

Stochastic thermostats: comparison of local and global schemes

Giovanni Bussi*, Michele Parrinello

Computational Science, Department of Chemistry and Applied Biosciences, ETH Zürich, USI Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

Available online 18 January 2008

Abstract

We show that a recently introduced stochastic thermostat [J. Chem. Phys. 126 (2007) 014101] can be considered as a global version of the Langevin thermostat. We compare the global scheme and the local one (Langevin) from a formal point of view and through practical calculations on a model Lennard–Jones liquid. At variance with the local scheme, the global thermostat preserves the dynamical properties for a wide range of coupling parameters, and allows for a faster sampling of the phase-space.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Molecular dynamics; Stochastic thermostat; Lennard–Jones

The most common approaches to isothermal molecular dynamics are perhaps those based on the introduction of an extended Lagrangian. The root of all these schemes is the Nosé algorithm [1], often used in the Hoover formulation [2]. This scheme can be rigorously shown to provide the correct Boltzmann distribution and has a conserved quantity, which can be used to check the integration timestep. A major drawback of the Nosé–Hoover method is that it is not ergodic in some difficult cases, such as harmonic systems. Several different extensions of the Nosé–Hoover method have been introduced, the most notable one being the so-called Nosé–Hoover chains [3], which addresses the ergodicity issue at the price of an increased complexity in the algorithm. Although the Nosé–Hoover scheme was originally written as a global thermostat, i.e. coupled only to the total kinetic energy of the system, it is sometimes implemented in a local manner (also called massive Nosé–Hoover), i.e. using an independent thermostat on each degree of freedom [4].

Another common choice is the weak-coupling method, introduced by Berendsen et al. [5]. This scheme is a continuous version of the velocity-rescaling scheme, thus it is a global thermostat. It is deterministic, stable and easy to implement, but it does not produce configurations in the canonical ensemble.

An alternative approach to canonical sampling is to use stochastic molecular dynamics. The most common form is

Langevin dynamics [6]. The Langevin thermostat is local, and its major feature is that ergodicity can be proven also in pathological cases. However, since the friction and noise terms alter significantly the Hamiltonian dynamics, it cannot be used to compute dynamical properties, unless an extremely small friction is used. Moreover, the effect of the friction and noise terms on the sampling efficiency is nontrivial. Even in applications where dynamical properties are not relevant it can be difficult to properly tune the friction in order to achieve an efficient sampling.

In a recent paper [7] we proposed a stochastic velocity rescaling which can be considered as Berendsen thermostat plus a stochastic correction leading to canonical sampling. We also showed that, in spite of its stochastic nature, one can define a conserved quantity. This scheme does not suffer of ergodicity problems in solids [7], has been used in practical applications for equilibration purposes [8] or to perform ensemble averages [9,10] and can be combined with variable-cell dynamics to perform simulations in the isothermal–isobaric ensemble [11]. In the present paper we present an alternative derivation of the same scheme, where stochastic velocity rescaling is obtained starting from Langevin dynamics and minimizing the disturbance of the thermostat on the Hamiltonian trajectory, nevertheless retaining the same thermalization speed of Langevin dynamics. This idea was also used by Berendsen et al. to derive their algorithm [5]. Moreover, we show how stochastic velocity rescaling can be considered as a global version of Langevin dynamics. Thus the relationship between the two schemes is

* Corresponding author.

E-mail address: gbussi@ethz.ch (G. Bussi).

similar to that between standard Nosé–Hoover and massive Nosé–Hoover. Finally, we compare in practical situations the efficiency of the local (Langevin) and global (rescaling) versions, and show that the disruption of Hamiltonian dynamics observed using Langevin thermostat is not due to the stochastic nature of the algorithm but to the use of a local thermostat.

