PROOF COPY 302439JCP

High-order averaging schemes with error bounds for thermodynamical properties calculations by molecular dynamics simulations

Eric Cancès[a)]
CERMICS, Ecole Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2, France
and MICMAC, INRIA Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay Cedex, France

Francois Castella[b)]
IRMAR, Université de Rennes 1, IPSO, INRIA Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France

Philippe Chartier[c)] and Erwan Faou[d)]
IPSO, INRIA Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France

Claude Le Bris[e)]
CERMICS, Ecole Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2, France
and MICMAC, INRIA Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay Cedex, France

Frédéric Legoll[f)]
CERMICS, Ecole Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2, France and MICMAC, INRIA Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay Cedex, France and EDF R&D, Analyse et Modèles Numériques, 1, avenue du Général de Gaulle, 92140 Clamart, France

Gabriel Turinici[g)]
CERMICS, Ecole Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2, France and MICMAC, INRIA Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay Cedex, France

(Received 14 July 2003; accepted 27 July 2004)

We introduce high-order formulas for the computation of statistical averages based on the long-time simulation of molecular dynamics trajectories. In some cases, this allows us to significantly improve the convergence rate of time averages toward ensemble averages. We provide some numerical examples that show the efficiency of our scheme. When trajectories are approximated using symplectic integration schemes (such as velocity Verlet), we give some error bounds that allow one to fix the parameters of the computation in order to reach a given desired accuracy in the most efficient manner. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1794611]

I. INTRODUCTION

The properties of a given physical system at thermodynamical equilibrium (radial distribution functions, free energies, transport coefficients, etc.) can be computed as averages of some observables over the phase space of a representative microscopic system.[1,2]

In most applications of interest, this microscopic system is composed of a high number of particles (typically more than 100 000 atoms for biological systems), so that the dimension of the phase space, which is six times the number of particles, is also very large. This makes the computation of averages a challenging issue.

Two families of methods are commonly used in order to sample the phase space according to the convenient probability density (which depends on the thermodynamic ensemble in which the calculation is performed: NVE,NVT,NPT, etc.): Monte Carlo methods[1,2] on the one hand, and molecular dynamics (MD) on the other hand. In the latter case, the time evolution of the microscopic system is simulated (possibly coupled with a bath); under the ergodicity assumption, the trajectory of the system samples the phase space and the time average of the observable over the trajectory converges toward the ensemble average when time goes to infinity.[3]

When time averages are estimated by MD calculations, a numerical scheme is needed to compute the dynamics of the system. The numerical trajectory covers a finite but large range of time and in addition may deviate significantly from the exact trajectory for large times. It is, however, of common belief and indeed commonly observed that, in many situations, results obtained by time averages are satisfactory.

In this article, we introduce and analyze a new method for calculating time averages, which can be used either to speed up the convergence at a fixed given accuracy, or to achieve a better accuracy at a given computational time.

The motivation for the introduction of such a method is the observation, complemented by a rigorous numerical

[a)]Electronic mail: cances@cermics.enpc.fr
[b)]Electronic mail: francois.castella@univ-rennes1.fr
[c)]Electronic mail: chartier@irisa.fr
[d)]Electronic mail: efaou@irisa.fr
[e)]Electronic mail: lebris@cermics.enpc.fr
[f)]Electronic mail: legoll@cermics.enpc.fr
[g)]Electronic mail: Gabriel.Turinici@inria.fr

 © 2004 American Institute of Physics

PROOF COPY 302439JCP

PROOF COPY 302439JCP

analysis on simple cases, that the error between the time average and the ensemble average often decreases as $1/T$ for long times when a standard discretization of $(1/T)\int_0^T A(t)dt$ is employed (here and in the following, $A(t)$ denotes the value of the observable under study at time t). On the contrary, if the time average is evaluated by the formula we introduce, namely $\int_0^T A(t)f_{k,T}(t)dt$, where $f_{k,T}(t)$ is some *filtering function* (to be made precise in the following, see formulas (2.10) and (2.11)), then the convergence rate can be in general improved to $1/T^k$, for arbitrary large k. This rate can again be rigorously established in simple cases, and the precise estimation of the error (as a function of the time integration step, the trajectory length, the order of the numerical scheme, and other parameters of the method) can be used in order to efficiently calibrate all these parameters.

Our method is presented in Sec. II. Under some strong hypotheses collected in the Appendix (in particular, we assume that the Hamiltonian system is integrable), and if the numerical scheme used to compute the trajectory is symplectic, we are able to rigorously establish the above-mentioned error estimate. The rather intricate and mathematically demanding proof of this estimate can be read in a companion article.[5]

We next turn in Sec. III to applications to examples of physical interest. Test cases are performed in order to illustrate the method and show the sharpness of the error bounds obtained analytically. We provide in Secs. III A and III B numerical examples on integrable systems that show that our estimate is indeed sharp. Of course, very few of the systems used in physics or chemistry are integrable. Fortunately, even if we are not able to prove it rigorously, our estimate seems to also hold true in more realistic and complex situations, as it is shown in Sec. III C, where convincing results on a system of particles interacting through a Lennard-Jones potential are provided, and in Sec. III D, where alkane chains are studied. In Sec. III E, we simulate a two-dimensional (2D) particle in a double well potential, in the regimes of low or high energy. This toy model mimicks larger systems whose potential energy surface presents several basins corresponding to metastable states.

For all the examples that we study in Sec. III, time averages computed by MD converge to ensemble averages at the rate $1/T$. In this case, we show that our method allows one to speed up this convergence rate to $1/T^k$. In Sec. IV, we continue the study of systems that present several metastable states, and we focus on the regime when transitions between these states are possible but they are rare events in comparison with the faster characteristic time scale of the system. This energy regime is outside the scope of our theoretical analysis, and we observe that the rate of convergence of time averages calculated by MD is slowed down to $1/\sqrt{T}$.

We wish to emphasize that we only deal here with the *NVE* thermodynamical ensemble. The equations of motion that correspond to this ensemble are the Newton equations (governing the "physical" dynamics of the system), which can be recast as a Hamiltonian dynamical system. Although we have no definite conclusions to date, similar results are likely to hold for other statistical ensembles, provided the underlying dynamical system is Hamiltonian; that is the case when Nosé-Poincaré thermostats[4] are used to perform computations in the *NVT* ensemble.

II. MAIN SETTING AND RESULT

Let us consider M particles in three dimensions (3D). Each particle i is described by its position $\mathbf{q}_i \in \mathbb{R}^3$, its momentum $\mathbf{p}_i \in \mathbb{R}^3$, and its mass m_i. Let $H(\mathbf{q},\mathbf{p})$ be the Hamiltonian of the system, defined on $\mathbb{R}^{3M}\times\mathbb{R}^{3M}$ by

$$H(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{M} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q}_1,\ldots,\mathbf{q}_M),$$

where V is the interaction potential, and $\mathbf{q}$ and $\mathbf{p}$ are notations for $(\mathbf{q}_1,\ldots,\mathbf{q}_M)$ and for $(\mathbf{p}_1,\ldots,\mathbf{p}_M)$, respectively. The Hamiltonian dynamical system associated to H is given by

$$\begin{aligned} \frac{d\mathbf{q}}{dt} &= \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}(t),\mathbf{p}(t)), \\ \frac{d\mathbf{p}}{dt} &= -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{q}(t),\mathbf{p}(t)). \end{aligned} \tag{2.1}$$

