PCCP

A random tunneling algorithm for the structural optimization problem

Haiyan Jiang, Wensheng Cai and Xueguang Shao*

Department of Chemistry, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China. E-mail: xshao@ustc.edu.cn

Received 28th June 2002, Accepted 30th July 2002
First published as an Advance Article on the web 28th August 2002

Based on the concept of subenergy transformation and the terminal repeller in the terminal repeller unconstrained subenergy tunneling (TRUST) algorithm, a global optimization algorithm, called the random tunneling algorithm (RTA), is proposed. RTA is a two-phase optimization method in which a global phase is carried out by random tunneling and a local phase by gradient optimization with the BFGS method. In RTA, the population of start points is generated randomly, and the similarity checking is performed during the cycles of the two phases. RTA was compared with other well-known stochastic methods in the optimization of the standard test functions. The potential energy function of Lennard-Jones clusters was also selected as an instance to investigate the performance of the RTA in high-dimensional optimization. All the global minima of LJ clusters containing up to 100 atoms were successfully located. The results indicate that RTA may be a good tool for the structural optimization problem.

1. Introduction

The global optimization problem is an active field with rapid growth in structural optimization, chemical engineering design and molecular biology. It is very important for chemists or bio-scientists to discover the lowest-energy structure of molecules such as proteins, atom clusters and crystals. In computational chemistry, this problem is notoriously dificult because there are many local minima in the energy landscape, and also a large number of parameters in the force field need to be optimized. As a result, a global optimization algorithm for this kind of problem should have good performance on multi- or high-dimensional continuous problems. Different optimization methods have been applied to solve the problem, such as genetic algorithms (GAs),[1,2] simulated annealing,[3,4] potential deformation,[5] fast annealing evolutionary algorithm (FAEA),[6,7] simple linkage (SL),[8] *etc.*

Different from the stochastic algorithms mentioned above, terminal repeller unconstrained subenergy tunneling (TRUST) is a deterministic optimization algorithm. The tunneling method for global optimization was introduced by Levy and Montalvo.[9] It is composed of a sequence of cycles, where each cycle has two phases: a local minimization phase and a tunneling phase. The schematic diagram of the tunneling is shown in Fig. 1. In the first phase, minimization algorithms are employed to optimize an objective function $f(\boldsymbol{x})$. Assuming that we start from an initial point $\boldsymbol{x}^0$, the minimization converges to the first local minimum $\boldsymbol{x}^{1(*)}$. In the second phase, the method searches for the zeros of the tunneling function, that is $f(\boldsymbol{x}^{1(*)}) = f(\boldsymbol{x}^1)$ and $\boldsymbol{x}^{1(*)} \neq \boldsymbol{x}^1$. Then the zero point is used as the starting point of the next cycle and the process is repeated sequentially. The process for exploring the other lower minimum from the current minimum is called tunneling.

TRUST has been successfully used in exploratory seismology.[10] In this method, the dynamic system has a global descent property and always escapes local minima valleys by incorporating the subenergy tunneling and terminal repeller effect.[10,11] Theoretically, TRUST is an effective and reliable algorithm for optimizing a one-dimensional problem, whereas a direct extension of the 1D scheme to multi-dimensional global optimization problems cannot guarantee that the global optimum will always be found.[10,12] Diller has studied the global optimization abilities of several global optimization algorithms including the TRUST algorithm for the purpose of molecular docking and found that the results of the TRUST algorithm were disappointing.[12]

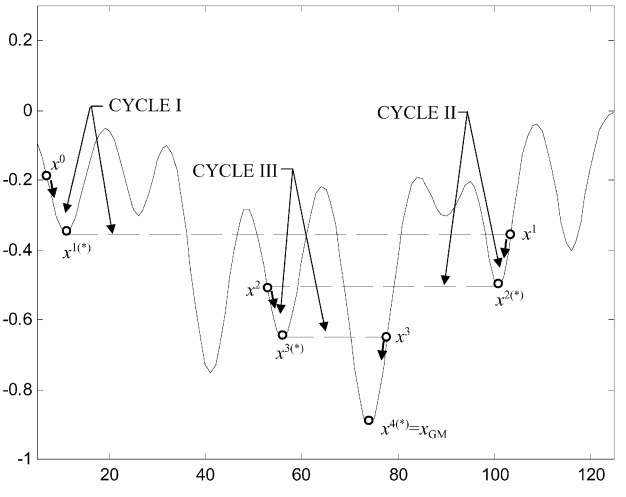

Fig. 1 Schematic diagram of tunneling operation.

Inspired by the subenergy tunneling technique of TRUST, a stochastic algorithm, called the random tunneling algorithm (RTA), was developed for the purpose of solving multi- or high-dimensional global optimization problems. In RTA, a population of start points is adopted in order to explore the solution space. RTA consists of two optimization phases, a global phase in which the global sampled points are generated by a random tunneling technique and a local phase in which the gradient method BFGS is applied to these points, yielding various local optima. During the cycles of the two phases, similarity checking is performed on the population of RTA. To assess the algorithm, RTA was applied to a set of standard

 DOI: 10.1039/b206251h

multi-dimensional test functions, and the results were compared with some well-known global optimization methods. Furthermore, the structural optimization of Lennard-Jones (LJ) clusters was also studied in order to investigate the performance of RTA in a high-dimensional case. Only using a pure RTA, the lowest known energies of LJ clusters containing up to 60 atoms were located. By moving the outside atoms of the best cluster[7] obtained by RTA, the optimization results reached a size of up to 100 atoms.

