

Improved importance sampling distribution for rate constant calculation

Massimo Mella^{a)}

School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 10 January 2005; accepted 15 March 2005; published online 24 May 2005)

An efficient method to compute the thermal rate constant for rare events within the correlation function $C(t)$ approach is presented. This method, which is based on a modification of the sampling function used to evaluate the dynamical correlation function $C(t)$, can be applied to high-dimensional systems having a rough energy landscape without previous knowledge on the transition states location. In this work, the sampling of a Boltzmann-like distribution for the linear momenta with a lower inverse temperature ($\beta^* = 1/kT^*$) than the correct one ($\beta = 1/kT$) is proposed as a way to increase the number of reactive trajectories. The mismatch between the β^* and β distributions is then corrected by a reweighting procedure which allows one to obtain the exact correlation function $C(t)$. The efficiency of this scheme in computing the rate of a particle jumping across the barrier of a simple 2D double well potential is improved by a factor 4–25 depending on the relative value of the original β and modified β^* temperatures. When merged with the “puddle potential” method [S. A. Corcelli, J. A. Rohman, and J. C. Tully, *J. Chem. Phys.*, **118**, 1085 (2003)], the new importance sampling function improves the efficiency of the correlation function approach by a factor 16–800 with respect to the unbiased sampling. To test the method in a more challenging case, the previous model system was extended by adding six harmonically restrained particles, each one interacting with the diffusing particle. This model introduces both the possibility of energy exchange and a rougher energy landscape. The new sampling function alone is found to produce an improvement in efficiency of, at least, an order of magnitude when compared with the unbiased case; when merged with the puddle potential method, a 400-fold saving in computer time is found. © 2005 American Institute of Physics.
[DOI: 10.1063/1.1903903]

I. INTRODUCTION

During a physical or chemical transformation, a system is very often required to surmount a free energy barrier to change its state; so the overall process needs an excess of energy to take place. At constant temperature, the energy excess required to pass the barrier is usually provided by the interaction with a thermal bath (sometimes the bath is composed by other degrees of freedom of the system not directly involved in the transformation), and its magnitude is quantified using familiar concepts such as the activation (free) energy. The latter allows us to predict and interpret the rate of a specific transformation as a function of the features of the potential energy surface of the system itself. Indeed, obtaining an estimate of the activation energy and the location of the stationary point connecting two states is often the very first step leading to the prediction of a reaction rate or a diffusion coefficient in a chemical system. Once these two pieces of information are obtained, transition state theory (TST) (Ref. 1) can be used as a practical tool to estimate the transformation rate.

Despite its practical utility in understanding and rationalizing reaction rates, the applicability of TST is limited to situations where its fundamental hypotheses are satisfied. More specifically, TST assumes that a transition state (TS, i.e., a stationary point of first order) can be located and that,

once the system has reached the TS and begun to “roll” downhill toward the product P state, there is no possibility of coming back to the reactant R region (hence no recrossing of the transition state geometry). The latter requirement can be relaxed if one allows the choice of the planar surface defining the common boundary between R and P to vary so to minimize the computed reaction rate; in fact, not accounting for the recrossing, the TST rate is always higher than the exact one (see Ref. 2 for a recent review on the state of TST). Alternatively, the exact transmission coefficient can be estimated by starting swarms of trajectories from the dividing surface and counting the number of recrossing events.³

Condensed phase reactions are among the processes for which the direct application of TST (or its variational variant VTST) is hampered with several difficulties. As an example, it is mentioned that it may become quite difficult to locate a TS for systems with a rough energy landscape; it is also quite likely that the no-recrossing requirement would be violated due, for instance, to the solvent cage around the reactants. Luckily, TST can be shown to represent only a particular case of a general approach based on the time-dependent picture of the reactive process,⁴ the so called “flux-flux correlation approach.” The latter dispenses altogether with locating the TS state between two stable states and focuses on the direct calculation of the dynamical trajectories that bring the system from the R state to the P one. These states are defined

^{a)}Electronic mail: MellaM@cf.ac.uk

as the volume of phase space containing the vast majority of the probability distribution for the P and R species during equilibrium simulations.

As explained in the following, in the flux-correlation approach one must evaluate the conditional probability for the system of reaching state P at time t when it was in state R at the time origin. However, one should bear in mind that any dynamical event requiring an excess energy is usually quite rare, and that it may be difficult to tackle this calculation simply relying on straightforward molecular dynamics (MD) simulations. A substantially similar difficulty is present even if one constrains the problem to a more qualitative and much less demanding question: what are the most likely events or transformations taking place from a given state as an effect of thermal excitation? To tackle these fundamental difficulties, several dynamical methods have been devised to overcome the critical separation between the time scale of atomic vibration and barrier passing. Among these, it is worth citing the “hyperdynamics” method,⁵ its simpler variant called the “boost potential,”⁶ and the “conformational flooding” procedure.⁷ The latter has been recently evolved to a new practical level by Laio and Parinello⁸ to facilitate the escape of the system from deep free energy wells. Sorensen and Voter also introduced a temperature-accelerated dynamics (TAD) to speed up rare events by increasing the simulated temperature;⁹ exploiting a similar idea, VandeVondele and Rothlisberger¹⁰ suggested that, by adiabatically decoupling subsets of degrees of freedom, a higher temperature could be used to speed up the rate of thermal events of a restricted number of active atoms [canonical adiabatic free energy sampling (CAFES)]. One of the advantage of TAD or similar methodologies is that the so computed transition rate can be successively converted back to lower temperature by assuming the validity of TST. This assumption is, however, as “risky” as the original assumption of TST itself, and two different strategies have been recently proposed to bypass these difficulties: the transition path sampling (TPS) (Ref. 11) and the “puddle potential” (PP) technique.¹² The latter is an ingenious extension of the boost potential proposed by Steiner *et al.*⁶ to the calculation of dynamical rates and it has already found application in tackling the ergodicity problem in molecular simulations.¹³

Despite their common goal, the TPS and PP methods use different strategies to tackle the rare event issue; whereas the TPS method focuses on harvesting a swarm of trajectories satisfying the requirement of starting from R and passing through P at least once in their lifetime, the PP technique tries to smooth the rough energy landscape experienced by a trajectory leaving R by applying a bias potential. As a consequence of their conceptual differences, quite different algorithmic implementations are produced, and the PP method appears, in our view, the simplest to code. Being felt that this simplicity represents an appealing feature, the work presented here is concerned with improving the understanding and performance of the PP method; its main goal is to present algorithmic modifications able to improve the efficiency of PP and, hopefully, its range of applicability.

The structure of the paper is organized as follow. Section II presents a brief introduction to the flux-flux correlation

method conceived to make the manuscript self-standing and to familiarize the reader with the notation. Also, some of the details of the PP method¹² are discussed together with a mechanistic understanding of its improved performance. In the last part of Sec. II, the modifications to the PP method are introduced and discussed on the basis of the general theory of the importance sampling technique in Monte Carlo simulations. Section III presents the results obtained by simulating two nontrivial model systems (composed by 2 and 14 degrees of freedom, respectively) using the original PP approach and the newly proposed modifications. A detailed comparison of the results and relative efficiency is also given. Finally, Sec. IV presents our conclusions and indicates a few fields of applications in which the improved methodology can be helpful in producing accurate transformation rates.

II. THEORY AND METHOD

The objective of this section is twofold. First, we wish to give a summary of the theory relevant to the “flux-flux correlation” (or correlation function) approach as an aid to the understanding of the developments presented in later sections. Second, the PP approach¹² and the new method are introduced as two extremal cases of a more general methodology based on importance sampling in phase space, a task that is considered easier after such a discussion.

