Taboo search by successive confinement: Surveying a potential energy surface

Sergei F. Chekmarev
Institute of Thermophysics, 630090 Novosibirsk, Russia
and Novosibirsk State University, 630090 Novosibirsk, Russia
(Received 27 November 2000; published 27 August 2001)

A taboo search for minima on a potential energy surface (PES) is performed by means of confinement molecular dynamics: the molecular dynamics trajectory of the system is successively confined to various basins on the PES that have not been sampled yet. The approach is illustrated for a 13-atom Lennard-Jones cluster. It is shown that the taboo search radically accelerates the process of surveying the PES, with the probability of finding a new minimum defined by a propagating Fermi-like distribution.

DOI: 10.1103/PhysRevE.64.036703 PACS number(s): 02.50.Ng, 82.20.Wt, 36.40.−c, 87.15.He

I. INTRODUCTION

Molecular dynamics (MD) and Monte Carlo (MC) methods have long become the basic tools of computational studies of multiatomic systems. Provided that the interatomic potential is known, these methods allow one to obtain all desirable information about a system described in terms of classical mechanics, including a knowledge of its potential energy surface (PES), equilibrium properties, and kinetics.

Potentialities of the MD and MC methods were increased substantially due to Stillinger and Weber (SW) [1], who suggested supplementing MD (MC) simulations with a diagnostic quenching of the system at regular intervals. Every quench leads to a minimum on the PES, in the vicinity of which the system moves at the present moment. Due to this, the current point of the MD (MC) trajectory can be related to a certain minimum, and the points related to the same minimum will represent the basin surrounding this minimum. Every minimum, in turn, can be associated with a certain inherent structure of the system, so that the PES is separated into basins for inherent structures. Following this approach, one can collect the phase points for each of the basins separately. Therefore, not only can the PES be surveyed, but the equilibrium properties of particular inherent structures and rates of the transitions between them can also be calculated.

The SW approach is very explicit and easy to implement, but computationally it is not as efficient as desirable. The reason for this is that conventional MD and MC methods reproduce a ''true'' behavior of a system that obeys the laws of inherent dynamics and statistics, and, as a consequence, they suffer from the limitations imposed by these laws. In particular, according to the Boltzmann distribution, a system predominantly samples the lower part of a basin, and, as a result, it may dwell in the basin for a long time. Moreover, in complex systems, the basins often cluster into superbasins (funnels) separated by high barriers. Therefore, on time scales accessible to current computers, the system may be trapped in a superbasin, which results in nonergodic behavior of the system.

For the last two decades, a variety of methods has been suggested to make sampling of PES's more uniform. Depending on the underlying concept, these methods can roughly be separated into three groups [2]: *a modification of the PES by introducing an appropriate biasing potential* (umbrella potentials [3–5], multicanonical sampling [6,7], a generalized ensemble approach [8,9], and hyperdynamics [10]), *biasing the temperature* (weighted histogram analysis [11,12] and jump-walking [13] methods, and simulated [14] and parallel [15] tempering), and *a modification of the method of surveying the PES* (the eigenmode method [16] and the activation-relaxation technique [17–19]). In the case of biasing the PES or temperature (the first two groups of the methods), the results for the true conditions are recovered from the results of the biased simulations by applying the formulas of statistical mechanics, specifically the Boltzmann law. In contrast to the SW approach, all these methods are not of multipurpose character; each of them pursues a specific goal: either *the calculation of equilibrium properties* (the weighted histogram analysis method [11,12], umbrella potentials [3–5], simulated [14] and parallel [15] tempering, multicanonical sampling [6,7], generalized ensemble approach [8,9]), or *the study of kinetics* (hyperdynamics [10]), or *surveying a PES* (the jump-walking method [13], the eigenmode method [16], and the activation-relaxation technique [17–19]).

Considerably less attention has been paid to the fact that the system, in accordance with the Boltzmann statistics, repeatedly visits the basins that were previously visited. At the same time, if the goal is to survey a PES, such events should be controlled, because many of them bear no new information about the PES. For this purpose one could employ the taboo search method, which has been elaborated for discrete surfaces [20,21] and extended to continuous surfaces [22].

Also note that the problem of surveying a PES is closely related to the global optimization problem (see, e.g., Ref. [23]). In the latter, the goal is the global minimum of the surface rather than the surface itself. This utilizes certain specific features in the above methods when they are applied to the search for the global minimum [23].

Of particular importance for a system is the knowledge of its PES, because the PES completely defines a system's behavior under specific conditions. The understanding of how the PES is built provides a deep insight into the dynamics and statistics of a system [24–26]. Moreover, a knowledge of the PES offers the possibility of writing a master equation, which governs the system's kinetics [27,28], and thus opens a way for replacing the time-consuming simulations by a relatively easy solution of the master equation.

To characterize a PES, one has to know, in the first instance, the local minima and the saddles that connect these minima [24]. Of most importance are the minima, particularly if an indication is given which of them are directly connected. With this, finding the saddles becomes a straightforward task, which can be fulfilled using one of the numerous methods suggested for this purpose, e.g. the TRAVEL algorithm [29], the ridge method [30] or the method of contangency curves [31] (also see Ref. [32], which contains a comprehensive bibliography on these and related methods, as well as a comparison of typical methods in efficiency).