2. Method

2.1. Problem formulation

The optimization problem can be stated as follows. Let the objective function $f(\boldsymbol{x}):D \rightarrow R$, where $\boldsymbol{x}$ is a state vector having a finite dimension D. D is termed the domain or the search space which can be represented as $D = \{x_i|\alpha_i^- \leqslant x_i \leqslant \alpha_i^+; \forall i \in [1,....,n]\}$; where α_i^- and α_i^+ denote the lower and upper bounds of the ith component of the state vector, and n quantifies the dimension of D. R is a subset of real number space, indicating the values of the objective function. The goal is to find the global minimum $\boldsymbol{x}^{\mathrm{gm}}$ that minimizes $f(\boldsymbol{x})$ in D.

2.2. Subenergy transformation and terminal repeller

The tunneling technique in the TRUST algorithm combines subenergy transformation and non-Lipschitz terminal repellers. With these two energy transformation techniques, the dynamic system will flow from one current local minimum to the next.

Let $f(\boldsymbol{x})$ be a function to be globally minimized over D that is presented in Fig. 2 as the solid line. The subenergy transformation term is defined as:[10,11]

$$E_{\mathrm{sub}}(\boldsymbol{x},\boldsymbol{x}^*) = \log(1/[1 + \exp(-(\hat{f}(\boldsymbol{x}) + \beta))]) \qquad (1)$$

where β is a positive constant and $\hat{f}(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x}^*)$, in which $\boldsymbol{x}^*$ is the current local minimum. As the dashed line shown in Fig. 2, the transformation converts those points whose value is greater than the current local minimum to approximately zero and preserves all the extremal points that lie below this local minimum in the same order as the initial energy landscape. The subenergy tunneling transformation can achieve its most desirable asymptotic behavior when the value of β is 2.[10,11]

The terminal repeller term is a penalty function that converts the current minimum to a local maximum,

$$E_{\mathrm{rep}}(\boldsymbol{x},\boldsymbol{x}^*) = -(3/4)\rho(\boldsymbol{x} - \boldsymbol{x}^*)^{4/3}\theta(\hat{f}(\boldsymbol{x})) \qquad (2)$$

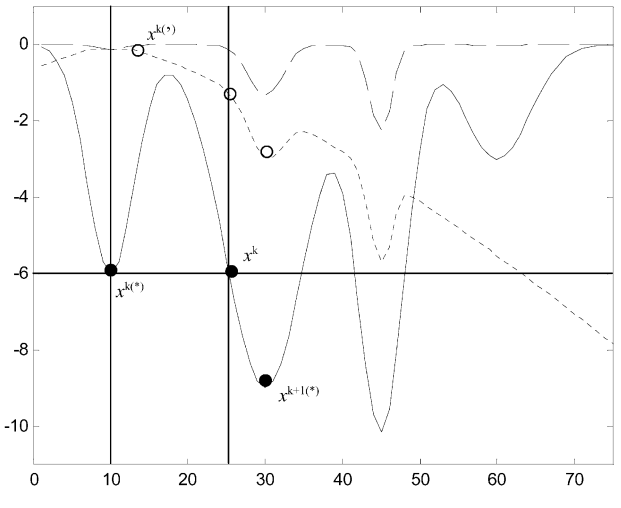

Fig. 2 Schematic diagram of subenergy transformation and terminal repeller.

where ρ is a positive constant and θ is the Heaviside function, which is equal to 0 when $\hat{f}(\boldsymbol{x})$ is negative and 1 when its argument is positive. Here the equilibrium point $\boldsymbol{x}^*$ is termed a terminal repeller. The Lipschitz condition ($|\partial g(\bar{x}_{\mathrm{eq}})/\partial \bar{x}| < +\infty$) is violated when $\boldsymbol{x}^*$ is perturbed to $\boldsymbol{x}'$, that is a hollow circle in Fig. 2. As a result, the dynamic system $\dot{\boldsymbol{x}} = g(\boldsymbol{x})$ will escape from the terminal repeller in finite time.

Finally, the virtual objective function that combines the above two transformation strategies is as follows:

$$E(\boldsymbol{x},\boldsymbol{x}^*) = E_{\mathrm{sub}}(\boldsymbol{x},\boldsymbol{x}^*) + E_{\mathrm{rep}}(\boldsymbol{x},\boldsymbol{x}^*). \qquad (3)$$

The two terms of this formula are the subenergy transformation term and the terminal repeller term. In Fig. 2, the dotted line represents the virtual objective function $E(\boldsymbol{x},\boldsymbol{x}^*)$. At the beginning of the tunneling phase in TRUST, the current local minimum $\boldsymbol{x}^{k(*)}$ is perturbed to $\boldsymbol{x}^{k(')}$. Starting from $\boldsymbol{x}^{k(')}$, the dynamic system will flow down on the transformed energy landscape until $\boldsymbol{x}^k$ is found, at which the value of the Heaviside function is changed to zero. Then the local minimization phase is started, and a gradient method is applied to find the next lower minimum $\boldsymbol{x}^{k+1(*)}$.