Most of the following is based on the arguments given in Chap. VIII of Ref. 14, so that the present discussion is restrained to present only the necessary information.

In classical mechanics, the state of a system composed by N particles at time t is indicated by the phase space point $[\mathbf{R}(t), \mathbf{P}(t)]$. Here, $\mathbf{R}(t) = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ is the vector containing all the particle positions while $\mathbf{P}(t) = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ is the vector describing the particle linear momenta. For a system in thermal equilibrium with a heat reservoir at temperature T the average value of a dynamical variable $A(\mathbf{R}, \mathbf{P})$ is given by

$$\langle A \rangle_{\mathcal{H}} = \frac{\int A(\mathbf{R}, \mathbf{P}) e^{-\beta \mathcal{H}(\mathbf{R}, \mathbf{P})} d\mathbf{R} d\mathbf{P}}{\int e^{-\beta \mathcal{H}(\mathbf{R}, \mathbf{P})} d\mathbf{R} d\mathbf{P}}, \quad (1)$$

where $\beta = (kT)^{-1}$ and $\langle \dots \rangle_{\mathcal{H}}$ indicates the canonical ensemble average for the system described by \mathcal{H} .

Let us now assume that the system was prepared in a slightly different state from the one characterized by the equilibrium Boltzmann distribution $e^{-\beta \mathcal{H}(\mathbf{R}, \mathbf{P})}$ by coupling a scalar field f with the dynamical variable A ; the Hamiltonian function for the system becomes $\mathcal{H}_p = \mathcal{H} - fA$. Once the perturbation (i.e., the scalar field f) is switched off at $t=0$, the system relaxes back to the canonical equilibrium state following a dynamics that is given by solving the unperturbed Hamilton equations (or the equivalent Liouville equations) with the off-equilibrium distribution as starting probability density and with a rate given by

$$\begin{aligned} \Delta A(t) &= \bar{A}(t) - \langle A \rangle_{\mathcal{H}} \approx \beta f [\langle A(t)A(0) \rangle_{\mathcal{H}} - (\langle A \rangle_{\mathcal{H}})^2] \\ &= \beta f \langle \delta A(t) \delta A(0) \rangle_{\mathcal{H}}. \end{aligned} \quad (2)$$

This is the classical form of the fluctuation-dissipation theo-

rem in which we defined $A(t)=A[\mathbf{R}(t),\mathbf{P}(t)]$ and $\delta A(t)=A(t)-\langle A \rangle_{\mathcal{H}}$.

This form of the fluctuation-dissipation theorem can be used to link the change in chemically relevant quantities (such as the concentration of R and P in the monomolecular process $R \leftrightarrow P$) to the microscopic dynamics that takes place at molecular level; the latter task that can be done by introducing the microscopic dynamical quantity $\bar{n}_R(t)$. To define it, we start by supposing that our chemical process has an intrinsic coordinate $q=q(\mathbf{R})$ which allows R and P to be identified by two different ranges of q . If this is the case, we define q^* as the value of q separating the two species. Therefore, indicating with $H[q]$ the Heaviside function such that $H[q]=1$ if $q \leq q^*$ and $H[q]=0$ if $q > q^*$, one can define \bar{n}_R as

$$\begin{aligned} \bar{n}_R(t) &= H_R[q(t)] = 1 \quad \text{if } q[\mathbf{R}(t)] < q^*, \\ \bar{n}_R(t) &= H_R[q(t)] = 0 \quad \text{if } q[\mathbf{R}(t)] > q^*, \end{aligned} \quad (3)$$

where the index R of H indicates we are defining the reactant state. It should be recognized that is not necessary for q to be a “real” coordinate but it can be any “variable” of the system that is able to discriminate between R and P during the process. Using this definition, the algebraic manipulation of Eq. (2) leads to

$$\frac{\partial C(t)}{\partial t} = \frac{\langle \dot{H}_P[q(t)]H_R[q(t=0)] \rangle_{\mathcal{H}}}{\langle H_R[q(t=0)] \rangle_{\mathcal{H}}} = k_{RP}e^{-t/\tau_{rlx}}, \quad (4)$$

where $\tau_{rlx}=(k_{RP}+k_{PR})^{-1}$, k_{RP} and k_{PR} are the direct and inverse rate constants, and the correlation function $C(t)$ has been defined as

$$C(t) = \frac{\int H_P[\mathbf{R}(t),\mathbf{P}(t)]H_R[\mathbf{R}(0),\mathbf{P}(0)]e^{-\beta\mathcal{H}[\mathbf{R}(0),\mathbf{P}(0)]}d\mathbf{R}(0)d\mathbf{P}(0)}{\int H_R[\mathbf{R}(0),\mathbf{P}(0)]e^{-\beta\mathcal{H}[\mathbf{R}(0),\mathbf{P}(0)]}d\mathbf{R}(0)d\mathbf{P}(0)}. \quad (7)$$

Once again, the trajectories $[\mathbf{R}(t),\mathbf{P}(t)]$ are obtained by solving the Hamilton equation for the unperturbed Hamiltonian. If the latter can be written as $\mathcal{H}[\mathbf{R},\mathbf{P}]=\mathcal{T}[\mathbf{P}]+\mathcal{V}[\mathbf{R}]$, the phase space integrals can be separated and Eq. (7) becomes

$$C(t) = \frac{\int H_P[\mathbf{R}(t)]H_R[\mathbf{R}(0)]e^{-\beta\mathcal{V}[\mathbf{R}(0)]}d\mathbf{R}(0) \int \frac{e^{-\beta\mathcal{T}[\mathbf{P}(0)]}}{\int e^{-\beta\mathcal{T}[\mathbf{P}(0)]}d\mathbf{P}(0)}d\mathbf{P}(0)}{\int H_R[\mathbf{R}(0)]e^{-\beta\mathcal{V}[\mathbf{R}(0)]}d\mathbf{R}(0)}, \quad (8)$$

where we have made use of the fact that the definition of q does not contain the linear momenta.

This factorization is, however, only formal because the position of the system in configuration space at time t depends on the specific choice of the initial momenta at $t=0$. Nevertheless, Eq. (8) is a useful starting point to propose possible algorithms to compute $C(t)$. For instance, one could sample the distribution $H_R[\mathbf{R}(0)]e^{-\beta\mathcal{V}[\mathbf{R}(0)]}$ in configuration space by using a Metropolis Monte Carlo algorithm, and randomly generate starting linear momenta with distribution $e^{-\beta\mathcal{T}[\mathbf{P}(0)]}/\int e^{-\beta\mathcal{T}[\mathbf{P}(0)]}d\mathbf{P}(0)$. The trajectory is then computed by integrating the Hamilton equations, and the overall algorithm is similar to the Hybrid Monte Carlo method proposed

$$C(t) = \frac{\langle H_P[q(t)]H_R[q(t=0)] \rangle_{\mathcal{H}}}{\langle H_R[q(t=0)] \rangle_{\mathcal{H}}}. \quad (5)$$

Equations (4) and (5) give us the link between the macroscopic observable k_{RP} and the microscopic events taking place during the reactive process. Focusing on a macroscopic time longer than the microscopic time τ_{mic} of the reactive events (i.e., the time necessary for an activated species to reach P starting from R) but still shorter than the system lifetime τ_{rlx} (i.e., $\tau_{mic} \ll \tau \ll \tau_{rlx}$), one can make use of $\tau/\tau_{rlx} \ll 1$ so that Eq. (4) becomes

$$\frac{\partial C(t)}{\partial t} = \frac{\langle \dot{H}_P[q(t)]H_R[q(t=0)] \rangle_{\mathcal{H}}}{\langle H_R[q(t=0)] \rangle_{\mathcal{H}}} \approx k_{RP}. \quad (6)$$

We can now attribute some meaning to the correlation function $C(t)$. This gives the conditional probability that a trajectory with starting condition $[\mathbf{R}(t=0),\mathbf{P}(t=0)]$ representing \mathbf{R} in the canonical ensemble, ends up in the product region after the elapsed time t . Such a correlation function can, in principle, be computed using a swarm of short trajectories spanning the time range $[0, \tau]$ and initiated by sampling the canonical distribution in coordinate space and randomly generating the initial velocities or momenta (e.g., see Ref. 11 for a description of the TPS approach and Ref. 12 for the PP one). In the following, efficient methodologies to tackle this task are introduced and discussed.

