Particle Swarm Optimization Applied to the Atomic Cluster Optimization Problem

R.J.W. Hodgson

Physics Department
University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
rhodgson@physics.uottawa.ca

Abstract

The Particle Swarm Optimization method has proven quite successful in treating a variety of applied problems. Here we further test its capabilities by studying its behavior when applied to a challenging problem, namely the search for energy conformations of atomic clusters. In its simplest form this is known as the Lennard-Jones Problem. Results are compared with those achieved using simple Genetic Algorithms.

1 INTRODUCTION

The problem of minimizing the potential energy function of clusters of atoms is generally known as the molecular conformation problem. Cluster sizes can range from a few atoms up to several hundred atoms. Physical and chemical characteristics vary with size. The determination of the global minima or ground states of these energy functions is of particular interest to researchers in chemistry, biology, physics and optimization methods. One particular class of these problems in molecular conformation is that where the interaction potential is the pure Lennard-Jones potential function. This turns out to be a very difficult problem to solve since the number of local minima has been estimated to increase exponentially with the number of atoms N (Tsai and Jordan, 1993; Stillinger, 1999). Studies have shown that at N = 13 there are 988 local minima, and for N = 98 the number grows to the order of 10^{40}. In spite of this formidable hurdle, success has been found in locating what are believed to be the global minima (ground states) for systems with N as large as 250 (Hartke, 1993 and 2001).

While Genetic Algorithms (GA) have played an important role in treating this problem (Barron et al. 1999; Deaven and Ho, 1995; Zeiri, 1995), the generic GA has not proven to be very successful beyond small numbers of atoms, requiring a large number of generations in order to locate the global minima. In place of it, modifications to the GA operations of crossover, and mutation have had to be made in order to bring about more rapidly converging sequences. These variations have been based on physical insight into the problem (Hartke, 2001) incorporating crossover and mutation procedures based on the physical geometry of the clusters. In addition, improvements have been achieved by incorporating local search algorithms into the GA as well (Deaven et al., 1996; Doye et al., 1999; Neisse & Mayne, 1996; Radcliffe and Surry, 1995; Wales and Doye, 1997). These investigations have shown that the problem can be treated using problem-specific GAs.

Here we investigate and compare the success of the Particle Swarm Optimization (PSO) method with a fairly generic GA when applied to the Lennard-Jones problem. There has already been some discussion of the similarities and differences between the PSO method and genetic algorithms (Eberhart and Shi, 1998; Angeline, 1998). The PSO method has proven to be successful in a variety of applications (Kennedy and Eberhart, 2001). Our objective is to see how well it can handle quite a challenging function in order to better understand its capabilities. Certainly the success achieved using special treatments tailored to the cluster problem will not be achieved using a simple approach such as the PSO. However, it's ability to seek out the global minimum in a function with a large number of local minima will be severely tested. If it can accomplish this test as well if not better than simple generic genetic algorithms, then it should merit recognition as a valuable tool for treating other global optimization problems.

2 THE LENNARD-JONES PROBLEM

The Lennard-Jones problem is concerned with determining the lowest-energy configuration of a cluster of neutral atoms interacting via the Lennard-Jones potential. The function to be minimized is the total energy (in reduced units) of the Lennard-Jones cluster as computed from

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(r_{ij}^{-12} - r_{ij}^{-6} \right)$$

where N is the number of atoms in the cluster, and r_{ij} represents the distance separating atom *i* and atom *j*. This problem has a long history (Hoare, 1979; Leary, 1997). It has served as a test-bed for a wide variety of optimization algorithms, primarily due to the exponentially increasing number of local minima. It is significant in the field of chemical physics for the insight gained by studying the structure of the clusters as the size increases. It also serves as a reasonably accurate mathematical model of a real physical system, namely that of low-temperature microclusters of heavy rare-gas atoms such as argon, krypton and xenon.

Much of the foundation for the study of the Lennard-Jones problem was laid by the well-known results of Northby (1987). He established multilayer icosahedral conformations as the dominant structural motif for the optimal microclusters and produced global optima for all $N \leq 147$. Other authors (e.g. Xue, 1994; Deaven et al. 1996; Leary & Doye, 1999) have since found new configurations having lower energies than those of Northby for some values of N. Many of these studies employed hybrid genetic algorithms with local search. More recently (Hartke, 2001) the PHENIX method, also based on a GA, has extended the values of N successfully treated up to 250. It is believed that the lowest-energy configurations have now been determined for all $N \leq 250$.

