Enhanced sampling of rare events

Simone Melchionna
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1 EW, United Kingdom
(Received 16 April 2000)

We present a method to enhance sampling of a given reaction coordinate by projecting part of the random thermal noise along a preferential direction. The approach is promising to study rough energy landscapes and highly activated barriers that can be overcome by increasing the attempt frequency. Furthermore it allows us to rescale a given reaction coordinate without biasing the configurational properties of the system. A major advantage of the method is its simplicity in the analytical derivation and numerical implementation.

PACS number(s): 02.70.−c, 05.20.Gg, 05.70.Ce

I. INTRODUCTION

Condensed matter in the quantum and classical regimes often undergoes transitions between stable states separated by free-energy barriers of low dimensionality and with heights of order kT or larger. Computer simulation is a powerful tool to study microscopic phenomena but in its simplest form is limited to small systems for rather short time windows. When dealing with rare events, the current computational limits do not allow for spanning for sufficiently long times many reactions of interest. The situation is worsened when a detailed knowledge of the reaction path is not known in advance. The limited ergodic attitude and the unknown preferential states are major hindrances for a detailed understanding of the phase-space landscape of a given system.

In past years, different strategies have been devised in order to enhance sampling of rare events and obtain accurate calculations of free-energy and/or rate constants. These methods can be grouped into two different classes. In one class the initial knowledge of the reaction coordinate (RC) and location of the relevant free energy minima allows the use of simulation techniques such as Monte Carlo or molecular dynamics. The methods rely on either introducing a biasing potential dependent on a preset value of a given RC, as with the umbrella sampling method [1], or on constraining the reaction coordinate, as with the blue moon ensemble method [2,3]. More often a suitable weighting procedure is not known *a priori* and must be constructed by trial and error. This is the first drawback of this approach. A very interesting technique has been recently introduced by Bolhuis *et al.* [4] that enables us to find the optimal reaction path once the location of the relevant free-energy minima is known, via a path integral approach.

Combining these approaches, once a good weighting procedure has been found, a stepwise walk along the reaction path can be performed for a defined set of values of the RC, and with a limited number of possible reaction channels. Consequently, sampling the reaction path can be computationally very expensive due to the large number of points to be sampled in the multidimensional phase space. In the case of umbrella sampling applied to highly activated processes, the computation can be even more difficult due to the steep biasing potential needed to achieve a good local sampling of the RC.

These approaches, although they bias the natural statistical distribution, can be rigorously inverted to compute free energies with high accuracy. Furthermore, by using molecular dynamics, the rate constant associated with the process can be computed as an average over initial states belonging to the biased distribution [2,5].

The second class of strategies includes a plethora of different techniques that enables us to explore the free-energy landscape and eventually locate the relevant minima, finding application both in optimization problems and in enhancing the sampling rate of complex systems. Among others, several techniques, such as multicanonical sampling [6], *J*-walking [7] and *S*-walking [8] schemes, and hyperdynamics [9], overcome energetic or entropic barriers in specific cases (see for example, Chaps. 6 and 7 in [10] and references therein). These approaches are usually based on altering either the potential function or the thermodynamic state of the system, typically by raising the temperature, to span rough energy landscapes. Most of these techniques are usually applied via Monte Carlo and its variants, by relying on a statistical, rather than dynamical, sampling.

In this paper, we describe a method that enhances the sampling of a complex phase space, while preserving the configurational distribution of the system. To this aim, Hamiltonian equations of motion will be used and the fluctuations of the RC will be enhanced by a kinetic mechanism, produced by rescaling the generalized mass associated with the RC. The underlying idea of the present approach is to work in Cartesian coordinates and project part of the random thermal noise available to the system (of order $gkT/2$ for a system composed of g degrees of freeedom) along one or more reaction coordinates so as to favor transitions between states of high probability separated by high-(free-) energy barriers. A similar idea was proposed several years ago by Bennett who used the independence from the masses of the configurational properties to accelerate equilibration [11].