A. The puddle potential approach

To proceed further, let us write the correlation function [Eq. (5)] so that the ensemble average becomes apparent,

by Duane and Kennedy.¹⁵ However, this simplistic algorithm is doomed to failure if the process is required to pass a barrier higher than a few $kT=\beta^{-1}$. In this case, only a few trajectories would have enough energy and/or be close enough to its top to surmount the barrier, so that the statistics collected even after a large number of spanned trajectories may be quite poor.

The solution to this problem proposed in the recent PP approach¹² is to add an additional term $\mathcal{V}_{bias}(\mathbf{R})$ in the potential energy with the same spirit of the “umbrella potential” method. This is chosen to bias the sampling away from the local minimum and toward regions of higher $\mathcal{V}(\mathbf{R})$ and has the analytical form

$$\mathcal{V}_{\text{bias}}[\mathbf{R}] = 0 \quad \text{if } \mathcal{V}[\mathbf{R}] > V_{\text{puddle}},$$

$$\mathcal{V}_{\text{bias}}[\mathbf{R}] = V_{\text{puddle}} - \mathcal{V}[\mathbf{R}] \quad \text{if } \mathcal{V}[\mathbf{R}] < V_{\text{puddle}}, \quad (9)$$

so that Eq. (8) can be rewritten as

$$C(t) = \frac{\int H_P[\mathbf{R}(t)] H_R[\mathbf{R}(0)] e^{-\beta(\mathcal{W}[\mathbf{R}(0)] - \mathcal{V}_{\text{bias}}[\mathbf{R}(0)])} d\mathbf{R}(0) \int \frac{e^{-\beta\mathcal{T}[\mathbf{P}(0)]}}{\int e^{-\beta\mathcal{T}[\mathbf{P}(0)]} d\mathbf{P}(0)} d\mathbf{P}(0)}{\int H_R[\mathbf{R}(0)] e^{-\beta(\mathcal{W}[\mathbf{R}(0)] - \mathcal{V}_{\text{bias}}[\mathbf{R}(0)])} d\mathbf{R}(0)}. \quad (10)$$

Here, $\mathcal{W} = \mathcal{V} + \mathcal{V}_{\text{bias}}$ and V_{puddle} is a constant chosen to improve the efficiency of the algorithm [note that $\mathcal{V}_{\text{bias}}$ in Eq. (10) has been defined in a way that is slightly different to the one used in Ref. 12]. Using Eq. (10), one could decide to sample $H_R[\mathbf{R}(0)] e^{-\beta\mathcal{W}[\mathbf{R}(0)]}$ instead of $H_R[\mathbf{R}(0)] e^{-\beta\mathcal{V}[\mathbf{R}(0)]}$, and correct for the difference in the two distributions by means of a reweighting procedure; at every point in configuration space sampled following $H_R[\mathbf{R}(0)] e^{-\beta\mathcal{W}[\mathbf{R}(0)]}$, one attaches the well-behaved weight $e^{-\beta\mathcal{V}_{\text{bias}}[\mathbf{R}(0)]}$.

Figure 1 schematically summarizes the effect of different choice for V_{puddle} on the sampled Monte Carlo distribution $e^{-\beta\mathcal{W}[\mathbf{R}(0)]}$ for a one-dimensional (1D) double well model potential. In this figure, five different values for V_{puddle} were used to show the change in the sampled distributions for $\beta = 4$, namely, $V_{\text{puddle}} = 0.00, 0.25, 0.50, 0.75, 1.00$. Upon increasing the V_{puddle} value, the probability of sampling configurations well inside the barrier region and close to $x=0$ is increased; in this position, one would ideally place the separatrix between the reactant and product regions. Such an increased sampling in the barrier region could generate two distinct algorithmic advantages: first, for a given amount of kinetic energy given to a sampled replica of the system, the number of trajectories able to surmount the barrier should be increased simply due to the increased sampling density close to $x=0$; second, since the sampled configurations are allowed to visit regions of higher potential energy more frequently, a larger number of started trajectories would possess enough kinetic energy to pass the maximum, ending up in the product region. In other words, a larger number of trajectories should be able to depart from the R region at time $t=0$ and visit the P region at time t improving the statistical estimate of $C(t)$.

However, these algorithmic improvements do not come without a price. Since one is forced to sample a different

distribution from the one that naturally emerges from Eq. (8), a reweighting procedure is needed in order to eliminate the distribution bias. This reweighting (indeed *any* weighting procedure) would increase the variance of the $C(t)$ estimator.¹⁶ The structure of Eq. (10) also imposes to estimate separately both numerator and denominator to properly compute an estimate for $C(t)$, a necessity that introduces a bias unless an infinite number of samples is used. The last problem is especially important when one is forced to estimate a quantity by using a relatively poor sample of events.

Despite these two methodological issues, the increased number of trajectories that reach the P region may in practice improve the poor sampling obtained when the barrier height is several kT , therefore allowing a more accurate, precise, and robust calculation of the correlation function $C(t)$. In other words, PP should help in effectively overcoming the ergodicity problem associated with this kind of calculations, and indeed, it has been shown to improve the precision of the rate constant calculation for two model systems,¹² a simple 2D double well potential and the 1D isomerization process of a bistable molecule immersed in a Weeks–Chandler–Andersen fluid.

B. The “temperature-accelerated puddle potential” dynamics

We shall now make a step beyond the PP method by noticing that its simple and effective change in the sampled distribution can be easily generalized by including also a biased sampling in the linear momentum space if $\mathcal{H}[\mathbf{R}, \mathbf{P}] = \mathcal{T}[\mathbf{P}] + \mathcal{V}[\mathbf{R}]$. To do so, let us introduce a new normalized phase space density $\rho(\mathbf{R}, \mathbf{P})$ and insert it into Eq. (7). Multiplying and dividing both numerator and denominator one gets

$$C(t) = \frac{\int H_P[\mathbf{R}(t), \mathbf{P}(t)] \mathbf{H}_R[\mathbf{R}(0), \mathbf{P}(0)] e^{-\{\beta\mathcal{H}[\mathbf{R}(0), \mathbf{P}(0)] + \ln(\rho[\mathbf{R}(0), \mathbf{P}(0)])\}} \rho[\mathbf{R}(0), \mathbf{P}(0)] d\mathbf{R}(0) d\mathbf{P}(0)}{\int H_R[\mathbf{R}(0), \mathbf{P}(0)] e^{-\{\beta\mathcal{H}[\mathbf{R}(0), \mathbf{P}(0)] + \ln(\rho[\mathbf{R}(0), \mathbf{P}(0)])\}} \rho[\mathbf{R}(0), \mathbf{P}(0)] d\mathbf{R}(0) d\mathbf{P}(0)}. \quad (11)$$

The mathematical structure of this equation is identical to the one of Eq. (10) with the only difference that we can now

sample a phase space distribution different from the canonical one. However, it is important to stress that the trajectories