One recently developed approach to surveying a PES invokes successive walks from one minimum to an adjacent saddle, then from this saddle to another minimum, and so on [16–19]. These walks can be referred to as activation and relaxation phases, respectively [17]. For the activation phase, one of the uphill climbing methods [32] can be used, e.g., the eigenmode-following method of Cerjan and Miller [33] (or its variations [16,34]), as in Refs. [16,19,35,36], or a more economical algorithm, which does not require an evaluation of the full Hessian matrix at each step, as in Refs. [17,18]. For the relaxation phase, one of the standard methods, such as the steepest-descent or conjugate gradient method, is employed. If the number of minima is not very large, all, or practically all, minima can be located by a direct walk from one minimum to another [16,35,36]. However, if the number of minima becomes too large, just a statistical search for the minima is possible. For this purpose the transition between any two nearby minima is taken as an elementary Metropolis step [37], that is, the move from one minimum to another is accepted with a probability of $\min[1,\exp(-\Delta E/k_BT)]$, where k_B is the Boltzmann constant, T is the temperature, and ΔE is the change in energy associated with this move. The latter can be either the energy difference between the minima or the height of the barrier between the minima [17,19].

In this paper we consider an approach to surveying a PES, which implements a taboo search for minima by means of the confinement MD [38–40]. The technique of the latter, in brief, is as follows. Suppose a system (namely, its representative point) is within one of the basins on the PES. As in the SW method, we quench the system at regular intervals (τ_{quench}) along the MD trajectory, but in this case in order to check if the system is still in a given basin or has left it for another basin. If the system is found in another basin, it is returned to the given basin. For this, a certain procedure of the reversal of the MD trajectory is employed. In contrast to conventional dynamics, the system can now be kept in a basin for an arbitrarily long time, and thus the thermodynamic functions for the basin can be accurately calculated. Further, though the system does not leave the basin for a time longer than τ_{quench}, both the neighboring basins and all attempts of the system to pass into them can be recorded. Therefore, the rates of the transitions into the neighboring basins can also be calculated, and thus the kinetics of the system can be determined.

Allowing an enhanced sampling of local regions of a PES, the confinement technique offers interesting possibilities for the study of many-body systems. Some of these possibilities were demonstrated in application to clusters: the calculation of the caloric curves and absolute densities of states of specific structures (by the confinement of the system to the basins for these structures [39,40]), the estimation of the rate of complex transitions between specific structures (by successive confinement to the basins for the initial and intermediate structures [38,40]), and creating a simple subsystem of a complex system that retains characteristic features of the system (by confinement to the area of the PES associated with this subsystem [40,41]). Originally designed for a detailed study of specific structures and the channels of transitions between them, the confinement technique was recently applied with success to the study of a system (a solvated alanine tetrapeptide) as a whole [42]. For this purpose, following a certain taboo search strategy of surveying the PES (which allowed the system to pass just into the basins that were not previously sampled), the MD trajectory of the system was successively confined to various basins on the PES. Note that no prior information about the PES landscape is required to sample the entire PES: being placed at an arbitrarily chosen point on the PES, the system, basin by basin, can explore the entire PES. One essential feature of this approach is that it retains the multipurpose character of the SW approach [1], allowing one to survey the PES and calculate both the equilibrium properties and kinetics in the course of a single MD (or MC) run. At the same time, the approach does not suffer from the disadvantages inherent to the conventional simulations.

In this paper the process of surveying a PES by successive confinement is examined in more detail. For a specific system we will consider a 13-atom Lennard-Jones cluster, which possesses a large number of isomers and at the same time allows a complete survey of its PES. Applying a simple taboo search strategy, we will explore the PES for the cluster, and compare the results with those obtained in the conventional (SW) simulations as well as with the results of previous studies [16,35,36,43]. It will be shown that the taboo search radically accelerates the process of surveying the surface, with the probability of finding a minimum defined by a propagating Fermi-like distribution.

II. APPROACH AND COMPUTATIONAL BACKGROUND

With atoms interacting through a pairwise additive Lennard-Jones (LJ) potential, the potential energy of the cluster is

$$U=\sum_{i=1}^{n}\sum_{j<i}4\epsilon[(\sigma/r_{ij})^{12}-(\sigma/r_{ij})^{6}].$$

Here n is the number of atoms in the cluster, r_{ij} is the distance between ith and jth atoms, and ϵ and σ are, respectively, the characteristic energy and length, which are specific for a specific system (for example, for Ar $\epsilon=119.8$ K and $\sigma=3.904\times10^{-8}$ cm). Together with the atomic mass (m), ϵ and σ constitute so called LJ systems of units; the time unit, in particular, is $\sqrt{m\sigma^2/\epsilon}$ (2.15×10^{-12} s for Ar). In what follows, all numerical values will be given in this system of units.

In this paper we use a constant-temperature MD method based on numerical integration of the coupled Langevin equations

$$m\frac{d^2\mathbf{r}_i}{dt^2}+\alpha\frac{d\mathbf{r}_i}{dt}=-\frac{\partial U}{\partial \mathbf{r}_i}+\mathbf{\Phi}_i(t), \tag{1}$$

where α is viscosity, and $\mathbf{\Phi}_i(t)$ are random (Langevin's) forces due to thermal fluctuations in the system at given temperature T. These forces have zero mean, and are related to the viscosity by the fluctuation-dissipation theorem

$$\langle \mathbf{\Phi}_i(t)\rangle=0, \quad \langle \Phi_i^j(t)\Phi_{i'}^{j'}(t+\tau)\rangle=2\alpha k_B T\delta_{ii'}\delta_{jj'}\delta(\tau), \tag{2}$$

where the upper index at Φ stands for a component of the vector $(j,j'=1,2,3)$, the angular brackets denote an ensemble average, and $\delta_{kk'}$ and $\delta(\tau)$ are the Kronecker and Dirac deltas, respectively. Issuing phase trajectories from different points of phase space, one can form a *canonical* ensemble of the systems that corresponds to a given temperature T. Equations (1) and (2) were numerically integrated using the algorithm by Biswas and Hamann [44]; the time step is equal to 0.01, and the friction coefficient to 3.0 (for Ar, 2.15×10^{-14} s, and 1.4×10^{12} s^{-1}, respectively).