2.3. Procedures of RTA

wRTA has two phases, a global phase and a local phase. In RTA, the global phase is also termed the tunneling phase, where a random tunneling technique is utilized to explore solution space. In the tunneling phase, the terminal repeller and subenergy transformation is adopted to provide the transformed energy hypersurface for random tunneling. In the local phase, a gradient optimization is carried out by the BFGS method. Furthermore, a population of sampled points and the diversifying strategy of the population are performed to ensure that the solution space is explored widely and efficiently. A schematic description of the algorithm is given in Fig. 3. The procedures of RTA are explained as follows:

1. Initialization. The population is initialized randomly in the solution space. Then, each individual of the population is optimized to a local minimum $\boldsymbol{x}^*$ by the limited memory BFGS (L-BFGS) method,[13,14] which was proven to converge very fast with low iteration cost. In Fig. 3(a), the initial individual (hollow circle) is optimized to the local minimum $\boldsymbol{x}^*$ (solid circle).

2. Random tunneling. Theoretically, the convergence of TRUST to a global minimum was not formally guaranteed in the multi-dimensional case due to the constant perturbation direction vector.[11] Therefore, several modifications including random perturbation were done on the subenergy tunneling method of the TRUST algorithm, and a random tunneling technique was developed.

To explain the characteristic of the random tunneling technique in minimizing multidimensional problems, a schematic diagram of the random tunneling on a two-dimensional function is shown in Fig. 4, where the landscape of the objective function is represented with a contour line. The start points of two individuals for random tunneling are located at two local minima LM1 and LM2, and the arrows represent the tunneling direction vectors that are determined by the random perturbation. As a result, the dynamic system will not rely on only one tunneling direction as in TRUST, and will have flexibility in exploring the global minimum.

As in other tunneling methods, the purpose of random tunneling is to find the next start point $\boldsymbol{x}^{\#}$ for local minimization as in Fig. 3(b). At the beginning of the random tunneling, new individuals are generated around the current local minimum $\boldsymbol{x}^*$ with random direction in a local range so as to break down

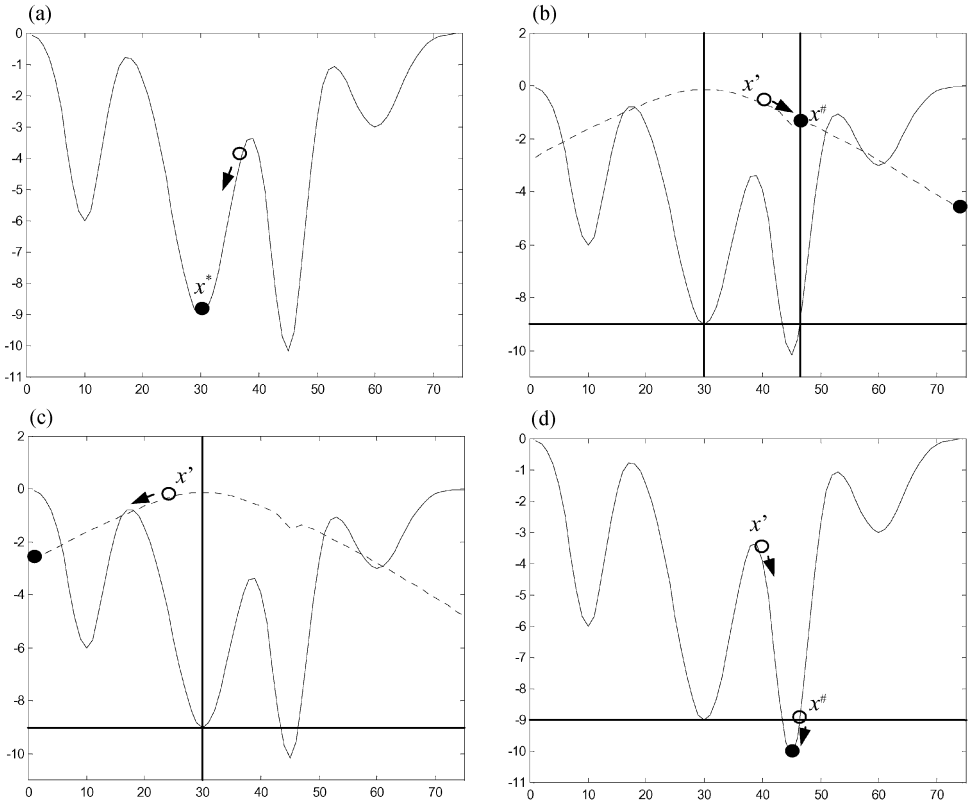

Fig. 3 A schematic description of the random tunneling algorithm. (a) Individual initialization. The hollow circle is the start point chosen randomly. It is optimized to the local minimum x^* (solid circle) by L-BFGS. (b) Positive tunneling on the transformed function when $s_i^1 = 1$. Initial function $f(x)$, the solid line, is transformed by the terminal repeller and subenergy transformation to $E(x,x^*)$, the dotted line. x', the hollow circle, is the result of positive perturbation. The gradient method with random step size is then applied to find the solid circle $x^{\#}$. If this point is not found, it will reach the solid circle at the right corner and overflow the upper limitation. (b) Negative tunneling on the transformed function when $s_i^1 = -1$. The hollow circle is the result of negative perturbation. In this case, $x^{\#}$ cannot be found and the dynamic system overflows the lower limitation, this solid circle will be relocated on the right of x^* as the hollow circle shown in (b). (c) Local minimization after tunneling. If $x^{\#}$ is found, optimize $x^{\#}$ to reach the next lower minimum or the global minimum, the solid circle. Otherwise, the start point for local minimization is x'.