3 PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization (PSO) method has evolved from a purely qualitative social optimization scheme to a truly numeric optimization scheme (Kennedy and Eberhart, 1995; Eberhart and Kennedy 1995). In its most applied form, the algorithm is designed to search for global optima in an n-dimensional search space of real numbers. An excellent review of the background and philosophy behind this method can be found in a recently published book (Kennedy and Eberhart 2001).

As in the more common Genetic Algorithm, a population of individuals is formed. Each individual in the PSO method is considered as a "particle" which is free to move about the search space. In the case of the Lennard-Jones problem each particle of the population is characterized by a set of 3N real numbers corresponding to the (x,y,z) positions of each atom. The fitness of a particle corresponds to the energy of the collection of N atoms, with the objective to make this energy as small as possible. In practice this energy is negative, corresponding to a 'bound' set of N atoms.

If particle i in the population, at time t, is represented as the vector $\vec{x}_i(t)$, then the change of position as the particle goes from one time step to the next is defined to be its velocity,

$$\vec{v}_i(t) = \vec{x}_i(t) - \vec{x}_i(t-1)$$

How this velocity changes at each time step is determined by the history of the particles past motion as well as that of its neighbours. This information is encapsulated in two parameters, the previous best position $\vec{p}_i$ for this individual particle, and the previous best position $\vec{p}_g$ for all those particles in the neighbourhood of this particular particle. We will define this neighbourhood later. The formulas to adjust the particle's velocity and position are then given by

$$\vec{v}_i(t) = w\vec{v}_i(t-1) + \varphi_1\left(\vec{p}_i - \vec{x}_i(t-1)\right) + \varphi_2\left(\vec{p}_g - \vec{x}_i(t-1)\right)$$

for the velocity, and

$$\vec{x}_i(t) = \vec{x}_i(t-1) + \vec{v}_i(t)$$

for the position. Here *w* is a weighting factor which is set to 0.9 at the beginning of the process, and decreases linearly to 0.4 at the end of the specified number of time steps (Shi and Eberhart 1998). The other two parameters, φ_1 and φ_2 are positive random numbers with upper bounds $\varphi_1^{\max}$ and $\varphi_2^{\max}$.

The other constraint on the system requires that the velocity is limited within a certain range so that all the particles will not escape from the search area. This is defined by limiting each velocity to $\pm v_{\max}$. Finally, the neighbourhood of each particle, mentioned above, refers to those particles which are adjacent in the population. For example a neighbourhood of 2 particles means that the particle in question has one neighbour on either side of it. A neighbourhood of 4 means that there are 2 particles on either side. It is understood that the population is considered to be a loop with the first and last particles in the population connected. The algorithm is then reasonably straightforward.

begin PSO
  g := 0 (generation counter)
  Initialize population P(g)
  Evaluate population P(g) (i.e. the cluster energy)
  while not done **do**
    g := g + 1
    Evaluate new velocities
    Evaluate new positions
    Determine local and global best
    Evaluate new population P(g)
  end while
end PSO

4 GENETIC ALGORITHMS

In order to do some sort of a comparison we have also carried out a number of calculations of the Lennard-Jones energies using a basic generic genetic algorithm, as well as a slightly more involved one. The first GA we refer to as BasicGA, which is described by the pseudocode shown below:

```
begin BasicGA
    g := 0 (generation counter)
    Initialize population P(g)
    Evaluate population P(g)
    while not done do
        g :=g+1
        Select P(g) from P(g-1)
        Crossover P(g)
        Mutate P(g)
        Evaluate P(g)
    end while
end BasicGA
```

This algorithm employed Roulette wheel selection, single-point crossover, and the mutation consisted of randomly modifying one parameter. The crossover probability was 0.8 and the mutation probability was 0.15.

In addition we ran a more general genetic algorithm based on the Genetic Algorithm Optimization Toolbox (GAOT) (Houck) which incorporated three different kinds of crossover functions and also three different mutation operators. Geometric selection was used, the crossover methods used were simple, arithmetic, and heuristic, and the mutation operators were boundary, uniform, and non-uniform (Michaleweiz 1992). In each generation two sets of parents were chosen for each crossover method, and four individuals were chosen for each of the mutation methods. Hence each generation is able to provide a wider search of the parameter space than in the case of the BasicGA above.