1. Continuous equations of motion

We consider a system described by coordinates q_i and momenta p_i , where i runs over the N_f degrees of freedom, and with q and p we indicate the set of coordinates q_i and p_i . We associate a mass m_i to each degree of freedom, and we define a Hamiltonian $H(p, q) = K(p) + U(q)$, where $U(q)$ is the potential energy, and $K(p) = \sum_i \frac{p_i^2}{2m_i}$ is the kinetic energy. We want to sample the canonical distribution $P(p, q) dp dq \propto e^{-\beta(K(p)+U(q))}$, where β is the inverse temperature, by means of equations of motion in the form

$$dp_i(t) = -\frac{\partial U}{\partial q_i} dt + g_i(t) dt, \quad (1a)$$

$$dq_i(t) = \frac{p_i(t)}{m_i} dt. \quad (1b)$$

Eqs. (1) are Hamilton equations plus a correction force $g_i(t)$ which artificially modifies the dynamics of the system. Since the total energy H is conserved in Hamiltonian dynamics, only $g_i(t)$ is responsible for its variations and leads to the system thermalization.

In standard Langevin dynamics, the correction force is

$$g_i(t) dt = -\gamma p_i(t) dt + \sqrt{\frac{2m_i\gamma}{\beta}} dW_i(t), \quad (2)$$

where γ is the friction coefficient, and $dW_i(t)$ is a vector of N_f independent Wiener noises, normalized as $\langle \frac{dW_i(t)}{dt} \frac{dW_j(t')}{dt'} \rangle = \delta(t-t')\delta_{ij}$. The thermalization speed can be quantified calculating the time derivative of the total energy from Eqs. (1) and (2). Using the Ito chain rule [12] we obtain

$$dH(t) = \sum_i \left(-\frac{\gamma p_i^2(t)}{m_i} dt + \sqrt{\frac{2\gamma}{\beta}} \frac{p_i^2(t)}{m_i} dW_i(t) \right) + \frac{\gamma N_f}{\beta} dt. \quad (3)$$

This expression can be further simplified defining the average kinetic energy $\bar{K} = N_f(2\beta)^{-1}$, a relaxation time $\tau = (2\gamma)^{-1}$, and exploiting the fact that the noise terms on different degrees of freedom are independent of each other:

$$dH(t) = -\frac{K(t) - \bar{K}}{\tau} dt + 2\sqrt{\frac{\bar{K} K(t)}{N_f \tau}} dW(t). \quad (4)$$

It is worth noting that while in Eqs. (1) and (3) there are N_f independent noise terms, in Eq. (4) a single noise term is present.

We now want to design a new correction force $\tilde{g}_i(t)$ which gives the same variation of the total energy as Langevin dynamics, thus the same thermalization speed, but minimizes the disturbance on the trajectory. This procedure is exactly the same used by Berendsen et al. [5], the only difference being that

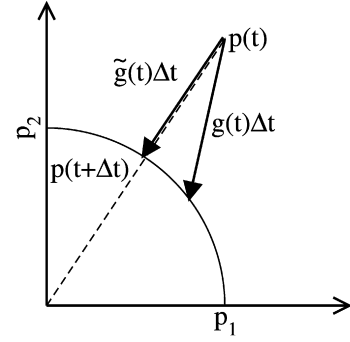


Fig. 1. Schematic representation of the momentum components at time t and at time $t + \Delta t$. $g(t)$ is a generic force applied to the system. $\tilde{g}(t)$ is a force which leads to the same change in the kinetic energy as $g(t)$, but minimizes the disturbance.

Eq. (4) on the total energy now contains a stochastic term. We first notice that, since the force only acts on the momenta, fixing a value for H is equivalent to fixing a value for K . Following Ref. [5] we quantify the disturbance as $\sum_i m_i^{-1} (\tilde{g}_i(t) dt)^2$. As it is seen in Fig. 1, the minimal disturbance for a fixed kinetic energy increment is obtained with a force $\tilde{g}_i(t)$ which is proportional to $p_i(t)$. Thus $\tilde{g}_i(t) = \lambda(t) p_i(t)$, where $\lambda(t)$ is chosen so as to enforce a given variation of the total energy. Since $\lambda(t)$ includes a stochastic part, and since the variation of the total energy depends on the momenta only through the kinetic energy K , this last relation can be written as

$$\tilde{g}_i(t) dt = p_i(t) [A(K(t)) dt + B(K(t)) dW(t)], \quad (5)$$

where $A(K)$ and $B(K)$ are arbitrary functions of the kinetic energy. The change in the total energy is then

$$dH = 2A(K)K dt + 2B(K)K dW + B^2(K)K dt. \quad (6)$$

Expressions for $A(K)$ and $B(K)$ can be obtained setting Eq. (4) equal to Eq. (6), resulting in the correction force

$$\tilde{g}_i dt = \frac{1}{2\tau} \left[\left(1 - \frac{1}{N_f} \right) \frac{\bar{K}}{K} - 1 \right] p_i dt + \sqrt{\frac{\bar{K}}{N_f K \tau}} p_i dW. \quad (7)$$