We suppose that the dynamical system (2.1) is integrable (see the Appendix) and that the numerical scheme we use to integrate it is symplectic (recall that, for instance, the velocity Verlet scheme is symplectic). Since the dynamical system is integrable, it has $3M$ invariant functions $I_j(\mathbf{q},\mathbf{p})$, $j=1,\ldots,3M$, which means that $I_j(\mathbf{q}(t),\mathbf{p}(t))$, where $(\mathbf{q}(t),\mathbf{p}(t))$ is solution of Eq. (2.1), is constant with respect to time. Of course, $H(\mathbf{q},\mathbf{p})$ is one of these invariants, since the energy is preserved by the dynamics. We denote by

$$S(\mathbf{q},\mathbf{p}) = \{(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{3M}\times\mathbb{R}^{3M}$$

such that

$$\forall j \in [1,3M], I_j(\mathbf{x},\mathbf{y}) = I_j(\mathbf{q},\mathbf{p})\} \tag{2.2}$$

the level set of the invariant functions $\{I_j\}_{1\leq j\leq 3M}$ containing the phase space point $(\mathbf{q},\mathbf{p})$.

Let $(\mathbf{q}_0,\mathbf{p}_0)$ be an initial condition, and let $A(\mathbf{q},\mathbf{p})$ be an observable whose average is well-defined in the *NVE* ensemble. As we suppose that the dynamical system has $3M$ invariants, the actual trajectory starting from the initial point $(\mathbf{q}_0,\mathbf{p}_0)$ remains for all times on the level set $S(\mathbf{q}_0,\mathbf{p}_0)$. Thus, the ensemble average of the observable A reads

$$\langle A\rangle = \frac{\int_{S(\mathbf{q}_0,\mathbf{p}_0)} A(\mathbf{q},\mathbf{p})d\mu(\mathbf{q},\mathbf{p})}{\int_{S(\mathbf{q}_0,\mathbf{p}_0)} d\mu(\mathbf{q},\mathbf{p})}, \tag{2.3}$$

where $d\mu(\mathbf{q},\mathbf{p})$ is the invariant measure[5] on $S(\mathbf{q}_0,\mathbf{p}_0)$.

Under some hypotheses collected in the Appendix, the so-called ergodic theorem[5,6] holds:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T A(\mathbf{q}(t),\mathbf{p}(t))dt = \langle A\rangle, \tag{2.4}$$

where $(\mathbf{q}(t),\mathbf{p}(t))$ is the trajectory of the system (2.1) starting from the initial point $(\mathbf{q}_0,\mathbf{p}_0)$. As an approximation to Eq. (2.3), it is a standard approach to use the discrete sum

$$\langle A\rangle_{\text{num}}^{\text{Rie}}(\delta t,T) = \frac{1}{\mathcal{N}}\sum_{j=0}^{\mathcal{N}-1} A(\mathbf{q}_j,\mathbf{p}_j), \tag{2.5}$$

PROOF COPY 302439JCP

PROOF COPY 302439JCP

where $(\mathbf{q}_j,\mathbf{p}_j)_{j=0}^{\mathcal{N}-1}$ is the numerical trajectory given by an integration scheme with time step $\delta t=T/\mathcal{N}$ applied on Eq. (2.1). However, as the convergence in Eq. (2.4) occurs[6] only at the rate $1/T$, we cannot expect a faster convergence rate of Eq. (2.5) toward $\langle A\rangle$.

In order to improve this asymptotic convergence rate, we suggest replacing the uniformly weighted time average $(1/T)\int_0^T A(\mathbf{q}(t),\mathbf{p}(t))dt$ by

$$\int_0^T A(\mathbf{q}(t),\mathbf{p}(t))f_T(t)dt, \tag{2.6}$$

where f_T acts as a filtering function of the signal $t\mapsto A(\mathbf{q}(t),\mathbf{p}(t))$. We present two different types of filtering functions, that both improve the convergence rate. The first one allows for a better understanding of our method, whereas the second one (see formulas (2.10) and (2.11)) is easier to implement. Denoting by $B_1(t_1)=(2/T)\int_{t_1}^{t_1+T/2}A(\mathbf{q}(t),\mathbf{p}(t))dt$ the time average in the interval $[t_1,t_1+T/2]$ (with $t_1\in[0,T/2]$), one can consider the average

$$\begin{aligned}\langle A\rangle^{(2*)}(T)&=\frac{2}{T}\int_0^{T/2}B_1(t_1)dt_1\\&=\left(\frac{2}{T}\right)^2\int_{t_1=0}^{T/2}\int_{t_2=0}^{T/2}A(\mathbf{q}(t_1+t_2),\mathbf{p}(t_1+t_2))dt_1\,dt_2.\end{aligned} \tag{2.7}$$

Let us define the function $\alpha_{[0,\mathcal{T}]}$ by

$$\alpha_{[0,\mathcal{T}]}:t\in\mathbb{R}\mapsto 1/\mathcal{T}\quad\text{if}\quad t\in[0,\mathcal{T}],\quad 0\ \text{otherwise}.$$

Then $\langle A\rangle^{(2*)}(T)$ can be written in the form (2.6) with $f_T\equiv f^*_{2,T}$, where $f^*_{2,T}$ is the convolution of $\alpha_{[0,T/2]}$ with itself. So $\langle A\rangle^{(2*)}(T)$ can be understood as a filtered average and also as a mean value, over initial conditions chosen along a MD trajectory (here, the trajectory $(\mathbf{q}(t_1),\mathbf{p}(t_1))$, $t_1\in[0,T/2]$), of standard MD averages (see the right-hand side of Eq. (2.7)). More generally, denoting by $f^*_{k,T}$ the kth convolution of the function $\alpha_{[0,T/k]}$ with itself, one can define the time average by $\langle A\rangle^{(k*)}(T)=\int_0^T A(\mathbf{q}(t),\mathbf{p}(t))f^*_{k,T}(t)dt$. We can prove,[5] under the assumptions collected in the Appendix, that

$$\langle A\rangle^{(k*)}(T)=\langle A\rangle+O\left(\frac{1}{T^k}\right). \tag{2.8}$$

Thus, for any *arbitrary* positive integer k, it is possible to design an averaging scheme which converges to the ensemble average at the rate $1/T^k$.

The discrete version of $\langle A\rangle^{(k*)}(T)$ is

$$\langle A\rangle^{(k*)}_{\text{num}}(\delta t,T)=\frac{1}{N^k}\sum_{j=0}^{k(N-1)}C(k,N-1,j)A(\mathbf{q}_j,\mathbf{p}_j), \tag{2.9}$$

where $\mathcal{N}=kN$ is the number of time steps that have been computed, $(\mathbf{q}_j,\mathbf{p}_j)_{j=0}^{kN}$ is the numerical trajectory given by an integration scheme with time step $\delta t=T/(kN)$, and where $C(k,N-1,j)$ is the number of k-tuples of positive integers $j_1,\ldots,j_k$ whose sum is fixed to be j and which all belong to $[0,N-1]$. These coefficients can be computed by recursion, in such a way that their computation is almost for free in comparison with the computation of the trajectory $(\mathbf{q}_j,\mathbf{p}_j)_{j=0}^{kN}$.