5 CALCULATIONS

In order to initially investigate the effect of the various parameters in the PSO calculation for the Lennard-Jones problem, we have focussed on a fairly simple case, that of the 8-atom cluster. As the best energies for clusters having up to 250 atoms are well known, we have examined the convergence of the iterations to the best known energy for this case (-19.82). Hence in the first part of this study we allow the iterations to proceed until the energy of the cluster (the best energy in the swarm) is within a small distance of the known value. Here we have required the value of this energy to converge to 4 significant figures.

The first calculations examined the effect of the neighbourhood size on the rate of convergence. We varied the size of the neighbourhood from 0 (meaning that all particles were in the neighbourhood) up to 14. The size of the swarm was kept at 30 particles. The maximum velocity was set at 0.2 and the bounds for $\varphi_1^{\max}$ and $\varphi_2^{\max}$ were both set at 2.0. If the swarm did not converge after 10,000 iterations, it was terminated. We repeated the runs 100 times and recorded the number of failures (i.e. the number of times that convergence was not achieved in the maximum allowed number of iterations), as well as the average of the number of iterations required for those cases where convergence was achieved. Figure 1 shows the behaviour of these two measures as a function of the neighbourhood size

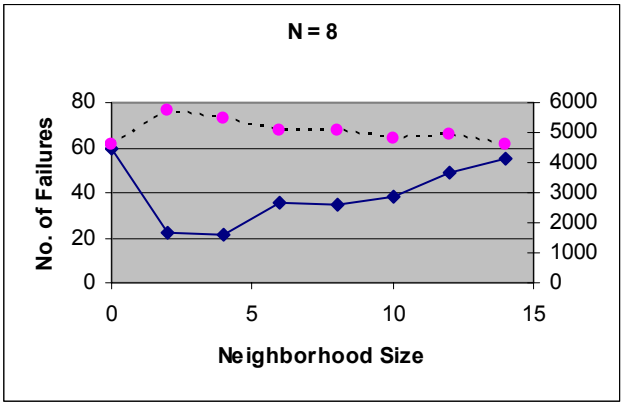

Figure 1: Variation with the neighbourhood size

The solid line in Figure 1 represents the number of failures, and the dashed line the average number of iterations. Here we see that the average number of iterations for the runs which converged did not vary

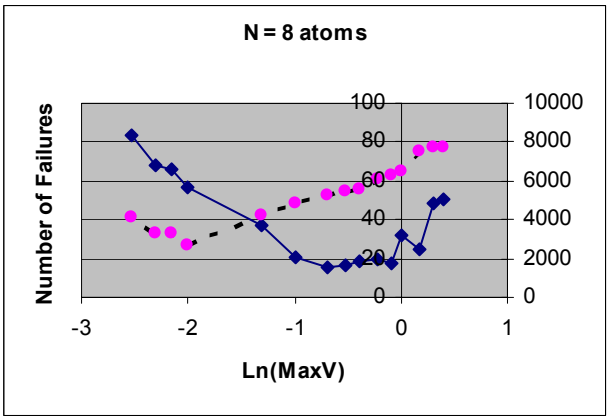

Figure 2: Variation with the log of Max V.

significantly from around 5,000. However there is a clear minimum for the number of failures when the neighbourhood size was either 2 or 4.

We proceeded next to investigate the effect of varying the value of the maximum velocity (MaxV).

The neighbourhood size was set to 4, and all other parameters were as above. The maximum velocity was varied between 0.003 and 2.5. With such a large range, we plotted the results as a function of the logarithm of the maximum velocity. The same two measures were used as above, and are shown in Figure 2.

The solid line, showing the number of failures for the 100 trials, shows a broad minimum with the number of failures to converge being less than or equal to 20 for MaxV in the range from 0.1 to 0.8. The number of average iterations has a minimum when MaxV equals 0.01. Hence again there is no common minimum. However, with the primary objective of achieving convergence, we have tended to focus more on the number of failures as opposed to the average number of iterations. The latter quantity does not vary over a large range. Selecting a maximum velocity anywhere in the range from 0.1 to 0.8 does not significantly affect the number of iterations to convergence. Hence for the next computation, we chose to select a value of 0.1 for MaxV.