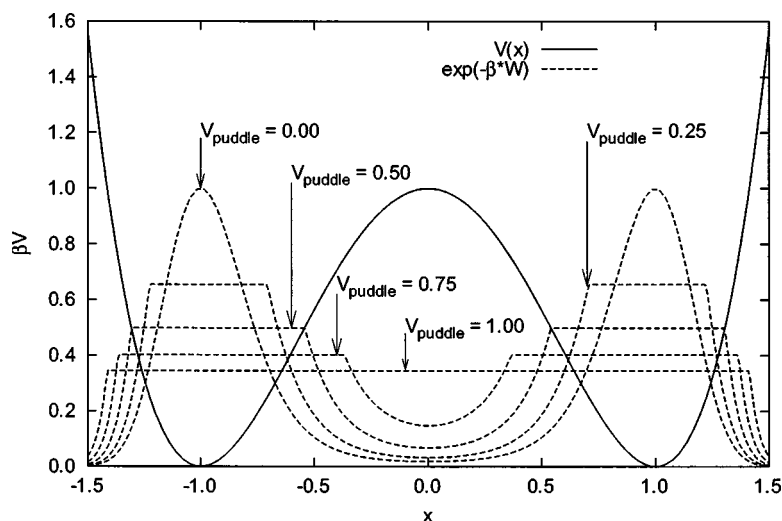


FIG. 1. Plot of βV (solid line) and of the normalized $\exp[-\beta W]$ probability density functions (dotted lines) for a simple 1D double well potential as a function of V_{puddle} . Upon increasing V_{puddle} the density increases its values in the region around the barrier.

$[\mathbf{R}(t), \mathbf{P}(t)]$ are obtained by solving the Hamilton equation for the unperturbed Hamiltonian and that $\rho(\mathbf{R}, \mathbf{P})$ modifies only the process by which one chooses the initial condition of the differential equations while the dynamics of every trajectory is left unchanged. In other words, whereas every computed trajectory is an exact representation of a possible short time dynamics for the systems, the relative probability of sampling any pair of trajectories is no longer given by the standard Boltzmann distribution but instead by a suitable distribution that one hopes may improve the efficiency of the $C(t)$ calculation. Once again, the bias introduced by using a different sampling distribution can be eliminated by re-weighting each trajectory “on the fly” with the weight $w_{\rho}[\mathbf{R}(0), \mathbf{P}(0)] = e^{-\beta \gamma t[\mathbf{R}(0), \mathbf{P}(0)] - \ln[\rho(\mathbf{R}(0), \mathbf{P}(0))]}$.

Clearly, the range of possible choices for the sampling distribution is quite large, but an obvious requirement is that the weight $w_{\rho}[\mathbf{R}(0), \mathbf{P}(0)]$ must behave properly and should not diverge faster than $(\rho[\mathbf{R}(0), \mathbf{P}(0)])^{-1/2}$ for $|\mathbf{P}|, |\mathbf{R}| \rightarrow \infty$. If this condition is satisfied, then the statistical average obtained by a Monte Carlo simulation possesses a well defined and finite variance. Although not formally necessary, the implementation of the method is simplified by choosing $\rho[\mathbf{R}(0), \mathbf{P}(0)]$ to be written as a normalized distribution in configuration space $\gamma[\mathbf{R}(0)]$ times a normalized distribution in momentum space $\mu[\mathbf{P}(0)]$. In this way, the sampling of the full $\rho[\mathbf{R}(0), \mathbf{P}(0)]$ can be split in two parts and carried out using any variant of the Metropolis scheme for the coordinate part and a simpler method (e.g., rejection method) to generate the momentum distribution.

Several issues should now be clarified in order to complete the presentation of the method. For instance, what are the possible advantages of using a different momentum distribution to compute the correlation function $C(t)$? Also, which specific analytical form $\mu[\mathbf{P}(0)]$ should be used to improve the efficiency of the method?

In order to answer these questions, we notice that it is possible to distinguish two possible limiting effects generated by an alternative momentum distribution as a function of the differences between $\mu[\mathbf{P}(0)]$ and the Boltzmann distribution. First, one could choose a broader momentum distribution so that an amount of kinetic energy per degree of

freedom larger than kT is assigned more often than when using the Boltzmann distribution. By doing so, an increase in the average number of trajectories that can surmount the energy barrier is produced. In other words, momenta distributions $\mu[\mathbf{P}(0)]$ with longer tails for $\|\mathbf{P}\| \rightarrow \infty$ may improve the statistics of the simulations by “energizing” the swarm of randomly initiated trajectories; this improvement should roughly follow the exponential law $\exp(\beta \Delta T)$, where ΔT is the difference between the average kinetic energy produced by the standard Boltzmann distribution and by the new $\mu[\mathbf{P}(0)]$. Second, despite the fact that the Boltzmann distribution for the momenta is isotropic, one may be better off choosing a skewed distribution in order to sample more often initial conditions that are directed toward the P region. For instance, one may expect that a low temperature (i.e., low kinetic energy) trajectory would reach the P region more easily if directed toward the transition state by the biased selection of initial momenta. Thus, it could be advantageous to bias the angular distribution in order to sample more often the “slice” of angles pointing in the direction of the transition states.

Before discussing the actual choice for the sampled distributions, a comment on the technicalities of the correlation function method is due. In a many-body system, it may not be straightforward to define q^* to represent the location of a TS for the process, and so P and R are often defined on the basis of the regions of q visited during a normal Monte Carlo (MC) or MD simulation of finite length. This means that the P and R regions may not be defined as having a shared boundary, and that there is a range of q that does not belong either to P or R for which we have no information. This range may even contain very high barriers. If this were the case, the PP may lose some of its effectiveness due to the fact that those barriers are not “smoothed” out by the puddle potential and that the sampled density is still far from the top of these barriers. Conversely, the possibility of increasing the average kinetic energy assigned to the trajectories could substantially improve the situation by effectively overcoming the limited knowledge of the energy landscape between R and P . This is done by “muscling” the system to surmount every energy barrier by means of the additional kinetic boost.

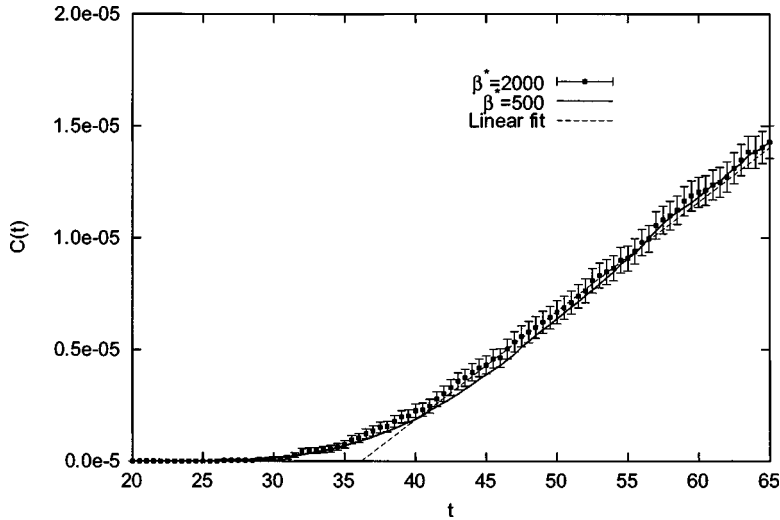


FIG. 2. Comparison between $C(t)$ functions at inverse temperature $\beta=2000$ computed using $\beta^*=2000$ and $\beta^*=500$ as kinetic temperature. The dashed line is the linear fit to the results obtained using $\beta^*=2000$. The excellent agreement between the two simulations indicates the absence of bias due to the use of a higher effective β^* .