However, formula (2.9) is not incremental, due to the presence of the coefficients $C(k,N-1,j)$ in the summation. In other words, the average after $k(N+1)$ steps is not simply deduced from the knowledge of that after kN steps. This difficulty leads us to use other filtering functions and to define the time average by

$$\langle A\rangle^{(k)}(T)=\int_0^T A(\mathbf{q}(t),\mathbf{p}(t))f_k\left(\frac{t}{T}\right)dt, \tag{2.10}$$

with

$$f_k(t)=\alpha_k t^{k-1}(1-t)^{k-1}, \tag{2.11}$$

where α_k is a normalizing constant ($k\geqslant 1$). Again, $\langle A\rangle^{(k)}(T)$ satisfies an estimate similar to Eq. (2.8). The discrete version of $\langle A\rangle^{(k)}(T)$ is

$$\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)=\frac{\sum_{j=0}^{\mathcal{N}-1}A(\mathbf{q}_j,\mathbf{p}_j)f_k\left(\frac{j}{\mathcal{N}}\right)}{\sum_{j=0}^{\mathcal{N}-1}f_k\left(\frac{j}{\mathcal{N}}\right)}, \tag{2.12}$$

where $T=\mathcal{N}\delta t$. Decomposing f_k as a sum of monomial functions, one can see that the averages after $\mathcal{N}+1$ steps can be computed from the averages after $\mathcal{N}$ steps and the observable values at time step $\mathcal{N}+1$. So, just like the standard formula (2.5), formula (2.12) allows one to compute time averages "on the fly." In the test cases to follow, we will make use of Eq. (2.12) with several values of k to compute an approximation of the ensemble average.

Formulas (2.9) and (2.12) are generalizations of (2.5) since $\langle A\rangle^{(k*=1)}_{\text{num}}(\delta t,T)=\langle A\rangle^{(k=1)}_{\text{num}}(\delta t,T)=\langle A\rangle^{\text{Rie}}_{\text{num}}(\delta t,T)$. To the best of our knowledge, formulas (2.9) or (2.12) do not seem to be used in molecular dynamics simulations, or at least are not reported on in the literature.

In addition, if a symplectic scheme of order r_0 is used to compute the trajectory ($r_0=2$ for the velocity Verlet scheme), we can prove[5] the following error estimate:

$$|\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)-\langle A\rangle|\leqslant C\left(\frac{1}{T^k}+(\delta t)^{r_0}\right) \tag{2.13}$$

for some constant C which does not depend on T or δt (a similar estimate holds for $\langle A\rangle^{(k*)}_{\text{num}}(\delta t,T)$).

We emphasize the fact that the polynomial decay of the error with respect to the total time of simulation T is made possible by the fact that the values $A(\mathbf{q}_j,\mathbf{p}_j)$ are correlated in time, in the sense that they are obtained by the simulation of one single trajectory. When one uses a Monte Carlo method, the values $A(\mathbf{q}_j,\mathbf{p}_j)$ are not correlated any longer, and the error between the computed empirical average and the ensemble average $\langle A\rangle$ only decays at the universal rate $1/\sqrt{T}$.

The proof of Eq. (2.13), together with details on the filter function used in formula (2.12), may be read in a companion article.[5]

PROOF COPY 302439JCP

PROOF COPY 302439JCP

III. NUMERICAL EXAMPLES

We now turn to numerical examples showing that:

(1) estimate (2.13) is sharp for integrable Hamiltonian systems satisfying (at least some of) the assumptions under which the estimate has been rigorously proved.
(2) in some particular conditions, estimate (2.13) seems to also hold for some nonintegrable systems of practical interest, although our proof does not carry through such nonintegrable systems.

We first present results in the case of a collection of independent harmonic oscillators (see Sec. III A), then on the Kepler problem (see Sec. III B). Both systems are integrable, and fall within the hypotheses of our theoretical result. In Sec. III C, we study a system of particles subjected to the Lennard-Jones potential, and in Sec. III D, we study alkane chains described by the "united atom" (UA) model.[7] Finally, in Sec. III E, we study a particle submitted to a double well potential in 2D, in the case when the particle energy is lower than or much larger than the saddle point energy (the study of the other cases is postponed to Sec. IV). In all these cases, time averages computed by Eq. (2.5) converge to ensemble averages at rate $1/T$, and we show the efficiency of the method (2.12) that we propose.

With a view to showing that our approach is insensitive to the numerical integration scheme, provided that it is symplectic, we have used, depending on the case at hand, three different algorithms,[8] of different orders, to integrate the dynamical system (2.1).

(1) The standard velocity Verlet algorithm, which is, as recalled earlier, symplectic and of order 2; we denote by $\phi_{\delta t}^{VV}$ a step of the algorithm of time step δt: $(\mathbf{q}_{n+1},\mathbf{p}_{n+1}) = \phi_{\delta t}^{VV}(\mathbf{q}_n,\mathbf{p}_n)$.

(2) A composition of the velocity Verlet algorithm with itself so that the order of the algorithm increases to 4; namely,

$$(\mathbf{q}_{n+1},\mathbf{p}_{n+1}) = \phi_{\delta t}^{(4)}(\mathbf{q}_n,\mathbf{p}_n) = \phi_{b_1\delta t}^{VV} \circ \phi_{b_0\delta t}^{VV} \circ \phi_{b_1\delta t}^{VV}(\mathbf{q}_n,\mathbf{p}_n)$$

with $b_1 = 1/(2-2^{1/3})$ and $b_0 = 1-2b_1$.

(3) A composition of the velocity Verlet algorithm with itself so that the order of the algorithm increases to 6; namely,

$$(\mathbf{q}_{n+1},\mathbf{p}_{n+1}) = \phi_{\delta t}^{(6)}(\mathbf{q}_n,\mathbf{p}_n) = \phi_{c_3\delta t}^{VV} \circ \phi_{c_2\delta t}^{VV} \circ \phi_{c_1\delta t}^{VV} \circ \phi_{c_0\delta t}^{VV} \circ \phi_{c_1\delta t}^{VV} \times \circ \phi_{c_2\delta t}^{VV} \circ \phi_{c_3\delta t}^{VV}(\mathbf{q}_n,\mathbf{p}_n)$$

with

$$c_1 = -1.177\,679\,984\,178\,87,$$

$$c_2 = 0.235\,573\,213\,359\,357, \quad c_3 = 0.784\,513\,610\,477\,560,$$

and $c_0 = 1-2(c_1+c_2+c_3)$.

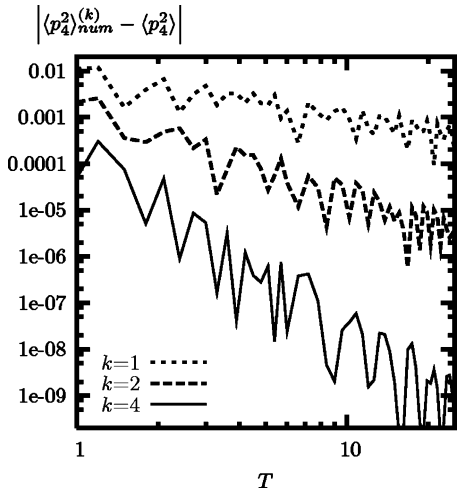

FIG. 1. Convergence of $\langle p_4^2\rangle_{\text{num}}^{(k)}(\delta t,T)$ to $\langle p_4^2\rangle$ for different values of k, in the case of a collection of five harmonic oscillators. Initial conditions are $q_1=4$, $p_1=-1.2$, $q_i=0$, and $p_i=1$ for $i=2,\ldots,5$.