This equation is stochastic, with the same noise term used on all the particles. It is also very similar to the expression of the force in the Berendsen algorithm.

The combination of Eqs. (1) and (7) results in a continuous, stochastic dynamics which can be shown to sample exactly the canonical ensemble. The effect of a \tilde{g}_i parallel to p_i is the same of a rescaling procedure and the enforced increment of the total energy in Eq. (4) is the same as in Ref. [7]. Thus, Eqs. (1) and (7) represents the continuous version of the velocity rescaling described in Ref. [7].

Notably, if $N_f = 1$, Eq. (7) becomes equivalent to Eq. (2). Thus when the thermostat is applied to a single degree of freedom, it is completely equivalent to a Langevin thermostat. One can perform Langevin molecular dynamics by applying a thermostat per degree of freedom, or stochastic rescaling by applying a single thermostat to the total kinetic energy. Intermediate schemes can be designed, where a thermostat is applied on each molecule or group of atoms.

2. Finite timestep algorithm

In the practical implementation, time is incremented in discrete steps, and the Trotter decomposition scheme [13,14] can be used to separate the integration of Hamilton equations and the update of the momenta due to \tilde{g} . The former is then integrated using standard velocity-Verlet, while for the latter we need to integrate Eq. (7). A possible approach is to consider the propagation of kinetic energy when the momenta evolve according to Eq. (7), as it is done in the appendix of Ref. [7]. The analytical solution of Eq. (7) for a finite time Δt is

$$p_i(t + \Delta t) = \alpha(t) p_i(t), \quad (8a)$$

where

$$\alpha^2(t) = c + \frac{(1-c)(S_{N_f-1}(t) + R^2(t))\bar{K}}{N_f K(t)} + 2R(t) \sqrt{\frac{c(1-c)\bar{K}}{N_f K(t)}}. \quad (8b)$$

Here $c = e^{-2\gamma\Delta t} = e^{-\Delta t/\tau}$, $R(t)$ is a Gaussian number with unitary variance and S_{N_f-1} is the sum of $N_f - 1$ independent, squared, Gaussian numbers. Eq. (8) has been obtained enforcing the evolution of the kinetic energy, thus, strictly speaking, it does not fix the sign of α . A more rigorous analysis shows that the sign of α should be chosen as

$$\text{sign}(\alpha(t)) = \text{sign}\left(R(t) + \sqrt{\frac{cN_f K(t)}{(1-c)\bar{K}}}\right), \quad (9)$$

to keep into account the finite probability to observe a flip of the momenta p_i when the force in Eq. (7) is applied. The Gaussian number in Eq. (9) needs to be the same that is used in Eq. (8). The probability to observe the flip is extremely small if N_f is large and $c \approx 1$, which is the usual case when the thermostat is used as global and $\tau > \Delta t$. This is the case in Ref. [7], where we set $\alpha > 0$. On the other hand, when the thermostat acts on a few degrees of freedom, the sign of α needs to be calculated by means of Eq. (9). This is always the case for Langevin dynamics. With simple manipulation, it can be shown that for $N_f = 1$ the integration scheme given by Eqs. (8) and (9), combined with velocity-Verlet, is completely equivalent to the integration scheme for Langevin dynamics introduced in Ref. [14].

