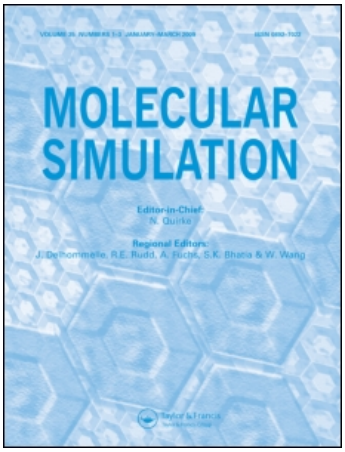


**Molecular Simulation**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

# Error Estimation in Molecular Dynamics Experiments with a Tabulated Intermolecular Interaction Potential

M. G. Kiselev[a]; B. G. Abrosimov[a]; I. I. Vaisman[a]; Y. M. Kessler[a]
[a] Institute of Non-Aqueous Solutions Chemistry of the Academy of Sciences of the USSR, Ivanovo, USSR

**To cite this Article** Kiselev, M. G. , Abrosimov, B. G. , Vaisman, I. I. and Kessler, Y. M.(1988) 'Error Estimation in Molecular Dynamics Experiments with a Tabulated Intermolecular Interaction Potential', Molecular Simulation, 1: 5, 321 — 326
**To link to this Article: DOI:** 10.1080/08927028808080953
**URL:** http://dx.doi.org/10.1080/08927028808080953

*Molecular Simulation*, 1988, Vol. 1, pp. 321–326
Photocopying permitted by license only
Reprints available directly from the publisher


# ERROR ESTIMATION IN MOLECULAR DYNAMICS EXPERIMENTS WITH A TABULATED INTERMOLECULAR INTERACTION POTENTIAL

M.G. KISELEV, B.G. ABROSIMOV, I.I. VAISMAN, Y.M. KESSLER

*Institute of Non-Aqueous Solutions Chemistry of the Academy of Sciences of the USSR, Ivanovo 153751, USSR*

(*Received August, 1987; in final form December, 1987*)

Stringent estimations of statistical errors in the computer simulation of a system with a tabulated interaction show no accumulation of errors in calculated binary distribution functions, potential of mean force and velocity autocorrelation functions.

KEY WORDS: Molecular dynamics, tabulated potential, statistical errors



## INTRODUCTION

The computer simulation of molecular liquids, solutions or solids using molecular or stochastic dynamics may require a large amount of computer time. The calculation of interparticle potentials and forces is the most time consuming part of the whole procedure. To increase the speed of the simulation it is possible to tabulate potentials and forces as functions of the distance between particles $i$ and $j$. The present paper is an attempt at a strict evaluation of statistical errors of various thermodynamic functions due to such tabulation.

## BASIC DEFINITIONS

Consider an n-dimension configurational space of identical particles with a binary interaction potential:

$$V(x_1, x_2), \quad x_1, x_2 \in R^n, \quad x_1 \neq x_2 \tag{1}$$

We denote $N(c)$ as the number of particles in a given configuration, $c \in R^n$, $H(c)$ the potential energy of the configuration:

$$H(c) = \sum_{\substack{x,y \in c \\ x \neq y}} V(x, y) \tag{2}$$

and introduce an indicator function [1] $\xi(x) > 0$ for a mainfold $c \in R^n$. Then the probability distribution relative to Lebegue measure is [1]:

$$P_{\xi,z,\beta}(c) = V^{N(c)} \left( \prod_{x \in c} \xi(x) \right) exp \ (-\beta H(c))/\Xi(\xi, z, \beta) \tag{3}$$

where $V$ is the volume, $\beta$ is a constant and the partition function $\Xi(\xi, 2, \beta)$ is defined

by

$$\Xi(\xi, z, \beta) = \int_{cR^n} V^{N(c)} \left( \prod_{x \in c} \xi(x) \right) exp\,(-\beta H(c))\, dc \tag{4}$$

For the correlation functions of distribution (3) we have:

$$\varrho_{\xi,z,\beta}(c) = V^{N(c)} \left( \prod_{x \in c} \xi(x) \right) V^{N(\bar{c})} \int_{cRn} \left( \prod_{x \in \bar{c}} \xi(x) \right) exp\,(-\beta H(c \cup \bar{c}))\, d\bar{c} \tag{5}$$

The assumed properties of the pair potential are as follows:

a) Translational invariance: $V(x_1, x_2) = \bar{V}(x_1 - x_2)$

b) Stability: there is a constant $B > 0$ such that for any configuration and $N \geqslant 0$ the inequality (6) is fulfilled.