To conclude this section, a particular choice for the distribution $\rho[\mathbf{R}(0), \mathbf{P}(0)]$ is now given, and this reads

$$\rho[\mathbf{R}(0), \mathbf{P}(0)] \propto e^{-\beta V_{\text{bias}}[\mathbf{R}(0)]} \prod_i^{N_{\text{dof}}} \left[\frac{2m_i \pi}{\beta^*} \right]^{1/2} e^{-\beta^* (p_i^2 / 2m_i)}. \quad (12)$$

Here, $\beta > \beta^*$, m_i are the particle masses, and N_{dof} is the number of degrees of freedom of the system. In other words, we choose $\rho[\mathbf{R}(0), \mathbf{P}(0)]$ to be similar to the distribution sampled by the puddle potential in the configuration space, and to be a higher temperature Boltzmann distribution for the momenta. The latter choice is motivated by the effectiveness of the temperature-accelerated dynamics⁹ and of the CAFES procedure¹⁰ in improving the exploration of the configuration space, and therefore the new method is dubbed temperature-accelerated puddle potential dynamics (TAPPD).

Bearing in mind the previous comments on the possible advantages of sampling a skewed distribution, we consider our choice for the momentum distributions a very “naive” one. However, we also consider interesting to evaluate the performance of the method while using an unsophisticated approach to fully appreciate its possible strengths and weaknesses. Also, the isotropic form chosen in Eq. (12) may represent a valid choice when no information is available on potential energy surface.

III. TESTS OF THE METHOD

In this section, the TAPPD method is applied to two test cases in order to quantitatively evaluate its performances. As first system, the 2D potential of Refs. 12 and 17 is used. This simple system has the advantage that any quantity computed in the simulations can be accurately monitored very quickly and with high statistical precision. Also, the sampled distributions can be plotted to make sure that one is indeed generating the correct probability densities.

In order to expose the method to a more severe test, six harmonically restrained particles were added to the previous 2D model system to mimic the chance of energy transfer, recrossing, and the possibility of having a change in activation free energy as a function of the simulation temperature.

Also, this model bears relevance to the atomic diffusion process on the surface of a metal catalyst, so that in our view it represents an interesting and challenging case study.

A. 2D double well potential

The analytical form for the 2D model potential reads

$$V_{2D}(x, y) = V_{\text{TS}} [4(1 - x^2 - y^2)^2 + 2(x^2 - 2)^2 + [(x + y)^2 - 1]^2 + [(x - y)^2 - 1]^2 - 2] / 6 \quad (13)$$

and, apart from the multiplicative constant V_{TS} , it is identical to the form used in Refs. 12 and 17. This potential has two symmetric minima located at $[(20/16)^{1/2}, 0]$ and $[-(20/16)^{1/2}, 0]$ with energy $-V_{\text{TS}}/12$. It also presents two symmetric transition states at $(0, -1)$ and $(0, +1)$ with energy V_{TS} , so that the barrier height is $E_{\text{barrier}} = 13V_{\text{TS}}/12$. In the following, we use an unit particle mass and choose $V_{\text{TS}} = 0.005$. We stress that only the product βE_{barrier} is important in testing the method and in defining its efficiency since it is

TABLE I. Computed standard errors $\sigma(\beta, \beta^*)$ for simulations carried out at different values of β and β^* without puddle potential. All the results were obtained sampling 25×10^6 trajectories.

β	β^*	$\sigma(\beta, \beta^*)$
2400	2400	2.1×10^{-7}
2400	2000	8.6×10^{-8}
2400	1000	4.3×10^{-8}
2400	500	4.4×10^{-8}
2000	2000	5.6×10^{-7}
2000	1000	1.5×10^{-7}
2000	700	1.3×10^{-7}
2000	500	1.5×10^{-7}
1600	1600	1.7×10^{-6}
1600	1000	6.4×10^{-7}
1600	700	5.0×10^{-7}
1600	500	4.3×10^{-7}
1000	1000	2.8×10^{-5}
1000	700	1.6×10^{-5}
1000	500	1.3×10^{-5}

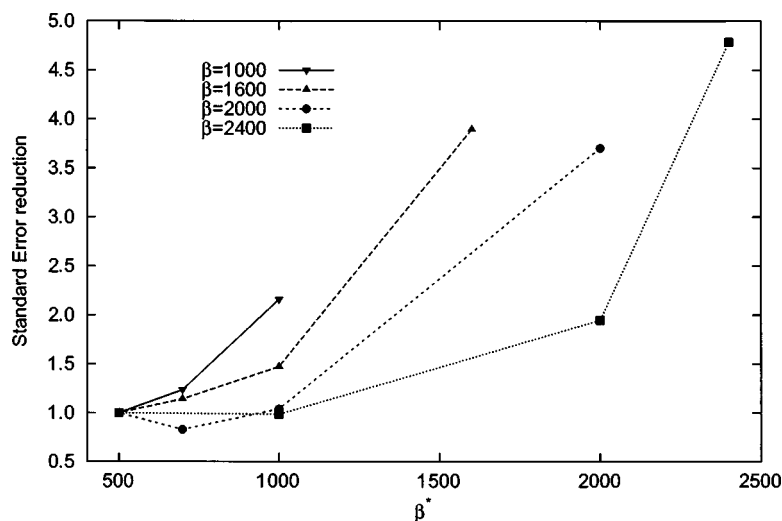


FIG. 3. Values of the standard error ratio $\sigma(\beta, \beta^*)/\sigma(\beta, 500)$ for the errors presented in Table I. The different lines represent different values of the Monte Carlo inverse temperature β used to sample the configuration space during the correlation function calculation.

the Boltzmann factor $e^{-\beta E_{\text{barrier}}}$ that dictates the number of trajectories having enough kinetic energy to pass the barrier.

As a first test, the calculation of $C(t)$ has been carried out for $V_{\text{puddle}} = -\infty$ (so no bias potential) and various values of β and β^* to check the correctness of the procedure. The simulation details are as follow. All the simulations were run sampling 25×10^7 MC configurations using a step size that gave an acceptance probability of 0.5. In order to eliminate statistical correlation, trajectories were started every 10 MC steps. So, our final ensemble average was done using 25×10^6 trajectories, a number that was found to guarantee a robust convergence of average values at all the temperatures. P and R regions were defined as two circles of radius $r = 0.7$ centered in $(1, 0)$ and $(-1, 0)$, and the MC simulation was constrained to sample only inside the R region by rejecting all the proposed configurations lying outside the circle. The Hamilton equations were integrated for 200 steps using the standard “leap-frog” algorithm and a time step of 0.5. The energy conservation was found to be better than 50 ppm at all the temperatures.

Figure 2 shows two such simulations carried out using $\beta = 2000$ as MC temperature and $\beta^* = 2000$ or $\beta^* = 500$ as a kinetic temperature. In the time range $[45, 60]$, the $C(t)$ for both simulations presents the expected linear regime characteristic of the activated process discussed by the theory presented in the preceding section (the dashed line is obtained by fitting the numerical results with a straight line in the range $[45, 55]$). More importantly, the results of both simulations agree within the statistical error associated to the $C(t)$ values. This is also true for simulations carried out using $\beta^* = 1500, 1000,$ and 700 , not shown in Fig. 2.