3. Examples

Up to now we simply established a theoretical relationship between Langevin dynamics and stochastic rescaling. The outcome is that the effect of the two algorithms on the total energy should be equivalent if the friction in the Langevin dynamics and the relaxation time in the stochastic scale are related by $\tau = (2\gamma)^{-1}$. However, the stochastic rescaling is expected to give better dynamical properties, since only the component of the force which changes the total energy is retained. We test here this affirmations on a simple test-case, namely a Lennard–Jones fluid with density $\rho = 0.8442$ and temperature $\beta^{-1} = 0.722$, which is close to the triple point. Throughout this

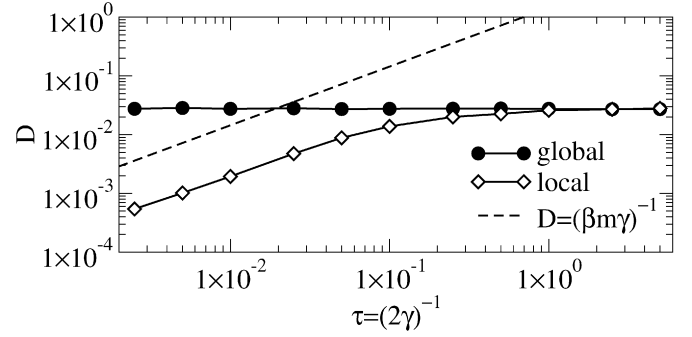


Fig. 2. Diffusion coefficient as obtained from thermostated simulations as a function of the thermostat relaxation time τ , for the local and global thermostats as indicated. The statistical error is smaller than the symbol size. As a reference, the diffusion coefficient of a free particle subject to the same Langevin equation is plotted (dashed line). All the quantities are in Lennard–Jones reduced units.

section we use reduced Lennard–Jones units for temperature, distance and time. We simulate a box containing 108 particles, with periodic boundary conditions, and we cut the interaction at distance 2.5. We set the timestep to $\Delta t = 0.005$, which leads to a reasonable conservation of the effective energy [7,14]. We then perform runs of 10^7 steps, with both the local scheme (Langevin dynamics) and the global one (stochastic rescaling), using a broad range of values of the thermostat relaxation time τ .

To quantify the disturbance on Hamiltonian dynamics, in Fig. 2 we show the diffusion coefficient as a function of τ , as obtained from the Einstein relations. When the local thermostat is used with a short relaxation time, the diffusion is strongly quenched. This happens when the typical collision time with the external bath γ^{-1} is shorter than the typical collision time between the particles, so that the former becomes the real bottleneck for the diffusion process. In the limit of short τ , the equations of motion tend to a high-friction Langevin dynamics. In this case, the observed diffusion coefficient D is proportional to $(\beta m \gamma)^{-1}$, which is the value of the diffusion coefficient for a free particle subject to the same Langevin dynamics. The prefactor is related to the difficulty in crossing the barriers between different liquid configurations. On the contrary, with the global thermostat D is almost independent on τ , indicating that the disturbance on the dynamics is very small and that good estimates of D in the canonical ensemble can be obtained also with a thermostated simulation.

To quantify the equilibration speed we calculate the autocorrelation time of a few global observables. The efficiency of a sampling algorithm is optimal when the autocorrelation time τ_X of the desired quantity X is minimal. If X is the total energy, τ_X also indicates how fast a simulation started from an unlikely configuration is equilibrated. We integrate the autocorrelation function using a windowing function

$$\tau_X = \int_0^T dt \frac{\langle \delta X(0) \delta X(t) \rangle}{\langle \delta X(0) \delta X(0) \rangle} \left(1 - \frac{t}{T}\right), \quad (10)$$

where $\delta X = X - \langle X \rangle$ and T is a large value. The windowing function in parenthesis helps the convergence because it