$$\sum_{\substack{x_1, x_2 \in c \\ x_1 \neq x_2}} U(x_1, x_2) \geqslant -2BN(c) \tag{6}$$

c) Regularity: a pair potential is regular if the relation (7) exists:

$$c(\beta) = \int_0^\infty (exp\,(-\beta H(c)) - 1)\, dx < \infty \tag{7}$$



## STATISTICAL ERRORS IN THE CALCULATION OF DISTRIBUTION FUNCTIONS

In this section we present formulae for estimating statistical errors in distribution functions caused by the use of a tabulated form of the pair potential. With the definitions given above, the $S$-particle correlation function is expressed as follows:

$$\varrho^s_{z,\beta}(c) = V^s \int_{cR^n} [exp\,(-\beta H(c))_{N(c)}]\, dc / \Xi(\xi, z, \beta) \tag{8}$$

We further assume that the potential energy of interaction can be represented as the sum of three terms:

$$U(x)_N = U(x)_S + U(x)_{N-S} + U(x)_{S,N-S} \tag{9}$$

where $U(x)_S$ is the interaction energy of $s$ particles arbitrarily distributed with respect to each other in the physical space, $U(x)_{N-S}$ is the same for $N$–$S$ particles and $U(x)_{N-S,S}$ is the interaction energy between particles $S$ and $N$–$S$.

With the use of the property (b) of the potential we obtain the estimate for the interaction energy

$$U_{S,N-S} > -2SB. \tag{10}$$

Then substituting (10) in (8) we obtain:

$$|\varrho^s(c)| < V^s\, exp\,(-\beta H(c))\, exp\,(2S\beta B) \int_{cR^n} exp\,(-\beta H(\bar{c})\, d\bar{c}$$

$$= exp\,(-\beta H(c))\, exp\,(2S\beta B)\, V^s\, \Xi_{N-1}/\Xi_N \cdot \Xi_{N-2}/\Xi_{N-1} \cdots \Xi_{N-S}/\Xi_{N-S+1} \tag{11}$$

By definition,

$$z = \lim_{N\to\infty}\left(\frac{N+1\ \Xi_N}{\Xi_{N+1}}\right), \quad v = V/N \tag{12}$$

where $z$ is the activity. In the thermodynamic limit we have:

$$\varrho_z^s(S) < (zv)^s \exp(-\beta v(x)) \exp(2S\beta B) \tag{13}$$

Equation (13) allows us to immediately estimate statistical errors caused by the use of tabulated potentials instead of their analytic expressions. To that end tabulated numerical values of a potential were presented in analytic form by virtue of the step-function $\sigma(x)$

$$\tilde{V}(\xi_{ij}) = \sum_{j=1}^{L} [\sigma(\xi_{ij} - \xi_k) - \sigma(\xi_{ij} - \xi_{k+1})]\, V(\xi_k) \tag{14}$$

where $L$ is a number of table steps. From (13) and (14) we have for the case of binary distribution functions:

$$|\Delta\varrho_z^2(c)| \equiv |\varrho_z^2(c) - \tilde{\varrho}_z^2(c)| \leqslant (zv)^2 |\exp(-\beta v(\xi_{12})) - \exp(-\beta v(\xi_k))| \tag{15}$$

By definition, the mean force potential $\Psi(c)$ is given by:

$$\Psi(c) = -kT \ln \varrho_z \tag{15a}$$

So we readily obtain the formula:

$$|\psi(c) - \tilde{\psi}(c)| \leqslant \sup|V(c) - \tilde{V}(c)| \tag{15b}$$

These inequalities show that statistical errors are not a function of number of steps or of the number of particles in configuration $c$. In other words, errors are not accumulating.

It seems worthwhile to obtain a still more explicit expression than (15) by employing a specific function $V(x_1, x_2)$. We have chosen the Lennard-Jones (LJ) potential due to its widespread use. The LJ potential has the form

$$V(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{15c}$$

and can be shown to obey conditions a–c.
The first condition is obviously fulfilled:

$$V(x_1, x_2) = 4\varepsilon\,[(\sigma/(x_1 - x_2))^{12} - (\sigma/(x_1 - x_2))^6] \tag{16}$$