As it will be shown, a detailed knowledge of the reaction path is not needed in advance, so that the system is not forced to follow a predefined reaction channel but it is permitted a certain freedom to find the best path to undergo the unfavorable transitions. Moreover, a knowledge of the reaction coordinate in itself is not needed, but only of the unit vector directed along the RC gradient.

The method has its potential application in molecular dynamics, where the continuous nature of the evolution equations allows us to study large and flexible molecules and

systems at high density, and in the hybrid Monte Carlo approach, where the kinetic term of the Hamiltonian is part of the propagation algorithm.

The paper is organized as follows. In Sec. I, the algebraic treatment of the method is outlined. Section II illustrates a numerical example. Section III contains some concluding remarks. Finally, the Appendix illustrates how to couple dynamical equations with a thermostat.

II. PROJECTED HAMILTONIAN

Let us consider a mechanical system defined by the set of $6N$ independent coordinates and momenta $\{r,p\}$. For the sake of simplicity, we will restrict ourselves to a single reaction coordinate function of positions only, $\xi(r)$, so that the probability of finding the system at position ξ^* is

$$\mathcal{P}(\xi=\xi^*)=\langle \delta(\xi(r)-\xi^*)\rangle. \tag{1}$$

If the system is described by the Hamiltonian $H(r,p)=\Sigma_i p_i^2/2m_i+V(r)$, the quadratic velocity associated to ξ reads

$$\dot{\xi}^2=\underline{\underline{Z}}:\underline{\underline{K}}, \tag{2}$$

where $\underline{\underline{A}}:\underline{\underline{B}}$ indicates the contraction of $\underline{\underline{A}}$ and $\underline{\underline{B}}$, $\underline{\underline{A}}:\underline{\underline{B}}=\mathrm{Tr}(\underline{\underline{A}}\underline{\underline{B}})$, and the two following symmetric matrices have been introduced:

$$\underline{\underline{Z}}=Z_{ij}=\frac{1}{\sqrt{m_i m_j}}\nabla_i\xi\nabla_j\xi, \tag{3}$$

$$\underline{\underline{K}}=K_{ij}=\frac{1}{\sqrt{m_i m_j}}p_i p_j. \tag{4}$$

The average quadratic velocity is thus found to be

$$\langle\dot{\xi}^2\rangle=\frac{\int drdp e^{-\beta[p^2/2m+V(r)]}\underline{\underline{Z}}:\underline{\underline{K}}}{\int drdp e^{-\beta[p^2/2m+V(r)]}}$$

$$=\frac{kT}{2}\frac{\int dr e^{-\beta V(r)}\underline{\underline{Z}}:\underline{\underline{1}}}{\int dr e^{-\beta V(r)}}=\frac{kT}{2}\langle \mathrm{Tr}\,Z\rangle. \tag{5}$$

Our goal is to alter the natural distribution in order to amplify the average quadratic velocity of the RC. Let us now introduce the projector $P_\perp$ along the $(1/\sqrt{m_i})\nabla_i\xi$ direction and acting on the $3N$-configurational space

$$\underline{\underline{P_\perp}}=\frac{1}{\mathrm{Tr}\,Z}\underline{\underline{Z}} \tag{6}$$

and its parallel counterpart

$$\underline{\underline{P_\parallel}}=\underline{\underline{1}}-\underline{\underline{P_\perp}}. \tag{7}$$

Clearly, any vector with components $v_i=c(1/\sqrt{m_i})\nabla_i\xi$, with c being a constant, is an eigenvector of $\underline{\underline{P_\perp}}$:

$$\underline{\underline{P_\perp}}v=c\frac{1}{\mathrm{Tr}\,Z}\frac{1}{\sqrt{m_i m_j}}\nabla_i\xi\nabla_j\xi\frac{1}{\sqrt{m_j}}\nabla_j\xi=v. \tag{8}$$

The projector $\underline{\underline{P_\perp}}$ is not invertible and obeys the rules $\mathrm{Tr}\ P_\perp=1$, $P_\perp^2=P_\perp$, while for its parallel counterpart $\mathrm{Tr}\ P_\parallel=3N-1$, and $P_\parallel^2=P_\parallel$.

