PHYSICAL REVIEW E **66**, 025701(R) (2002)

Hierarchical global optimization of quasiseparable systems: Application to Lennard-Jones clusters

Sergei V. Krivov*
Novosibirsk State University, 630090 Novosibirsk, Russia
and Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, 67000 Strasbourg, France
(Received 30 May 2002; published 27 August 2002)

An unbiased hierarchical global optimization (GO) method for quasiseparable systems is presented. In such systems the coordinates of a set of nearby local minima approximately describe the coordinates of a much larger set of surrounding local minima. This allows one to reduce the original GO problem to a much simpler GO subproblem that uses the coordinates of the local minima to reduce the search space and simplify its landscape. The algorithm showed excellent performance in tests on ''difficult cases'' of Lennard-Jones (LJ) clusters. Putative global minima of LJ_{500} and LJ_{1000} are obtained.

DOI: 10.1103/PhysRevE.66.025701 PACS number(s): 02.60.Pn, 36.40.Mr, 82.20.Kh

Global optimization (GO) is one of todays rapidly growing fields of science with many important practical applications. In a general statement of the problem, the global optimization of an arbitrary function requires a search through the whole of configurational space. The problem is nondeterministic polynomial-time (NP)-hard due to the fact that the space grows exponentially with the problem size. However, in many interesting problems, the functions to be optimized are not arbitrary and possess some characteristic properties. This allows one to devise effective heuristic algorithms to find global minima. However, the problem of proving that the determined minimum is a global one usually still remains NP hard.

For brevity we will consider the problem of an unbiased GO of a potential energy surface (PES) $U(\boldsymbol{x})$ of a many-body system (e.g., cluster), where $\boldsymbol{x}$ represents the system coordinates. A number of general methods have addressed this problem. Simulated annealing [1], the first announced generally applicable unbiased optimization technique, was not very successful at the GO of these systems if the PES was rugged. Genetic algorithms [2] (GA) were the first methods that reproduced, in an unbiased search, all the known global minima of moderate Lennard-Jones (LJ) clusters as well as finding new global minima. According to the literature, the ''basin-hopping'' method [3] (which is similar to the Monte Carlo plus minimization method [4]) is one of the fastest unbiased global optimization methods for moderate sized systems. It was applied to various cluster systems (LJ, Morse, water TIP4P, etc.) with results being entered into the Cambridge Cluster Database (CCD) [5]. A review of other GO methods can be found in [6].

GO by these methods becomes very time consuming for larger systems because of the huge search space, e.g., global optimization of LJ_{98} with a variant of basin hopping takes on average, 30 CPU hours on a Sun Ultra II 333 MHz [7]. Here we will exploit a property of some large systems (quasiseparability) that can significantly improve the efficiency of GO methods. As the system size increases, distant parts become almost independent of each other; i.e., the interaction of the atoms in one subsystem with atoms in another subsystem is weak compared to the interactions inside the subsystem: it is strong enough to make a notable contribution to the total energy, but it is weak with respect to the change of the configuration inside the subsystem, which is stabilized via much stronger local interaction.

For such systems, the coordinates of a set of nearby local minima approximately describe the coordinates of a much larger set of surrounding local minima, and allow one to reduce the original GO problem over that region to a new GO subproblem (hierarchical step), which operates directly with the local minima. This significantly simplifies the original problem by reducing its search space and simplifying its landscape (Fig. 1). Such hierarchical reduction of the problem is continued until the last problem can easily be solved by a simple GO algorithm.

Consider a local minimum structure of the system. If we randomly perturb a small part of the structure and apply local minimization, we can find a new local minimum nearby. Since distant parts of the system are independent, the new structure will, most likely, differ significantly from the original in only a small (perturbed) part, with the rest of the structure being almost the same. We call such systems quasiseparable. More rigorously, let $\boldsymbol{x}=\{x_1,\ldots,x_n\}$ and $\boldsymbol{x}'=\{x_1',\ldots,x_n'\}$ be the coordinates of two nearby local minima on the PES. Given a threshold ϵ, we call the coordinate x_i' (or x_i) *changed* if $\|x_i-x_i'\|>\epsilon$ [8]. The system is quasiseparable if, for almost every pair of nearby local minima, the number of *changed* coordinates is much smaller