To reveal a possible dissociation (evaporation) of the cluster, the following criterion was used: the cluster was considered to be in a dissociated state if it divided into fragments (possibly consisting of a single atom) spaced at the distance of 2.6 or larger. In the course of the simulations the following characteristics of each inherent structure (the minimum of the basin) were calculated: the energy $U_{\min}$, the normal mode frequencies $\omega_1,\ldots,\omega_{3n}$, the principal momenta of inertia I_α $(\alpha=1,2,3)$, and the order of the point group h. In the general case of nonlinear atomic configurations, six of the above frequencies (corresponding to the translational and rotational degrees of freedom) are equal to zero, and the others are real. $U_{\min}$ was employed to compare the structures, and the other characteristics were used in theoretical estimates. Permutational structures were identified by calculating the distance in configuration space after a superposition of the structures.

Both conventional and confinement simulations were conducted. The conventional simulations followed the SW approach: the system freely explored the PES, and information for each basin was collected separately. In the confinement simulations the system was successively confined to various basins on the PES. To initiate the simulations, the system was placed into one of the basins, and the MD run was began. Quenching the system at regular intervals and comparing the quenched structure with that corresponding to the minimum of the current basin, we determined if the system was still in this basin or had left it for another one. If the system was found in another basin, then, depending upon the strategy of surveying the surface (see below), it either was allowed to continue its motion in the new basin or was returned to the given basin. In the latter case, a new trajectory had to be initiated in the given basin. In previous works [38–40] this was accomplished by the reversal of the trajectory at the boundary of the basin. However, this method sometimes fails, leading to a "cycling" of the trajectory in the neighborhood of a saddle point of high order [40]. In addition, since many quenches are required to locate the boundary, the method is time consuming. Therefore, a new trajectory in a basin was initiated immediately at the point where the system was found in the basin last time. The atomic velocities required to initiate the trajectory were taken as random values distributed according to the Maxwell law at temperature T.

This procedure of initiating the trajectories was carefully tested [42,45]. First, in the same manner as for the reversal of the trajectory at the boundary of the basin [39], the confinement simulations were checked for the reproduction of statistical characteristics of the system in an individual basin. For this, the MD trajectory was confined to a basin for the ground-state isomer in the LJ_{13} cluster, and the data obtained in the confinement simulations were compared with those obtained in the conventional simulations [45] (both the Newton and Langevin dynamics were examined, which are applicable to generating microcanonical and canonical ensembles, respectively). The distributions of the microstates over the potential and total energies (for the microcanonical and canonical ensembles, respectively) and the distributions of lifetimes were considered. In addition, to see how accurately the attempts of the system to escape the basin are distributed among different channels, the rates of the transitions into adjacent basins were compared. In all cases the results were found to be in good agreement, except for a slight difference in the distributions of lifetimes smaller than τ_{quench} [45]. Further, in the study of the solvated alanine tetrapeptide [42], the confinement and conventional simulations were also compared for the transition probabilities, and the results were found in good agreement as well. Therefore, the employed way of initiating new trajectories in basins is completely consistent with determining both the equilibrium properties and kinetics of a system.

III. STRATEGIES OF SUCCESSIVE CONFINEMENT

The strategy of how to explore the PES is one of the most essential elements in the confinement simulations. It prescribes to which basin the system should be confined next, and for how long a time. The strategy may be based on different criteria, depending on the problem under consideration and the goal of the study [42].

In this paper we examine one of the simplest strategies, which allows the system to pass into each new basin which has been encountered (when the system was confined to preceding basins) but not yet sampled. This strategy is applicable to the systems that are small enough so that a complete sampling of the entire PES is possible, as it was for the LJ_{13} cluster. In the underlying concept, it is very close to the taboo search [20,21] mentioned in Sec. I. To realize this strategy, different protocols are possible. We will consider the following two.

Protocol I. This corresponds to a direct taboo search in the terminology of Refs. [20,21]. The system is kept in each of the basins until, among the basins into which it attempts to

pass, a new basin appears that has not been previously sampled. To embody this protocol, a list of the basins is made and updated that indicates the basins that have been sampled. Each basin is characterized by its minimum (i.e., by the inherent structure).

Protocol II. This corresponds to a partial stochastic sampling [20,21]. The system is kept in each of the basins until it makes a certain number of attempts k_{att} to pass into adjacent basins. In this case two lists of the basins are made and updated: one indicates the basins that have been sampled, as in protocol I, and the other contains information about the basins that have been encountered but not yet sampled. Such information includes the characteristics of the minima of the basins (to recognize the basins) and the atomic coordinates and velocities of the system in these basins (to start the sampling of the basins). The basins to pass into are then chosen from this second list of the basins. The simplest choice is to take them in the order they were encountered. However, preference may be given to specific basins, for example, to basins with lower values of U_{min} [42]. If no new minima are found while the system is confined to each of the basins encountered, it is returned into the first basin of the second list of the basins, and the exploration of the PES is continued until the system visits each basin k_{ret} times (a new trajectory is thus initiated in each of the basins $k_{att}k_{ret}$ times in total). If an isomer whose permutational copy has already been sampled is encountered, in both cases (protocols I and II) the transition into the basin corresponding to this isomer is forbidden.

Optimal values of τ_{quench}, k_{att}, and k_{ret} are generally dependent upon many factors: the nature and size of the system, the conditions under consideration, and the goal of the study. Therefore, specification of these parameters is, to a large degree, a matter of trial and comparison. In other words, one should try different values of the parameters and compare the results for convergence. Nevertheless, certain recommendations can be given. For example, to efficiently survey a PES, τ_{quench} should evidently be of the order of the lifetime of the system in a basin. This is applicable to both protocols I and II. As for k_{att} and k_{ret} in protocol II, the results of the search are dependent upon the product of these quantities rather than upon their individual values. Moreover, in a wide range of $k_{att}k_{ret}$, provided that none of the factors is unreasonably small, this dependence is rather weak. In the simulations presented below the following values of these parameters were typically used: $\tau_{quench}=50$ time steps, $k_{att}=20$, and $k_{ret}=20$.