Let us now introduce the following invertible operator

$$\underline{\underline{\chi}}=\underline{\underline{P_\parallel}}+\alpha\underline{\underline{P_\perp}}=\underline{\underline{1}}+(\alpha-1)\underline{\underline{P_\perp}}, \tag{9}$$

where α is a free parameter, with $\alpha\neq 0$.

The following formula applies for a generic function of the matrix $\underline{\underline{\chi}}$:

$$f(\underline{\underline{\chi}})=f(1)\underline{\underline{1}}+[f(\alpha)-f(1)]\underline{\underline{P_\perp}} \tag{10}$$

as it is seen by expanding $f(\underline{\underline{\chi}})$ in powers of $\underline{\underline{\chi}}$, and observing that $\underline{\underline{\chi}}^m=\underline{\underline{1}}+(\alpha^m-1)\underline{\underline{P_\perp}}$. Therefore the following equations apply:

$$\underline{\underline{\chi}}^{-1}=\underline{\underline{1}}+\left(\frac{1}{\alpha}-1\right)\underline{\underline{P_\perp}},$$

$$\mathrm{Tr}\ \chi=3N-1+\alpha,$$

$$\det\chi=e^{\mathrm{Tr}\ln\chi}=e^{\mathrm{Tr}\ln(\alpha)P_\perp}=\alpha. \tag{11}$$

By construction, given a generic vector in configurational space, the operator $\underline{\underline{\chi}}$ scales its component parallel to $(1/m_i)\nabla_i\xi$ by the factor α, while the orthogonal components are untouched by the transformation.

The matrix $\underline{\underline{\chi}}$ is now applied to define the Lagrangian

$$L_b=\frac{1}{2}\sum_{ij}\chi_{ij}\sqrt{m_i}\dot{r}_i\sqrt{m_j}\dot{r}_j-V(r) \tag{12}$$

containing a position-dependent kinetic term. The associated momenta are $p_i=\Sigma_j\chi_{ij}\sqrt{m_i m_j}\dot{r}_j$ and the Hamiltonian is

$$H_b=\frac{1}{2}\underline{\underline{\chi}}^{-1}:\underline{\underline{K}}+V(r), \tag{13}$$

where $\underline{\underline{K}}$ is defined by Eq. (4).

The presence of the $1/\alpha$ scaling term in the Hamiltonian corresponds to a scaling of the generalized mass associated to the variable ξ. Therefore the projectors will redirect part of the thermal noise on the kinetic term associated to the RC. We notice here that H_b is reminiscent of the Hamiltonian proposed by Forrest and Suter to enhance the sampling of polymeric chains in the framework of the hybrid Monte Carlo simulation method [12].

The standard Lagrangian and Hamiltonian functions are recovered if $\alpha=1$, since $\underline{\underline{\chi}}=\underline{\underline{1}}$. On the other hand the Hamiltonian for $\alpha=0$. that intuitively would correspond to a constrained RC, is pathological since $\underline{\underline{\chi}}$ is not invertible anymore. Thus the Hamiltonian (13) with $\alpha=0$ is singular and unable to produce a constrained dynamics [13].

If we now suppose that the system obeys a canonical distribution $f_b(r,p) \propto e^{-\beta H_b}$, the quadratic velocity associated to ξ is

$$\dot{\xi}^2 = \underline{\underline{\chi}}^{-1} \underline{\underline{Z}} \underline{\underline{\chi}}^{-1} : \underline{\underline{K}}, \tag{14}$$

so that its average value over the distribution $f_b(r,p)$ reads

$$\langle \dot{\xi}^2 \rangle_b = \frac{\int dr dp e^{-\beta H_b} \underline{\underline{\chi}}^{-1} \underline{\underline{Z}} \underline{\underline{\chi}}^{-1} : \underline{\underline{K}}}{\int dr dp e^{-\beta H_b}}$$

$$= \frac{kT}{2} \frac{\int dr e^{-\beta V(r)} \underline{\underline{Z}} : \underline{\underline{\chi}}^{-1} \det \chi^{1/2}}{\int dr e^{-\beta V(r)} \det \chi^{1/2}}$$