A. Collection of harmonic oscillators

In this section, we consider the dynamics of a system of five particles in one dimension (1D) described by the Hamiltonian

$$H = \sum_{i=1}^{5} \frac{p_i^2}{2} + \omega_i^2 \frac{q_i^2}{2},$$

with the following frequencies:

$$\omega_1 = 1.0, \quad \omega_2 = 8.012\,56, \quad \omega_3 = 12.252\,45, \quad \omega_4 = 17.234, \quad \omega_5 = 20.987\,65. \tag{3.1}$$

As an instance, we study the average of $A = p_4^2$ with the algorithm of order 6. The exact value $\langle A\rangle = 0.5$ of the ensemble average can be obtained by analytical calculations.

All but one hypothese needed to prove our estimates are satisfied: the integration algorithm is symplectic and the analyticity and integrability assumptions on the Hamiltonian function (see Sec. 1 of the Appendix) are satisfied. The frequencies (3.1) do not satisfy the diophantine assumption (see Sec. 2 of the Appendix) for they are rational numbers, but the level set (defined by Eq. (2.2)) of the invariant functions $I_j(\mathbf{q},\mathbf{p}) = p_j^2/2 + \omega_j^2(q_j^2/2)$ is well-enough sampled, in comparison to the precision needed for the computation of $\langle A\rangle$.

We use formula (2.12) with different values of k ($k=1$ to 4), for a given δt ($\delta t = 0.025$). The numerical error as a function of the length of the MD trajectory is displayed in Fig. 1. A least-squares fit provides the following result: the error $\left|\langle A\rangle_{\text{num}}^{(k)}(\delta t,T) - \langle A\rangle\right|$ decreases with rate $1/T^{\tilde{k}}$, with

$$\tilde{k}(k=1) = 0.978, \quad \tilde{k}(k=2) = 1.91,$$

$$\tilde{k}(k=3) = 3.05, \quad \tilde{k}(k=4) = 3.87.$$

So, the time average defined by Eq. (2.12) converges to the ensemble average with a rate very close to $1/T^k$.

PROOF COPY 302439JCP

PROOF COPY 302439JCP

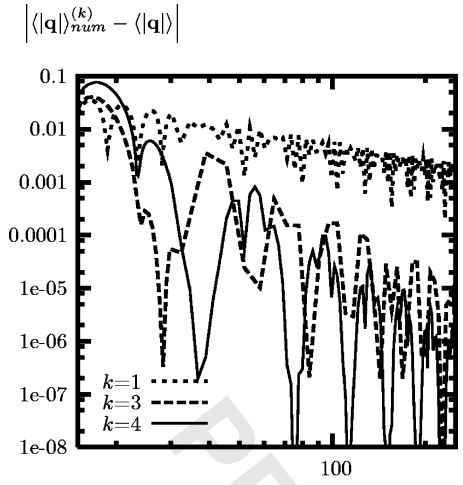

FIG. 2. Convergence of $\langle|\mathbf{q}|\rangle^{(k)}_{\text{num}}(\delta t,T)$ to $\langle|\mathbf{q}|\rangle$ for different values of k, in the case of the Kepler problem. Initial conditions are $\mathbf{q}=(0.9,0,0)$ and $\mathbf{p}=(0,1.1,0.5)$.

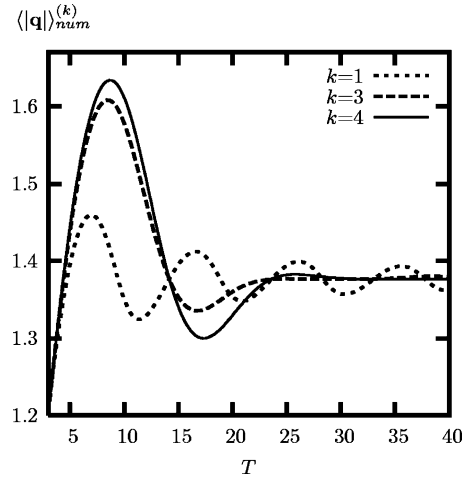

FIG. 3. Oscillations of $\langle|\mathbf{q}|\rangle^{(k)}_{\text{num}}(\delta t,T)$ at the beginning of the simulation (Kepler problem), for different values of k. Initial conditions are $\mathbf{q}=(0.9,0,0)$ and $\mathbf{p}=(0,1.1,0.5)$.

B. The Kepler problem

We next turn to the simulation of the Kepler problem. We consider the Hamiltonian

$$H=\frac{\mathbf{p}^2}{2}-\frac{1}{|\mathbf{q}|},$$

describing a 3D particle in an attractive Coulomb potential.

Let us for instance study the average of $A=|\mathbf{q}|$. As in the previous example, the ensemble average can be computed analytically: $\langle|\mathbf{q}|\rangle=(3+2E_0L_0^2)/(4|E_0|)$, where E_0 is the initial (negative) energy and L_0 is the norm of the initial kinetic momentum. The simulation time step is $\delta t=0.01$, and we use the algorithm of order 6. As for the harmonic oscillators, the Hamiltonian function is analytic and integrable and the integration scheme is symplectic.

We compute the averages (2.12) with k ranging from 1 to 4. We see that, for $T\geqslant 40$ (approximately beyond 4 revolution periods), the larger k, the larger the convergence rate (see Fig. 2; the convergence rate is quantitatively studied in the following). Remark that, as k increases, the amplitude of the oscillations at the beginning of the simulation (which is not the focus here) increases (see Fig. 3). This fact seems to be generic, and not to be restricted to the Kepler problem.

Let us now check that the estimate (2.13) is sharp. For this purpose, we study the dependence of the error $e^{(k)}(\delta t,T)$ with respect to T and δt in order to check whether there exist constants $C_1(k)$ and $C_2(r_0)$ such that

$$e^{(k)}(\delta t,T)=\left|\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)-\langle A\rangle\right|\approx\frac{C_1(k)}{T^k}+C_2(r_0)(\delta t)^{r_0}.$$

The initial conditions are such that $\langle A\rangle=1.37$. We first determine $C_2(r_0)$. To do this, we choose several values of the time step δt in the interval $[0.025;0.2]$, and for each of them, we compute a long-time trajectory. We estimate that the trajectory is long enough when the oscillations of the function $T\mapsto\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)$ are both smaller than 10^{-8} and more than 100 times as small as the difference between $\langle A\rangle$ and $\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)$. So the leading term of the error is $C_2(r_0)\times(\delta t)^{r_0}$. We observe indeed that $\ln e^{(k)}(\delta t,T)$ is a linear function of $\ln\delta t$ of slope $-r_0$, and find values for $C_2(r_0)$ which *do not depend on* k and are given by

$$C_2(r_0=6)=0.163,\quad C_2(r_0=4)=1.09,$$

$$C_2(r_0=2)=0.622.$$

In order to determine the values of $C_1(k)$, we work with very small values of δt (respectively, $\delta t=10^{-2}$, 5×10^{-3} and 5×10^{-4} when we use the algorithm of order 6, 4, and 2) so that the term $C_2(r_0)(\delta t)^{r_0}$ is lower than 10^{-7}. We also observe that the time average converges to the exact ensemble average up to an error of 10^{-7}. So, as long as the error is larger than 10^{-7}, its leading term is $C_1(k)/T^k$. A least-squares fit of $\ln e^{(k)}(\delta t,T)$ as a function of $\ln T$ shows that the error decreases with rate $1/T^{\tilde{k}}$, with

$$\tilde{k}(k=1)=1.02,\quad \tilde{k}(k=2)=2.01,$$

$$\tilde{k}(k=3)=2.82,\quad \tilde{k}(k=4)=3.85.$$

Likewise, we observe that the rescaled error $T^k\left|\langle A\rangle^{(k)}_{\text{num}}(\delta t,T)-\langle A\rangle\right|$ is bounded independently of T, and we also notice that the bound is independent of r_0. We obtain

$$C_1(k=1)\approx0.6,\quad C_1(k=2)\approx14.4,$$

$$C_1(k=3)\approx212,\quad C_1(k=4)\approx8270.$$

Not surprisingly, the values of $C_1(k)$ quickly increase with k (this is in agreement with the fact that oscillations at the beginning of the simulation are larger for larger values of k).