IV. RESULTS AND DISCUSSION

The total number of geometrically different minima found for the LJ_{13} system grew with every work: Hoare and McInnes [43], employing a growth algorithm to seek for candidate structures, found 988 minima. Later Tsai and Jordan [16], who used a direct search by walking on the PES with the help of variation of the eigenmode-following method [33], increased this number up to 1328. Finally, Doye, Miller, and Wales [35] and Ball and Berry [36], who also employed a direct search (with different variations of the eigenmode-following method), found 1467 and 1478 minima, respectively. The search for the total number of minima is complicated by the fact that there is no definite criterion to decide when the exploration of the PES can be regarded to be complete. Also note that the search for specific minimum is related to the class of NP-hard problems (in application to clusters, e.g., see Ref. [46]), since the number of minima grows with cluster size exponentially, the search generally requires not less than an exponential number of steps.

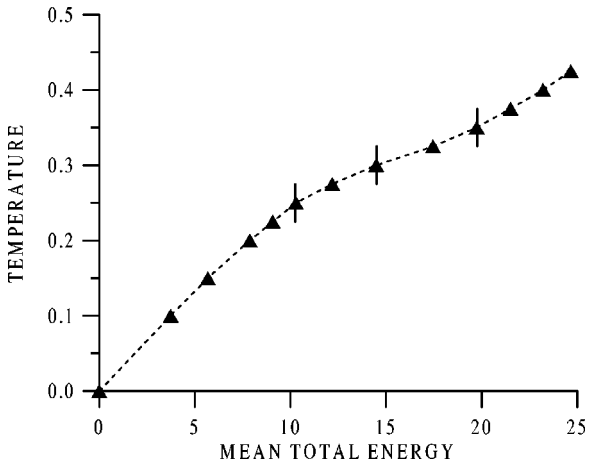

FIG. 1. Caloric curve for the LJ_{13} canonical ensemble.

In this paper simulations were conducted at three characteristic temperatures, which are indicated by bars at the (canonical) caloric curve shown in Fig. 1. These temperatures correspond to the solid-like state region (more definitely, to approximately the upper bound of this region, $T=0.25$), to the melting region ($T=0.3$), and to the liquid-like state region ($T=0.35$).

Figure 2 compares protocols I and II for the number of minima located with time in individual runs. The temperature corresponds to the liquidlike state ($T=0.35$). In both cases the system was initially placed into a basin corresponding to the ground-state isomer, which is the icosahedron:

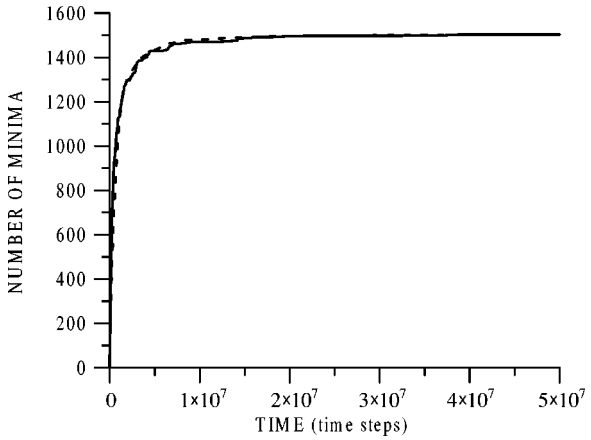

FIG. 2. Number of minima located with time in the confinement simulations: protocol I (solid line) and protocol II (dashed line). $T=0.35$.

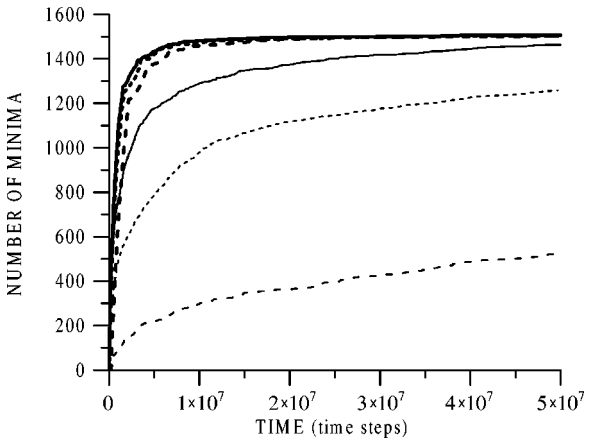

FIG. 3. Number of minima located with time in the confinement simulations under protocol II (thick curves), and in conventional simulations (thin curves): the long-dashed lines correspond to $T=0.25$, the short-dashed lines to $T=0.3$, and the solid lines to $T=0.35$.

$U_{\min}=-44.3268$. Though $k_{\rm ret}$ (protocol II) did not reach its maximum value ($k_{\rm ret}=20$) for each of the basins, both runs were terminated at 5×10^7 time steps, where in the case of protocol II no new minima were found for approximately 1×10^7 last steps. The total numbers of minima found under protocols I and II are 1504 and 1506, respectively. In both cases these exceed the maximum number of minima reported in previous works.

As can be seen from Fig. 2, in the initial stage of the search, when the list of sampled minima is almost empty, the number of minima in protocol I grows a bit faster than in protocol II. However, in the final stage, when that list of minima is almost complete, protocol I becomes less efficient because the system has to dwell in a current basin for a long time, until a basin which has not been sampled yet is encountered. Moreover, it may happen that the minima missed in the initial and intermediate stages of the search are not accessible from the basins sampled in the final stage. Therefore, protocol II seems to be more preferable for surveying a PES.