$$= \frac{kT}{2\alpha} \langle \mathrm{Tr}\, Z \rangle, \tag{15}$$

where we have used Eq. (11) and the fact that

$$\underline{\underline{Z}} : \underline{\underline{\chi}}^{-1} = \underline{\underline{Z}} : \left[\underline{\underline{1}} + \left(\frac{1}{\alpha} - 1 \right) \underline{\underline{\underline{\underline{P_\perp}}}} \right] = \frac{1}{\alpha} \mathrm{Tr}\ Z, \tag{16}$$

since

$$\underline{\underline{Z}} : \underline{\underline{\underline{\underline{P_\perp}}}} = \frac{\sum_i \sum_j (\nabla_i \xi \nabla_j \xi)^2 / (m_i m_j)}{\sum_k (\nabla_k \xi)^2 / m_k} = \mathrm{Tr}\ Z. \tag{17}$$

Consequently, we obtain

$$\langle \dot{\xi}^2 \rangle_b = \frac{1}{\alpha} \langle \dot{\xi}^2 \rangle \tag{18}$$

and in general, by applying a similar calculation to the statistical moments of the RC velocity, $\langle \dot{\xi}^m \rangle_b = (1/\alpha^{m/2}) \langle \dot{\xi}^m \rangle$. Therefore if we choose the parameter $\alpha < 1$, the effect of the kinetic bias in the Hamiltonian H_b is to amplify the fluctuations of the RC velocity by a factor $1/\alpha$.

An important consequence of the projected Hamiltonian is that the configurational distribution function $f_b(r) = \int dp f_b(r,p)$ is unbiased by the presence of the projectors

$$f_b(r) = \frac{\int dp e^{-\beta H_b}}{\int dr dp e^{-\beta H_b}} = \frac{e^{-\beta V(r)}}{\int dr e^{-\beta V(r)}} = f(r). \tag{19}$$

Hence, for any observable dependent on positions only,

$$\langle A(r) \rangle_b = \langle A(r) \rangle. \tag{20}$$

This is an extension of the well-known result of statistical mechanics that configurational averages do not depend on the numerical value of masses, now valid for a generalized mass associated to an RC. Therefore, at variance with other free-energy sampling methods, one can compute configurational averages without any reweighting procedure.

The previous argument applies as well to the case of the microcanonical ensemble since the configurational distribution is computed to be

$$f_b(r) = \frac{\int dp \delta[H_b - E]}{\int dr \int dp \delta[H_b - E]}$$

$$\propto \det \chi^{-1/2} [E - V(r)]^{(N-2)/2}$$

$$\propto [E - V(r)]^{(N-2)/2} \tag{21}$$

and therefore $f_b(r) = f(r)$. However, for applications it is often desirable to work in the canonical ensemble [14]. In the Appendix it is shown how the generalized mass rescaling method can be implemented in the canonical ensemble by employing the Nosé-Hoover thermostat [15,16].

III. NUMERICAL ILLUSTRATION

As a simple test case of the proposed method, we implement the projector mechanism to the study of a pair of tagged particles indexed 1 and 2, embedded in a soft sphere fluid. The system, treated with periodic boundary conditions, has, for a density $\rho \simeq 0.0248$ Å^{-3}, a cubic simulation box of length $L = 27.2$ Å and contains 500 atoms of equal mass. The atoms interact via a truncated Lennard-Jones potential $V = \Sigma_{i<j} 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6]$, where $\sigma = 3.4050$ Å, $\epsilon = 0.995\,581$ KJ/mol, and with cutoff at $r_c = 2^{1/6}\sigma$, where the potential is shifted to avoid discontinuities. Furthermore, for particles 1 and 2, we introduce a confining potential of the form $4\epsilon[\{\sigma/(r_{12} - l)\}^{12} - \{\sigma/(r_{12} - l)\}^6]$ with $l = 15$ Å, for the reasons that will become clear in the following. The external temperature is set to 80 K, corresponding to an average pressure of 4.6 ± 0.1 kbar. The coupling time of the Nosé-Hoover thermostat is set to 0.5 ps.

