Journal of Global Optimization **18:** 367–383, 2000.

Global Optimization on Funneling Landscapes ⋆

ROBERT H. LEARY
San Diego Supercomputer Center, University of California, San Diego, P.O. Box 85608, San Diego, CA 92186-9784, USA (e-mail: leary@sdsc.edu)

Abstract. Molecular conformation problems arising in computational chemistry require the global minimization of a non-convex potential energy function representing the interactions of, for example, the component atoms in a molecular system. Typically the number of local minima on the potential energy surface grows exponentially with system size, and often becomes enormous even for relatively modestly sized systems. Thus the simple multistart strategy of randomly sampling local minima becomes impractical. However, for many molecular conformation potential energy surfaces the local minima can be organized by a simple adjacency relation into a single or at most a small number of funnels. A distinguished local minimum lies at the bottom of each funnel and a monotonically descending sequence of adjacent local minima connects every local minimum in the funnel with the funnel bottom. Thus the global minimum can be found among the comparatively small number of funnel bottoms, and a multistart strategy based on sampling funnel bottoms becomes viable. In this paper we present such an algorithm of the basin-hopping type and apply it to the Lennard–Jones cluster problem, an intensely studied molecular conformation problem which has become a benchmark for global optimization algorithms. Results of numerical experiments are presented which confirm both the multifunneling character of the Lennard–Jones potential surface as well as the efficiency of the algorithm. The algorithm has found all of the current putative global minima in the literature up to 110 atoms, as well as discovered a new global minimum for the 98-atom cluster of a novel geometrical class.

Key words: Global optimization, Lennard–Jones clusters, Basin-hopping, Energy landscape, Folding funnel, Molecular conformation

1. Introduction

1.1. SCOPE AND OUTLINE OF PRESENTATION

This paper describes the application of a new variant of the highly successful basin-hopping algorithm of (Wales and Doye, 1997) to the Lennard–Jones (LJ) microcluster problem, an important and intensively studied global optimization problem. Section 1.2 provides a basic description and historical background of the LJ problem, including a summary of the current state of knowledge concerning global minima. Section 1.3 introduces some concepts related to the landscape of the LJ potential energy function, particularly the organization of local minima into ‘monotonic sequence basins’ and ‘funnels’ on a graph structure induced by an

⋆ The research of this author was supported in part by National Science Foundation Cooperative Agreement ACI-9619020

adjacency relation relation between local minima. Section 2.1 is an overview of the original basin-hopping algorithm as a Metropolis Monte-Carlo search (similar to simulated annealing, but with a constant temperature parameter) combined with a local minimization at each step. The new basin-hopping variant is introduced in Section 2.2 by setting the temperature parameter to zero. This in essence changes the Metropolis acceptance procedure of occasionally accepting uphill steps to a rule that only accepts downhill moves, thus producing a monotonic sequence of local minima which terminates in a funnel bottom. After sequence termination, a new sequence is initialized from a fresh random starting local minimum. Thus the new variant can be viewed as a multistart algorithm that samples funnel bottoms. In contrast, the original algorithm produces a single long sequence which samples many local minima and relies on the Metropolis acceptance criterion to escape from funnel bottoms.

Sections 2.3 and 2.4 describe a numerical implementation of the new variant and some extensive computational results, respectively. Comparison with the original algorithm reveals quite distinctive behaviors, particularly on the more difficult cases where the new variant has a marked advantage. Surprisingly, in light of the considerable analysis the problem has received from chemists as well as the extensive numerical searches conducted previously with many different types of algorithms, the new variant has found a new global minimum of a novel geometrical class for the 98-atom LJ cluster.

The paper concludes in Section 3 with a summary of the performance and feature characteristics of the new algorithm in comparison with the original. Different types of energy landscapes where each might be preferable are identified.

1.2. THE LJ MICROCLUSTER PROBLEM

Molecular conformation problems are global optimization problems in computational chemistry where the objective function is an energy function modeling the interactions between atoms or other small sub-units of a molecule. This function can represent either a potential energy arising from the forces between atoms, or a thermodynamic ‘free energy’, which also includes temperature and entropy dependencies. Here we will be primarily concerned with the simpler potential energy case, which is more amenable to analytical models. We note that at zero temperature the distinction between potential and free energy vanishes.

The global minimum of an appropriate energy function is of interest because it is believed to determine the *native conformation* or most probable configuration of the molecule. Other conformations corresponding to local minima can be viewed as lower probability alternatives in an equilibrium among various conformations, or sometimes as intermediates in a dynamical process which proceeds from a high energy starting conformation to the global minimum. This latter viewpoint is often taken for the *protein folding problem*, a molecular conformation problem of great scientific importance that is concerned with finding the 3-dimensional structure of

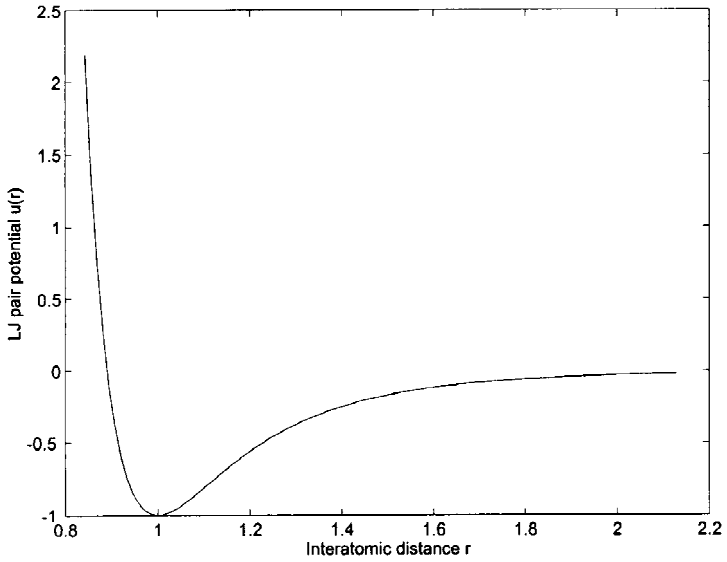

Figure 1. The Lennard–Jones pair potential functions.

a protein given a one-dimensional sequence of component amino acids. A comprehensive review oriented toward applied mathematicians can be found in (Neumaier, 1997).