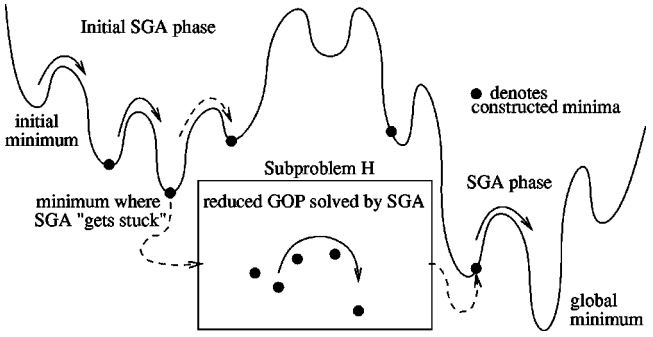

FIG. 1. Sketch of the HGA on a model PES. The search space of subproblem *H* is much simpler than the original problem.

*Electronic address: krivov@quentin.u-strasbg.fr

than the total number of degrees of freedom of the system.

Note that quasiseparability depends on the region of the PES considered. In high energy regions the system is usually very unstable with respect to minimization, and the perturbation of a small part of the structure could lead to the reorganization of a considerable part of it.

Consider the local minimum $\boldsymbol{x}^0=\{x_1^0,\ldots,x_n^0\}$ together with m surrounding nearby local minima $\boldsymbol{x}^1,\ldots,\boldsymbol{x}^m$. We call this the basis set of local minima. For every minimum $(\boldsymbol{x}^j)$ in the basis set we store the set of *changed* coordinates with respect to $\boldsymbol{x}^0$ as $\delta x^j=(\{x_{i_1}^j,\ldots,x_{i_k}^j\},\{i_1,\ldots,i_k\})$, where $\{x_{i_1}^j,\ldots,x_{i_k}^j\}$ and $\{i_1,\ldots,i_k\}$ are the values and the indices of the *changed* coordinates, respectively. Let $T_j\boldsymbol{x}=\boldsymbol{x}'$ denote the following transformation (substitution): if $i\in\{i_1,\ldots,i_k\}$ of δx^j, then $x_i'=x_i^j$, else $x_i'=x_i$, i.e., $T_j\boldsymbol{x}^0$ is equal (with accuracy ϵ) to the $\boldsymbol{x}^j$ local minimum. If we further apply such a transformation T_k to $T_j\boldsymbol{x}^0$, which acts solely on nonchanged coordinates in $T_j\boldsymbol{x}^0$, we find the approximate position of a new local minimum $T_kT_j\boldsymbol{x}^0$, because the local surroundings of each coordinate are the same as those of the coordinate in its respective local minimum of the basis set and the distant interactions are weak. Consequently, applying successive transformations one obtains the following set of local minima: $T_{i_1}\ldots T_{i_s}\boldsymbol{x}^0$, with $s\leqslant m$. To simplify things, we may apply the entire set of transformations, though some of them would conflict with others (if they share the coordinates). The total number of local minima could be estimated as 2^m (which demonstrates the well-known fact that in clusters, the number of local minima grows exponentially with the number of atoms [9]). We call this the constructed set of local minima.

The above construction approximately describes the coordinates of an exponentially large number of local minima in a region of the PES in compact form. It can also incorporate the coordinates of saddle points and be used for the description of landscapes of important regions of the PES of large systems, where the simple description as a database of local minima and saddle points is impossible due to their huge number.

Using this construction, we can significantly improve the general purpose, simple greedy algorithm (SGA), which, when implemented in local minima terms, is as follows (see Fig. 1). Applying local minimization (quenching) to a random configuration, one finds an initial local minimum. The result is improved by randomly exploring nearby local minima. These minima can be found, for example, by confinement simulation [10], i.e., a random walk in configurational space with subsequent quenching. If no lower minima are found during the specified number of quenches n_{try} (SGA ''gets stuck''), the SGA is terminated. The algorithm is suitable for systems with simple (''one funnel'') PESs, but could be applied to systems with more complex landscapes if the described procedure is repeated a number of times (multistart) with the hope of finding a funnel that leads to a global minimum (right one in Fig. 1). For the PES in Fig. 1, nearly half of all starts will find the global minimum.