These numerical experiments therefore confirm the sharpness of the estimate (2.13).

C. Particles in a truncated Lennard-Jones potential

We study here a system of $M=288$ particles in 3D interacting through a truncated Lennard-Jones potential. The Hamiltonian of the system reads

PROOF COPY 302439JCP

PROOF COPY 302439JCP

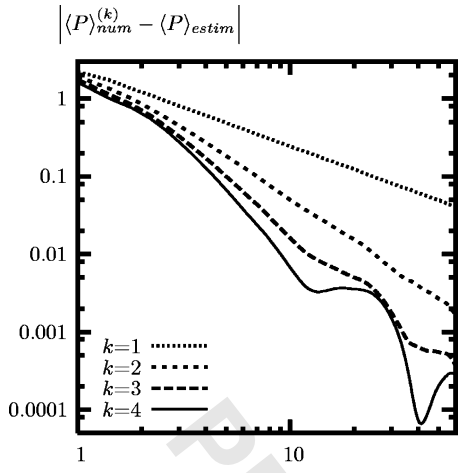

FIG. 4. Convergence of the average pressure $\langle P\rangle_{\text{num}}^{(k)}(\delta t,T)$ to $\langle P\rangle_{\text{estim}}$ for different values of k, in the case of the Lennard-Jones system.

$$H=\sum_{i=1}^{M}\frac{\mathbf{p}_i^2}{2}+\sum_{i=1}^{M}\sum_{j>i}V(|\mathbf{q}_j-\mathbf{q}_i|),$$

with $V_{\text{LJ}}(z)=4/z^{12}-4/z^6$ and $V(z)=V_{\text{LJ}}(z)-V_{\text{LJ}}(z_c)-V'_{LJ}(z_c)(z-z_c)$ if $z\leqslant z_c$, 0 otherwise. Thus, both the potential and the forces are continuous at the cutoff radius $z_c=3.08$. We apply periodic boundary conditions (the simulation box is $[-3.2;3.2]^3$) and use the minimum image convention.[1] The initial conditions correspond to a solid state and are such that the density of the system is 1.12 and the average kinetic temperature is 0.884. Such an initial condition ensures that the system remains in the solid phase. The hypotheses of the Appendix are *not satisfied* (in particular, there are not enough invariant functions). The observable under examination is the pressure:

$$A(\mathbf{q},\mathbf{p})=P=\frac{1}{3V}\sum_{i=1}^{M}\left(\mathbf{p}_i^2+\sum_{j>i}\mathbf{q}_{ij}\cdot\mathbf{F}_{ij}\right),$$

where $\mathbf{q}_{ij}=\mathbf{q}_i-\mathbf{q}_j$ and $\mathbf{F}_{ij}=-\partial H/\partial\mathbf{q}_{ij}$. We work with the velocity Verlet algorithm, using a time step of $\delta t=5\times10^{-3}$.

The results are displayed in Fig. 4. We have checked that the error $\left|\langle A\rangle_{\text{num}}^{(k)}(\delta t,T)-\langle A\rangle\right|$ decreases with the rate $1/T^{\tilde{k}}$, with

$$\tilde{k}(k=1)=0.996,\quad \tilde{k}(k=2)=1.85,$$

$$\tilde{k}(k=3)=2.67,\quad \tilde{k}(k=4)=3.76.$$

So, as in the previous examples, the rate of convergence is improved from $1/T$ to almost $1/T^k$, even if the system at hand is not integrable. Indeed, even if the system remains in the solid phase, the potential cannot be approximated by a harmonic potential: along the trajectory of the system, we have observed that some eigenvalues of the Hessian of the potential are negative, and that some other ones vary by 20%.

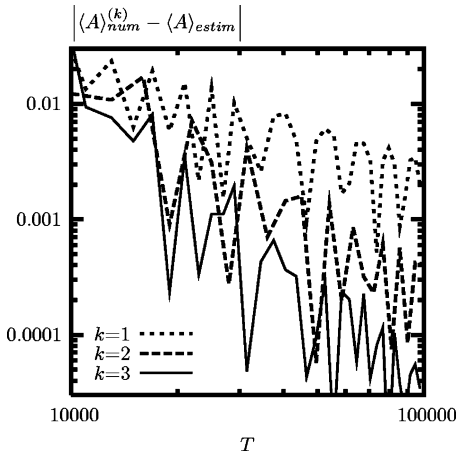

FIG. 5. Convergence of the end-to-end distance of an alkane chain ($M=40$ particles) for different values of k, in the low temperature case ($T=2.13$ K).

D. Alkane chains

We now study chains of particles interacting through the Ryckaert–Bellemans united atom (UA) model.[7] These chains represent alkane molecules. The observable under examination is the normalized end-to-end distance

$$A(\mathbf{q})=\frac{|\mathbf{q}_1-\mathbf{q}_M|^2}{Md_{C-C}^2},$$

where $d_{C-C}=1.53$ Å is the constant bond length. We work with the velocity Verlet algorithm. As in the previous example, the hypotheses of the Appendix are not satisfied.

The first example is a chain of $M=40$ particles. For an extremely low temperature ($T=2.13$ K), we obtain the convergence of the time averages (see Fig. 5, obtained with a time step of $\delta t=0.1$ fs) with the rate $1/T^{\tilde{k}}$, with the following values for $\tilde{k}$:

$$\tilde{k}(k=1)=0.98,\quad \tilde{k}(k=2)=1.70,$$

$$\tilde{k}(k=3)=3.02,\quad \tilde{k}(k=4)=4.21.$$

So we again observe convergence at a rate close to $1/T^k$. We have also studied a 100 particles chain: results on this larger system are similar to those for the 40 particles chain.

The case of higher temperatures will be studied in Sec. IV.

E. One particle in a double well potential

In this section, we deal with the Hamiltonian

$$H=\frac{p_x^2}{2}+\frac{p_y^2}{2}+V(q_x,q_y),$$

where

$$V(q_x,q_y)=(q_x^2-1)^2+(q_y+q_x^2-1)^2,$$

describing a 2D particle in a double well potential. The potential V goes to $+\infty$ at infinity and has three critical points:

PROOF COPY 302439JCP

PROOF COPY 302439JCP

two global minima located at $(\pm 1,0)$, at which $V(\pm 1,0)=0$, and one saddle point located at (0,1), at which $V(0,1)=1$.