A remarkably simple molecular conformation model that has attracted considerable scientific, mathematical, and computational research is the *LJ cluster problem*. LJ clusters consist of identical atoms that interact pairwise via the LJ pair potential defined for two atoms i and j by

$$u(r_{ij}) = r_{ij}^{-12} - 2r_{ij}^{-6} \tag{1}$$

where r_{ij} is the Euclidean distance $\|x_i - x_j\|$ between the atomic coordinate vectors $x_i,\ x_j \in \Re^3$. As seen in Figure 1, the pair potential $u(r)$ is unimodal and has been scaled so that the minimum occurs at $r = 1$ with corresponding potential energy $u(1) = -1$. For $N > 2$ atoms, pairwise additivity is assumed so the cluster potential energy is given by

$$E(X) = \sum_{j>i} u(r_{ij}) \tag{2}$$

where $X = (x_1, \ldots, x_N)$ is the $3N$-dimensional vector of atomic coordinates.

The LJ cluster potential energy function is believed to be a reasonably accurate model of low temperature clusters of heavy rare gas atoms such as argon or xenon, physical systems which are accessible to experimental measurement. Thus results from global optimization computations can be compared in various respects with laboratory measurements. LJ pair potential terms are also an important component in many protein folding model energy functions, and the LJ cluster problem can be viewed as a test problem for algorithms potentially applicable to protein folding.

These factors, combined with the simplicity of the formulation and computational ease of implementation of the LJ cluster model, have led to considerable interest and a substantial literature over the past thirty years from both the chemical physics and global optimization communities. Recent reviews can be found in Wales and Scheraga (1999) and Leary (1997). The LJ cluster problem has also become a computational benchmark for testing and comparison of a wide variety of global optimization methods such as simulated annealing (Wille, 1987), genetic algorithms (Niesse and Mayne, 1996; Deaven et al., 1996; Wolf and Landman, 1998), smoothing and hypersurface deformation techniques (Kostrowicki et al., 1991; Pillardy and Piela, 1995), lattice methods (Northby, 1987; Xue, 1994b), growth sequence analysis (Hoare and Pal, 1971; Leary, 1997), tunneling (Barron et al., 1996), and basin-hopping (Wales and Doye, 1997). An online list of the best current putative LJ global minima in the range $N \leqslant 150$ is maintained in the Cambridge Cluster Database (CCD) at http://brian.ch.cam.ac.uk (Wales et al., URL).

Clusters in this range are often referred to as *microclusters*. Microclusters are characterized by a large number of surface atoms relative to interior atoms; hence bulk interior effects do not dominate and microclusters exhibit strikingly different overall structures than bulk samples of the same atomic composition. For example, bulk solid argon and xenon exhibit a crystalline structure based on a face-centered cubic (fcc) lattice, whereas most LJ microcluster global minima are based on a non-crystalline icosahedral structural motif known as a Mackay icosahedron (Mackay, 1962), first systematically explored by (Northby, 1987) in the context of LJ microclusters. A few LJ microcluster global minima (LJ_{75}, LJ_{76}, LJ_{77}, and LJ_{102}, LJ_{103}, LJ_{104}) are based on another non-crystalline structure known as a Marks decahedron (Doye et al., 1995), which is derived from a pentagonal bipyramid. One notable fcc global minimum, LJ_{38}, assumes the shape of a truncated octahedron (Barron et al., 1996), and recently a new global minimum at $N = 98$ with unusual tetrahedral symmetry and a partly fcc interior has been found (Leary and Doye, 1999) with the new algorithm described here. Pictures, energies, and coordinate files for all of these structures can be found in the CCD.

1.3. LANDSCAPE OF THE LJ MICROCLUSTER PROBLEM

Although deceptively simple in its algebraic description, the LJ potential energy surface (PES) defined by Equations 1 and 2 is extremely complex in terms of the number of local minima. Current data and theory (Tsai and Jordan, 1993; Stillinger, 1999) suggest that the number of energetically distinct (e.g., excluding permutational isomers) local minima follows a simple exponential growth law, leading to estimates of more than 10^{40} local minima for $N = 100$.

Given a local minimization procedure, we can associate with each local minimum $X_i \in \Re^{3N}$ a *catchment basin* $\mathcal{C}(X_i) \subseteq \Re^{3N}$ such that the local minimization algorithm initiated at any point $X \in \mathcal{C}(X_i)$ converges to X_i. Together, the catch-

ment basins partition configuration space. The simple *multistart* global optimization strategy defines a compact subset of configuration space, such as a ball or box of suitable size centered at the origin, that contains all bound configurations of interest. This region is then sampled uniformly, and the local minimization algorithm is initiated at each sample point to arrive at the corresponding local minimum. The putative global minimum is selected as the best local minimum found among the samples.

