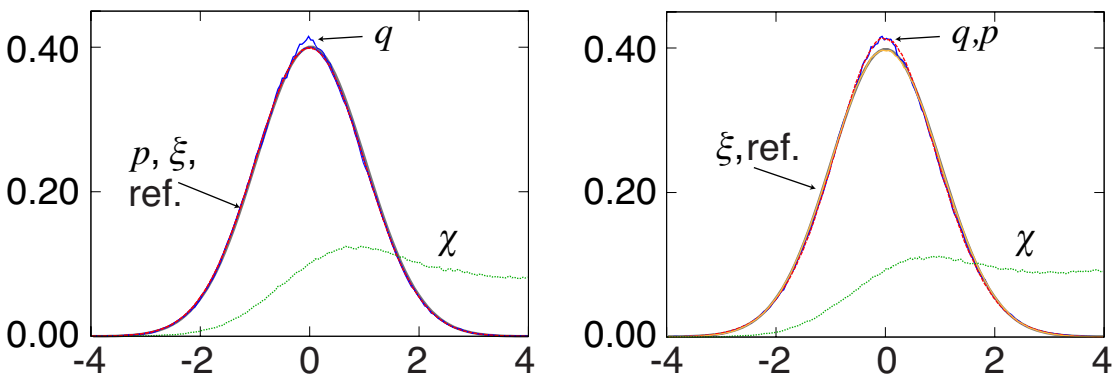

FIG. 6: Ad-NHL was able to approximately maintain the canonical distribution in the presence of smooth driving perturbations, with minor defects visible in the distribution for q for both slow and fast driving, and slightly more significant defects as the system is driven more rapidly.

wave pulse, with period 4τ, where

$$\sigma_{\text{heat}}(t) = \begin{cases} 0, & 0 \le t < 3\tau \\ 4k_BT, & 3\tau \le t < 4\tau \end{cases}$$

Then, depending on time parameter τ, very different behaviors were observed.

We applied each of Ad-NHL, Ad-NHC and Ad-Langevin to this problem. The Ad-NHL results are shown at the left in Fig. 7 for $\tau = 100$. In this case it appears that the method is able to maintain an approximately correct distribution as enough time is available for equilibration following driving events, with a slight deviation from the Gaussian distribution just apparent. When the period between driving pulses is reduced ($\tau = 10$) the thermostat fails to cope and the distribution forms a sharp nonphysical peak. Similar results

are obtained for the Ad-NHC method, as shown in Fig. 8. By comparison, the Ad-Langevin method ($\gamma = 1$) gives a much better distribution for both slow and fast pulses (see Fig. 9).

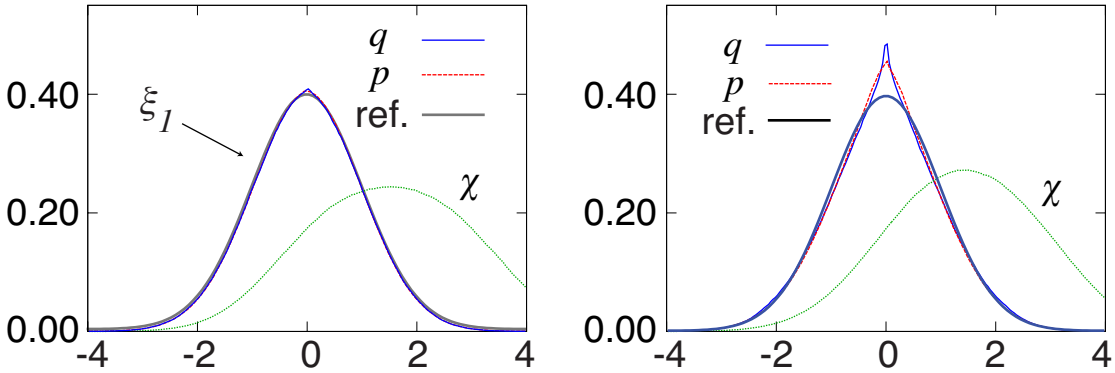

FIG. 7: Ad-NHL distributions for slow (left) and fast (right) square wave pulsed driving forces. When the stochastic driving force is sufficiently slow, Ad-NHL is able to adjust the distribution by eliminating excess heat. For fast driving forces the results are poor.