Figure 3 shows typical results of surveying the PES for the temperatures indicated in Fig. 1. Three upper (thick) curves correspond to a successive confinement taboo search (SCTS) under protocol II, and the other (thin) curves to a direct search by the SW method (SWDS); in both cases the quenching interval was the same, 50 time steps. First, as could be expected, the SCTS is faster than the SWDS. Second, and somewhat surprising, the SCTS results are slightly dependent upon the temperature: by the time the last of 1506 minima was located at $T=0.35$ ($t=4\times10^7$ time steps), there were found 1502 minima and 1499 minima found at $T=0.3$ and $T=0.25$, respectively; the total numbers of minima in the two latter cases were equal to 1502. This is particularly surprising in comparison with the SWDS simulations, where the decrease of the temperature from the liquidlike state ($T=0.35$) to the solidlike state ($T=0.25$) results in a radical reduction of the number of minima.

The energy spectrum of the minima located in these simulations is presented in Fig. 4. The disconnectivity graph, which shows how the most significant minima in the LJ_{13} cluster are connected, can be found in Ref. [35].

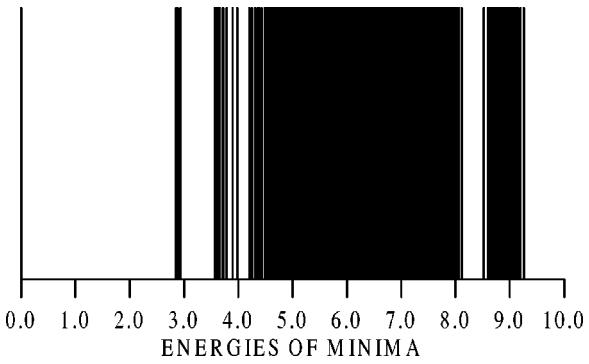

FIG. 4. Energy spectrum of minima for the LJ_{13} cluster. The energy is counted from the ground-state isomer energy $U_{\min}=-44.3628$.

It should be noted that the conventional (SWDS) simulations presented in Fig. 3 are not truely conventional: if an act of dissociation (evaporation) of the cluster was indicated, the MD run was continued in the same manner as in the confinement simulations, that is, a new trajectory was initiated at the last point where the system was found in bound state. Otherwise, at higher temperatures the trajectories have to be terminated in a short time: in Fig. 3, in particular, at approximately 6×10^4 and 3×10^4 time steps for $T=0.3$ and 0.35, respectively.

To learn more about SCTS simulations, consider the distributions of the number of quenches over the minima to which these quenches lead. In the SWDS simulations such a distribution originates from the Boltzmann distribution. For the PES divided into basins, the total partition function of the cluster of n atoms is written [1] as

$$Z=\sum_i \exp(-U_{\min,i}/k_BT)Z_i(T), \tag{3}$$

where

$$Z_i=\frac{1}{(2\pi\hbar)^{3n}}\frac{1}{h_i}\int \exp\{-H_i/k_BT\}d\mathbf{r}^{(3n)}d\mathbf{p}^{(3n)} \tag{4}$$

is the partition function for ith basin. Here $\hbar$ is the Planck constant, h_i is the order of the point group, and $H_i=\sum_{j=1}^{n}\mathbf{p}_j^2/2m+U_i$ is the Hamiltonian of the cluster associated with the ith basin (the potential energy U_i is counted from the basin minimum energy $U_{\min,i}$). According to Eqs. (3) and (4), the mean number of the quenches leading to the minimum of the ith basin can be estimated as

$$\langle N_i\rangle=N_{\rm tot}\exp(-U_{\min,i}/k_BT)Z_i/Z, \tag{5}$$

where $N_{\rm tot}$ is the total number of quenches, and the angular brackets denote the ensemble average.

Figure 5 compares the SWDS distribution with Eq. (5), in which Z_i are estimated according to the harmonic approximation [47]

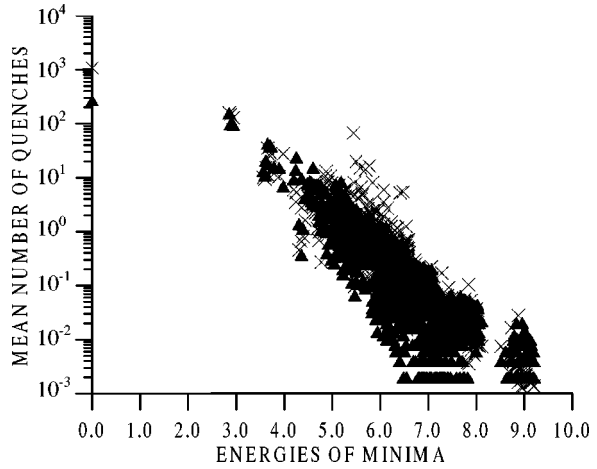

FIG. 5. Distribution of the mean number of quenches over the minima in the conventional simulations, $T=0.35$. Triangles correspond to the simulation data, and crosses to Eq. (5), with Z_i estimated in the harmonic approximation [Eq. (6)].

$$Z_i(T)=\left(\frac{k_BT}{\hbar\bar{\omega}_i}\right)^{3n-6}\frac{(2k_BT)^{3/2}(\pi I_{1i}I_{2i}I_{3i})^{1/2}}{h_i\hbar^3}, \qquad (6)$$

where $\bar{\omega}_i$ and $I_{\alpha i}$ $(\alpha=1,2,3)$ are, respectively, the geometrical mean of normal frequencies and the principal momenta of inertia for the inherent structure corresponding to the minimum of the ith basin. These characteristics, and also h_i, were determined in the course of the simulations. The first and second multipliers on the right-hand side of Eq. (6) correspond to the vibrational and rotational degrees of freedom, respectively; a contribution of the translational degrees of freedom is omitted, since it is the same for all isomers. A discrepancy between the simulation data and the estimate of Fig. 5 could be attributed to anharmonic effects, because, in the liquidlike region, to which the value of $T=0.35$ corresponds, the anharmonicity is appreciable (see, e.g. Ref. [48]).