For molecular conformation problems with exponential local minima growth curves, the multistart strategy quickly becomes overwhelmed as N increases. This is essentially the observation behind the 'Levinthal Paradox' (Neumaier, 1997) for the protein folding problem. Naturally occurring proteins with approximately 10^{300} local minima are observed to reliably move from an unfolded state to the folded native conformation in times typically less than one second. However, it is known on physical grounds that transitions between successive conformations require at least 10^{-15} seconds. Thus significant fractions of the conformation space cannot be sampled in times consistent with the observed folding rate. The 'paradox' can be resolved by the current view that protein folding proceeds not by random sampling, but rather by following any one of many possible energetically descending sequences of conformations, all of which converge upon the global minimum. Thus the process can be described in terms of a 'folding funnel' (Bryngelson et al., 1995), in which every conformation is connected to the global minimum by at least one chain of descending local minima.

It has become common to describe the PES as a landscape with topographical features associated with critical points of the potential function. Particular emphasis is placed on local minima and first order (exactly one negative Hessian eigenvalue) saddle points, also known as *transition points* by computational chemists since such points typically lie on the boundary between adjacent catchment basins on the most probable pathway for a conformational change from one local minimum to the other. Larger scale groupings of local minima into *monotonic sequence basins* and *funnels* on the landscape can be defined in terms of connectivity via monotonic sequences, as described below.

The set of transition points defines an adjacency relation between the associated pairs of local minima and hence a graph with local minima as nodes and transition points as arcs. We will call this the transition point landscape graph to distinguish it from the landscape graph induced by other possible adjacency relations between local minima, such as simple proximity in $\Re^{3N}$. A (descending) *monotonic sequence* of local minima on a landscape graph has the defining property that each successive node is adjacent to its predecessor and lower in energy. Clearly every monotonic sequence on a finite landscape graph must terminate at a distinguished botttom node b^* from which no further descent is possible. Note there may be more than one bottom node on the graph.

If there is only one bottom node, then the entire landscape graph can be regarded, as in the idealized protein folding funnel case, as a 'funnel' where every

monotonic sequence started from any local minimum must terminate at the global minimum. More general landscapes, however, may have several or even many bottom nodes. However, following monotonic sequences to bottom nodes may be a viable global optimization strategy as long as the number of bottom nodes remains manageable and the difficulty of generating the monotonic sequences is not too great. This is in fact the strategy adopted by the algorithm introduced in this paper for the LJ microcluster problem.

In the case of landscape graphs with more than one bottom node, each such node b^* defines two subgraphs, a *monotonic sequence basin* (Berry and Breitengraser-Kunz, 1995) and a *funnel*. (Note the distinction between a monotonic sequence basin, which is a subgraph of the landscape graph, and a catchment basin, which is a subset of $\Re^{3N}$.) The monotonic sequence basin is the set of local minima and arcs which lie on at least one monotonic sequence that terminates at b^*. Thus a local minimum, particularly if it is relatively high on the landscape graph, may belong to more than one monotonic sequence basin. The funnel is the subgraph of the monotonic sequence basin, usually lying near the bottom, such that *all* monotonic sequences initiated from a local minimum in the funnel terminate at b^*. Thus a local minimum may belong to at most one funnel, and it makes equal sense to refer to b^* as a funnel bottom or a monotonic sequence basin bottom – we will use the funnel bottom terminology. The distinction between funnels and monotonic sequence basins disappears when there is only one bottom on the landscape graph. However, the LJ microcluster problem displays energy landscapes with multiple funnel bottoms. Particularly for the larger LJ microclusters, a randomly selected starting local minimum usually belongs to many monotonic sequence basins, and a monotonic sequence typically takes several steps before entering a funnel.

Although Berry and Breitengraser-Kunz (1995) defined monotonic sequence basins using sequences in which adjacent members share a common transition point, the concept is equally applicable for any adjacency relation. Both variants of the basin-hopping algorithm presented in the next section use an adjacency relation implicitly defined by proximity in 3N-dimensional space, thus avoiding the computational difficulties of computing transition points in high-dimensional spaces. It should be noted that large scale organizational features of landscapes such as monotonic sequence basins and funnels depend explicitly on the adjacency relation used, while the simpler point features defined directly by critical points of $E(X)$ obviously do not.

2. The Monotonic Sequence Basin-Hopping Algorithm

2.1. BASIN-HOPPING AND THE PLATEAU TRANSFORMATION

Simulated annealing and related methods such as Monte Carlo search for the global minimization of $E(X)$ operate on as simple perturbation-conditional acceptance scheme. The current point X is conditionally succeeded by a perturbed point $Y = X + S$, where S is a random perturbation, depending on an acceptance rule (usually

the Metropolis criterion described below) based on the objective function change $E(Y) - E(X)$. In simulated annealing, the acceptance rule also depends on a 'temperature' parameter T which is lowered according to an annealing schedule as the search proceeds. In contrast, Monte Carlo search treats T as an adjustable algorithmic parameter that remains constant for any given search – i.e. no annealing is done.

Li and Scheraga (1987) proposed an algorithm called 'Monte Carlo-minimization' (MCM) which employs a Monte Carlo search in which a local minimization is performed at each step before the Metropolis acceptance rule is applied. Thus acceptance is based on the function change $E(Y_{\min}) - E(X_{\min})$ where $X_{\min}$ and $Y_{\min}$ are the local minima obtained by applying the local minimization procedure starting at X and Y, respectively. If the acceptance criterion is satisfied, the MCM algorithm resets the accepted point (and starting point for the next perturbation) to $Y_{\min}$ from Y, so a sequence of local minima is produced.