If the number of funnels is large, the SGA with a multistart procedure becomes very inefficient, since only a tiny fraction of attempts will succeed; e.g., for the GO of LJ_{98} only six of 1000 attempts were successful [11] (the variant of basin hopping described in this work is SGA). The fraction of successful attempts can be improved by broadening the search region when looking for lower minima. This can be achieved by considering not only the neighboring minima, but also the second neighbor minima, the third neighbor minima, and so on. Though this process seems rather elaborate for an arbitrary system, it can be done effectively for the quasi separable systems, where the nearby minima approximately describe the minima of the constructed set. We state the GO subproblem (call it subproblem H): find the lowest minimum in the constructed set, i.e., find $i_1\ldots i_s$ with the lowest $U(i_1\cdots i_s)$, where $U(i_1\cdots i_s)$ denotes the energy of the minimum corresponding to $T_{i_1}\cdots T_{i_s}\boldsymbol{x}^0$. Subproblem H is significantly simpler than the original problem because (see Fig. 1); first, we significantly narrow the search space (now we operate only on coordinates of local minima instead of the full space) and second, we simplify the landscape (instead of a basin with barriers we have just one point). If subproblem H is still complex we can, in turn, reapply the same idea, consequently obtaining a hierarchy of GO subproblems with decreasing complexity, until the last subproblem can effectively be solved by SGA. We cannot rigorously call this a recursion algorithm, since not every problem in this hierarchy can be solved with the same implementation of SGA; rather, we call it a *hierarchical greedy algorithm* (HGA).

The local minimum energy $U(i_1\ldots i_s)$ can be found by quenching, starting from $T_{i_1}\ldots T_{i_s}\boldsymbol{x}^0$, or estimated as $U(T_{i_1}\ldots T_{i_s}\boldsymbol{x}^0)$, or in a more complex manner.

Let us see how HGA is related to other unbiased GO techniques. Like GA, HGA uses parts of different structures to construct new structures. However, unlike GA, HGA requires these structures (parents in GA terms) to be very similar (number of changed coordinates should be much less then the total number of degrees of freedom). This seems more in line with natural evolution, where not every species crosses every other. HGA can be considered as a limiting case of GA (if we accept the requirements of very similar parents), where instead of searching for a very similar configuration, it is easier to perturb the current one. Next, in GA the ''child'' is often produced stochastically, whereas in HGA it is a separate GO subproblem that needs to be solved. This peculiarity of HGA resembles the genetic renormalization algorithm (GRA) [12], where the problem of selecting the optimal ''child'' is reformulated as the original problem with lower size (renormalization step). However, the GRA could be applied only to renormalizable systems (it was applied to the traveling salesmen and spin glass problems [12]), where the problem of selecting the optimal child is again the same problem of lower size. HGA uses another hierarchy to construct subproblems, namely, the hierarchy of local minima, i.e., each subproblem optimizes the set of local minima found by the preceding problem, consequently reducing the search space. We require the subproblem H to be just a GO problem that is simpler than the original problem, and that can be solved by any general purpose GO algorithm. Another

difference between the algorithms is that GRA, like GA, does not pay attention to the similarities between the structures and works with the entire configurational space. This seems inefficient when the PES of the systems has different deep minima with different structures (e.g., in fcc and icosahedral packing schemes in LJ clusters), because it is unlikely that the lowest minimum consists of parts of different structures (i.e., the ''genetic code'' that one deep minimum possesses is useless for other distant deep minima). In comparison with the basin-hopping method [3], HGA finds the lowest constructed minima (if subproblem H has been solved) in spite of the actual landscape of the PES in that region; the former techniques are sensitive to the landscape of the deformed PES, since they rely on the ergodicity of Monte Carlo sampling of the landscape.

Let us apply the HGA to LJ clusters, which are widely accepted as a benchmark for GO algorithms. In dimensionless units, the problem is to find a configuration of n atoms $(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$, with the lowest value of the following potential energy:

$$U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n) = 4 \sum_{i<j} r_{ij}^{-12} - r_{ij}^{-6}, \quad r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|.$$

The set of nearby local minima $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^m$ found during the SGA phase are superposed with the lowest one $(\boldsymbol{x}^1)$ in the set. From atom positions $\{\mathbf{x}_1^1, \ldots, \mathbf{x}_n^1, \mathbf{x}_1^2, \ldots, \mathbf{x}_n^m\}$ we form a pool of distinguished atoms (SGA pool), whose distance from each other is more then $\epsilon = 0.3$ (with total number N). Subproblem H is to select from the pool of N atoms a set of n atoms with the lowest potential energy (transformations T_i move atoms between different positions). The total number of constructed minima could be estimated as C_N^n. Subproblem H is solved again by SGA in local minima terms $(n_{try} \sim 100)$; the energies of the constructed local minima are estimated. The output is refined by the minimization of a number of lowest minima, giving us the true global minimum of the constructed set with high probability. For the considered range of atoms (0–1000), SGA was able to solve subproblem H. However, in the limit of large N, it was not very efficient, and to solve the subproblem H in more difficult cases, HGA should have more hierarchical steps.