In the following, we estimate the relative statistical efficiency of simulations having the same length and β but using different β^* values by choosing a value of t within the linear regime of $C(t)$ and directly comparing the computed standard errors $\sigma(\beta, \beta^*)$ or their ratios at this time. In this way, one does not have to deal with the intricacy of the statistical correlation between subsequent $C(t)$ values.¹⁸

The standard error $\sigma(\beta, \beta^*)$ results for several values of β and β^* are presented in Table I and their ratio $\sigma(\beta, \beta^*)/\sigma(\beta, 500)$ is plotted in Fig. 3. The latter clearly

shows the statistical improvement generated by using a higher temperature ($\beta^* < \beta$) in sampling the initial momenta for the trajectories. For instance, using $\beta^* = 500$ gives an improvement in efficiency (proportional to the $[\sigma(\beta, \beta^*)/\sigma(\beta, 500)]^2$) of a factor 4–22 depending on the value of β .

To test the ability to reproduce the correct behavior with respect to the temperature, simulations for several different values of β ($\beta^* = 500$) were run. The rate constants $k_{R \rightarrow P}$, extracted by fitting the linear regime of $C(t)$, produced the Arrhenius-like plot $\ln(k_{R \rightarrow P})$ vs β shown in Fig. 4. The linear fitting of the numerical results with $\ln(A) - \beta E_{\text{barrier}}$ gives $E_{\text{barrier}} = 0.00536(5)$, which is in good agreement with the theoretical value $E_{\text{barrier}} = 13V_{\text{TS}}/12 = 0.00541$.

As a last comment on these tests, we mention that the estimator of $C(t)$ is unbiased because the value of the denominator in Eq. (11) is simply given by the number of trajectories generated during the simulation.

The next step is to merge the approach just discussed with the puddle potential.¹² The results are presented in Table II; shown are the standard error $\sigma(\beta, \beta^*)$ obtained simulating the 2D double well potential using $\beta = 2000$ and various values of β^* and V_{puddle} . The computational details of these simulations are identical to the previous one.

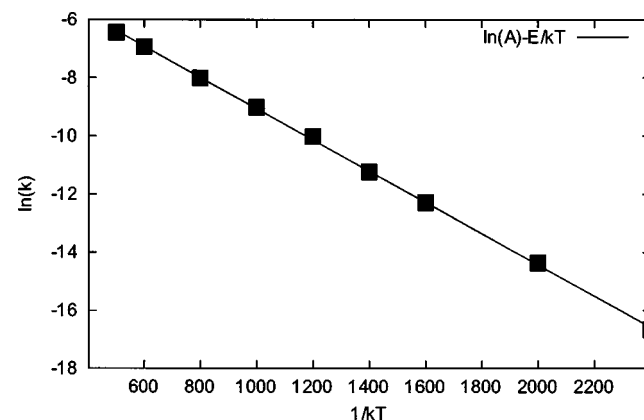


FIG. 4. $\ln(k_{RP})$ as a function of $\beta = 1/kT$ for the two 2D potential Eq. (13). The straight line is obtained by fitting $\ln A - \beta E_{\text{barrier}}$ to extract the Arrhenius parameter A and the activation energy E_{barrier} .

TABLE II. Standard errors $\sigma(\beta, \beta^*)$ obtained simulating the 2D model potential at $\beta=2000$ as function of β^* and V_{puddle} . All the results were obtained sampling 25×10^6 trajectories.

$\beta^*/V_{\text{puddle}}$	0.000 00	0.001 25	0.002 50	0.005 00
2000	5.6×10^{-7}	5.0×10^{-7}	4.7×10^{-7}	5.1×10^{-7}
1500	2.5×10^{-7}	1.4×10^{-7}	1.2×10^{-7}	1.5×10^{-7}
1000	1.5×10^{-7}	6.8×10^{-8}	4.4×10^{-8}	4.5×10^{-8}
500	1.5×10^{-7}	5.9×10^{-8}	2.6×10^{-8}	2.0×10^{-8}

The results obtained using different V_{puddle} values for the $\beta=\beta^*=2000$ case (Table II) show that using $V_{\text{puddle}} > 0$ has only a weak effect on the simulation efficiency. This outcome that can be explained remembering that, in contrast to Ref. 12, the R and P regions are defined by two circular regions around the potential minima. Due to this choice, the MC sampling of the configuration space has not direct access to the TSs even for $V_{\text{puddle}}=0.005$, and only a few more trajectories starting on the border of the R region have enough kinetic energy to surmount the barrier. Conversely, reducing β^* produces a substantial effect on the simulation efficiency and it is apparent this is more effective than increasing the V_{puddle} value (compare the results in the first column— $V_{\text{puddle}}=0$ and $\beta^* < \beta$ —of Table II with the one in the first row—i.e., $V_{\text{puddle}} > 0$ and $\beta^* = \beta$).

Despite the weak effect of V_{puddle} on $\sigma(\beta, \beta^*)$ for $\beta = \beta^* = 2000$, simultaneously decreasing β^* and increasing V_{puddle} produces a consistent reduction of $\sigma(2000, \beta^*)$ that cannot simply be explained by invoking the increased effective temperature $T^* = (k\beta^*)^{-1}$ (e.g., compare the first and second columns of Table II by computing the ratio $\sigma(2000, 2000)/\sigma(2000, \beta^*)$) and that strongly suggest some cooperation between the lower β^* and V_{puddle} . A likely explanation of this finding is the following: the configurational advantage produced by introducing V_{puddle} (i.e., the larger number of MC configurations that get closer to the TSs upon increasing the bias potential) is more effectively exploited by the larger number of trajectories starting with high kinetic energy produced by the biased sampling ($\beta^* < \beta$). Indeed, a substantial improvement in efficiency is obtained even juxtaposing a mild increase in temperature with a moderate V_{puddle} value; as a consequence, the computational cost can be reduced by a factor 16–784 depending on the parameters of the simulation.

This global improvement is even more striking at larger β (lower temperature). For instance, Table III shows the $\sigma(\beta, \beta^*)$ values for $\beta=4000$ ($\beta V_{\text{TS}}=20$) obtained with the same simulation protocol as before. To converge an unbiased

TABLE III. Standard errors $\sigma(\beta, \beta^*)$ obtained simulating the 2D model potential at $\beta=4000$ as functions of β^* and V_{puddle} . All the results were obtained sampling 25×10^6 trajectories, apart from the (4000, 4000) for which 25×10^8 were used.

$\beta^*/V_{\text{puddle}}$	0.001 25	0.002 50	0.005 00
2000	1.8×10^{-11}	9.5×10^{-12}	1.7×10^{-11}
1000	1.3×10^{-11}	2.7×10^{-12}	1.6×10^{-12}
500	9.1×10^{-12}	2.4×10^{-12}	8.0×10^{-13}

simulation ($\beta=4000$ and $V_{\text{puddle}}=-\infty$) for comparison purpose, 100 times more trajectories were needed to obtain a standard error of 1.1×10^{-11} . As a result, the effective improvement factor in efficiency spans the range 61–18 900, clearly suggesting that is possible to improve the statistical efficiency of the correlation function method by as much as five orders of magnitude for this simple 2D double well potential.

B. Augmented-2D double well potential

The analytical form for the second model potential reads

$$V_{14}(\mathbf{R}) = V_{2D}(x_d, y_d) + \sum_{i=1}^6 \frac{k_i}{2} (\mathbf{r}_i - \mathbf{r}_i^{(0)})^2 + \sum_{i=1}^6 \frac{d_i}{(\mathbf{r}_i - \mathbf{r}_d)^8}. \quad (14)$$

Here, $\mathbf{R}=(\mathbf{r}_d, \mathbf{r}_1, \dots, \mathbf{r}_6)$ is a vector indicating the position of all the particles in the system, $\mathbf{r}_d=(x_d, y_d)$ is the position of the diffusing particle sitting in the 2D well used in the previous test case, and \mathbf{r}_i is the position of the i th harmonically restrained particle. The restraining potential for the i th particle is centered in $\mathbf{r}_i^{(0)}=(\pm 2, \pm 2)$ and $(0, \pm 1.6)$ and has a force constant equal to $k_i=0.005$ for all the particles. All the d_i values were chosen equal to 0.1 on the basis of preliminary test runs that showed substantial changes in the effective potential experienced by the diffuser from the original 2D model. In particular, the free energy barrier for the hopping process was found to be larger (roughly 0.0075–0.0078, depending on the value of β) than the original $V_{\text{TS}}=0.005$.