the Lipschitz condition. Perturb $\boldsymbol{x}^*$ with

$$x_i' = x_i^* + S_i^1 \lambda_1 r(\alpha_i^+ - \alpha_i^-), \tag{4}$$

where $\forall i \in [1,....,n]$, n is the dimension of the problem. The second term of eqn. (4) is the perturbation on the terminal repeller $\boldsymbol{x}^*$. S_i^1, a sign variable determining the direction of tunneling with value -1 or 1, is generated randomly on each dimension. In the one-dimensional case, the dynamic system will tunnel toward the positive direction when S_i^1 is 1, whereas it will tunnel toward the negative direction when S_i^1 is -1 as shown in Fig. 3(b) and Fig. 3(c). For a multi-dimensional problem, each dimension will have a different perturbation direction. Here α_i^- and α_i^+ are the upper and lower limits of parameter x_i. λ_1 is a positive value that determines the range of perturbation and should be lower than 0.5, r is a random number between 0 and 1 for each dimension.

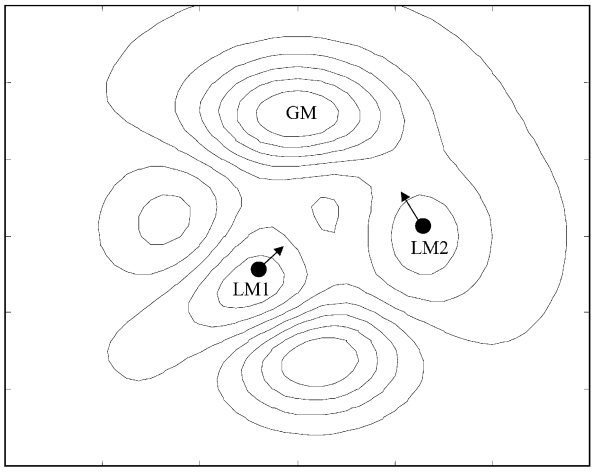

Fig. 4 The random tunneling behavior of the RT algorithm. GM: Global minimum.LM1,LM2: Local minima 1 and 2. Solid circles are the current local minima and the start points for random tunneling. The arrows show the tunneling directions of the dynamic systems.

After the perturbation, the energy difference is computed by $\hat{f}(\boldsymbol{x}') = f(\boldsymbol{x}') - f(\boldsymbol{x}^*)$. If $f(\boldsymbol{x}) \leqslant f(\boldsymbol{x}^*)$, *i.e.*, $\hat{f}(\boldsymbol{x}') \leqslant 0$, the current solution $\boldsymbol{x}'$ will be accepted as the next start point $\boldsymbol{x}^{\#}$ and go to step (3) for local optimization. Otherwise the energy function $f(\boldsymbol{x})$ is replaced with the transformed function $E(\boldsymbol{x},\boldsymbol{x}^*)$ in eqn. (3) and $\boldsymbol{x}^*$ is used as the terminal repeller in the kth minimization phase, denoted as $\boldsymbol{x}^{k(*)}$. The tunneling procedure is started on the transformed energy hyper-surface $E(\boldsymbol{x},\boldsymbol{x}^*)$ as shown in Fig. 3(b).

As is shown in Fig. 3(b), the basins of the transformed energy surface are usually very weak when the difference between the current local mimimum and the next lower minimum is small. To avoid missing the next or global minimum, a minute step size is helpful but is very costly. Considering both aspects, random step sizes were introduced. In random tunneling, the descent on each dimension was controlled by eqn. (5).

$$x_i = x_i' + \Delta t_i \dot{x}_i \tag{5}$$

in which $\dot{x}_i = -(\partial f(\boldsymbol{x})/\partial x_i)(1/[1+\exp(\hat{f}(\boldsymbol{x})+\beta)]) + \rho(x_i - x_i^*)^{1/3}\theta(\hat{f}(\boldsymbol{x}))$, the start point of each dimension is x_i' that is

Table 1 Test functions

Abbr.	Full name	No. of vars.	Upper and lower bound	Global optimum
BR	Branin	2	x_1: (−5,10), x_2: (0,15)	0.3978 (5/4 π)
CA	Camelback	2	(−5,5)	−1.0316285
GP	Goldstein–Price	2	(−2,2)	3
RA2, RA5	Rastrigin	2, 5	(−5.12,5.12)	0
SH	Shubert	2	(−10,10)	−186.7309
H3, H6	Hartman	3, 6	(0,1)	−3.3223
GW2, GW8	Griewank	2, 8	(−600,600)	0

the result of perturbation, and the time step of each dimension is

$$\Delta t_i = \lambda_2 r(a_i^+ - a_i^-) \tag{6}$$

where λ_2 is a small value, and r is a random number between 0 and 1 for each dimension, so as to provide random time steps of the integration on each dimension. During the integration, the energy difference $\hat{f}(\boldsymbol{x})$ is observed simultaneously, if $\hat{f}(\boldsymbol{x}) \leqslant 0$, the current solution $\boldsymbol{x}$ will be accepted as $\boldsymbol{x}^{\#}$ and go to step (3) for local minimization. Otherwise the random tunneling will be continued.

As the integration continues, if one variable x_i overflows the upper or lower limits, as the solid circle shown in Fig. 3(c), the following conversion will be performed. Let

$$x_i' = x_i^* + S_i^2 \varepsilon(\alpha_i^+ - \alpha_i^-), \tag{7}$$

where the sign variable S_i^2 is −1 or 1 respectively when $x_i > \alpha_i^+$ or $x_i < \alpha_i^-$, and ε is a constant value 5.0×10^{-3}. The purpose of this procedure is to place the current overflow variable on the opposite side of the latest local minimum. The tunneling phase will terminate when the number of the overflow variables reaches a limit that is $n/5$ in this paper.