The taboo search changes SWDS distributions in the part of interbasin statistics, or more definitely, in the part of sampling of the basins that were sampled: once the system has visited a basin, it is not allowed to pass into this basin again. Such a situation is characteristic of Fermi statistics. To find mean occupation numbers for the basins in this case, we can consider the collection of the isomers as an ideal gas. Since the intrabasin statistics remain Boltzmann statistics, the ith isomer will contribute to the partition function as $[Z_i(T)\exp(-U_{\min,i}/k_BT)]^{N_i}$, where N_i is equal to 0 or 1. The total partition function will then be

$$Z_{\mathrm{FL}}(T,N_{\mathrm{tot}})=\sum_{\{N_i\}}\prod_i\left[Z_i(T)e^{-U_{\min,i}/k_BT}\right]^{N_i}, \qquad (7)$$

where the summation is taken over all possible $\{N_i\}$ that satisfy the condition

$$\sum_i N_i=N_{\mathrm{tot}}. \qquad (8)$$

The calculation of $Z_{\mathrm{FL}}(T)$ and related quantities is hampered by condition (8). Therefore, as is usually done in such cases (see, e.g., Ref. [49]), consider the grand canonical ensemble instead of the canonical one. The total partition function for the grand canonical ensemble will be

$$\Theta_{\mathrm{FL}}(T)=\sum_{N_{\mathrm{tot}}=0}^{\infty}\lambda^{N_{\mathrm{tot}}}Z_{\mathrm{FL}}(T,N_{\mathrm{tot}}), \qquad (9)$$

where λ is the activity. Now condition (8) drops out, and the summation on the right-hand side of Eq. (9) is easily taken, and yields

$$\Theta_{\mathrm{FL}}(T)=\prod_i\left[1+\lambda Z_i(T)\exp(-U_{\min,i}/k_BT)\right].$$

The average value of N_{tot} is then calculated as

$$\langle N_{\mathrm{tot}}\rangle=(\partial\ln\Theta_{\mathrm{FL}}/\partial\ln\lambda)_T=\sum_i\frac{1}{1+Z_i^{-1}(T)e^{(U_{\min,i}-\mu)/k_BT}},$$

where $\mu=k_BT\ln\lambda$ plays a role of chemical potential. The quantities under the sum in the latter equation can be considered the mean occupation numbers for the basins, and we eventually have

$$\langle N_{\mathrm{FL},i}\rangle=\frac{1}{1+Z_i^{-1}(T)e^{(U_{\min,i}-\mu)/k_BT}}, \qquad (10)$$

where μ is determined by the condition $\Sigma_i\langle N_{\mathrm{FL},i}\rangle=\langle N_{\mathrm{tot}}\rangle$. In the case of small $\langle N_{\mathrm{FL},i}\rangle$, corresponding to the Boltzmann statistics, the second term in the denominator prevails, and this equation transforms into Eq. (5).

Equation (10) describes an equilibrium distribution in an infinite system. Therefore, it is applicable to the SCTS distributions under two conditions: first, the SCTS distributions should be of quasiequilibrium character; and, second, the front of the distribution should not be close to the upper bound of the energy spectrum. Fortunately, both these conditions hold well for a considerable range of elapsed times.

Figure 6 compares one of the SCTS distributions ($T=0.35$, $t=8\times10^4$ time steps) with Eq. (10). To be consistent with the latter, the distribution does not include the quenches leading to the minima which were previously sampled. The distribution was obtained by averaging over 500 runs started in randomly chosen basins; the choice of the basins did not affect the results noticeably. Z_i in Eq. (10) are calculated in the harmonic approximation [Eq. (6)], as previously in Eq. (5) (Fig. 5), and μ is determined from the condition that $\langle N_{\mathrm{tot}}\rangle$ is equal to the mean total number of quenches included into the distribution.

The front of the SCTS distribution proceeds with time, as Fig. 7 shows. To avoid an overlap of the data for different elapsed times, the curves are presented in a smoothed form: each point represents an average number of quenches falling into one of 25 boxes into which the energy range was uniformly divided. All curves, except for curve 6, are of quasiequilibrium pattern, i.e., the corresponding distributions are

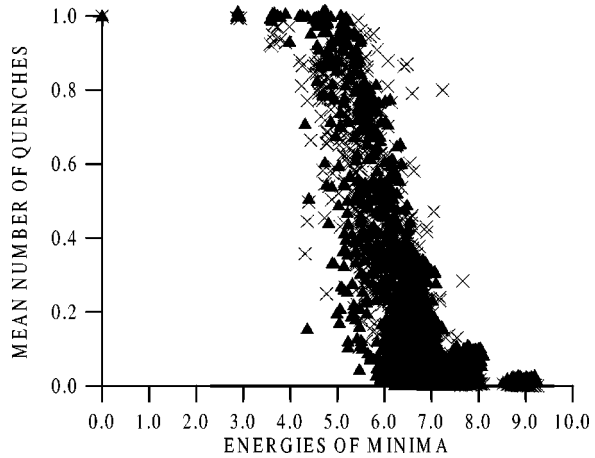

FIG. 6. Distribution of the mean number of quenches over the minima in the confinement simulations (protocol I), $T=0.35$, and $t=8\times10^4$ time steps. Triangles correspond to the simulation data, and crosses to Eq. (10) with Z_i estimated in the harmonic approximation [Eq. (6)].

described by Eq. (10) with approximately the same accuracy as in Fig. 6. Curve 6 clearly demonstrates the effect of finite number of isomers.