Bearing in mind the experience obtained by simulating the simpler 2D model, it was chosen to run only two sets of simulations covering various (β, β^*) pairs, with ($V_{\text{puddle}}=0.005$) or without ($V_{\text{puddle}}=0$) puddle potential. In the $V_{\text{puddle}}=0.005$ case, only the V_{2D} potential energy experienced by the diffuser was modified using V_{puddle} (Ref. 12) (i.e., the local version of the puddle potential was used), and only the diffuser momentum was sampled using a distribution having $\beta^* < \beta$. The advantage of using this local procedure is twofold: first, the energy boost produced by the puddle potential is not diluted over all the degrees of freedom of the system;¹² second, the spread of the weight values due to the sampling of a different momentum distribution is reduced and the final average converges more rapidly. An even better simulation procedure would be represented by using the PP approach on the extended quantity $V_{2D}(x_d, y_d) + \sum_i d_i / (\mathbf{r}_i - \mathbf{r}_d)^8$ where the repulsion potential with the restrained particles is also included. However, a similar procedure may be quite cumbersome in practice when simulating more complicated systems, so that it seems reasonable to explore this less general but simpler options as a first step.

Table IV shows a comparison between the standard errors obtained by simulating the model system defined by Eq. (14). These results were obtained running simulations with parameters similar to the 2D test case and using a total of 5×10^6 MC steps (i.e., 5×10^5 trajectories) for each simulation.

As a first step in the discussion of the results, let us focus on the simulations carried out at $\beta=2000$ with and without

TABLE IV. Computed standard errors $\sigma(\beta, \beta^*, V_{\text{puddle}})$ for simulations carried out at different values of β , β^* ; and V_{puddle} .

β	β^*	$\sigma(\beta, \beta^*, 0)$	$\sigma(\beta, \beta^*, 0.005)$
4000	4000	Not converged	Not converged
4000	3000	Not converged	Not converged
4000	2000	Not converged	Not converged
4000	1000	2.2×10^{-16}	3.6×10^{-16}
4000	500	7.1×10^{-17}	1.0×10^{-16}
3000	3000	Not converged	Not converged
3000	2000	Not converged	Not converged
3000	1000	1.3×10^{-12}	1.2×10^{-12}
3000	500	6.4×10^{-13}	8.9×10^{-13}
2000	2000	8.9×10^{-8}	1.4×10^{-8}
2000	1500	4.1×10^{-8}	1.0×10^{-8}
2000	1000	3.1×10^{-8}	5.0×10^{-9}
2000	500	1.7×10^{-8}	4.5×10^{-9}
1500	1500	6.0×10^{-7}	9.1×10^{-7}
1500	1000	4.3×10^{-7}	4.1×10^{-7}
1500	700	4.3×10^{-7}	3.3×10^{-7}
1500	500	4.2×10^{-7}	2.9×10^{-7}
1000	1000	4.6×10^{-6}	5.3×10^{-6}
1000	700	3.5×10^{-6}	3.7×10^{-6}
1000	500	3.7×10^{-6}	3.5×10^{-6}

the puddle potential; for these, converged results were still obtainable without the temperature boost (i.e., without using $\beta^* < \beta$). From the results in Table IV, it is clearly seen that lowering β^* with or without using V_{puddle} has a substantial effect on the efficiency. For instance, the computational cost is reduced by a factor of 27 using $\beta^* = 500$ and $V_{\text{puddle}} = 0$, and by a factor of roughly 390 when employing $\beta^* = 500$ and $V_{\text{puddle}} = 0.005$. It is also noticed that the efficiency slightly worsen upon introducing the potential boost for the cases $(\beta, \beta^*) = (1500, 1500)$ and $(1000, 1000)$, perhaps due to an increase in weight fluctuations. As for the low temperature $(\beta = 4000, 3000)$ simulations, it was not possible to obtain converged results for $C(t)$ without resorting to very small β^* values in spite of the potential boost or a tenfold increase in the number of trajectories.

The possibility of extracting from the simulations information other than the process rates was also explored. For instance, the behavior of $dC(t)/dt$ as a function of t was studied to monitor the dynamics involved in the reactive process, and its evaluation was implemented within a finite difference approach. Results for the extended system, obtained at $\beta = 1000$ and 500, are shown in Fig. 5. In the $\beta = 500$ case, whereas the plateau for $160 < t < 200$ represents the value of the reaction rate at this temperature, the structure in the $40 < t < 160$ region indicates the presence of a substantial amount of recrossing during the reactive process. Similar features are strongly suppressed at $\beta = 1000$, therefore suggesting that the thermal excitations of the restrained particles could play an important role, either in dissipating the excess energy of a reactive trajectory at low temperature or in “knocking back” the diffuser along its path after crossing the transition state at higher temperature.

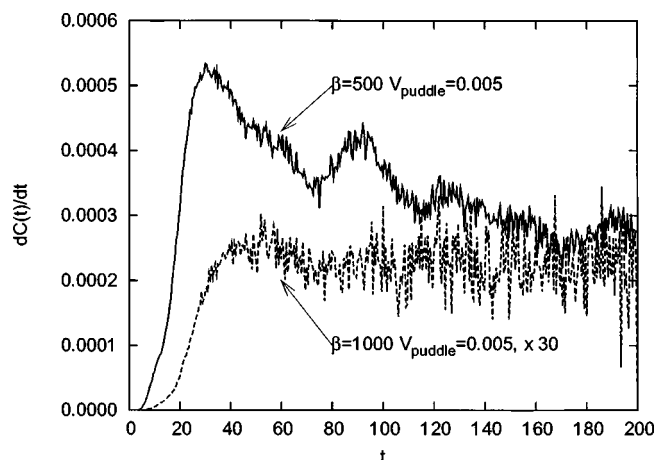


FIG. 5. Values of $dC(t)/dt$ as a function of t for two simulations of $V_{14}(\mathbf{R})$ at $\beta = 500$ and $\beta = 1000$. The oscillatory behavior at $\beta = 500$ is a result of a frequent recrossing of the diffusing particle, a feature that is largely suppressed at lower temperatures.

Other interesting observations concerning the mechanism of “jumping” across the barrier can be made accumulating the number of reactive trajectories $p(\alpha)$ as a function of the angle α formed by the direction of the initial diffuser momentum and the straight line joining the two minima. Figure 6 presents typical $p(\alpha)$ results for the 2D and the extended model potential; the latter were obtained at $\beta = 500$ with and without the use of V_{puddle} . The four distributions were normalized so that $\int p(\alpha) d\alpha = 1$.

By comparing the panels of Fig. 6, two striking differences are seen: first, the $p(\alpha)$ of the extended system presents a smoother behavior than in the 2D case, this feature suggesting a weaker dependency on the starting trajectory direction of the reactive events; second, the minimum at $\alpha = 0$ and the maximum around $\alpha = 55$ shown by the 2D model are replaced by a shallow maximum around $\alpha = 0$ in the case of the extended system. From the latter finding, one can infer that different mechanisms are operating in the two distinct cases: whereas the diffuser in the extended system shows to weakly prefer the most direct path (i.e., parallel to the x axis) to reach the other well, panel (a) indicates that it is quite unlikely for the 2D system to follow the same path and that the alternative route passing through the two TSs is preferred.¹⁹ In both cases a substantial probability of generating a reactive trajectory is also found for $\alpha = 180$, therefore indicating that the diffuser can be “backscattered” by the repulsive wall of the V_{2D} potential.