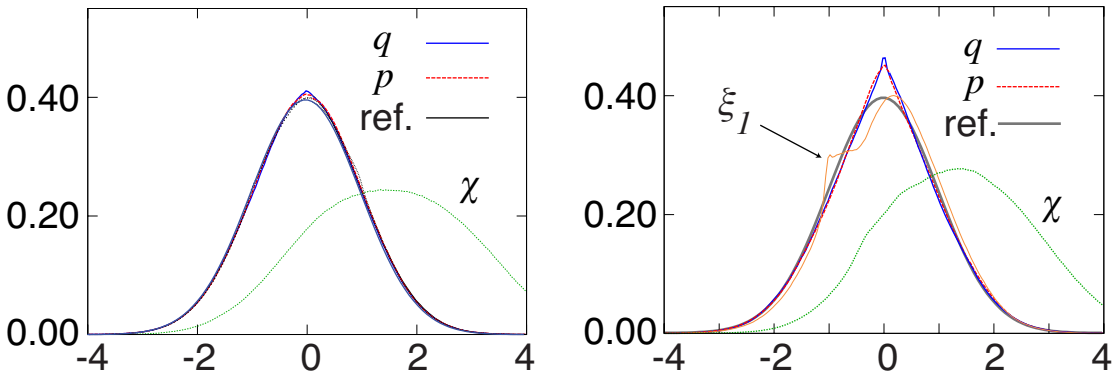

FIG. 8: Ad-NHC distributions for slow (left) and fast (right) square wave pulsed driving forces. The results here are similar to those of 7, i.e. the method does not seem to be able to cope with a rapid pulsed stochastic driving term.

C. Nonlinear model problem

Finally, we considered a 2-dof model of a pendulum with Hamiltonian

$$H(q,p) = p_1^2/2 + p_2^2/2 + \frac{K}{2}((q_1^2 + q_2^2)^{1/2} - 1)^2.$$

When K is large, the system oscillates in the vicinity of the unit circle, with both vibrational and rotational motions visible. Into this model, we add a Brownian perturbation in just one

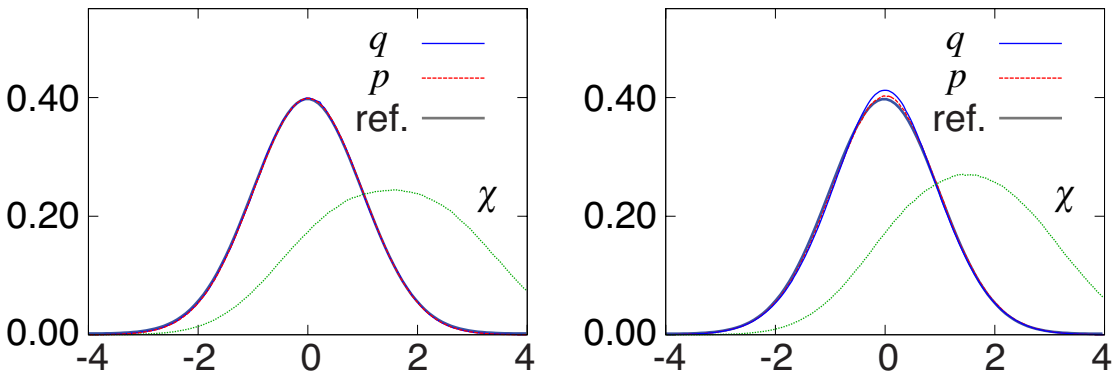

FIG. 9: Ad-Langevin distributions for slow (left) and fast (right) square wave pulsed driving forces. The results using Ad-Langevin are much better than for either of the Nosé-Hoover based methods.

of the two degrees of freedom

$$F_{\text{heat}} = \begin{bmatrix} 0 \\ \sigma_{\text{heat}} \dot{w}_{\text{heat}} \end{bmatrix}.$$

This creates two problems for the thermostat: it must correctly equilibrate the two degrees of freedom while dissipating the excess heat.