Equation (10) is formally applicable to protocol I. However, at large elapsed times, when the time required to execute k_{att} attempts to leave a current basin becomes small in comparison with the elapsed time, it can be applied to protocol II as well.

If τ_{quench} is not small relatively to the lifetimes of the system in the basins, any two successive quenches rarely lead to the same minimum. Therefore, the sum of $\langle N_i(t)\rangle$ over minima ($\langle N(t)\rangle=\Sigma_i\langle N_i(t)\rangle$) gives an estimate for the number of minima that will be found to time t in an individual run. Figure 8 compares $\langle N(t)\rangle$ calculated from the SCTS distributions of Fig. 7 with the data of Fig. 2 (protocol I). Only a part of the curve from Fig. 2 is reproduced, corresponding to the times for which reasonably good statistics were gained for the SCTS distributions.

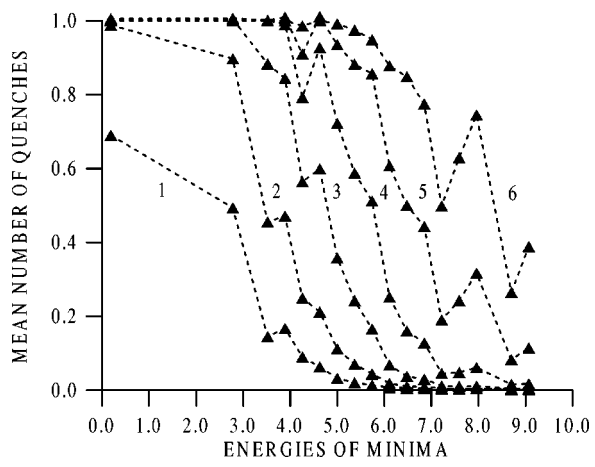

FIG. 7. Evolution of the SCTS distributions (protocol I) with time, $T=0.35$. Triangles indicate the simulation data, and the dashed lines are to guard the eye. Labels 1, 2, 3, 4, 5, and 6 correspond to the elapsed times 1.25×10^3, 5×10^3, 2×10^4, 8×10^4, 3.2×10^5, and 1.28×10^6 time steps, respectively.

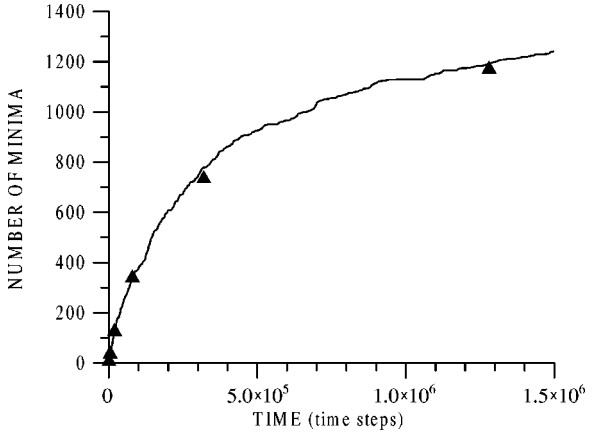

FIG. 8. Number of minima located with time. The curve corresponds to an individual run (Fig. 2, protocol I), and the triangles to the mean total numbers of minima for different elapsed times (Fig. 7).

Using Eq. (10), we can also calculate $Z_i(T)$ from $\langle N_{\mathrm{FL,i}}(t)\rangle$:

$$Z_i(T)=[\langle N_{\mathrm{FL,i}}\rangle/(1-\langle N_{\mathrm{FL,i}}\rangle)]e^{(U_{\mathrm{min,i}}-\mu)/k_BT}. \qquad (11)$$

The accuracy of the calculation of Z_i for different i is strongly dependent upon the elapsed time: For small t, such as for curve 1 in Fig. 7, a good accuracy is achievable only for low-lying isomers, simply because the other isomers have not yet been sampled. In contrast, for reasonably large t, as for curves 3 and 4, most of the isomers have been sampled but the low-lying isomers are occupied irrespective of their Boltzmann weight; therefore, a good accuracy can be achieved only for high-lying isomers. Figure 9 compares the values of $Z_i(T)e^{-U_{\mathrm{min,i}}/k_BT}$ calculated from the SCTS distri-

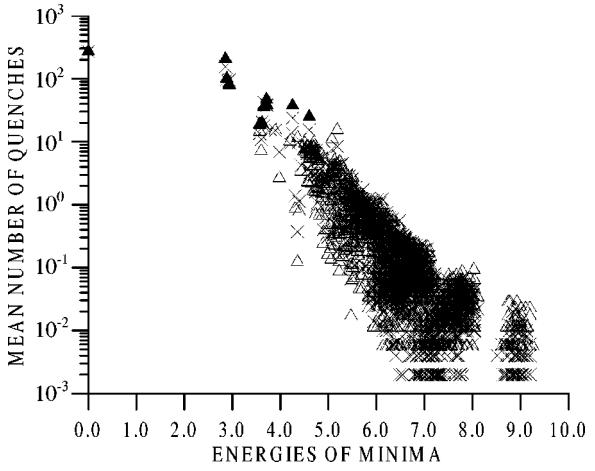

FIG. 9. Distribution of the mean number of quenches over the minima. Crosses present the SWDS distribution from Fig. 5, and solid and empty triangles indicate the numbers of quenches estimated through Eq. (11) from the SCTS distributions (they correspond, respectively, to curves 1 and 3 of Fig. 7). For more details, see the text.

butions corresponding to curves 1 and 3 of Fig. 7 ($t = 1.25\times10^3$ and 2×10^4 time steps, respectively), with the SWDS distribution of Fig. 5. The data are normalized to fit the SWDS distribution. For $t = 1.25\times10^3$ time steps we picked up the points for which $\langle N_i\rangle \geqslant 20$, and for 2×10^4 time steps those at $\langle N_i\rangle < 20$.