The other variables which we looked at were those for the bounds φ_1^{max} and φ_2^{max}. Once again for the 8-atom cluster we took the neighbourhood size as 4, and the maximum velocity was set at 0.1. Setting the two bounds to be equal to each other, we obtained the results shown in Figure 3.

The value of 2.0 gives the best results in the sense that the number of failures is the least and the average number of iterations appears to increase with increasing φ_{max}.

While we recognize that there could be some dependence on the number of atoms and the nature of the energy surface in determining these results, we took these best parameters and ran calculations of cluster sizes for N varying from 4 to 15. We wanted to test the ability of the

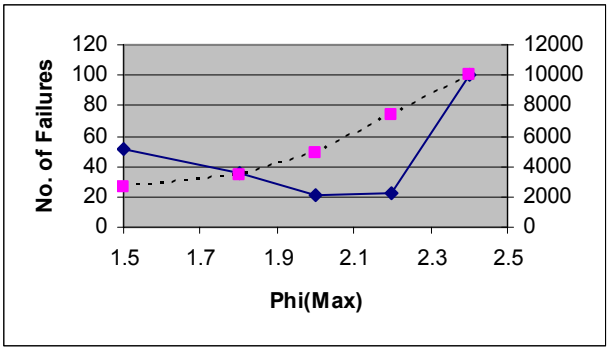

Figure 3: Variation with φ_{max}

PSO to achieve convergence over this range, and compare the results with fairly simple Genetic Algorithms applied to the same problem. In addition to looking for convergence, we also kept track of the number of function evaluations required for each run. In all cases we took the number of particles to be 30 and ran for 20,000 iterations. 100 trials were run for each value of N and the iterations were terminated if the energy matched the best known value for that value of N, up to 4 significant figures. If

Table 1.

		PSO			GA			
N	**Cluster Energy**	**No. Of Failures**	**Av. No. iterations**	**Av. No. Function Evaluations**	**No. Of Failures**	**Av. No. Generations**	**Av. No. Function Evaluations**	**Best Energy (GA)**
4	-6	0	2500	75008	57	11634	150016	
5	-9.104	0	5982	179487	95	16464	212581	
6	-12.71	96	7439	223177	100	20000	257500	-12.69
7	-16.51	52	8547	256424	100	20000	256720	-16.49
8	-19.82	29	9346	280388	100	20000	257482	-19.76
9	-24.11	67	10143	304293	100	20000	258916	-23.99
10	-28.42	91	9428	282846	100	20000	256877	-28.16
11	-32.77	95	11309	339294	100	20000	257706	-30.68
12	-37.97	96	10728	321855	100	20000	258336	-36.89
13	-44.33	100	20000	600000	100	20000	256950	-41.94
14	-47.84	99	16704	501120	100	20000	258426	-42.95
15	-52.32	97	16885	506550	100	20000	258444	-48.92

the run did not converge to this limit under the 20,000 iteration limit, then the run was counted as a failure.

Table 1 provides a summary of these results for N = 4 to 15. The second column is the known lowest energy for each of these cluster sizes (Leary, 1997). The next three columns show the PSO results: the number of failures to converge out of the 100 total trials; the average number of iterations required for those trials which did converge; and the average number of function evaluations used in the converged runs. The number of failures to converge is also graphed in Figure 4 as a function of N. Here two plots are shown, the solid line shows runs which were terminated after 10,000 iterations if convergence was not achieved, whereas the dashed line shows the same results for 20,000 iterations. It can be seen that there is not a great deal of difference between these two results.