The behavior of time averages strongly depends on the value of the particle energy E_0 with respect to the saddle point energy (here equal to 1). Three different regimes can be identified:

(1) For E_0 smaller than 1, the trajectory only explores a single basin. Time averages converge to a limit and using the filtering function f_k and formula (2.12) improves the convergence rate from $1/T$ to $1/T^k$ (for an arbitrary k) as in the previous examples. These results hold even if the energy is large enough (for instance, $E_0=0.8$) so that, on the region sampled by the particle, the potential cannot be approximated by a harmonic potential.
(2) For E_0 much larger than 1 (say $E_0 \geqslant 5$), the particle is so energetic that it does not really "feel" the barrier and the convergence of the time averages is similar to the one observed in the first case; once again, we have checked that we are far from the harmonic approximation.
(3) For E_0 a little larger than 1, the system really "feels" the presence of two distinct basins; we postpone the study of this case to Sec. IV.

Let us end this section by a discussion on the ergodicity of the 2D double well system under study. It is easy to check that the isoenergetic surface $H(q_x,q_y,p_x,p_y)=E_0$ can be parametrized by three angles:

$$\begin{aligned} q_x &= \pm\sqrt{1+\sqrt{E_0}\cos(\theta)}, \\ q_y &= \sqrt{E_0}(\sin(\theta)\cos(\phi)-\cos(\theta)), \\ p_x &= \sqrt{2E_0}\sin(\theta)\sin(\phi)\cos(\psi), \\ p_y &= \sqrt{2E_0}\sin(\theta)\sin(\phi)\sin(\psi), \end{aligned} \tag{3.2}$$

with $(\theta,\phi,\psi)\in[0,\theta_M]\times[0,\pi]\times[0,2\pi]$, where $\theta_M=\pi$ if $E_0\leqslant 1$ and $\theta_M=\arccos(-1/\sqrt{E_0})$ if $E_0>1$. It is therefore possible to compute numerically the *NVE* ensemble average of any (regular) observable with a high accuracy. It reads

$$\langle A\rangle(E_0)=\int_{[0,\theta_M]\times[0,\pi]\times[0,2\pi]} A(\theta,\phi,\psi) \times w(\theta,\phi)\,d\theta\,d\phi\,d\psi, \tag{3.3}$$

where

$$w(\theta,\phi)=C\frac{\sin(\phi)\sin(\theta)^2}{\sqrt{1+\sqrt{E_0}\cos(\theta)}}$$

is the relevant invariant probability density of the dynamical system on (θ,ϕ,ψ) (C is a normalizing constant). In Table I, we compare this ensemble average with several time averages computed from trajectories with various initial conditions that all correspond to the energy $E_0=0.5$. In general, these time averages are different from the ensemble average. This difference can be explained either by the fact that the isoenergetic surface has invariant subsets (this may be due to the existence of a second invariant that we did not manage to identify), or by the fact that the system is ergodic but on very long time scales, much longer than the simulation times we can afford.

TABLE I. Computed values for the ensemble average and the time average of different observables (double well potential). Ensemble averages are computed from Eq. (3.3). For the time averages, computed from Eq. (2.12), initial conditions are $(\mathbf{q}_1,\mathbf{p}_1)=(-1,0.5;0.5,-0.5)$, $(\mathbf{q}_2,\mathbf{p}_2)=(-0.9,0.772\ldots;0,-0.5)$ and $(\mathbf{q}_3,\mathbf{p}_3)=(-1.05,0.549;0.3606\ldots,0)$, corresponding to the energy 0.5, and the time step is $\delta t=0.005$. We checked that a simulation of length $T=5000$ was long enough and that time average values do not depend on k.

Observable A	$\langle A\rangle$ $\times(E_0=0.5)$	$\langle A\rangle_{\text{num}}^{(k=5)}$ $(\mathbf{q}_1,\mathbf{p}_1)$	$\langle A\rangle_{\text{num}}^{(k=5)}$ $(\mathbf{q}_2,\mathbf{p}_2)$	$\langle A\rangle_{\text{num}}^{(k=5)}$ $(\mathbf{q}_3,\mathbf{p}_3)$
q_x	−0.944 59	−0.931 18	−0.955 85	−0.923 86
q_x^2	0.928 43	0.900 77	0.951 70	0.885 65
q_x^4	0.989 64	0.928 55	1.040 74	0.895 65
q_y	0.071 562	0.099 221	0.048 292	0.114 346
q_y^2	0.255 17	0.142 98	0.347 25	0.085 192
p_x^2	0.244 82	0.339 49	0.167 94	0.387 17
p_y^2	0.244 82	0.146 15	0.325 74	0.095 278
$V(q_x,q_y)$	0.255 17	0.257 17	0.253 15	0.258 78

From a practical point of view, one way to solve this issue is to compute the mean value of the time averages over several different initial conditions. First, as in a Monte Carlo method, we choose many initial conditions $(\mathbf{q}_i,\mathbf{p}_i)_i$ on the surface of constant energy E_0, according to the probability density $w(\theta,\phi)$ (to do so, one can use the Acceptance-Rejection method[9]). For all these initial conditions, we compute a MD trajectory and a time average, and then the mean value $\mu(E_0)$ and the standard deviation $\sigma_{\text{MC+MD}}(E_0)$ of these data. Results are reported in Fig. 6: this method yields a good approximation of the ensemble average.

Thus, to compute an ensemble average, we can suggest to choose several different initial conditions on the constant energy surface, and to use each of them to compute a trajectory and a time average. A good approximation of the ensemble average is the mean value of these time averages.

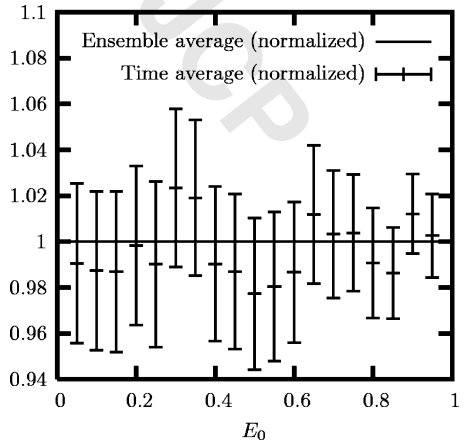

FIG. 6. Comparison between the ensemble average of $A=q_y^2$, computed from Eq. (3.3), and the expected time average (error bars correspond to a 95% confidence interval; 1000 initial conditions have been used), on the double well potential, for different energy levels E_0. All quantities have been normalized by the ensemble average $\langle q_y^2\rangle(E_0)$.

PROOF COPY 302439JCP

PROOF COPY 302439JCP

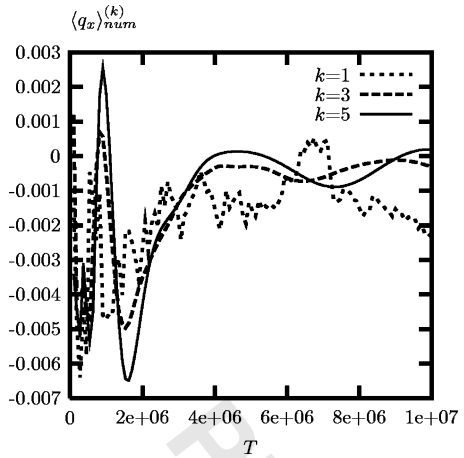

FIG. 7. Evolution of $\langle q_x \rangle_{\text{num}}^{(k)}(\delta t, T)$ as a function of T (double well potential, energy of 1.25 close to the barrier energy), for a time step $\delta t = 0.05$ and 2×10^8 time steps.