Finally note that such a simple picture of surveying a PES, expressed in terms of propagating Fermi-like distribution, is probably characteristic only of systems whose PES's consist of a single funnel, as for a LJ_{13} cluster and some other clusters (e.g., LJ_{19} and LJ_{55} clusters [35]). For multifunnel PES's this picture will evidently be destroyed because of the overlap of the energy spectra in different funnels. However, the employed strategy, by itself, turned out to be quite efficient for surveying multifunnel PES's [42].

V. CONCLUSION

We have shown that a taboo search carried out by successive confinement of the MD trajectory of the system to the basins on the PES radically accelerates the process of surveying the PES. Due to this acceleration, we were able to perform an exhaustive survey of the PES for a 13-atom Lennard-Jones cluster.

We also found that the probability for finding a new minimum is defined by a Fermi-like distribution, with the front of the distribution propagating with time over energies of the minima. The appearance of such a distribution in classical simulations may seem somewhat unusual. However, in fact, there is no a great surprise in this, since the quantum character of the Fermi statistics in its conventional applications is due to the Pauli exclusion principle and not to the statistics itself.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, Grant No. 99-03-33299, and by the Department of Education of Russian Federation, the Program ''Basic Problems in the Natural Sciences.''

[1] F. H. Stillinger and T. A. Weber, Phys. Rev. A **25**, 978 (1982).
[2] The references given below do not pretend to cover numerous publications in this field. We cite only a small part of them—either an original paper, where the underlying concept is described, and/or one of the recent papers, which shows the current state of the art and contains the corresponding references.
[3] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[4] M. Mezei, J. Comput. Phys. **68**, 237 (1987).
[5] C. Bartels and M. Karplus, J. Phys. Chem. B **102**, 865 (1998).
[6] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[7] S. Ono, N. Nakajima, J. Higo, and H. Nakamura, Chem. Phys. Lett. **312**, 247 (1999).
[8] M. R. Lemes, C. R. Zacharias, and A. Dal Pino, Phys. Rev. B **56**, 9279 (1997).
[9] U. H. E. Hansmann, F. Eisenmenger, and Y. Okamoto, Chem. Phys. Lett. **297**, 347 (1998).
[10] A. F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[11] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[12] P. Labastie and R. L. Whetten, Phys. Rev. Lett. **65**, 1567 (1990).
[13] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[14] A. P. Lyubartsev, A. A. Martinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[15] F. Calvo, J. P. Neirotti, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **112**, 10350 (2000).
[16] C. J. Tsai and K. D. Jordan, J. Phys. Chem. **97**, 11227 (1993).
[17] G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).
[18] N. Mousseau and G. T. Barkema, Phys. Rev. E **57**, 2419 (1998).
[19] J. P. K. Doye and D. J. Wales, Z. Phys. D: At., Mol. Clusters **40**, 194 (1997).
[20] F. Glover, ORSA J. Comput. **1**, 190 (1989); **2**, 4 (1990).
[21] R. Battiti R., in *Modern Heuristic Search Methods*, edited by V. J. Rayward-Smith, I. H. Osman, C. R. Reeves, and G. D. Smith (Wiley, New York, 1996), p. 61.
[22] D. Cvijović and J. Klinowski, Science **267**, 664 (1995).
[23] D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[24] R. S. Berry, Chem. Rev. **93**, 2379 (1993); J. Phys. Chem. **98**, 6910 (1994).
[25] C. M. Dobson, A. Šali, and M. Karplus, Angew. Chem. Int. Ed. Engl. **37**, 868 (1998).
[26] J. N. Onichic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 545 (1997).
[27] R. S. Berry and R. Breitengraser-Kunz, Phys. Rev. Lett. **74**, 3951 (1995).
[28] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[29] S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).
[30] I. V. Ionova and E. A. Carter, J. Chem. Phys. **98**, 6377 (1993).
[31] A. Ulitsky and D. Shalloway, J. Chem. Phys. **106**, 10099 (1997).
[32] F. Jensen, J. Chem. Phys. **102**, 6706 (1995).
[33] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).
[34] D. J. Wales, J. Chem. Phys. **101**, 3750 (1994).
[35] J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **111**, 8417 (1999).
[36] K. D. Ball and R. S. Berry, J. Chem. Phys. **111**, 2060 (1999).
[37] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. J. Teller, J. Chem. Phys. **21**, 1087 (1953).
[38] S. F. Chekmarev and S. V. Krivov, Chem. Phys. Lett. **287**, 719 (1998).
[39] S. F. Chekmarev and S. V. Krivov, Phys. Rev. E **57**, 2445 (1998).
[40] S. F. Chekmarev and S. V. Krivov, Eur. Phys. J. D **9**, 201 (1999).
[41] S. V. Krivov and S. F. Chekmarev, Eur. Phys. J. D **9**, 205 (1999).

[42] S. V. Krivov, S. F. Chekmarev, and M. Karplus (unpublished).
[43] M. R. Hoare and J. A. McInnes, Adv. Phys. **32**, 791 (1983).
[44] R. Biswas and D. R. Hamann, Phys. Rev. B **34**, 895 (1986).
[45] S. F. Chekmarev, in *Atomic Clusters and Nanoparticles*, Les Houches Summer School, Session No. LXXIII (Springer-Verlag and EDP Sciences, Les Ulis, in press).
[46] G. W. Greenwood, Z. Phys. Chem. (Munich) **211**, 105 (1999).
[47] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1965).
[48] S. F. Chekmarev and I. H. Umirzakov, Z. Phys. D: At., Mol. Clusters **26**, 373 (1993).
[49] A. Isihara, *Statistical Physics* (Academic, New York, 1971).