To find a new nearby minimum in the SGA phase, we randomly shifted an atom and applied local minimization (using the limited memory variant of the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) algorithm [13]).

During the SGA phase the algorithm collects the obtained local minima (forms the SGA pool) to build the set of transformations (local moves) to walk over the set of constructed minima. One can also use available *a priori* information about the system to expand the set of transformations and the set of constructed minima. For example, it is well known that the low energy structures of the homogeneous clusters possess some spatial symmetries, though not always complete (not every atom has a corresponding symmetrical one). Consider various symmetrically transformed copies (STC) of the structure. If the number of changed coordinates between STC and the original structure is small, we can consider the former as a new nearby local minimum in a proper coordinate basis. From all such STC we form a pool (STC pool) of atoms and proceed similarly to described the above.

TABLE I. Nine most frequently found minima. t and n_q denote the total computational time and the number of quenches in the run, respectively. U and n_f denote the energy of the minimum (in LJ units) and the number of times it was found.

$LJ_{75}, t = 450$ min, $n_q = 2.5 \times 10^5$		$LJ_{77}, t = 450$ min, $n_q = 2.5 \times 10^5$		$LJ_{98}, t = 765$ min, $n_q = 3 \times 10^5$	
U	n_f	U	n_f	U	n_f
−397.492331	27	−409.083517	19	−543.665361	53
−396.282249	62	−408.518265	42	−543.642957	63
−396.238512	166	−408.499443	73	−541.894959	8
−395.734052	39	−408.486640	23	−541.869748	11
−394.637268	9	−408.463320	25	−541.436260	22
−394.560154	18	−408.311778	13	−540.362599	7
−393.435288	4	−408.303313	25	−539.720452	110
−393.435065	4	−408.203358	51	−539.698759	62
−392.531982	3	−407.756902	15	−538.883972	11

The overall structure of one attempt of the HGA is (1) generate an initial random local minimum (2) SGA phase $(n_{try} = 100)$; (3) solve subproblem H with the SGA pool of atoms, if a lower minimum is found, go to (2); (4) solve subproblem H with the STC pool of atoms, if a lower minimum is found, go to (2).

The algorithm was tested on a number of LJ clusters. For all clusters HGA found the known lowest minima given at the CCD [5]. To illustrate how the algorithm works, we choose the clusters to be difficult cases for unbiased global optimization [11]: LJ_{75}, LJ_{77}, and LJ_{98}. We issued 500 attempts for each of the clusters. The nine most frequently found minima of the clusters are given in Table I. HGA found all known putative global minima. In the table other ''global'' (lowest) minima of structures with different morphologies (e.g., lowest icosahedral structure of LJ_{75} with energy −396.282 249) are also shown. This confirms that HGA is unbiased. Since subproblem H does not always give the true global minimum of the constructed set, not every minimum in the table represents a structure with different morphology. For example, the following pairs of LJ_{77} minima, −408.518 265 and −408.303 313, −408.499 443 and −408.203 358, differ by only two atoms. Comparing the mean time to find the global minimum for the LJ_{98}, $t \simeq 14.4$ min (PII 333 MHz), with the value reported in [7], $t \simeq 30$ h (Sun Ultra II 333 MHz), we found that HGA is faster than a variant of SGA. Another manifestation of a more efficient optimization of HGA is the fraction of successful attempts ~10.6%, compared to SGA's ~0.6% [7].