We set $K = 50$ and $\sigma_{\text{heat}} = 1$. The target temperature was $kT = 1$. The Ad-NHL method was applied to this case, first with parameters for a damped oscillation of kinetic energy with rate $r = 1$ (we chose $Q = 1, \mu = 0.5, \gamma = 4$). The resulting incorrect momentum distributions (for p_1 and p_2) are shown at left in Figure 10. The distributions for χ and ξ are seen at right.

We then repeated the simulation with parameters corresponding to the critical damping condition ($Q = 4, \mu = .103Q, \gamma = 5.8$), and for a number of other cases, with more or less similar poor results as in the first case. In no case were we able to get correct distributions of the physical variables. Apparently the Ad-NHL method is not able to reallocate the excess heating in the q_2 variable rapidly enough so heat builds up in this variable.

The Ad-Langevin thermostat proved to be more effective in this application, when properly tuned. We initially used the parameter selection suggested in Section V, with $\gamma = 2$, $Q = 0.25$ which should give an exponential rate of $r = 1$ for the convergence of kinetic energy. However, these choices proved to be inadequate to provide a proper distribution in the setting of the anisotropic heating (see the left panel of Figure 11. Only when a much larger value of $\gamma = 20$ was used, did we see a good control of equipartition between q_1 and q_2 (as shown in the right panel of Figure 11). The configurational distributions for Ad-Langevin

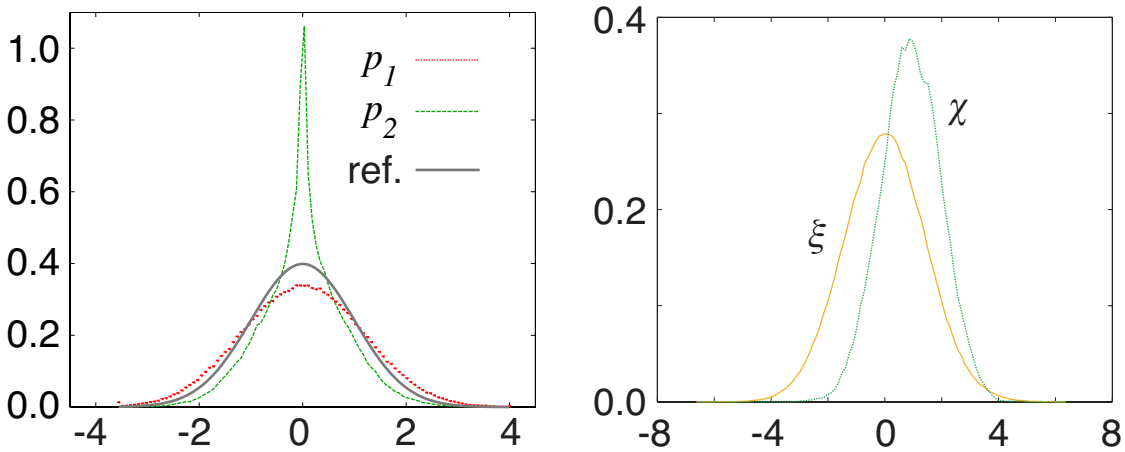

FIG. 10: Left: momentum distributions in the stiff pendulum model computed using Ad-NHL for $Q = 1, \mu = 0.5, \gamma = 4$. The method was unable to dissipate the excess heat. Right: the distributions for χ and ξ. Note that the χ distribution correctly reflects the shift as expected, but this is not enough to give the correct equilibrium distribution.

at $\gamma = 20$ and the χ-distribution are shown in Figure 12. When γ is large, the formula for the eigenvalues derived in Section V suggests that one eigenvalue is near γ and the other is near 0. The first eigenvalue has to do with the rate of internal equilibration of the system at equilibrium; the second controls the thermostat variable itself. Our observation suggests that the problem in this example is not with the adjustment of χ, but with obtaining a sufficiently rapid equilibration between the two degrees of freedom so that the model problem discussed in Section III is relevant. When this is achieved, the adjustment by a uniform shift of χ is sensible and leads to the desired equilibrium distribution.