At the point of minimum we have:

$$\frac{dV}{dr^*} = 0; \quad \frac{6}{\sigma}\cdot(\sigma/r^*)^7 - \frac{12}{\sigma}(\sigma/r^*)^{13} = 0 \tag{17}$$

where $r^* = (r)_{\min}$
Hence we get an inequality:

$$\sigma V(x_1 - x_2) \geqslant -2\varepsilon N(c) \tag{18}$$

where $\varepsilon$ is the depth of the minimum. Equation (18) verifies that condition (b) is also

fulfilled. The 3rd condition demands that the definite integral

$$\int_0^\infty dx| \; exp\; (-\beta(4\varepsilon((\sigma/r)^{12} - (\sigma/r)^6))) - 1| \tag{19}$$

would be less than infinity which is obviously so from (19). Therefore (15) can be used with (15a) to obtain an estimate for statistical errors in binary distribution functions

$$|\Delta\varrho_\lambda^2(c)| \leqslant (zv)^2 \; exp\; (4\beta\varepsilon) \cdot 4\varepsilon\delta \left| \left(\frac{\sigma}{r^* + \delta}\right)^{12} - \left(\frac{\sigma}{r^* + \delta}\right)^6 \right| \tag{20}$$

Equations (15b) and (20) lead us to a statement that structural information is neither deteriorated or distorted by exploiting a tabulated form of the potential. More than that, they show how statistical errors can be decreased by a proper choice of the step $\delta$.

*Estimates of Ursell Functions*

Under the conditions (A)–(C) we can write for a volume

$$ln\Xi(\lambda, z, \beta) = \int_{c\lambda} \psi_{z,\beta}(c)\, dc \tag{22}$$

where $\psi_{z,\beta}$ is an Ursell function.
Estimates for Ursell functions are given in [1]:

$$|\psi_{z,\beta}(c)| \leqslant exp\; (\beta B) \sum_{\gamma \in S_c} | \; exp\; (-\beta \tilde{V}(x_2 - x_1)| \tag{23}$$

where $S_c$ are all non-cycled graphs for configuration $c$. If $\psi_{z,\beta(c)}$ is the Ursell function for a tabulated form of potential, then starting from (23) we derive an inequality:

$$\frac{\psi_{z,\beta(c)} - \psi_{z,\beta(c)}}{N_c \psi_{z,\beta(c-1)}} \leqslant z\; exp\; (2\beta B) \cdot \delta \tag{24}$$

*Estimates of statistical errors in autocorrelation functions*

Let $C_0(x_1, x_2)$ be a configuration of particles at time $t_0$ and $C_t(x_1, x_2)$ be one at time $t$. Then we introduce an autocorrelation function in usual way:

$$\zeta(t) = 1/T \int_0^T \langle\, \omega(t)\omega(0) \,\rangle\, dt = 1/T \int_0^T \int_\lambda \varrho_\lambda(x)_2 \omega(t) w(0) dt d\Lambda \tag{25}$$

It was found in reference [2] that:

$$\int_\lambda K(\lambda)\delta(\lambda)\, d\Lambda \leqslant sup\; |K(\lambda)| \int_\lambda \delta(\lambda)\, d\lambda \tag{25a}$$

Starting from Equations (15) and (25a) we come after some trival but lengthy algebraic exercises to the desire results, Equation (26)

$$\Delta\zeta(t) \leqslant \delta(zv)^2\; exp\; (4\beta B)\; 1/T \left\{ \int_0^T \langle v(t) \rangle \int_{C_{t_0}} v(o)\, dc_0 + \right.$$

$$\left. + \langle v(o) \rangle \int_{C_t} v(t)\, dc_t \right\} dt + O(\delta^2) \tag{26}$$

We have found statistical errors for integral characteristics of a system, viz., Ursell functions, Equation (24), autocorrelation functions, Equation (26). These errors have the same dependence on $\delta$ as the errors in the distribution functions (Equation (20)). We can safely conclude that in both cases there is no error accumulation due to the use of a tabulated potential in the course of a molecular dynamics experiment.

## STATISTICAL ERRORS ARISING IN THE INTEGRATION OF NEWTON'S EQUATION

The use of a tabulated potential in a molecular dynamics simulation will lead to statistical errors. If the scheme of integration is as follows:

$$x^i_{n+1} = 2x^i_n - x^i_{n-1} + \Delta t^2 \sum_{j=1}^{N} F(x_i, x_j) \tag{27}$$

Then it should be readily apparent that the error accumulation is determined by the force errors. For the latter we have:

$$\Delta \bar{F}_i = \sum_i (F(\xi_{ij}) - F(\xi_k))\,[\sigma(\xi_{ij} - \xi_k) - \sigma(\xi_{ij} - \xi_{k+1})] \leqslant \sum_i \frac{dF}{d\xi_{ij}}\,\delta \tag{28}$$