Xue applied a similar idea (without resetting) in a simulated annealing context, calling it 'two-level simulated annealing' (Xue, 1993; Xue, 1994a) to emphasize that the perturbation is done at one level and the acceptance rule applied at a second level after a local minimization has been performed. He used the procedure to search icosahedral lattices for global minima of LJ clusters.

Wales and Doye (1997) observed that use of a minimization procedure before applying the acceptance criterion is equivalent to searching a transformed objective function, where the transformation is the operation of applying the local minimization procedure to the original function at each point in its domain. They introduced the term 'basin-hopping' to describe the general class of algorithms that search the transformed 'plateau' energy function $P(X)$, where

$$P(X) = E(X_{\min}) \tag{3}$$

on the catchment basin $\mathcal{C}(X_{\min})$ surrounding each local minimum $X_{\min}$. Thus the landscape of $P(X)$ consists of a large number of interlocking plateaus of various heights, each plateau covering a particular catchment basin. 'Basin-hopping' simply refers to the fact that members of this class of algorithms jump between catchment basins.

As noted by Wales and Scheraga (1999), several algorithms that have been applied to the LJ cluster problem, including the MCM approach of Wales and Doye (1997), Xue's two-level simulated annealing (Xue, 1993; Xue, 1994a), the genetic algorithms of Deaven et al. (1996) and Wolf and Landman (1998) and the 'exponential tunneling' approach of Barron et al. (1996) can be viewed as members of the basin-hopping family. However, for the remainder of this paper, we will narrow the scope of 'basin-hopping' to refer to the MCM-based version described immediately below and the new variant described in Section 2.2 Wales and Doye (1997) very successfully applied the original MCM version to the Lennard–Jones microcluster problem up to $N = 110$, obtaining all known putative global minima,

including the Marks decahedra which had previously not been found by any global optimization algorithm.

Starting from any given local minimum, X_k, a random perturbation S is proposed and a local minimization is initiated at $(X_k + S)$ to yield the local minimum Y. If $E(Y) < E(X_k)$, this downhill step is accepted immediately and the algorithm proceeds with $X_{k+1} = Y$ as the next iterate. If $E(Y) \geqslant E(X_k)$, the Metropolis acceptance criterion is applied, i.e. the uphill step is accepted with probability

$$M = e^{(E(X_k)-E(Y))/T} \tag{4}$$

by generating a random number U uniformly in [0,1] and accepting the step if $U \leqslant M$. Here $T > 0$ is a preselected 'temperature' parameter that remains fixed (i.e., no annealing is done). If the step is rejected, new random perturbations S are generated until a step is accepted.

If the step S is limited in range to nearby catchment basins, the algorithm performs a downwardly biased random search over successive adjacent plateaus of $P(X)$. Equivalently, basin-hopping can be viewed as a downwardly biased random walk on the landscape graph of $E(X)$ defined by the proximity adjacency relation induced by the limited range of S. Much of the power of the method derives from the fact that, unlike simulated annealing or Monte Carlo methods applied to $E(X)$ such as described in Wille (1987), there are no energetic barriers for moving downhill between plateaus, and plateau crossing can easily take place at any point along the plateau boundaries. Moreover, passage between catchment basins takes place in a single step.

On the other hand, the MCM basin-hopping method can become trapped or delayed by reaching a funnel bottom which is not the global minimum and is separated from it by a high energy barrier. This indeed has been found to be the case for the decahedral global minima LJ_{75}, LJ_{76} and LJ_{77} (Doye et al., 1999b) and, less severely, the LJ_{38} fcc truncated octahedron (Doye et al., 1999a). Here the largest and most easily accessible funnels correspond to icosahedral forms, and most random walks initially descend to an icosahedral funnel bottom. From there they must ascend via a relatively low probability sequence of Metropolis steps to gain sufficient altitude to have a chance of entering the proper funnel. A much more favorable landscape occurs when the global optimum is at the bottom of a wide, dominant funnel and/or the barriers between funnels are small. Then the random walk has a good chance of entering the correct funnel on the initial descent or rapidly exploring a number of funnels until the correct one is found without significant delays to overcome large interfunnel barriers.

2.2. MONOTONIC SEQUENCE BASIN-HOPPING

An interesting variant of basin-hopping is obtained if the Metropolis acceptance criterion is abandoned in favor of only accepting downhill steps. Note this is equivalent to setting the temperature parameter $T = 0$. Then the random walk will

follow a strictly descending sequence of local minima until a funnel bottom is encountered. The original basin-hopping algorithm depends on the Metropolis acceptance criterion of occasionally accepting uphill steps to escape from funnel bottoms. However, if the funnel bottom can be recognized, as for example, by the lack of downward progress in a large number of move attempts, an alternate escape mechanism is simply to restart the sequence from a fresh initial random local minimum. This simple idea forms the basis of the *monotonic sequence basin-hopping* (MSBH) variant presented here. MSBH can be viewed as a multistart method that samples funnel bottoms rather than local minima. Even if the number of local minima on the landscape grows exponentially, the method can be quite efficient if the funnel containing the global minimum has a sufficiently large 'cross-section' for entry from a random starting local minimum.

More formally, we can state the MSBH algorithm as follows:

Step 0 (initialization): Generate a random starting local minimum X_0. Set the consecutive rejection counter R to 0.