VII. CONCLUSION

We have demonstrated that Nosé-Hoover dynamics is adaptive (i.e. can adapt to a uniform stochastic perturbation of arbitrary strength while retaining a target canonical distribution) although it is not ergodic. On the other hand the rigorously ergodic Langevin and Nosé-Hoover-Langevin dynamics, as well as the popular Nosé-Hoover-Chains of length 2, are not adaptive. Ad-Langevin, Ad-NHL and Ad-NHC provide alternatives for nonequilibrium modelling with provable Gibbsian stationary solutions in the case of steady Brownian heating. Using numerical experiments we have also shown that each of these methods able to

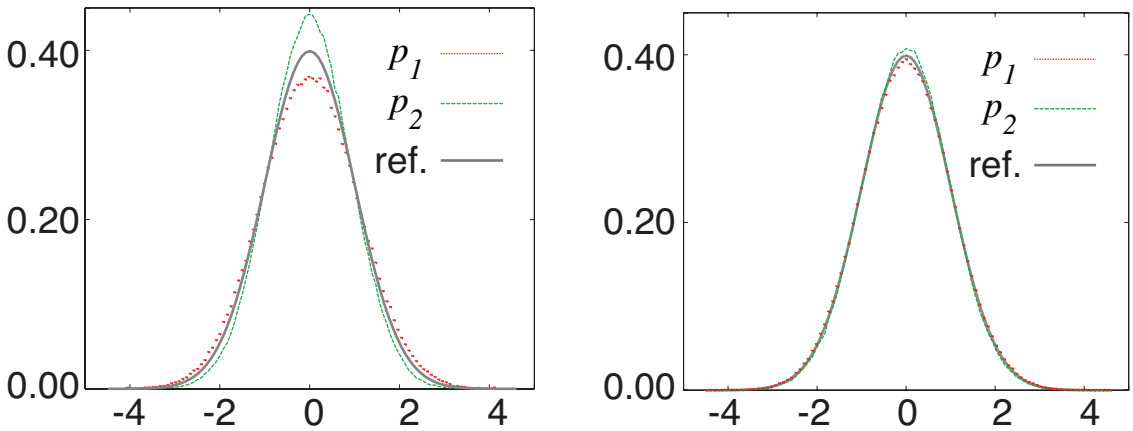

FIG. 11: Momentum distributions computed using Ad-Langevin. Left: $\gamma = 2$, Right: $\gamma = 20$. Ad-Langevin can preserve the canonical distribution if a large collision coefficient γ is used, effecting a rapid internal equilibration.

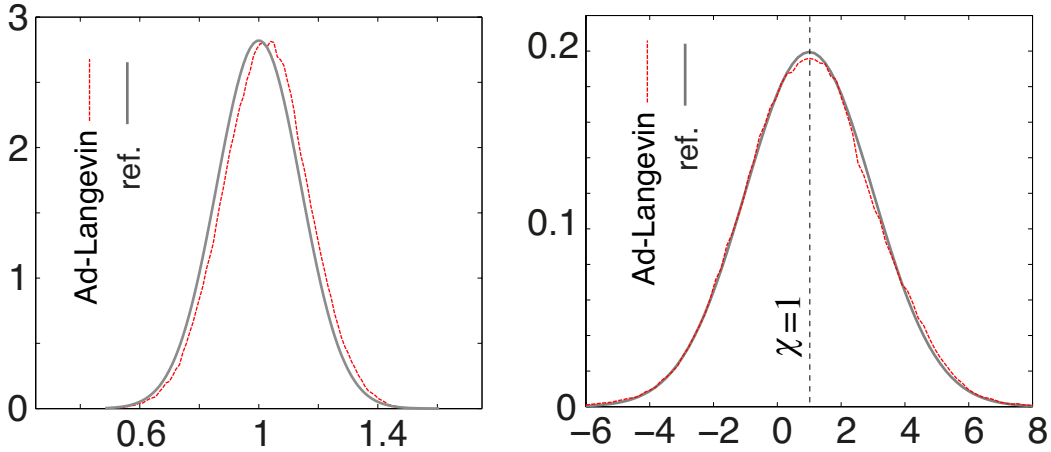

FIG. 12: Distributions for $\gamma = 20$ in Ad-Langevin. On the left: radial distribution $\rho \propto e^{-\beta(r-1)^2/2}$ compared to the exact, right: χ distribution $\tilde{\rho}(\chi) \propto e^{-\beta\mu(\chi-\chi_{\rm heat})^2/2}$. Good approximations of all the distributions were obtained.