We choose a reaction coordinate that is a generic function of the distance between particles 1 and 2, $\xi(r_1, r_2) = \xi(|r_{12}|)$. For this general choice the matrices $\underline{\underline{\underline{\underline{P_\perp}}}}$ and $\underline{\underline{\chi}}^{-1}$ take the form

$$\underline{\underline{\underline{\underline{P_\perp}}}} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix},$$

$$\underline{\underline{\chi}}^{-1} = [\underline{\underline{1}} + \gamma \underline{\underline{\underline{\underline{P_\perp}}}}], \tag{22}$$

where $\gamma = 1/\alpha - 1$, and $\underline{\underline{A}}$ is the 6×6 block matrix

$$\underline{\underline{A}} = \frac{1}{2} \begin{pmatrix} \hat{r}_{12} \\ \hat{r}_{21} \end{pmatrix} (\hat{r}_{12} \quad \hat{r}_{21}). \tag{23}$$

The corresponding equations of motion are

$$\dot{r}_1 = \frac{p_1}{m} + \frac{\gamma}{2m} (\hat{r}_{12} \cdot p_{12}) \hat{r}_{12},$$

$$\dot{r}_2 = \frac{p_2}{m} - \frac{\gamma}{2m} (\hat{r}_{12} \cdot p_{12}) \hat{r}_{12},$$

$$\dot{r}_n = \frac{p_n}{m},$$

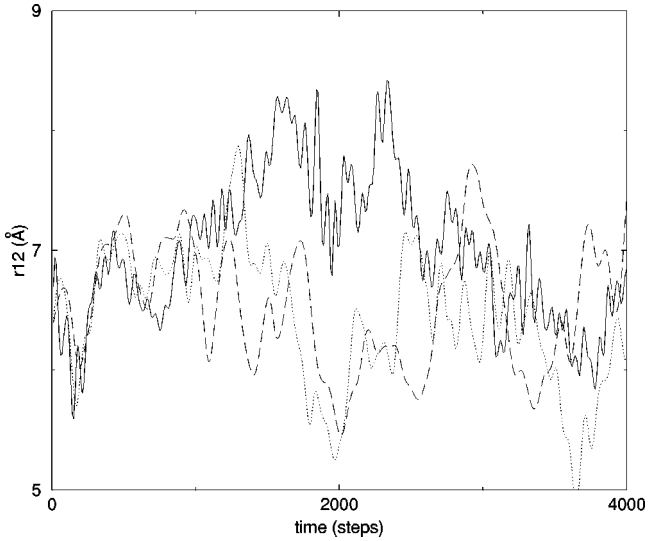

FIG. 1. The time dependence of the distance between particles 1 and 2 computed during a time span of 2 ps, for $\alpha=0.1$ (solid line), $\alpha=0.5$ (dotted line), and for the unperturbed trajectory (dashed line).

$$\dot{p}_1 = F_1 - \frac{\gamma(\hat{r}_{12}\cdot p_{12})}{2mr_{12}}[p_{12}-(\hat{r}_{12}\cdot p_{12})\hat{r}_{12}] - \zeta\left(p_1 - \frac{mP_o}{M}\right),$$

$$\dot{p}_2 = F_2 + \frac{\gamma(\hat{r}_{12}\cdot p_{12})}{2mr_{12}}[p_{12}-(\hat{r}_{12}\cdot p_{12})\hat{r}_{12}] - \zeta\left(p_2 - \frac{mP_o}{M}\right),$$

$$\dot{p}_n = F_n - \zeta\left(p_n - \frac{mP_o}{M}\right),$$

$$\dot{\zeta} = \frac{2}{\tau^2 3NkT}\left[\sum_{i=1}^{N}\frac{p_i^2}{2m} + \frac{\gamma(\hat{r}_{12}\cdot p_{12})^2}{2m} - \frac{P_o^2}{2M} - \frac{3(N-1)kT}{2}\right],$$

$$\dot{s} = 3(N-1)s\zeta, \tag{24}$$

where $n \geqslant 3$. From the expression of $\underline{\underline{\chi}}^{-1}$ and the equations of motion it follows that, if the pair 1-2 diffuses at distance larger than $L/2$, the biasing term with periodic boundaries is discontinuous and we are unable to accurately control the quality of the integrated trajectory. In the present test, to avoid any possible problem we introduce the confining potential between particles 1 and 2 mentioned above.