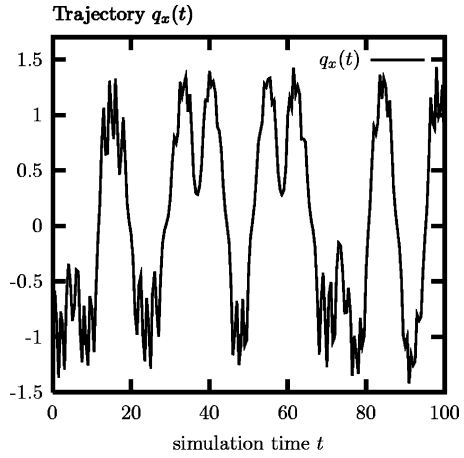

FIG. 8. Evolution of $q_x(t)$ as a function of time t (double well potential, energy of 1.25 close to the barrier energy, time step $\delta t = 0.01$).

IV. THE CASE OF SYSTEMS WITH MULTIPLE METASTABLE STATES

Alkane chains and the double well example are systems presenting several metastable states. In Sec. III, we have studied them in the regimes of either low or high energy (see Secs. III D and III E). In such regimes, either a unique well is explored (because the system does not have enough energy to visit other wells), or the system is so energetic that it does not really "feel" the barriers. We now focus on the case of an intermediate energy: transitions between metastable states are possible but they are rare events in comparison with the faster characteristic time scale of the system. In this regime, we show that time averages converge to ensemble averages at the slower rate $1/\sqrt{T}$.

A. One particle in a double well potential

In this section, we simulate the same system as in Sec. III E, but we now choose to work with an energy $E_0 = 1.25$ that is slightly larger than the saddle point energy (which is equal to 1). In this case, one does not observe any convergence of the time averages, even for simulation times 10 000 times as large as those which lead to convergence in the cases of low or high energy (see Fig. 7).

Actually, as shown in Fig. 8, the particle spends "a long time" in one basin, then quickly undergoes a transition into the other basin, in which it spends another "long time," and so on. In this simple example, the basin locations are known and the mean exit time can be easily estimated from MD simulations. These data can be used to parametrize a two-state Markov chain model[10] mimicking the transition between the two basins. A comparison between the convergence rate of time averages (2.12) computed by MD on the one hand and by the so-obtained Markov chain model on the other hand, for the observable q_x (the average of which is zero due to the symmetry of the potential), is reported in Fig. 9. The good quantitative agreement between these two convergence rates confirms that the bottleneck in the computation of averages by brute force MD simulations is the presence of several well-separated basins; in this case, the convergence rate is no longer of $1/T$ or better, but falls down to $1/\sqrt{T}$, at least in the range of accessible simulation times to date.

B. Alkane chains

In this section, we simulate the same chains of particles as in Sec. III D (with $M = 40$ particles), but we now work at higher temperature.

We first simulate the chain at an energy such that the temperature is $T = 135$ K, and study the end-to-end distance. We observe that, even if we simulate the system for a time 400 times longer that in the case studied in Sec. III D, time averages have not yet reached their infinite time limit. We study in Fig. 10 their convergence rate: it is very close to $1/\sqrt{T}$.

The reason for this slow convergence is the same as in the double well potential case: the system has enough energy to explore many wells, but it spends a long time in each well (in comparison with the time needed to go from one well to another one). Thus, once again, time average convergence rate is slowed down to $1/\sqrt{T}$.

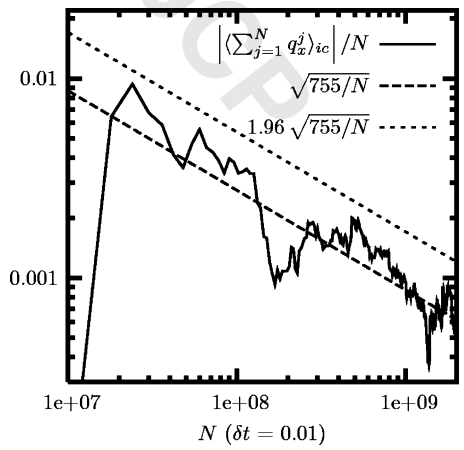

FIG. 9. Estimation of the convergence of $(1/N)\,\Sigma_{j=1}^{N} q_x^j$ to 0 with respect to N with a Markov chain model (double well potential, energy of 1.25 close to the barrier energy, $\delta t = 0.01$; $\langle \cdot \rangle_{\text{ic}}$ denotes average over 16 initial conditions).

PROOF COPY 302439JCP

PROOF COPY 302439JCP

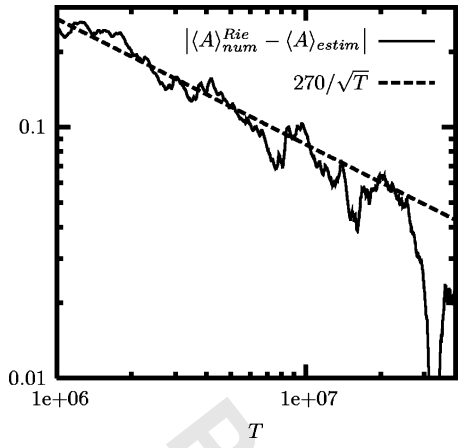

FIG. 10. Estimation of the convergence rate of the averaged end-to-end distance (alkane chain with $M=40$ particles, temperature of $T=135$ K).

Let us now simulate the chain at a temperature somewhat higher ($T=4.16$ K) than that in the case studied in Sec. III D (where $T=2.13$ K), but still very low. The situation is exactly the same as that in the case $T=135$ K: neither time averages computed by Eq. (2.5) nor by Eq. (2.12) converge (see Fig. 11).

V. CONCLUSIONS

In order to compute ensemble averages of observables, one can use molecular dynamics and compute a time average of these observables over one (or several) trajectory of a dynamical system. When the ergodic assumption is satisfied, this time average converges, as the trajectory length goes to infinity, toward the ensemble average. We have presented a numerical method which allows, when this convergence occurs at the rate $1/T$, to speed up the convergence up to $1/T^k$ (k arbitrary) so that a trajectory of reduced length (and thus needing less computational effort) is sufficient to achieve a given precision. Together with this numerical method, we have presented some error bounds that may be used to optimize the simulation parameters. We have demonstrated by test cases the sharpness of these bounds. However, a limitation of our method must be pointed out.

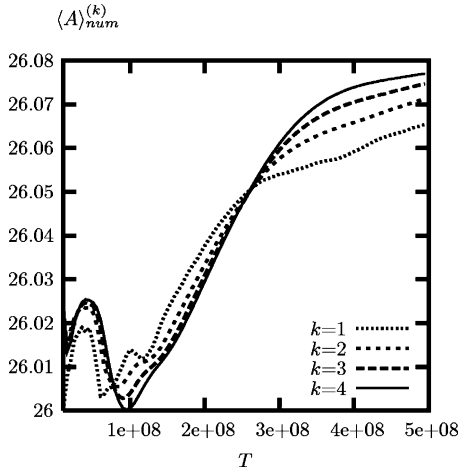

FIG. 11. Evolution of the end-to-end distance time average as a function of T (alkane chain with $M=40$ particles, temperature of $T=4.16$ K).

A typical problem in molecular dynamics is the exploration of the whole phase space when the potential energy of the system has several basins corresponding to distinct metastable states, and when the energy is such that transitions between these states are possible but are rare events. In this case, the convergence rate of time averages is very slow (it is close to $1/\sqrt{T}$). The method we have presented in this paper does not solve this issue (it does not improve the phase space exploration, but the computation of the time average). Other methods have then to be used, for instance those suggested in Refs. 11–14, and may be combined with our method.