Step 1 (iteration): Starting with current local minimum X_k, generate a random perturbation $S \in \Re^{3N}$ and let Y be the local minimum attained by initiating the local minimization procedure at $(X_k + S)$.

a) if $E(Y) < E(X_k)$, accept $X_{k+1} = Y$ as the next member of the monotonic descent sequence. Reset the consecutive rejection counter R to 0 and return to the beginning of Step 1.

b) If $E(Y) \geqslant E(X_k)$, reject Y and increment the consecutive rejection counter R by 1. If $R \leqslant R_{\max}$, maintain X_k as the current local minimum and go to the beginning of Step 1 for a new random step. If $R > R_{\max}$ go to Step 2.

Step 2 (termination): Terminate the sequence with the current X_k as the funnel bottom.

The overall algorithm simply adds as many restarts as desired, thus becoming a multistart algorithm that samples funnel bottoms.

2.3. NUMERICAL IMPLEMENTATION

A numerical implementation of MSBH has been written in Fortran 77 and implemented on a 16-processor parallel SUN 333 MHz Ultra 2 computer. Specific implementation parameters and details are as follows:

a) The local minimization procedure is the PARTAN conjugate gradient algorithm (Luenberger, 1969) using analytical gradients.

b) The sequence initialization step consists of generating a random $3N$-dimensional coordinate vector X with independent identically distributed components from a Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.25$. This distribution is spherically symmetric, with the root mean square distance between atoms given by $3^{1/2}\sigma = 0.43$, as compared with typical near neighbor distances of approximately 1.0 for most local minima. Thus the initial configuration

is somewhat compressed; it is then expanded to the starting local minimum by application of the conjugate gradient algorithm. The details are not critical and any other reasonable initialization procedure should work as well.

c) The random $3N$-dimensional perturbation vector S is generated with independent identically distributed components from a Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma = 0.21$. The choice $\mu = 0$ is obvious since there is no reason to bias the step in any direction. However, the choice of σ is critical since it determines the average step size. If σ is chosen too small, the step typically will not escape from the catchment basin of the starting local minimum and no progress will be made. If σ is chosen too large, there will be too many possible catchment basins within range of the step. Such large steps will in effect randomly sample large portions of conformation space, an unproductive strategy when the number of local minima is huge. This is particularly true when the sequence has already progressed to relatively low energy conformations near the funnel bottom, where large steps may be expected to almost always reach higher energy plateaus. The chosen value $\sigma = 0.21$ was determined by varying σ until roughly half of the steps attempted escape the starting catchment basin. Somewhat surprisingly, this value works acceptably well for all cluster sizes and seems to be uniformly effective over the entire conformation space for each cluster size. If necessary, this parameter can be adaptively adjusted as the algorithm proceeds for potentials for which this is not the case. Note that $N^{1/2}\sigma$, which is the square root of the expected value of $\|S\|^2$, can be loosely interpreted as the typical 'radius' of a catchment basin for cluster size N.

We also experimented with uniform distributions for S, as used in the original basin-hopping implementation in Wales and Doye (1997), and found little difference in performance of the algorithm as long as similar standard deviations were used.

d) The value of $R_{\max}$ was set equal to 1000. Thus a local minimum was not declared to be a funnel bottom until 1000 consecutive move attempts failed to locate a lower nearby local minimum. This value is excessive for the smaller clusters but necessarily large at the high end of the cluster size range. Generally, adjacent lower local minima are easily found early in the sequence, but are progressively more difficult to find as the funnel bottom is approached, particularly for the larger clusters. It is not uncommon for successive iterates to be separated by hundreds of move attempts for the larger clusters.

2.4. COMPUTATIONAL RESULTS

A series of numerical experiments followed 1000 randomly initialized monotonic sequences to the corresponding funnel bottoms for each cluster size in the range $13 \leqslant N \leqslant 110$. In most cases a large number of sequences terminated at the global minimum and reasonable statistics could be generated regarding the fractional global minimum hit rate. It should be noted that experiments on this scale

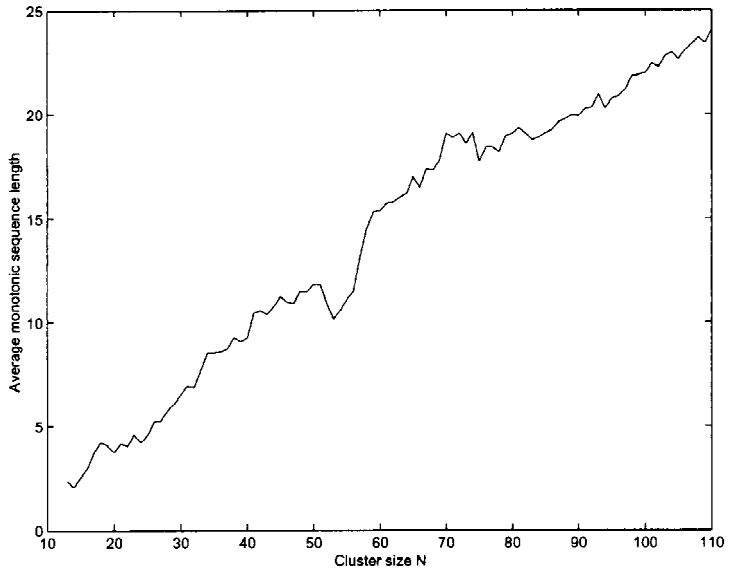

Figure 2. Average sequence length from random starting local minimum to funnel bottom as a function of cluster size.

are extremely time consuming in terms of computational resources, with the entire series requiring in excess of 5000 hours of CPU time. Thus access to a parallel computer proved invaluable.