The equations of motion are integrated with a time step of 4×10^{-4} ps for a total of 5×10^6 steps, corresponding to 2 ns. The propagation procedure, which needs to be properly addressed in the presence of velocity-dependent forces as for our case, is casted in the form of the velocity Verlet integration algorithm [1] together with an iterative scheme to propagate accurately the velocity-dependent forces. The details of the procedure will be presented elsewhere.

In Fig. 1, we report the time dependence of the distance between particles 1 and 2, computed from three trajectories with $\alpha=0.1$, $\alpha=0.5$, and $\alpha=1.0$ (unperturbed equations of motion), for a limited time span. The fluctuations are visible and exhibit faster motion for the RC as the parameter α decreases.

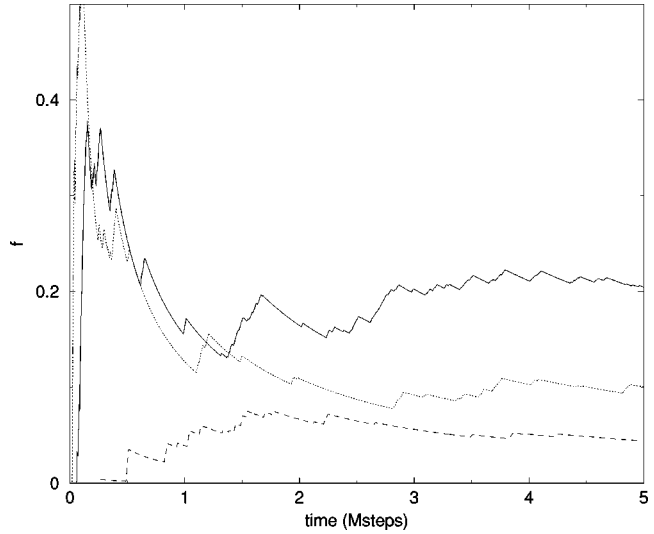

FIG. 2. The cumulative average of the encounter frequency f defined for $r12(t)<4$ Å. Symbols as in Fig. 1 and 1 Mstep $=10^6$ steps $=0.4$ ns.

Figure 2 shows the cumulative average of the frequency of encounter of the tagged pair. Given a distance threshold of $d=4$ Å, the encounter is detected when $r_{12}(t)<d$ and $r_{12}(t-h)>d$, where the condition is checked every $h=0.02$ ps time interval. To encounter, the two particles have to compete with the large entropic barrier given by the available volume and with layering of the surrounding fluid. The data exhibit that the encounter frequency of the unperturbed trajectory converges over a 1.6-ns time span, while for $\alpha=0.5$ and 0.1 the convergence is reached over a period of about 1.0 ns. Overall the encounter frequency has increased by around 100% for $\alpha=0.5$ and around 200% for $\alpha=0.1$ with respect to the unperturbed dynamics.

In Fig. 3, we report the normalized histogram of finding the pair at distance r, $h(r)$, and multiplied by the factor $4\pi r^2 V(N-1)/N$, for comparison with the radial distribution function of the purely repulsive Lennard-Jones fluid. The panels correspond to time averages over 0.4, 1.2, and 2.0 ns. The data show a slow convergence for the unperturbed trajectory, while for $\alpha=0.5$ and 0.1 the convergence is faster, particularly for the latter value, as seen from averaging over the whole trajectory and inside the first coordination shell.

IV. CONCLUSIONS

We have presented a method to enhance sampling of a given reaction coordinate by applying a rescaling of its generalized mass and projecting part of the random thermal noise along a preferential direction. The method makes use of a continuous form of the equations of motion, and therefore is well-suited for molecular dynamics and hybrid Monte Carlo schemes.