3. Local optimization. If $\boldsymbol{x}^{\#}$ is found, then optimize $\boldsymbol{x}^{\#}$ to the next minimum $\boldsymbol{x}^{k+1(*)}$, let $k = k+1$ and go to step (2). As shown in Fig. 3(d), if $\boldsymbol{x}^{\#}$ is not found, the result of perturbation $\boldsymbol{x}'$ is optimized to the local minimum by L-BFGS. This new local optimum is accepted as $\boldsymbol{x}^{k+1(*)}$ if the result is lower than $\boldsymbol{x}^{k(*)}$. Otherwise the next tunneling phase is done with $k = k+1$.

4. Similarity checking of population. To diversify the population, the similarity checking is carried out after M iterations of the tunneling phase and the local minimization. In the program, M is 100. To denote two different configurations of the population, I_i and I_j are used with the index i and j no larger than the population size. The similarity degree between two configurations I_i and I_j is calculated according to eqn. (8).

$$\text{Similarity}_{i,j} = 1 - \frac{\sum_{k=1}^{n} (I_{ik} - I_{jk})^2}{\sum_{k=1}^{n} \text{Range}_k^2} \tag{8}$$

where I_{ik} and I_{jk} are the value of the kth parameter in I_i and I_j, respectively, and Range_k is the range of the kth parameter. If the similarity is larger than a corresponding threshold, similar individuals will be regenerated as in the initialization procedure.

In general, the program will stop when the global minimum is found. Otherwise these procedures will continue until the circle count k reaches a maximal integer, then the best individual is output as the result.

3. Results and discussion

3.1. Optimization results of test functions

To assess the optimization ability of the RTA algorithm, a set of well-known multidimensional functions were adopted as benchmark test problems. Table 1 summarizes the functions investigated. The results are listed in Table 2. In the optimization of these functions, the population size was set to 1 because the problems are comparatively simple, and the achievement of the global minimum with a desired accuracy 10^{-6} was taken as the termination criterion. The range of local perturbation λ_1 in eqn. (4) was optimized simply with a series of values as 0.5, 0.1, 0.05, 0.01 and 0.005. A typical value of the parameter λ_1 is 0.1 for all test functions except for H6 and GW, where 0.5 and 0.005 are used respectively. The other important parameter for random tunneling λ_2 in eqn. (6) is set to 0.005 by experience. The parameters β and ρ in eqn. (5) for the terminal repeller and subenergy transformation strategy were 2 and 20, respectively.

Table 2 Number of function calls (including energy evaluations and gradient calls) required by RTA and other global optimization methods to reach the global minima of the test functions

	Test function									
Method	BR	CA	GP	RA2	RA5	SH	H3	H6	GW2	GW8
PRS[a]	4850	—	5125	5964	—	6700	5280	18 090	—	—
MGA[b]	—	—	—	—	9000	7200	—	—	29 251	—
TUN[c]	—	1469	—	—	—	12 160	—	—	—	—
TS[d]	492	—	486	540	—	727	508	—	—	—
MLSL[e]	206	—	148	—	—	—	197	—	—	—
FAEA[f]	394	303	490	544	2762	446	488	2229	7804	—
TRUST[g]	55	31	103	59	—	72	58	—	—	—
RTA	23	135	113	383	687	202	60	196	281	465

[a] Pure random search method by Andressen.[15] [b] The modified genetic algorithm by Cai *et al.*[16] [c] The tunneling method by Levy and Montalvo.[9] [d] The taboo search scheme by Cvijovic and Klinowski.[17] [e] The multiple-level single-linkage method by Kan and Timmer.[18] [f] The fast annealing evolutionary algorithm by Cai and Shao.[6] [g] The terminal repeller unconstrained subenergy transformation algorithm.[10]

In Table 2, the performance of RTA is compared with the global optimization methods that were widely reported with reproducible results for the test functions, and TRUST is the only deterministic algorithm. The criterion for comparison is the number of function calls that include the energy evaluations and all gradient calls. All data shown in Table 2 are the average of 100 runs. It is obvious that, although TRUST outperforms RTA for most of the 2- and 3-dimensional test functions, RTA is substantially faster than the stochastic methods listed in Table 1, especially for the test functions with higher dimension such as H6, RA5 and GW8. The efficiency of RTA is not bad. However, in such a comparison, the difficulty is that it is not really fair to compare global optimization algorithms that use different local minimization subroutines. Consequently, the comparison can only provide knowledge about the general performance of RTA, but it cannot be concluded from the results that the faster algorithm is the superior.