The average path length in terms of number of conformations in the monotonic sequence is shown in Figure 2 as a function of cluster size. The average is seen to increase roughly linearly with N to an average sequence length of approximately 24 conformations at $N = 110$. Thus the scaling behavior of sequence length with cluster size is relatively benign. Similarly, scaling as measured by the average number of local minimizations per sequence (not including the final fixed overhead of $R_{\max}$ at the sequence end) does not appear to increase faster than linearly with N, with an average of about 286 local minimizations at $N = 30$, 844 at $N = 80$, and 1256 at $N = 110$.

For all values of N the MSBH algorithm found the known putative global minimum (and at $N = 98$ improved upon it, as discussed below) within the sample size of 1000 sequences. Observed fractional hit rates for the global minimum are plotted in Figure 3 as a function of cluster size N. The fractional hit rates range from a high of 1000/1000 for many of the small icosahedral clusters to a low of 2/1000 for the LJ_{77} decahedral cluster, with similar low rates of 4/1000 and 8/1000 for the neighboring LJ_{75} and LJ_{76} decahedral clusters, respectively. Surprisingly, the hit rates for the larger decahedral clusters at LJ_{102}, LJ_{103} and LJ_{104} are roughly an order of magnitude higher than those for the smaller decahedral global minima and not much different than similarly sized icosahedral global minima. Up to about $N = 60$, the hit rate is consistently over fifty percent, with exceptional dips to 387/1000 at LJ_{30},

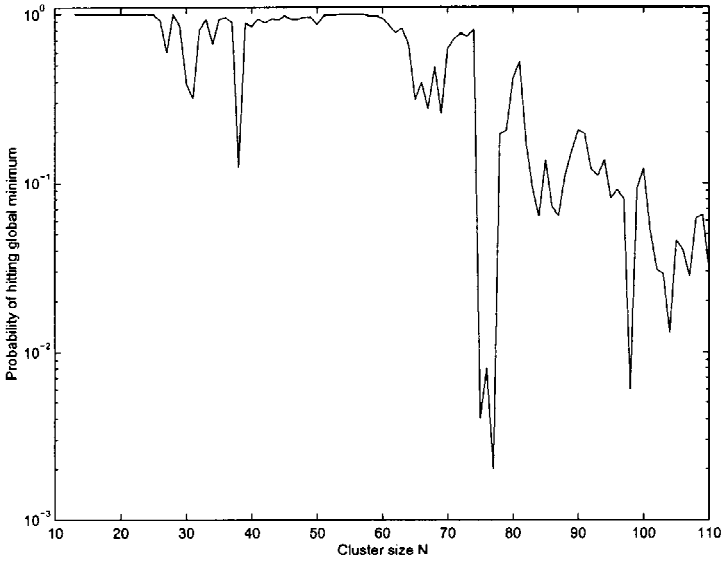

Figure 3. Observed probability of hitting the global minimum from a randomly initialized sequence, based on 1000 sequences at each cluster size.

316/1000 at LJ_{31} and 123/1000 for the LJ_{38} fcc truncated octahedron. Thus the funnel corresponding to the global minimum is relatively wide and easily entered for the smaller icosahedral clusters. However, even for the icosahedral clusters, the hit rate begins to drop systematically after $N = 70$, suggesting difficulty in finding the correct funnel as the number of funnels increases.

For $N \leqslant 25$, only a single funnel is observed and all monotonic sequences terminate at the (icosahedral) global minimum. The first instance of a multiple funnel landscape is found at LJ_{26}, where 75 out of 1000 sequences terminate at the best icosahedral structure with a 'Mackay' outer layer, as opposed to the 'anti-Mackay' (Doye et al., 1995) surface layer at the global minimum (these surface variants are also called FC and IC, respectively, in Northby (1987)). A particularly interesting case occurs at $N = 34$, where four funnels are observed, each corresponding to a different geometrical class: 667 sequences terminate at the Mackay icosahedral global minimum, 130 at the anti-Mackay icosahedral funnel bottom, 198 at a partly fcc structure with tetrahedral symmetry (an analog to the new tetrahedral LJ_{98} local minimum described below), and 5 at the best decahedral structure. Note that these four funnel bottoms are not the four lowest local minima, as there are several intermediate icosahedral structures. Rather these funnel bottoms are the lowest members of their respective classes. Thus the MSBH algorithm not only finds global minima, but also identifies the funneling properties of the landscape as well as the 'best-of-class' local minima at the funnel bottoms.

For $N \leqslant 60$, the total number of funnels on the landscape (i.e., the number of energetically distinct funnel bottoms observed in the 1000 sequences) is observed

Table 1. Performance of original MCM and new MSBH basin-hopping algorithms in terms of average number of local minimizations (LM) required to find the global minimum, number of energetically distinct funnels found by the MSBH algorithm, and global minimum hit rate for MSBH. $N = 38$, 75, 98, and 102 are non-icosahedral global minima; the others are icosahedral

N	MCM LM	MSBH LM	MSBH funnels	MSBH hit rate
20	35	34	1	1.000
30	1140	739	2	0.387
38	2674	2875	3	0.124
40	208	279	9	0.849
50	251	460	7	0.868
60	384	388	12	0.948
70	1527	1526	90	0.630
75	$\sim 10^7$	152000	75	0.004
80	2540	2009	132	0.420
90	3024	4699	198	0.206
98	$> 10^6$	180000	210	0.006
100	7610	9128	221	0.122
102	$> 10^6$	36028	245	0.031
110	11362	40420	288	0.031

to range from 1 to 12, with the global minimum usually lying at the bottom of the first or second most commonly entered funnel. However, the observed number of funnels for the larger clusters quickly climbs to approximately 300 at $N = 110$, and a larger sequence sampling size would undoubtedly have found considerably more. The general trends can be seen in the data in Table 1.