The approach is promising for the study of rough energy

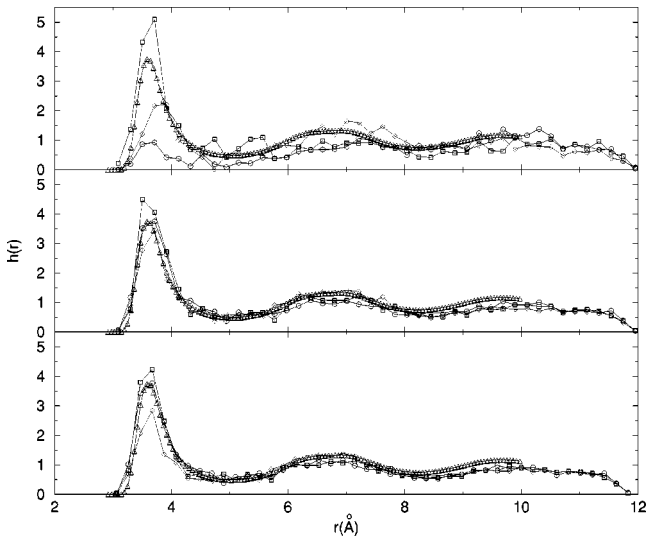

FIG. 3. The histogram of occurrence of particles 1 and 2 at distance r, $h(r)$, multiplied by $4\pi r^2 V(N-1)/N$. The panels correspond to averaging over the first 0.4 (upper panel), 1.2 (middle panel), and the overall trajectory of 2 ns (lower panel). Symbols correspond to $\alpha=0.1$ (circles), $\alpha=0.5$ (squares), the unperturbed trajectory (diamonds), and the radial distribution function of the purely repulsive Lennard-Jones fluid (triangles).

landscapes and highly activated barriers that can be overcome by an enhanced attempt frequency, i.e., by a ''hotter'' reaction coordinate. Useful applications can be for studying systems at high density: from the investigation of phase transitions, to complex macromolecular systems, as polymers and biomolecules, and finally *ab initio* systems, where a speedup of nuclear and/or electronic degrees of freedom is essential. To this aim, further attention needs to be given to investigate the effect of the inertial parameter α to obtain optimal sampling of the unfavorable phase space in specific cases.

A major advantage of the proposed method is the simplicity of the formalism and numerical implementation. A number of further extensions can be foreseen. First, the approach can be extended to rescaling multiple (vectorial) reaction coordinates, allowing the investigation of complex free-energy landscapes by selecting a set of relevant reaction coordinates. Second, an appealing perspective is to combine the projected Hamiltonian formalism with the multiple time-step formulation of molecular dynamics [17]. These subjects are currently under investigation and will be part of a forthcoming article.

ACKNOWLEDGMENTS

The author gratefully acknowledges inspiring discussions with G. Ciccotti regarding the connection between the projector algebra and holonomic constraints, and D. Frenkel for important comments during the early stage of this project. This work was supported by the Leverhulme Trust.

APPENDIX: EQUATIONS OF MOTION VIA THE NOSÉ-HOOVER THERMOSTAT

In order to produce a canonical distribution function, the Nosé-Hoover thermostatting dynamics [15,16] can be used once it is properly adapted to the projector mechanism. First of all, let us write the Hamiltonian equations of motion derived from H_b,

$$\dot{r}_i=\frac{1}{\sqrt{m_i}}\sum_j (\chi^{-1})_{ij}\frac{p_j}{\sqrt{m_j}},$$
$$\dot{p}_i=F_i-\frac{1}{2}\nabla_i(\underline{\underline{\chi}}^{-1}:\underline{\underline{K}}) \tag{A1}$$

so that we introduce the Nosé-Hoover dynamic friction via the following non-Hamiltonian equations of motion:

$$\dot{r}_i=\frac{1}{\sqrt{m_i}}\sum_j (\chi^{-1})_{ij}\frac{p_j}{\sqrt{m_j}},$$
$$\dot{p}_i=F_i-\frac{1}{2}\nabla_i(\underline{\underline{\chi}}^{-1}:\underline{\underline{K}})-\zeta\left(p_i-\frac{m_i P_0}{M}\right),$$
$$\dot{\zeta}=\frac{2}{\tau^2 3NkT}\left(\frac{1}{2}\underline{\underline{\chi}}^{-1}:\underline{\underline{K}}-\frac{P_0^2}{2M}-\frac{3(N-1)kT}{2}\right),$$
$$\dot{s}=3(N-1)s\zeta, \tag{A2}$$

where $P_0=\Sigma_i p_i$ is the total momentum, $M=\Sigma_i m_i$ is the total mass, ζ and s are auxiliary dynamical variables that thermalize the system to the temperature T, and finally, τ is a coupling time regulating the thermostat efficiency.

The set of equations (A2) conserves total momentum if the RC and the potential are invariant by translation. In fact, in this case $\Sigma_i F_i=\Sigma_i\nabla_i\xi=\Sigma_i\nabla_i(\underline{\underline{\chi}}^{-1}:\underline{\underline{K}})=0$ and a truly canonical distribution is produced, as discussed in [18]. We notice that in the equations of motion (A2), the terms derived from the Hamiltonian H_b do not introduce any metric factor in the distribution function in the sense of Tuckerman *et al.* [19].

The conserved quantities derived from the equations of motion (A2) are

$$H_c=H_b+\frac{\tau^2 3NkT}{2}\zeta^2+kT\ln s=\text{const},$$
$$P_0=\text{const} \tag{A3}$$

and the associated distribution function is

$$f_b(r,p,\zeta)\propto\int ds\,\delta(H_c-\text{const})\prod_\nu \delta(P_{0\nu}-\text{const})$$
$$=\int ds\,\delta(H_b+\tau_T^2 3NkT_0\zeta^2/2+kT_0\ln s-\text{const})$$
$$\times\prod_\nu \delta(P_{0\nu}-\text{const})$$
$$\propto\exp[-\beta_0(H_b+\tau_T^2 3NkT_0\zeta^2/2)]$$
$$\times\prod_\nu \delta(P_{0\nu}-\text{const}). \tag{A4}$$

We note here that the Nosé-Hoover thermostat achieves thermalization in agreement with the equipartition theorem that takes the form

$$\frac{1}{2}\left\langle\sum_i p_i\frac{\partial H_b}{\partial p_i}\right\rangle_b=\frac{1}{2}\langle\underline{\underline{\chi}}^{-1}:\underline{\underline{K}}\rangle_b=\frac{3NkT}{2}. \tag{A5}$$

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, London, 1996).
[2] E.A. Carter, G. Ciccotti, J.T. Hynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[3] M. Sprik and G. Ciccotti, J. Chem. Phys. **109**, 7737 (1998).
[4] P. Bolhuis, C. Dellago, and D. Chandler, Faraday Discuss. **110**, 421 (1998).
[5] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[6] B. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[7] D.D. Frantz, D.L. Freeman, and J.D. Doll, J. Chem. Phys. **97**, 5713 (1990).
[8] R. Zhou and B. Berne, J. Chem. Phys. **107**, 9185 (1997).
[9] A.F. Voter, J. Chem. Phys. **78**, 1997 (1997).
[10] *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B.J. Berne, G. Ciccotti, and D.F. Coker (World Scientific, Singapore, 1997).
[11] C.H. Bennett, J. Comp. Physiol. **19**, 267 (1975).
[12] M.B. Forrest and U.W. Suter, Mol. Phys. **82**, 393 (1994).
[13] G.R. Kneller and T. Mulders, Phys. Rev. E **54**, 6825 (1996).
[14] H. A. Kramers, Physica (Utrecht) **7**, 284 (1940).
[15] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[16] W.G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[17] M.E. Tuckerman, G.J. Martyna, and B.J. Berne, J. Chem. Phys. **97**, 1990 (1992).
[18] S. Melchionna, Phys. Rev. E **61**, 6165 (2000).
[19] M.E. Tuckerman, C.J. Mundy, and G.J. Martyna, Europhys. Lett. **45**, 149 (1999).