ACKNOWLEDGMENTS

We would like to thank Gilles Zérah for his very helpful suggestions and Christophe Chipot for stimulating discussions. We also gratefully acknowledge the financial support of INRIA through the contract grant "Action de Recherche Coopérative PRESTISSIMO."

APPENDIX

In this Appendix, we detail the hypotheses needed to prove estimates (2.4) and (2.8). As we have mentioned in Sec. II, we consider a system of M particles in 3D, described by the Hamiltonian $H(\mathbf{q},\mathbf{p})$. Let $(\mathbf{q}_0,\mathbf{p}_0)$ be the initial condition.

1. Assumptions on the Hamiltonian function *H*

We suppose that the Hamiltonian H is an analytic function, and that it is *completely integrable*, that is to say that (see Ref. 6, pp 214 and 272):

(1) there exist $3M$ invariant functions $I_j(\mathbf{q},\mathbf{p})$, $j=1,\ldots,3M$, of the dynamical system (2.1) (the definition of an invariant function is recalled in Sec. II); we denote by $S(\mathbf{q},\mathbf{p})$ the level set of the invariant functions $\{I_j\}_{1\leqslant j\leqslant 3M}$ (see Eq. (2.2));
(2) these invariants I_j are in involution, i.e., satisfy the condition $\forall(\mathbf{q},\mathbf{p})\in\mathbb{R}^{3M}\times\mathbb{R}^{3M},\forall j_1,j_2,\nabla_{\mathbf{q}}I_{j_1}(\mathbf{q},\mathbf{p})\cdot\nabla_{\mathbf{p}}I_{j_2}(\mathbf{q},\mathbf{p})=\nabla_{\mathbf{p}}I_{j_1}(\mathbf{q},\mathbf{p})\cdot\nabla_{\mathbf{q}}I_{j_2}(\mathbf{q},\mathbf{p})$, where $\nabla_{\mathbf{q}}$ is the gradient with respect to the position variables and $\nabla_{\mathbf{p}}$ is the gradient with respect to the momentum variables;
(3) there exists a neighborhood of the initial condition $(\mathbf{q}_0,\mathbf{p}_0)$ such that, for all $(\mathbf{q},\mathbf{p})$ in this neighborhood, (i) the level set $S(\mathbf{q},\mathbf{p})$ is compact and connected,[15] and (ii) the gradients ∇I_j of the invariant functions are linearly independent.

2. Diophantine assumption

Let $\mathbb{T}$ be the torus $\mathbb{R}/2\pi\mathbb{Z}$. Under the above hypotheses, it is possible[6,8] to find a bounded open set $B\subset\mathbb{R}^{3M}$ containing $(I_1(\mathbf{q}_0,\mathbf{p}_0),\ldots,I_{3M}(\mathbf{q}_0,\mathbf{p}_0))$, a bounded open set $B'\subset\mathbb{R}^{3M}\times\mathbb{R}^{3M}$ containing the initial condition $(\mathbf{q}_0,\mathbf{p}_0)$, and a local change of variables ψ,

$$\psi:(\mathbf{a},\theta)\in B\times\mathbb{T}^{3M}\mapsto(\mathbf{q},\mathbf{p})\in B'\subset\mathbb{R}^{3M}\times\mathbb{R}^{3M},$$

PROOF COPY 302439JCP

PROOF COPY 302439JCP

such that the new variables $(\mathbf{a}(t),\theta(t))=\psi^{-1}(\mathbf{q}(t),\mathbf{p}(t))$ obey the simple dynamics

$$\frac{d\mathbf{a}}{dt}=0, \quad \frac{d\theta}{dt}=\omega(\mathbf{a}(t)), \tag{A1}$$

supplied with initial condition $(\mathbf{a}(0),\theta(0))=(\mathbf{a}_0,\theta_0)=\psi^{-1}(\mathbf{q}_0,\mathbf{p}_0)$. In Eq. (A1), ω is a function from $\mathbb{R}^{3M}$ to $\mathbb{R}^{3M}$. The trajectory in the new variables is $\mathbf{a}(t)=\mathbf{a}_0$, $\theta(t)=\theta_0+\omega(\mathbf{a}_0)t$.

We then make the following additional assumption (the so-called diophantine assumption) on the vector $\omega(\mathbf{a}_0)$: there exist a constant $C_0>0$ and an exponent $\gamma_0>0$ such that

$$\forall \alpha\in\mathbb{Z}^{3M}\backslash\{0\}, \quad |\alpha\cdot\omega(\mathbf{a}_0)|\geqslant\frac{C_0}{|\alpha|^{\gamma_0}}, \tag{A2}$$

where we note $|\alpha|=\alpha_1+\cdots+\alpha_{3M}$ and $\alpha g q(\mathbf{a}_0)=\alpha_1\omega_1(\mathbf{a}_0)+\cdots+\alpha_{3M}\omega_{3M}(\mathbf{a}_0)$. This assumption means that the quantity $\alpha\cdot\omega(\mathbf{a}_0)$ can get close to zero only if $|\alpha|$ goes to infinity.

Under the assumption that the Hamiltonian function H is analytic and integrable, and the diophantine assumption, one can prove that time averages converge to ensemble averages at the rate $1/T$, and that this rate can be improved up to $1/T^k$ by using filtering functions. In addition, if the numerical scheme used to integrate the equations of motion is symplectic, and if a nonresonance condition is added to the diophantine assumption (A2), then one can prove estimate (2.13) (see Sec. II and Refs. 5 and 6).

[1] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science, ■, 1987).
[2] D. Frenkel and B. Smit, *Understanding Molecular Simulation, from Algorithms to Applications*, 2nd ed. (Academic, New York, 2002).
[3] M. E. Tuckerman and G. J. Martyna, J. Phys. Chem. B **104**, 159 (2000).
[4] S. D. Bond, B. J. Leimkuhler, and B. B. Laird, J. Comput. Phys. **151**, 114 (1999).
[5] E. Cancès, F. Castella, Ph. Chartier, E. Faou, C. Le Bris, F. Legoll, and G. Turinici, Numer. Math. (submitted).
[6] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1989).
[7] J. P. Ryckaert and A. Bellemans, Faraday Discuss. **66**, 95 (1978); M. G. Martin and J. I. Siepmann, J. Phys. Chem. B **102**, 2569 (1998).
[8] E. Hairer, Ch. Lubich, and G. Wanner, *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer, Berlin, 2002).
[9] R. Y. Rubinstein, *Simulation and the Monte Carlo Method* (Wiley, New York, 1981).
[10] Ch. Schütte and W. Huisinga, in *Handbook of Numerical Analysis, Vol. X, Special volume: Computational Chemistry* (North-Holland, Amsterdam, 2003).
[11] M. R. Sorensen and A. F. Voter, J. Chem. Phys. **112**, 9599 (2000).
[12] W. E. W. Ren and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[13] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[14] Z. Zhu, M. E. Tuckerman, S. O. Samuelson, and G. J. Martyna, Phys. Rev. Lett. **88**, 100201 (2002).
[15] G. F. Simmons, *Topology and Modern Analysis* (McGraw-Hill, New York, 1963).

PROOF COPY 302439JCP

PROOF COPY 302439JCP