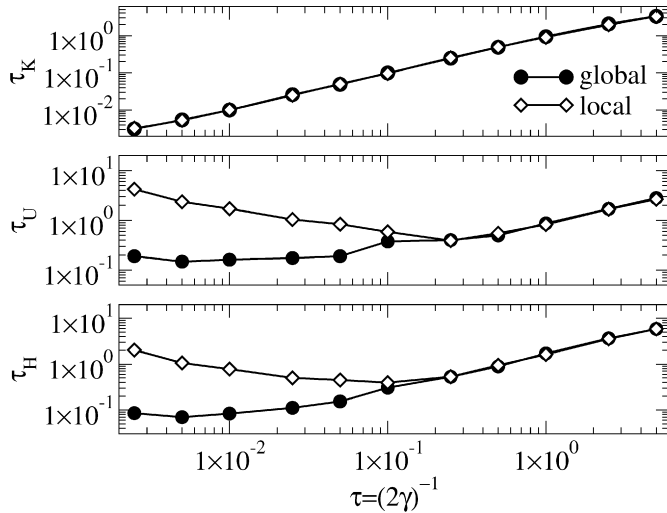


Fig. 3. Autocorrelation time of the kinetic (τ_K), potential (τ_U), and total (τ_H) energy, as a function of the thermostat relaxation time τ , for the local and global thermostats as indicated. The statistical error is smaller than the symbol size. All the quantities are in Lennard–Jones reduced units.

weights less the points with larger statistical error. Moreover, Eq. (10) is exactly equal to $T\epsilon^2(T)/(2\langle\delta X^2\rangle)$, where $\epsilon^2(T)$ is the mean square error from a time average of length T . The relative accuracy in evaluation of τ_X is approximately $\sqrt{2T/T'}$, where $T' = 5 \times 10^4$ is the total run length. In the following we choose $T = 50$, and we expect a relative accuracy on τ_X on the order of 5%. Since $T \gg \tau_X$, τ_X is also a good approximation for the autocorrelation time.

In Fig. 3 we plot the autocorrelation time of the kinetic energy τ_K , of the potential energy τ_U and of the total energy τ_H , as a function of the thermostat relaxation time τ . The autocorrelation time of the kinetic energy is completely dictated by the thermostat relaxation time, and independent on the choice of a local or a global scheme. On the contrary, the autocorrelation time of the total energy and of the potential energies are proportional to τ only in the limit of large τ . In the local scheme, when τ is smaller than 0.2 the disturbance of the trajectory becomes

so large that the phase-space exploration turns out to be slower. Comparing Figs. 2 and 3, it is seen that the optimal value for τ is the smallest one that still does not affect the diffusion coefficient. When the global scheme is adopted, even small values of τ can be safely used, resulting in a faster decorrelation of the total energy.

4. Conclusions

In conclusion, we have presented an alternative derivation of the global thermostat introduced in Ref. [7]. This derivation allows to write continuous equations of motion and shows the analogy between this scheme and the standard Langevin thermostat. Namely, the new scheme can be considered as a global version of the Langevin thermostat, that minimizes the disturbance of the original Hamiltonian dynamics. Finally, we have discussed these properties on a simple test case, showing that the global scheme preserves the dynamical properties. Moreover, using as a measure the autocorrelation time of the total energy, we have shown that the global scheme allows for a faster sampling.

References

- [1] S. Nosé, J. Chem. Phys. 81 (1984) 511.
- [2] W.G. Hoover, Phys. Rev. A 31 (1985) 1695.
- [3] G.J. Martyna, M.L. Klein, M. Tuckerman, J. Chem. Phys. 97 (1992) 2635.
- [4] D.J. Tobias, G.J. Martyna, M.L. Klein, J. Phys. Chem. 97 (1993) 12959.
- [5] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, J. Chem. Phys. 81 (1984) 3684.
- [6] T. Schneider, E. Stoll, Phys. Rev. B 17 (1978) 1302.
- [7] G. Bussi, D. Donadio, M. Parrinello, J. Chem. Phys. 126 (2007) 014101.
- [8] D. Donadio, G. Galli, Phys. Rev. Lett. 99 (2007) 255502.
- [9] F. Bruneval, D. Donadio, M. Parrinello, J. Phys. Chem. B 111 (2007) 12219.
- [10] A. Barducci, G. Bussi, M. Parrinello, Phys. Rev. Lett. 100 (2008) 020603.
- [11] T. Zykova, G. Bussi, M. Parrinello, in preparation.
- [12] C.W. Gardiner, Handbook of Stochastic Methods, third ed., Springer, Berlin, 2003.
- [13] M. Tuckerman, B.J. Berne, G.J. Martyna, J. Chem. Phys. 97 (1992) 1990.
- [14] G. Bussi, M. Parrinello, Phys. Rev. E 75 (2007) 056707.