3.2. Application to LJ cluster optimization

The LJ model of inert gas clusters has been investigated intensively and provides a useful testing ground for a putative global optimization algorithm. For the purpose of investigating the performance of RTA in a high-dimensional energy minimization system, the structures of LJ clusters were optimized. If N is the number of atoms that are positioned at $P_1,....,P_N \in \mathscr{R}^3$ in the cluster, the LJ potential energy is defined by

$$E = 4\varepsilon \sum_{i<j} \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} \right]$$

where r_{ij} is the Euclidean distance between points P_i and P_j, ε and $2^{1/6}\sigma$ are the pair well depth and the equilibrium pair separation, and $\varepsilon = \sigma = 1$ with reduced units. This problem is very difficult because there are a very large number of local minima and the number of local minima grows exponentially with atom number N. Hoare and McInnes have shown that the number of local minima in the potential energy surface of an LJ cluster becomes extremely large, even for quite small systems.[19,20] For example, the number of local minima for LJ_{13} is at least 1467.[21] A number of successful algorithms have been applied to optimize LJ clusters such as basin-hopping,[22] genetic algorithm[23] and phenotype algorithm.[24]

To test the optimization performance of RTA in a high-dimensional case, no related technique that exploits the preliminary knowledge about the LJ clusters, such as biased search on icosahedral lattice[25] and seeding,[7] was utilized in this article even though they are very effective for this problem. In RTA, the position of each atom is defined by Cartesian coordinates (x,y,z). Therefore, there are 3 N parameters to specify the configuration of the N-atom cluster. At the beginning of the program, the coordinates are initialized in the range $[-(n/(4\pi\sqrt{2}))^{1/3},(n/(4\pi\sqrt{2}))^{1/3}]$ where $n = 3N$. The two parameters of random tunneling λ_1 and λ_2 are 0.1 and 0.05 respectively. The values of β and ρ were the same as in the former case. The algorithm is stopped if $|E_{best} - E_{gm}| < 5 \times 10^{-7}$, where E_{best} is the lowest energy calculated by RTA, and E_{gm} is the known global minimum,[26] or when the tunneling count k reaches 500.

In the study of cluster optimization, it was found that many results have a similar structure to the global minimum with only several misplaced atoms on the surface. However repairing the final fault takes a lot of time. Angular movement of outside atoms was designed to repair the final fault and was proved to be effective.[7] Angular moves were first used in ref. 22, in which angular displacements were chosen randomly. The technique of angular moves in ref. 22 is different from that in this study, because here the angular displacements are randomly chosen from a given set. When the RTA procedure is finished and the lowest energy has not been reached, the outside atoms of the best cluster are then moved in succession by spherical angles. First, the spherical coordinates for encoding each atom of the best cluster, r, θ and ϕ are calculated. The atoms with r greater than 1/2 of r_{max} (the maximum distance from the center of a cluster to each molecule) are considered as the outside atoms. Then the corresponding angular displacements ($\Delta\theta$ and $\Delta\phi$) are randomly chosen from the given set. After moving an outside atom, the L-BFGS energy minimization is performed, starting from the new atomic coordinates. When the better solution is accepted, it will be restarted. This procedure is repeated until all outside atoms have been moved and no movement will decrease the energy of the cluster. The angular movement of outside atoms is only performed on the best individual because repeating this procedure is time consuming.

Table 3 Comparison of the average number of function calls (including energy evaluations and gradient calls) and success ratio in the optimization of LJ atomic clusters

N_{atom}	FAEA[a]	SL[b]	RTA[c]	N_{atom}	FAEA[a]	SL[b]	RTA[c]
2	77	—	177	12	45 407	—	1004
	100%		100%		97%		100%
3	314	—	240	13	59 609	38 074	1563
	100%		100%		93%	100%	100%
4	1214	—	500	14	—	23 586	1095
	100%		100%			100%	100%
5	2778	—	379	15	—	45 811	1429
	100%		100%			100%	100%
6	36 672	—	1944	16	—	50 359	1929
	100%		100%			100%	100%
7	24 737	—	1956	17	—	242 546	15 478
	93%		100%			100%	100%
8	31 941	—	500	18	—	1 307 620	34 027
	99%		100%			80%	96%
9	41 241	—	1145	19	—	234 669	8458
	100%		100%			100%	100%
10	34 227	—	1470	20	—	358 160	3933
	94%		100%			100%	100%
11	35 246	—	1581	38	—	—	173 667
	94%		100%				7%

[a] The average results over the successful runs out of 100 runs are from ref. 6. [b] The average results over the successful runs out of 10 runs are from ref. 8. [c] The number of population is 2.

The comparison of RTA with other algorithms for this problem was studied based on the published results we could obtain. The comparison is shown in Table 3. FAEA is an annealing evolutionary algorithm with a fast annealing schedule.[6] SL is a two-phase stochastic algorithm proposed by Schoen for high-dimensional global optimization problems.[9] 100 independent runs for the LJ potential minimization were performed with N ranging from 2 to 20. It was shown that RTA has the highest success percentage and the smallest number of function calls. The statistic result of LJ_{38} is also provided.

Moreover, the optimization performance of RTA for the larger size of LJ clusters is tested and the results are shown in Table 4. The pure RTA has successfully located the global minima of LJ clusters containing up to 60 atoms. In the optimization of LJ_{61-100}, the best individual obtained by RTA was further optimized by moving the outside atoms.[7] All the lowest-energy structures of LJ atomic clusters with size up to 100 atoms were located successfully. The test was stopped when the cluster size reached 100 and the average CPU time of ten runs for the optimization of LJ_{100} was about 6320 s on a PIII866 workstation.