Taking the ergodicity hypothesis as granted we obtain:

$$\Delta \bar{F}_i \leqslant \delta \int \sum_{ij} dF/d\xi_{ij}\, \varrho_i(x)_2 (dx)_2 \leqslant$$

$$\delta^2\, 64\varepsilon^2 \pi\beta (zv)^2\, exp\,(4\beta B) \left(6\,\frac{\sigma^6}{\delta^6} - 12\,\frac{\sigma^{12}}{\delta^{12}}\right) \times \tag{29}$$

$$((\sigma/(\sigma^* + \delta))^{12} - (\sigma/(\sigma^* + \delta))^6)$$

and hence we can derive expression for statistical errors associated with the integration procedure:

$$\Delta \bar{x} \sim 4\,\Delta F\,\Delta t^2 \tag{30}$$

It is evident that $\Delta \bar{x}$ should not exceed the thermodynamic fluctuations of $x$. The opposite would indicate that the simulation is liable to procedure artifacts. This provides the basis for a table step upper limit estimate for the LJ potential.

## STATISTICAL ERROR ESTIMATION FOR SOME OTHER FUNCTIONS

We define a function $F(x)_s$ on a differentiable subspace $R^s \subset R^n$ by Equation (31)

$$\zeta(x)_s \equiv \zeta(x_1, \cdots, x_s) \tag{31}$$

Then the statistical errors in $\zeta(x)_s$ appearing as the result of the MD experiment are as follows:

$$\Delta\zeta(x_1, \cdots, x_s) = \sum_{i=1}^{s} \frac{\partial \zeta}{\partial x_i}\,\Delta x_i \tag{32}$$

With Equations (20) and (32) and under the ergodicity hypotheses we obtain:

$$\bar{\Delta}\zeta \leqslant sup\,|\Delta x| \int_{C_R} \sum_{i=1}^{s} \frac{\partial \zeta}{\partial x_i}\, \varrho_s(x)\, d(x)_s \leqslant$$

$$sup\,|\Delta x| \cdot sup\,|\,\varrho_s(x)| \int_{C_{Rs}} \sum_{i=1}^{s} \frac{\partial \zeta}{\partial x_i}\,(x_1, \cdots, x_s)\, d(x)_s \leqslant \tag{33}$$

$$sup\,|\Delta x| \cdot sup \left| \sum_{i=1}^{s} V(\xi_{ij}) - \sum_{i=1}^{s} V(\xi_k) \right| \int_{C_{Rs}} \sum_{i=1}^{s} \frac{\partial \xi}{\partial x_i}\,(x_1, \cdots, x_s)\, d(x)_s$$

In a specifical case when s = 2 and for an LJ potential, Equation (15a), with Equations (20, 29–30), leads to:

$$\bar{\Delta}\xi \leqslant \Delta t^2 \cdot\ 1024\, \delta^3 \varepsilon^3 \pi \beta (zv)^4\, exp\,(8\beta\varepsilon) \left( 6\,\frac{\sigma^6}{\delta^6} - 12\,\frac{\sigma^{12}}{\delta^{12}} \right) *$$

$$\{((\sigma/\sigma^* + \delta))^{12} - ((\sigma/\sigma^* + \delta))^6\}^2 \tag{34}$$

## CONCLUSIONS

Various integral properties obtainable through the medium of an MD experiment on molecular systems have been considered from the point of view of statistical error accumulation deriving from the use of tabulated values of potentials and forces. It has been found that the properties concerned do not accumulate errors and that statistical errors depend on, and are of the order of, $\delta$. These findings offer considerable scope for further improvements on MD technique both in the reduction of errors and the saving of computing time.



### *References*

[1] R.A. Minlos, S.K. Pogosyan "Estimation of Ursell functions, group functions, and their derivatives", *Theor. & Math. Phys.*, **31**, 408, (1977), translation of *Theor. & Mat. Fyz.* **31**, 199 (1977).
[2] N.M. Korobov *The theoretically numerical methods in analysis*, Moscow, 1974.
[3] T.L. Hill *Statistical mechanics. Principles and selected applications* McGraw-Hill, 1956, New York.
[4] D.M. Ceperley The relative performances of several scientific computers for a liquid molecular dynamic simulation in "Super computer in chemistry", *ACS Symp.* **173**, 125 (1981).
[5] D. Ruelle *Statistical mechanics. Rigorous results*, New York, Amsterdam 1969.