Perhaps the most impressive success of the algorithm is the discovery of a new global minimum LJ_{98} with tetrahedral symmetry at an energy $E(X) = -543.665361$ that is 0.022404 units below the previous putative icosahedral global minimum found by Deaven et al. (1996) using a genetic algorithm. The discovery is quite surprising, given the extensive previous investigation of the LJ microcluster problem by a wide variety of methods. Details of the structure can be found in Leary and Doye (1999). The new structure is among the most difficult to find, with only 6 of 1000 sequences terminating there, and required an average of approximately 180,000 local minimizations per encounter. The new structure has subsequently been found by the original MCM basin-hopping algorithm (Wales, 1999) and a new procedure that uses geometrical penalty functions (Locatelli and Schoen, 1999).

Our implementation of MSBH is easily modified to implement the original MCM basin-hopping algorithm. Using the values presented in Wales and Scheraga

(1999) and Wales and Doye (1997) as well as some experimentation to determine reasonable values of the parameter T, a series of direct comparisons was made for a variety of cases spanning the microcluster range (Table 1). The original version was terminated whenever it hit the global minimum, so the reported number of local minimizations LM is an average first passage time to the global minimum over many diffferent random walks. The LM value for MSBH is simply the total number of local minimizations performed for the 1000 sequences divided by the number of global minimum encounters, excluding the final $R_{\max}$ local minimizations at the end of each sequence.

Neither basin-hopping variant shows a systematic relative advantage in terms of average CPU time or number of local minimizations between encounters with the icosahedral global minima. Both are also similarly successful on the LJ_{38} fcc truncated icosahedron, requiring an average of 4 CPU minutes and about three thousand local minimizations per global minimum hit. However for the LJ_{75}, LJ_{76} and LJ_{77} decahedral clusters, MSBH is much faster and more successful. The exact amount is statistically difficult to quantify since our implementation of the original algorithm, for example, only encountered the decahedral LJ_{75} global minimum once in fifty long simulations with a total time equivalent to following 10,000 MSBH sequences. In contrast MSBH found the LJ_{75} global minimum a total of 42 times in an extended sample of 10,000 sequences, with an average of 152,000 local minimizations between encounters. Similar results for the original MCM algorithm are noted in Doye et al. (1999b), who report that these decahedral conformations are found so infrequently by their implementation that good statistics for the first passage time have not yet been obtained. Their analysis using disconnectivity graphs found that the decahedral funnel is narrow and separated from the icosahedral region by a large energy barrier. Such a landscape will clearly penalize the MCM strategy severely.

However, as N increases and the number of funnels becomes larger, the probability of hitting the correct funnel on any given MSBH descent falls correspondingly. This may present a more difficult scaling problem for MSBH than the original version. If barriers between funnels are small, the MCM strategy of climbing over barriers, but still remaining in a relatively low energy region, gains in efficiency relative to restarting at a high energy. This is perhaps becoming evident for the larger icosahedral clusters in Table 1, where the original algorithm appears to be somewhat better than MSBH. Also, despite extensive computation, MSBH has not been able to reach the lowest energy value $E(X) = -1236.124253$ for LJ_{201} reported by Wales and Doye (1997) with the original MCM basin-hopping algorithm.

3. Conclusion

The MSBH variant of basin-hopping presented here has found all of the previously known LJ putative global minima in the Cambridge Cluster Database, in addition to locating a new LJ_{98} putative global minimum of a novel geometrical class. MSBH

appears to have an efficiency advantage over the original MCM basin-hopping algorithm for cases such as the LJ_{75}, LJ_{76}, and LJ_{77} decahedral structures, where the global minimum lies at the bottom of a relatively narrow, diffficult-to-find decahedral funnel which is separated from the more dominant icosahedral region by a large energy barrier. In this case the simple restart strategy employed by MSBH is a more efficient way of escaping from a funnel bottom and gaining altitude on the landscape graph than waiting for a low-probability sequence of uphill Metropolis steps. However, on landscapes in which a large number of funnels are separated by relatively low barriers, the original version may be expected to spend most of its time in low energy regions of the landscape graph, passing relatively easily between funnels. The monotonic sequence version must descend many times from high energy starts, a possibly much longer process. We hope to investigate this possibility in future research using parametrically controlled energy landscapes from other molecular conformation problems.

Both versions of basin-hopping are unbiased in the sense that no prior geometrical knowledge of optimal structures is included in the algorithms. This contrasts, for example, with methods that search lattices of a particular geometrical class, such as Mackay icosahedra, and therefore cannot find global minima outside of that class. Similarly, many of the genetic algorithms applied to the LJ cluster problem have used seeding strategies that introduce structures within the classes of known global minima. Unbiased methods may have the best chance of generalizing to more complex potential functions such as those in the protein folding problem.