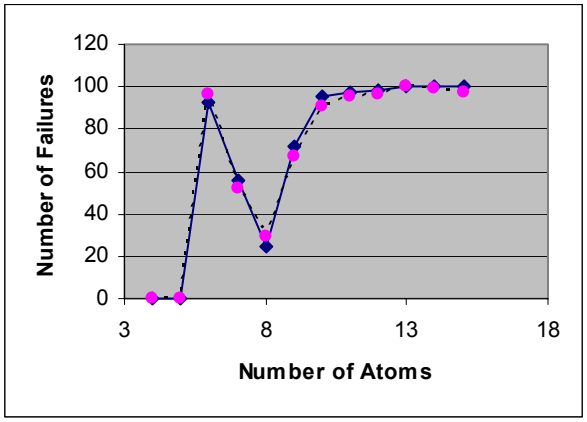

Fig. 4. Failures vs N

Table 1 also shows the results of runs using the modified GA. The basic GA was not successful in the sense that after 500,000 generations, convergence could not be achieved even for N = 4 atoms. Limiting the runs to 20,000 generations and using a population size of 30, the modified GA did find the known energies for N=4 and N=5, but was unsuccessful in 100 trials for the other values of N. The best energies achieved are shown in the table, and for the smaller values of N are relatively close to the known energies, indicating that allowing the GA to continue would likely achieve convergence.

The PSO method was able to achieve convergence in at least one of the trials for all values of N studied here, except for N = 13. In many cases only a few of the trials converged, seeming to indicate that the particle "swarm" could not get close enough to the fixed point in the set number of iterations.

6 SUMMARY AND CONCLUSIONS

The Particle Swarm Optimization method has been applied to a variety of areas since its introduction only a few years ago. As far as we are aware this is the first time that it has been applied to the energy conformation problem for atomic clusters. There has also been some other work on comparing the PSO with genetic algorithms (Eberhart and Shi 1998; Angeline 1998).

Genetic algorithms incorporate selection, crossover and mutation schemes in order to search the parameter space. In the PSO there is no specific selection process, however each individual carries with it a copy of its personal best value, which serves a somewhat similar role to that of a parent. The offspring of an individual is a function of this best value. The PSO is the only evolutionary algorithm that does not incorporate selection of the fittest.

The role of the crossover function in the GA is to select information from parents (usually two) to create offspring. In the PSO the influence on one particle by the others comes only in the value of the best position of the particles in the defined neighbourhood.

Mutations are an important aspect of the GA process in that they help to break out of the genetic sequence generated by the parents and offspring. However a common limitation of the GA is that as the population converges, the average fitness value becomes high so that mutations will usually result in a low-fitness chromosome which will be rejected by the selection process. The result is that the process may converge to a local optimum instead of finding the global one. There are variations which try to circumvent this, including the incorporation of local optimization methods.

Particle swarm uses a highly directional mutation operation as each individual's velocity vector is modified using a vector whose direction lies between the personal best and the neighbourhood best. As a consequence the PSO may have difficulties when the average local gradients point away from the global optima or are constantly changing.

Our calculations here have shown that the PSO is indeed an effective optimization method. Significantly better results have been found for the location of the global optimum energy value for a cluster of atoms interacting via the Lennard-Jones potential than for the case of a relatively generic GA. The potential energy surfaces for these problems are known to contain large numbers of local minima, often very close to the global minima for certain values of N. Hence it is not unexpected to find that many of the runs converge to one of these local minima.

The use of a low value of the maximum velocity in our calculations undoubtedly resulted in slow convergence in most cases, however it was somewhat necessary in the sense that the parameter values for the sizes of clusters studied here are in the range [-1, 1]. Small variations in these parameter values can shift the energy significantly, so that it is important to search over a relatively fine mesh.

The runs with the genetic algorithms quickly converged to a limit, and then tended to stay near that value for most of the subsequent generations. Here it would help to have

more flexible mutation and selection schemes to prevent this from happening.

Future work will focus on studying possible variations to the PSO which can help improve the success rate for this type of problem.

Acknowledgments

This investigation was supported in part by the National Science and Engineering Research Council of Canada (NSERC).

References

Angeline, P.J. "Particle Swarm - Evolutionary Optimization Versus Particle Swarm Optimization: Philosophy and Performance Differences", Lecture Notes in Computer Science **1447** (1998) 601-610.

Barron, C., Gomez, S., Romero, D. and Saaverdra, A. "A genetic algorithm for Lennard-Jones atomic clusters", App. Math. Letts. **12** (1999) 85-90.

Deaven, D.M. and Ho, K.M. "Molecular Geometry Optimization with a Genetic Algorithm", Phys. Rev. Letts. **75** (1995) 288-291.

Deaven, D.M., Tit, N., Morris, J.R. and Ho, K.M. "Structural optimization of Lennard-Jones clusters by a genetic algorithm", Chem. Phys. Letts. **256** (1996) 195-200.