adapt to steady or slowly-varying stochastic heating. Only the Ad-Langevin method coped well with a square wave pulse or an anisotropic perturbation in the experiments shown here, and the other adaptive methods were unable to sample the target distribution in these cases. Future work should be performed to gain a deeper theoretical understanding of this behavior.

Acknowledgements

We acknowledge relevant and valuable input from Glenn Martyna and Michael Tretyakov. Partial support for A. Jones's work was made possible by funding from the EPSRC Centre for Numerical Algorithms and Intelligent Software under grant EP/G036136/1.

---

[1] G. Martyna, M. Klein, and M. Tuckerman. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635, August 1992.

[2] A. Brünger, C. B. Brooks, and M. Karplus. Stochastic boundary conditions for molecular dynamics simulations of st2 water. *J. Chem. Phys. Lett.*, 105(5):495–500, 1984.

[3] B. Leimkuhler, N. Noorizadeh, and F. Theil. A gentle stochastic thermostat for molecular dynamics. *J. Stat. Phys.*, 135(2):261–277, 2009.

[4] A. Jones. *Quantum Drude oscillators for accurate many-body intermolecular forces.* PhD thesis, University of Edinburgh, UK, 2010.

[5] D. Marx. An introduction to ab initio molecular dynamics dimulations. *Computational Nanoscience: Do It Yourself!*, 31:195–244, 2006.

[6] C. Phillips and P. Crozier. An energy-conserving two-temperature model of radiation damage in single-component and binary lennard-jones crystals. *J. Chem. Phys.*, 131:074701, 2009.

[7] A.C. Brańka. Nosé-hoover chain method for nonequilibrium molecular dynamics simulation. *Phys. Rev. E*, 61:47694773, 2000.

[8] S.P.A. Gill. Nonequilibrium molecular dynamics and multiscale modeling of heat conduction in solids. In Traian Dumitrica, editor, *Trends in Computational Nanomechanics*, volume 9 of *Challenges and Advances in Computational Chemistry and Physics*, pages 85–134. Springer Netherlands, 2010.

[9] F. Legoll, M. Luskin, and R. Moeckel. Non-ergodicity of the Nose-Hoover thermostatted harmonic oscillator. *Arch. Rational Mech. Anal.*, 184:449–463, 2007.

[10] D. Quigley. *Constant Pressure Langevin Dynamics: Theory and Application to the Study of Phase Behaviour in Core-Softened Systems.* PhD thesis, University of York, UK, 2005.

[11] G. Martyna, M. Tuckerman, D. Tobias, and M. Klein. Explicit reversible integration algorithms for extended systems. *Mol. Phys.*, 87:1117–1157, 1996.

[12] B. Oksendahl. *Stochastic differential equations: an introduction with applications.* Springer, 6th edition, 2010.

[13] G. Bussi and M. Parrinello. Stochastic thermostats: comparison of local and global schemes. *Computer Physics Communications*, 179:26–29, 2008.

[14] A. Samoletov, M.A.J. Chaplain, and C.P. Dettmann. Thermostats for "slow" configurational modes. *J. Stat. Phys.*, 128:1321–1336, 2007.

[15] S. P. Meyn and R.L. Tweedie. Markov chains and stochastic stability. *Springer-Verlag, London,*, 1993.

[16] J.C. Mattingly, A.M. Stuart, and D.J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[17] B. Leimkuhler, N. Noorizadeh, and O. Penrose. Comparing the efficiencies of stochastic isothermal molecular dynamics methods. *J. Stat. Phys.*, page to appear, 2011.