As a last observation, it is highlighted that for both model systems the use of the puddle potential reduces the $p(\alpha)$ dependency on α , producing more uniform reactive probabilities over the whole range. This is particularly evident for the 2D model, the latter featuring probability ratios up to six to seven when no puddle potential is used. In turn, the weaker dependency on α when using the TAPPD approach seems to justify the choice of an isotropic form for $\mu[\mathbf{P}(0)]$.

IV. DISCUSSION AND CONCLUSIONS

In this work, an improvement of the puddle potential method for computing the rate of activated reactions has

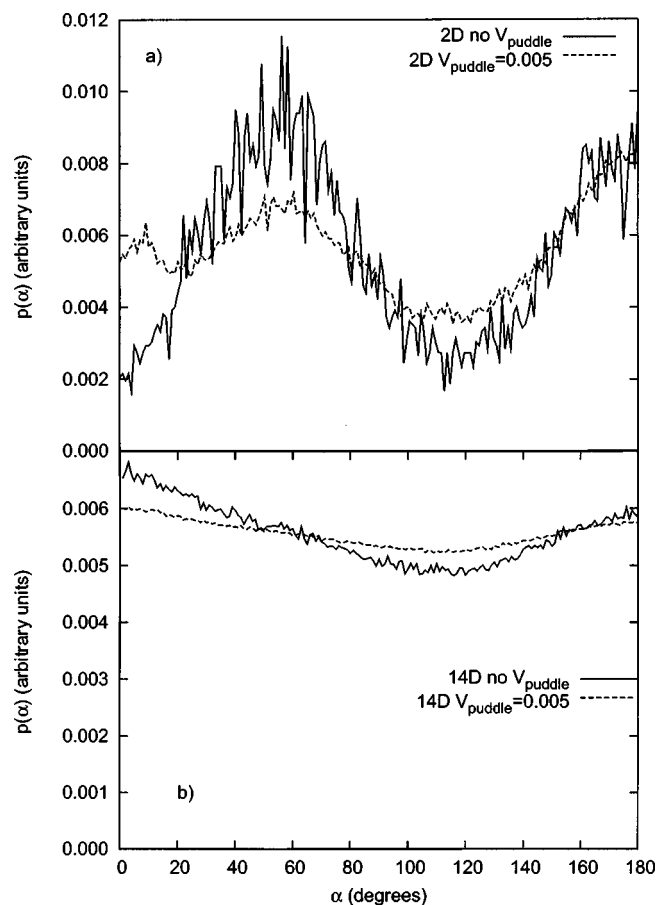


FIG. 6. Normalized probability density function $p(\alpha)$ for a trajectory of being reactive as a function of the departing angle (i.e., the direction of the speed vector with respect to the straight line connecting the two minima) with and without using V_{puddle} . (a) 2D model system, $\beta=500$. (b) Extended system, $\beta=500$.

been discussed and implemented. The proposed method is based on the use of a modified version of the importance sampling technique where not only the coordinate space is sampled with a less rough probability density, but also the probability density employed to sample the momentum space differs from the standard Boltzmann distribution. This allows more frequent sampling of higher kinetic energies.

Once corrected by using a reweighting procedure, this choice for the sampled distributions provides an efficient way to compute classical canonical reaction rates for the two model systems used as examples. Indeed, the improvement in efficiency with respect to the standard puddle potential in the two test cases spans several order of magnitude, depending largely on the ratio between original temperature of the system β and the simulation parameter β^* . Indeed, the increased efficiency of the proposed method makes it an ideal tool to study activated processes in condensed phase as well as in clusters.

Another interesting feature of the new method is represented by the fact that the trajectories needed for $C(t)$ are obtained by solving the unperturbed Hamilton equations for the system under study, and so they represent a correct description of the short time dynamics of the system itself. This fact can be exploited to gain some insight in the mechanism

of the transformation under study. As an example, a change of the diffusion mechanism in the two model systems on changing the number of interacting particles was discussed. It should also be possible to collect the initial condition of *all* the reactive trajectories to compute relevant statistical quantities or to directly monitor the system dynamics by visualizing the atomic and molecular motion in real time. This trick can be useful in discovering the presence of intermediate states that can act as attractors for the dynamics themselves or to select a better definition for the P and R regions of space. Indeed, improving the definition of P and R could play an important role in defining the overall efficiency of the method, especially if P and R may be defined in such a way that they share a common domain boundary. If this were possible, the puddle potential approach would be expected to show the best efficiency due to the fact that trajectories can be started very close to the separatrix of the two domains, making them more likely to fall in the P region and improve the statistics.

ACKNOWLEDGMENTS

The author acknowledges Steven Corcelli for making available the original test code for the 2D model, David J. Willock for carefully reading and commenting on the manuscript, and the EPSRC for an Advanced Research Fellowship (Grant No. GR/R77803/01).

- ¹E. P. Wigner, *J. Chem. Phys.* **5**, 720 (1937).
- ²D. G. Truhlar, B. C. Garret, and S. J. Klippenstein, *J. Phys. Chem.* **100**, 12771 (1996).
- ³J. B. Anderson, *J. Chem. Phys.* **58**, 4684 (1973).
- ⁴T. Yamamoto, *J. Chem. Phys.* **33**, 281 (1960).
- ⁵A. F. Voter, *Phys. Rev. Lett.* **78**, 3908 (1997).
- ⁶M. M. Steiner, P.-A. Genilloud, and J. W. Wilkins, *Phys. Rev. B* **57**, 10236 (1998).
- ⁷H. Grubmüller, *Phys. Rev. E* **52**, 2893 (1995).
- ⁸A. Laio and M. Parinello, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12562 (2002).
- ⁹M. R. Sorensen and A. F. Voter, *J. Chem. Phys.* **112**, 9599 (2000).
- ¹⁰J. VandeVondele and U. Rothlisberger, *J. Phys. Chem. B* **106**, 203 (2002).
- ¹¹C. Dellago, P. G. Bolhuis, and P. L. Geissler, *Adv. Chem. Phys.* **123**, 1 (2002).
- ¹²S. A. Corcelli, J. A. Rahman, and J. C. Tully, *J. Chem. Phys.* **118**, 1085 (2003).
- ¹³J. A. Rahman and J. C. Tully, *J. Chem. Phys.* **116**, 8750 (2002).
- ¹⁴D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, Oxford, 1988).
- ¹⁵S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Phys. Lett. A* **195**, 216 (1987).
- ¹⁶M. H. Kalos and P. Withlock, *Monte Carlo Methods, Basics* Vol. 1 (Wiley-Interscience, New York, 1986).
- ¹⁷C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *J. Chem. Phys.* **108**, 1964 (1998).
- ¹⁸Using this procedure, one might find that the estimate of the efficiency improvement is dependent on the specific value of t chosen. Indeed, as clearly seen in Fig. 2, the $\sigma(\beta, \beta^*)$ values increase with t . However, the dependency on β^* is weak (data not shown), so that we believe it is sound to use a value of t in the middle of the linear regime to estimate the simulation relative efficiency.
- ¹⁹The reason for the two different mechanisms must be sought in the relative importance of the maximum of the potential (free energy) located in (0,0) with respect to the TSs for both systems. Whereas for the 2D model the difference in height between the TSs and the maximum is quite large (roughly 100% of the TS barrier), in the extended model this is just 25% due to the presence of two repulsive particles close to the TSs location. This reduced energy difference tends to equalize the probability of passing over the maximum and through the TSs.