Figure 2, illustrates the performance of the HGA (we choose an attempt when the global minimum of LJ_{98} was found). It shows the number of changed atoms of the new local minimum compared to the current one (solid line), and its potential energy (dotted line) versus the current quenching. At the beginning, SGA progresses quickly. Near the 350th quench it slows down and at the 440th quench, SGA gets stuck at energy −537.65. At the 540th quench (n_{try}

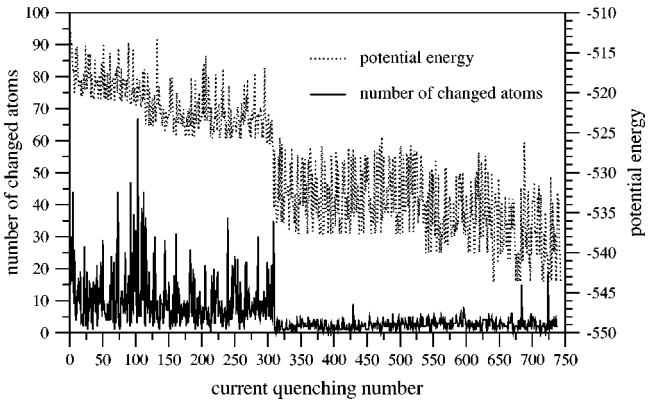

FIG. 2. Properties of minima during the HGA run on LJ_{98}.

$=100$), subproblem H was solved and a new minimum with energy -540.065 was found. At the 640th quench, the algorithm solves the subproblem H with the STC pool and finds the global minimum with energy -543.665. Beyond this point, the algorithm hopelessly tries to improve the obtained global minimum.

The solid line illustrates, in accordance with above, that in the early stages of the minimization process (the system is in a high lying region of the PES), the number of changed coordinates tends to be large. When the system is sufficiently minimized (310th quench), this number becomes small for almost all minima and the system becomes quasiseparable; HGA can then be applied.

To illustrate the possibilities of HGA, we apply it without modification to a global optimization of LJ_{500} and LJ_{1000}.

TABLE II. Three lowest minima. See Table I for notations.

$LJ_{500}, t=40\ h, n_q=3.7\times10^5$		$LJ_{1000}, t=150\ h, n_q=10^6$	
U	n_f	U	n_f
-3382.693487	6	-7128.821828	2
-3379.194148	2	-7127.127060	3
-3378.108229	1	-7126.146796	1

The three lowest minima found are given in Table II. We did not find any references in the literature to compare these minima, but since the lowest energy structures were found more then once, we can conclude that they are good candidates for the global minima. Due to the high pressure of external layers [14] the lowest energy structure of the LJ_{1000} has an icosahedral core without a central atom. The structure with a complete icosahedral core is the second lowest in Table II.

We applied HGA to some LJ clusters within the 149–309 atoms range. In all cases we found the global minima reported in Ref. [15]. For LJ_{186} we found a lower minimum with energy -1132.67.

In summary, we have exploited the issue of quasiseparability of large systems and proposed a new unbiased hierarchical GO algorithm (HGA). HGA showed excellent performance in the GO tests on LJ clusters. We hope that this method will allow one to improve the GO technique for a family of quasiseparable systems.

I wish to thank M. Karplus, M. Meuwly, and R. Stote for helpful discussions.

[1] S. Kirkpatrick, J.C.D. Gellat, and M. Vecchi, Science **220**, 671 (1983).
[2] D.M. Deaven and K.M. Ho, Phys. Rev. Lett. **75**, 288 (1995).
[3] D.J. Wales and J.P.K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
[4] Z. Li and H.A. Sheraga, J. Mol. Struct.: THEOCHEM **179**, 333 (1988).
[5] D. J. Wales, J. P. K. Doye, A. Dullweber, and F. Y. Naumkin, *The Cambridge Cluster Database*; URL http://brian.ch.cam.ac.uk
[6] D.J. Wales and H.A. Scheraga, Science **285**, 1368 (1999).
[7] R.H. Leary and J.P.K. Doye, Phys. Rev. E **60**, R6320 (1999).
[8] If the system possesses symmetry that does not change the potential energy (e.g., for clusters, translational, rotational, and permutational), then structures of these local minima should be superposed to minimize the number of *changed* coordinates.
[9] F.H. Stillinger and T.A. Weber, Science **225**, 983 (1984).
[10] S.V. Krivov, S.F. Chekmarev, and M. Karplus, Phys. Rev. Lett. **88**, 038101 (2002).
[11] R.H. Leary, J. Global Optim. **18**, 367 (2000).
[12] J. Houdayer and O.C. Martin, Phys. Rev. Lett. **83**, 1030 (1999).
[13] D. Liu and J. Nocedal, Math. Program. **B45**, 503 (1989).
[14] L.L. Boyer and J.Q. Broughton, Phys. Rev. B **42**, 11 461 (1990).
[15] D. Romero, C. Barron, and S. Gomez, Comput. Phys. Commun. **123**, 87 (1999).