The monotonic sequence variant also has the interesting and potentially useful property of clearly exposing the funneling nature of the energy landscape, and identifying a variety of funnel bottoms in the multifunnel case. Thus MSBH identifies not only global minima, but also other ‘best-of-class’ structures on the energy landscape. This property may be of interest for protein folding problems, where an accurate energy function with a global minimum corresponding to the correct physical native conformation is still beyond the state of the art. Rather a simplified approximate model function is often used, which may not necessarily produce the single funnel landscape hypothesized for naturally occurring proteins. In this case the identification of a variety of interesting funnel bottoms may in fact be a more appropriate goal than simply finding the global minimum.

References

Barron, C., Gomez, S. and Romero, D. (1996), Archimedean poplyhedron structure yields a lower energy atomic cluster, *Applied Mathematics Letters* 9: 75–78.

Berry, R.S. and Breitengraser-Kunz, R.E. (1995), Topography and dynamics of multidimensional interatomic potential surfaces, *Physical Review Letters* 74: 3951–3954.

Bryngelson, J.D.Onuchic, J.N., Socci, N.D. and Wolynes, P.G. (1995), Funnels, pathways, and the energy landscape of protein folding: a synthesis, *Proteins: Structure, Function, and Genetics* 21: 167–195.

Deaven, D.M., Tit, N., Morris, J.R. and Ho, K.M. (1996), Structural optimization of Lennard–Jones clusters by a genetic algorithm, *Chemical Physics Letters* 256: 195–198.

Doye, J.P.K., Wales, D.J. and Berry, R.S. (1995), The effect of the range of the potential on the structure of clusters, *Journal of Chemical Physics* 103: 4234–4249.

Doye, J.P.K., Miller, M.A. and Wales, D.J., (1999a), The double-funnel energy landscape of the 38-atom Lennard–Jones cluster, *Journal of Chemical Physics* 110: 6896–6906.

Doye, J.P.K., Miller, M.A. and Wales, D.J., (1999b), Evolution of the potential energy surface with size for Lennard–Jones clusters, *Journal of Chemical Physics* 111: 8417–8428.

Hoare, M.R. and Pal, P. (1971), Physical cluster mechanics: statics and energy surfaces for monatomic systems, *Advances in Physics* 20: 161–196.

Kostrowicki, J., Piela, J., Cherayil, B.J. and Scheraga, H.A. (1991), Performance of the diffusion equation method in searches for optimum structures of Lennard–Jones atoms, *Journal of Physical Chemistry* 95: 4113–4119.

Leary, R.H. (1997), Global optima of Lennard–Jones clusters, *Journal of Global Optimization* 11: 35–53.

Leary, R.H. and Doye, J.P.K. (1999), Tetrahedral global minimum for the 98-atom Lennard–Jones cluster, *Physical Review E* 60: R6320–R6322.

Li, Z. and Scheraga, H.A. (1987), Monte Carlo-minimization approach to the multiple minima problem, *Proceedings of the National Academy of Sciences, USA* 84: 6611–6615.

Locatelli, M. and Schoen, F. (1999), Fast global optimization of diffficult Lennard–Jones clusters (submitted).

Luenberger, D.G. (1969), *Optimization by Vector Space Methods*, Wiley, New York.

Mackay, A.L. (1962), A dense non-crystallographic packing of equal spheres, *Acta Crystallographica* 15: 916–918.

Neumaier, A. (1997), Molecular modeling of proteins and mathematical prediction of protein structure, *SIAM Review* 39: 407–460.

Niesse, J.A. and Mayne, H.R. (1996), Genetic algorithms for structural cluster optimization, *Journal of Chemical Physics* 105: 6129–6137.

Northby, J.A. (1987), Structure and binding of Lennard–Jones clusters: $13 \leqslant N \leqslant 147$, *Journal of Chemical Physics* 87: 6166–6178.

Pillardy, J. and Piela, L. (1995), Molecular dynamics on deformed potential energy hypersurfaces, *Journal of Physical Chemistry* 99: 11805–11812.

Stillinger, F.H. (1999), Exponential multiplicity of inherent structures, *Physical Review E* 59: 48–51.

Tsai, C.J. and Jordan, K.D. (1993), Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters, *Journal of Physical Chemistry* 97: 11227–11237.

Wales, D.J. (1999), personal communication.

Wales, D.J. and Doye, J.P.K. (1997), Global optimization by basin-hopping and the lowest energy structures of Lennard–Jones clusters containing up to 110 atoms, *Journal of Physical Chemistry A* 101: 5111–5116.

Wales, D.J., Doye, J.P.K., Dullweber, A. and Naumkin, F.Y. (URL), The Cambridge Cluster Database, URL htp://brian.ch.cam.ac.uk/CCD.html

Wales, D.J. and Scheraga, H.A. (1999), Global optimization of clusters, crystals, and biomolecules, *Science* 285: 1368–1372.

Wille, L.T. (1987), Minimum-energy configurations of atomic clusters: new results obtained by simulated annealing, *Chemical Physics Letters* 133: 404–410.

Wolf, M.D. and Landman, U., (1998), Genetic algorithms for structural cluster optimizati9on, *Journal of Physical Chemistry A* 102: 6129–6137.

Xue, G.L. (1993), Parallel two-level simulated annealing, 1993 ACM International Conference on Supercomputing, ACM Press, 357–366.

Xue, G.L. (1994a), Molecular conformation on the CM-5 by parallel two-level simulated annealing, *Journal of Global Optimization* 4: 187–208.

Xue, G.L. (1994b), Improvement of the Northby algorithm for molecular conformation: better solutions, *Journal of Global Optimization* 4: 425–440.