Doye, J.P.K., Miller, M.A. and Wales, D.J., "The double-funnel energy landscape of the 38-atom Lennard-Jones cluster", J. Chem. Phys. **110** (1999) 6896-6906.

Eberhart, R.C. and Kennedy, J. "A new optimizer using particle swarm theory", Proc. Sixth Int. Symp. on Micro Machine and Human Science (Nagoya, Japan) (1995) 39-43.

Eberhat, R.C. and Shi, Y. "Comparison between Genetic Algorithms and Particle Swarm Optimization", Lecture Notes in Computer Science **1447** (1998) 611-618.

Hartke, B. "Global Geometry optimization of clusters using genetic algorithms", J. Phys. Chem. **97** (1993) 9973-9976

Hartke, B. "Global geometry optimization of atomic and molecular clusters by genetic algorithms", Proc. Of the Genetic and Evolutionary Computation Coference (GECCO) 2001, 1284-1291.

Hoare, M. " Structure and dynamics of simple microclusters", Adv. Chem. Phys. **40** (1979) 49-135.

Hoare, M.R. and Pal, P. Adv. Physics **20** (1971) 161.

Houck, C.R., Joines, J.A. and Kay, M.G. Genetic Algorithm Optimization Toolbox (GAOT) for Matlab. http://www.ie.ncsu.edu/mirage/GAToolBox/gaot/

Kennedy, J. "The particle swarm: social adaptation of knowledge", Proc. 1997 IEEE Int. Conf. on Evolutionary Computation, 303-308.

Kennedy, J. and Eberhart, R.C. "Particle swarm optimization", Proc. 1995 IEEE Int. Conf. on Neural Networks (Perth, Australia) 1942-1948.

Kennedy, J. and Eberhart, R.C. "Swarm Intelligence", Morgan Kaufmann, 2001.

Leary, R.H. "Global Optima of Lennard-Jones Clusters", J. Global Optim. **11** (1997) 35-53

Leary, R.H. and Doye, J.P.K. "Tetrahedral global minimum for the 98-atom Lennard-Jones cluster", Phys. Rev. **E 60** (1999) R6320-22.

Michalewicz, Z. (1992) "Genetic Algorithms + Data Structures = Evolution Programs" Springer-Verlag,.

Niesse, J.A. and Mayne, H.R. "Global geometry optimization of atomic clusters using modified genetic algorithm in space-fixed coordinates", J. Chem. Phys. **105** (1996) 4700-4706.

Northby, J.A., "Structure and binding of Lennard-Jones clusters: 13 ≤ N ≤ 147", J. Chem. Phys. **87** (1987) 6166-6177.

Peterson, C., Sommelius, O. and Soderberg, B. "Variational appraoch for minimizing Lennard-Jones energies", Phys. Rev. **E 53** (1996) 1725-1731.

Radcliffe, N.J. and Surry, P.D., "Formal Memetic Algorithms", Evolutionary Computing. AISB Workshop. Selected Papers. Springer-Verlag. 1994, pp.1-16. Berlin, Germany.

Shi, Y. and Eberhart, R.C. "Particle Swarm - Parameter Selection in Particle Swarm Optimization", Lecture Notes in Computer Science **1447** (1998) 591-600.

Stillinger, F.H., Phys. Rev. E **59**, 48 (1999).

Tsai, C.J. and Jordan, K.D., J. Phys. Chem. **97**, 11 227 (1993).

Wales, D.J. and Doye, J.P.K. "Global optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters up to 110 Atoms", J. Phys. Chem. A 101 (1997) 5111-5116.

Wille, L. and Vennik, J., "Computational Complexity of the ground-state determination of atomic clusters", J. Phys A 18 (1985) L419-L422.

Wolf, M.D. and Landman, U. "Genetic Algorithms for Structural Cluster Optimization", J. Phys. Chem. A 102 (1998) 6129-6137.

Xue, G. "Improvement on the Northby Algorithm for Molecular Conformation: Better Solutions", J. Global Optim. **4** (1994) 425-440.

Zeiri, Y., "Prediction of the lowest energy structure of clusters using a genetic algorithm", Phys. Rev. E **51**, R2769 (1995).