Increasing the population size will improve the efficiency of RTA. By comparing the results of $N = 39$–45, it was shown that the large population corresponds to the high successful

Table 4 Global minima (E) of LJ clusters obtained by RTA

N_{atom}[a]	N_{pop}[b]	E/ε	N_{succ}[c]	N_{atom}[a]	N_{pop}[b]	E/ε	N_{succ}[c]
21	2	−81.684571	10	61	30	−312.008896	7
22	2	−86.809782	10	62	30	−317.353901	5
23	2	−92.844472	10	63	30	−323.489734	7
24	2	−97.348815	10	64	30	−329.620147	6
25	2	−102.372663	10	65	30	−334.971532	3
26	2	−108.315616	10	66	30	−341.110599	2
27	2	−112.873584	9	67	30	−347.252007	4
28	2	−117.822402	10	68	30	−353.394542	4
29	2	−123.587371	10	69	30	−359.882566	2
30	2	−128.286571	4	70	30	−366.892251	3
31	2	−133.586422	3	71	30	−373.349661	4
32	2	−139.635524	10	72	30	−378.637253	1
33	2	−144.842719	9	73	30	−384.789377	6
34	2	−150.044528	3	74	30	−390.908500	7
35	2	−155.756643	5	75	240	−397.492331	1/50
36	2	−161.825363	8	76	480	−402.894866	2/50
37	2	−167.033672	4	77	480	−409.083517	1/50
38	30	−173.928427	9	78	30	−414.794401	2
39	2, 30	−180.033185	8, 10	79	30	−421.810897	2
40	2, 30	−185.249839	7, 10	80	30	−428.083564	4
41	2, 30	−190.536277	2, 10	81	30	−434.343643	2
42	2, 30	−196.277534	4, 10	82	30	−440.550425	2
43	2, 30	−202.364664	2, 10	83	30	−446.924094	1
44	2, 30	−207.688728	1, 10	84	30, 60	−452.657214	0, 1
45	2, 30	−213.784862	0, 10	85	30	−459.055799	1
46	30	−220.680330	10	86	30, 60	−465.384493	0, 1
47	30	−226.012256	10	87	30	−472.098165	1
48	30	−232.199529	8	88	30	−479.032630	1
49	30	−239.091864	9	89	30, 60	−486.053911	0, 1
50	30	−244.549926	5	90	30	−492.433908	1
51	30	−251.253964	6	91	30	−498.811060	2
52	30	−258.229991	10	92	30	−505.185309	1
53	30	−265.203016	10	93	30	−510.877688	1
54	30	−272.208631	10	94	30	−517.264131	2
55	30	−279.248470	6	95	30	−523.640211	2
56	30	−283.643105	10	96	30, 60	−529.879146	0, 3
57	30	−288.342625	5	97	30	−536.681383	1
58	30	−294.378148	1	98	240	−543.665361	2/20
59	30	−299.738070	1	99	30, 60	−550.666526	0, 1
60	30	−305.875476	1	100	30	−557.039820	2

[a] N_{atom}: the atom number of LJ cluster. [b] N_{pop}: the population size used. [c] N_{succ}: the number of successes out of ten independent runs, except for LJ_{75-77} and LJ_{98}.

ratio. For the more difficult cases LJ_{84}, LJ_{86}, LJ_{89}, LJ_{96} and LJ_{99}, the optimizations succeeded with a population size of 60 whereas they failed at a population size of 30. For the most difficult cases LJ_{75-77} and LJ_{98}, a larger population size was adopted.

Except for the five particular clusters LJ_{38}, LJ_{75-77} and LJ_{98},[27,28] the structures of LJ clusters containing fewer than 100 atoms at the global minima are based on the icosahedral motif. The two clusters with the complete icosahedral structure, LJ_{13} and LJ_{55}, are presented in Fig. 5(a) and (c). Other clusters are packed around the icosahedral core of LJ_{13} or LJ_{55} and tend to form the next complete icosahedron. It is shown clearly in Fig. 5(b) and (d) that the structures of LJ_{26} and LJ_{100} have an icosahedral core of LJ_{13} and LJ_{55} respectively.

LJ_{38}, LJ_{75-77} and LJ_{98} are the exceptional clusters and serve as particular test cases.[27,28] At these sizes, the lowest-energy minimum based on icosahedral packing acts as a trap and is widely separated from the true global minimum.[27] The global minimum of LJ_{38} was successfully located with a high success ratio 9/10 when the population was 30. The global minimum structures of LJ_{75-77} are based on Marks' decahedron,[22] and LJ_{98} on tetrahedron that was found by Leary.[28] In the optimization of these clusters, the population was divided evenly into 8 sub-populations on which the RTA was performed with the sub-population size of 30 for LJ_{75} and LJ_{98}, and 60 for LJ_{76} and LJ_{77}. At the end of RTA, the angular movement of outside atoms was carried out on the best individual of each sub-population. Then the cluster with the lowest

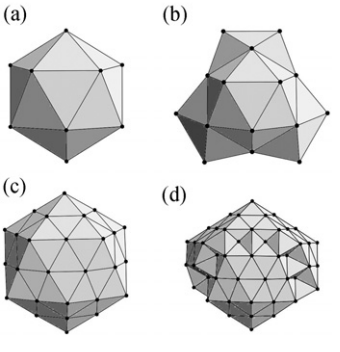

Fig. 5 The icosahedral structures of LJ clusters with atoms $N = 13$ (a), 26 (b), 55 (c) and 100 (d).

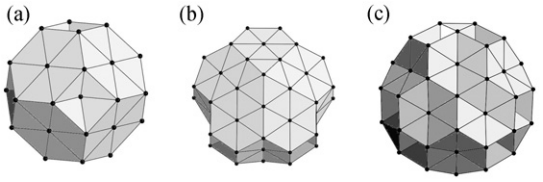

Fig. 6 The lowest-energy structures of (a) LJ_{38} (truncated octahedron), (b) LJ_{75} (Marks' decahedron), and (c) LJ_{98} (tetrahedron).

energy was selected as output. The global minima were successfully located in 1 out of 50 runs for LJ_{75} and LJ_{77}, 2 out of 50 runs for LJ_{76}, and 2 out of 20 runs for LJ_{98}. The exceptional structures of LJ_{38}, LJ_{75} and LJ_{98} are shown in Fig. 6.