[18] S. Nosé. A unified formulation of the constant temperature molecular dynamics method. *J. Chem. Phys.*, 81:511–519, 1984.

Appendix

A. Numerical Methods

The methods consist of Hamiltonian dynamics supplemented by auxiliary thermostatting terms. Here we give the formulas for the half steps needed to propagate the auxiliary terms; these should be composed with a Verlet step for the positions and momenta and a second, symmetrically reversed half step for the auxiliary terms. That is we write the mapping describing a full timestep as

$$\Phi_\tau = \tilde{\Phi}^{\text{Aux}}_{\tau/2} \circ \Phi^{\text{Verlet}}_\tau \circ \Phi^{\text{Aux}}_{\tau/2}$$

where $\tilde{\Phi}^{\text{Aux}}_\tau$ represents the same steps indicated for Φ^{Aux}_τ taken in reverse order and $\Phi^{\text{Verlet}}_\tau$ represents a one-step implementation of the velocity Verlet method

$$p := p - \frac{\tau}{2}F(q); \quad q := q + \tau M^{-1}p; \quad p := p - \frac{\tau}{2}F(q)).$$

We define $\Delta K(p) := \left(\sum_{i=1}^n \frac{p_i^2}{m} - nk_BT\right)/2$.

$\Phi^{\rm Aux}_{\tau/2}$	
NHC	NHL
$\xi_2 := \xi_2 + \tau(Q_1\xi_1^2 - kT)/(4Q_2)$	$\xi := \xi + \tau\Delta K(p)/(2\mu)$
$\xi_1 := e^{-(\tau/4)\xi_2}\xi_1$	$p := p \times e^{-(\tau/4)\xi}$
$\xi_1 := \xi_1 + \tau\Delta K(p)/(2Q_1)$	$\xi := e^{-\gamma\tau/2}\xi + \sqrt{k_BT\frac{1-e^{-\gamma\tau}}{\mu}}$Gaussian()
$p := p \times e^{-(\tau/2)\xi_1}$	$p := p \times e^{-(\tau/4)\xi}$
$\xi_1 := \xi_1 + \tau\Delta K(p)/(2Q_1)$	$\xi := \xi + \tau\Delta K(p)/(2\mu)$
$\xi_1 := e^{-(\tau/4)\xi_2}\xi_1$	
$\xi_2 := \xi_2 + \tau(Q_1\xi_1^2 - kT)/(4Q_2)$	

Notes:

• Ad-NHC, Ad-NHL: as above, plus $\chi := \chi + \tau\Delta K(p)/Q$

• The notation "Guassian()" refers to a call to a function producing a normally distributed random number with mean zero and standard deviation one.

• The motivation for sandwiching the p update in NHL between two ξ updates is to improve the accuracy of the latter variable which appears to have a strong beneficial effect on performance for small μ.

• For Langevin and Ad-Langevin we used a simple splitting into the Hamiltonian part and the remaining terms (with dissipation and stochastic forces). For Ad-Langevin the additional formula to propagate χ is the same as for Ad-NHC/Ad-NHL.

B. Eigenvalue analysis for Ad-NHL

There is always at least one real eigenvalue λ_R, and as the coefficients are positive, this must be negative when $\gamma_\ell > 0$. For $\lambda < -\gamma_\ell$ we have, following a straightforward calculation

$$p(\lambda) < -\hat{\mu}^{-1}\gamma_\ell$$

hence the real eigenvalues lie in $(-\gamma_\ell, 0)$ for all positive γ_ℓ.