From the results reported above, it is seen that RTA behaved with good optimization ability for high-dimensional problems. In the study, it was found that the effective local search method L-BFGS plays an important role in RTA. At the beginning of RTA, the initial population is optimized to local minima. Then, global tunneling is employed and always followed by local minimization. The effective local search method L-BFGS avoids wasting a large amount of time in transitional stages. From this point of view, RTA has some resemblance to one of the successful structural optimization methods: basin-hopping, in which every point in the catchment basin of each local minimum become the energy of that minimum.[22,27]

4. Conclusion

A random tunneling algorithm (RTA) is proposed based on the terminal repeller and subenergy transformation strategy of TRUST. It is a two-phase optimization algorithm with the global phase carried out by a random tunneling technique and the local phase by L-BFGS. Furthermore, a population with similarity checking is employed to provide different dynamic systems in searching for the global minimum. With the help of a local minimization technique, it is shown that the algorithm is efficient compared with the reported results of other global optimization methods. The algorithm was also applied to the structural optimization of LJ clusters, all the lowest-energy configurations containing atom number $N = 2$–100 were successfully located. Therefore, with further improvement RTA may become an effective tool for energy minimization or structural optimization problems in chemistry and biochemistry.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (No. 29975027 and No. 20172048), and the Teaching and Research Award Program for Outstanding Young Teachers (TRAPOYT) in higher education institutions of the Ministry of Education (MOE), P. R. China.

References

1 D. M. Deaven, N. Tit, J. R. Morris and K. M. Ho, *Chem. Phys. Lett.*, 1996, **256**, 195.
2 W. J. Pullan, *J. Comput. Chem.*, 1997, **18**, 1096.
3 R. F. Gutterres, M. A. Menezes, C. E. Fellows and O. Dulieu, *Chem. Phys. Lett.*, 1999, **300**, 131.
4 F. M. Torres, E. Agichtein, L. Grinberg, G. W. Yu and R. Q. Topper, *J. Mol. Struct. (THEOCHEM)*, 1997, **419**, 85.
5 L. Piela, J. Kostrowicki and H. A. Scheraga, *J. Phys. Chem.*, 1989, **93**, 339.
6 W. S. Cai and X. G. Shao, *J. Comput. Chem.*, 2002, **23**, 427.
7 W. S. Cai, Y. Feng, X. G. Shao and Z. X. Pan, *J. Mol. Struct. (THEOCHEM)*, 2002, **579**, 229.
8 F. Schoen, *Eur J. Oper. Res.*, 1999, **119**, 345.
9 A. Levy and A. Montalvo, *SIAM J. Sci. Stat. Comput.*, 1985, **6**, 15.
10 J. Barhen, V. Protopopescu and D. Reister, *Science*, 1997, **276**, 1094.
11 B. C. Cetin, J. Barhen and J. W. Burdick, *J. Optim. Theory Appl.*, 1993, **77**, 97.
12 D. J. Diller and C. L. M. J. Verlinde, *J. Comput. Chem.*, 1999, **20**, 1740.
13 D. C. Liu and J. Nocedal, *Math. Program. B*, 1989, **45**, 503.
14 http://www.netlib.org/.
15 R. S. Andressen, *Optimization*, University of Queensland Press, St. Lucia, Australia, 1972, p. 27.
16 W. S. Cai, F. Yu, X. G. Shao and Z. X. Pan, *Chin. J. Chem.*, 2000, **18**, 475.
17 D. Cvijovic and J. Klinowski, *Science*, 1995, **267**, 664.
18 H. G. R. Kan and G. T. Timmer, in *Numerical Optimization*, eds. P. T. Boggs, R. H. Byrd and R. B. Schnabel, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1985, p. 245.
19 M. R. Hoare, *Adv. Chem. Phys.*, 1979, **40**, 49.
20 (*a*) M. R. Hoare and M. McInnes, *Faraday Discuss. Chem. Soc.*, 1976, **61**, 12; (*b*) M. R. Hoare and M. McInnes, *Adv. Phys.*, 1983, **32**, 791.
21 J. P. K. Doye, M. A. Miller and D. J. Wales, *J. Chem. Phys.*, 1999, **111**, 8417.
22 D. J. Wales and J. P. K. Doye, *J. Phys. Chem. A*, 1997, **101**, 5111.
23 C. Barron, S. Gomez, D. Romero and A. Saavedra, *Appl. Math. Lett.*, 1999, **12**, 85.
24 B. Hartke, *J. Comput. Chem.*, 1999, **20**, 1752.
25 J. A. Northby, *J. Chem. Phys.*, 1987, **87**, 6166.
26 D. J. Wales, J. P. K. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin, F. Calvo, J. HernûÀndez-Rojas and T. F. Middleton, *The Cambridge Cluster Database*, available at http://www-wales.ch.cam.ac.uk/CCD.html.
27 D. J. Wales and H. A. Scheraga, *Science*, 1999, **285**, 1368.
28 R. H. Leary and J. P. K. Doye, *Phys. Rev. E*, 1999, **60**, 6320.