For $\gamma_\ell = 0$, the other two roots η, $\bar{\eta}$ are on the imaginary axis $\lambda_0^\pm = \pm i\sqrt{\hat{\mu}^{-1} + \hat{Q}^{-1}}$ and these move into the left half plane as γ_ℓ is increased. The sum of the eigenvalues is the trace,

hence $\lambda_R + 2\mathrm{Re}\,\eta = -\gamma_\ell$, so

$$\mathrm{Re}\,\eta = \frac{1}{2}(-\gamma_\ell - \lambda_R) < 0$$

Viewing the eigenvalues as functions of the parameter γ_ℓ we differentiate $w = \mathrm{Re}\,\eta$

$$w' = \frac{1}{2}(-1 - \lambda_R'),$$

to look for turning points. Setting $w' = 0$, we have $\lambda_R' = -1$. Differentiating $p(\lambda_R) = 0$ with respect to the parameter, we have

$$3\lambda_R^2\lambda_R' + 2\gamma_\ell\lambda_R\lambda_R' + \lambda_R^2 + (\hat{\mu}^{-1} + \hat{Q}^{-1})\lambda_R' + \hat{Q}^{-1} = 0.$$

Hence, for $\lambda_R' = -1$,

$$\lambda_R^2 + \gamma_\ell\lambda_R + \frac{1}{2}\hat{\mu}^{-1} = 0. \tag{2}$$

Multiplying this equation by λ_R, and subtracting from $p(\lambda_R) = 0$, we find

$$(\hat{\mu}^{-1} + \hat{Q}^{-1})\lambda_R + \hat{Q}^{-1}\gamma_\ell - \frac{1}{2}\hat{\mu}^{-1}\lambda_R = 0,$$

which gives a formula for λ_R and hence also a formula for $\mathrm{Re}\,\eta$ at the turning point:

$$\lambda_R = -\frac{1}{\varphi/2 + 1}\gamma_\ell, \quad \varphi = \hat{Q}/\hat{\mu}.$$

Now we must reintroduce this expression into (2) to find γ_ℓ. Setting $a = \frac{1}{\varphi/2+1}$, we have

$$a^2\gamma_\ell^2 - a\gamma_\ell^2 + \frac{1}{2}\hat{\mu}^{-1} = 0.$$

So,

$$\gamma_\ell = \sqrt{\frac{\hat{\mu}^{-1}/2}{a - a^2}} = (1 + \varphi/2)\sqrt{\frac{2gkT}{Q}}$$

We verified this formula using numerical computations, i.e. calculating the eigenvalues and observing the turning point in the conjugate pair for this choice of γ_ℓ. With γ_ℓ as above, we find

$$\lambda_R^\dagger = -\frac{1}{\varphi/2 + 1}\gamma_\ell = -\hat{Q}^{-1/2} = -\sqrt{\frac{2gkT}{Q}}$$

and

$$\mathrm{Re}\,\eta^\dagger = -\frac{\varphi}{4} \cdot \sqrt{\frac{2gkT}{Q}}$$

The quantity $\alpha = \sqrt{2gkT/Q}$ can be identified with the "natural frequency" of the Nosé-Hoover thermostat [18]. By choosing φ appropriately, we may adjust the damping properties

in proportion to this natural frequency. One way of choosing the parameters might be to fix $\varphi = 2$, i.e. $Q = 2\mu$, and adjust Q to achieve a desired rate of convergence; in this case $\operatorname{Re}\eta = -\alpha/2$, $\lambda_R = -\alpha$, $\gamma_\ell = 2\alpha$. This choice is diagrammed in the left panel of Figure 1.

Another option is shown at right in Figure 1. Here we choose the parameters so that all three eigenvalues are real and negative. Clearly this occurs if $\operatorname{Im}\eta^\dagger = 0$. To achieve this, we use synthetic division to remove the factor $\lambda - \lambda_R^\dagger$ from $p(\lambda) = 0$, resulting in a quadratic equation which may be written

$$\lambda^2 + \frac{1}{2}\varphi\alpha\lambda + (1 + \varphi/2)\alpha^2 = 0,$$

for the other two roots. For these to be real, we must have

$$\left(\frac{1}{2}\varphi\alpha\right)^2 - 4(1 + \varphi/2)\alpha^2 \geq 0.$$

Dividing out by α^2, we get a condition for φ of the form

$$\varphi \geq 4(1 + \sqrt{2}) \Rightarrow Q \geq 4(1 + \sqrt{2})\mu$$

Thus, to place all the eigenvalues on the real axis, we would need to use a relatively small value of $\mu \approx Q/10$; such a choice may influence numerical stability.