

Generalized simulated annealing: Application to silicon clusters

M. R. Lemes, C. R. Zacharias,* and A. Dal Pino, Jr.

Instituto Tecnológico de Aeronáutica, Centro Tecnológico da Aeronáutica, São José dos Campos, Brazil 12228-900

(Received 12 June 1997)

We have compared the recently introduced generalized simulated annealing (GSA) with conventional simulated annealing (CSA). GSA was tested as a tool to obtain the ground-state geometry of molecules. We have used selected silicon clusters (Si_n , $n=4-7,10$) as test cases. Total energies were calculated through tight-binding molecular dynamics. We have found that the replacement of Boltzmann statistics (CSA) by Tsallis's statistics (GSA) has the potential to speed up optimizations with no loss of accuracy. Next, we applied the GSA method to study the ground-state geometry of a 20-atom silicon cluster. We found an original geometry, apparently lower in energy than those previously described in the literature. [S0163-1829(97)06239-5]

To predict ground-state geometries of molecular systems, one must determine the global minimum of the total energy of the system. It is well known that, even for simple systems, the number of local minima increases exponentially with the number of atoms. For instance, a molecule composed of 20 atoms, interacting through a Lennard-Jones pair potential, has an estimated number¹ of 10^8 local minima. Clearly, geometry optimization of realistic systems represents a very difficult optimization problem. Local minimization procedures such as conjugate gradient² and steepest descent³ are not efficient to solve this class of problems because they are not designed to overcome potential barriers.

Methods based on the simulation of natural processes, such as simulated annealing⁴ and genetic algorithm⁵ have emerged as promising tools to solve NP optimization problems.⁶ The simulated annealing (SA) algorithm is a numerical simulation method based on the dynamics of crystallization. Consider heating a solid until its constituents can move freely and it melts. Then the melt is allowed to cool very slowly until it solidifies in a certain arrangement. The heating and slow cooling can be repeated many times to enhance the chance to find the configuration corresponding to the global energy minimum. SA was created by Kirkpatrick, Gellar, and Vecchi⁴ to search for the global extremum of generic objective functions by simulating the annealing process. Geman and Geman⁷ proved that if the temperature is reduced slowly enough conventional simulated annealing is guaranteed to find the global minimum. Unfortunately, such a slow pace makes the method inapplicable to search for the ground-state geometry of molecules with even a small number of atoms.

Recently, Penna⁸ has introduced another method, generalized simulated annealing (GSA). This procedure is based on Tsallis's⁹ statistics. Penna proved that the recently developed method could find a near optimum path for the "traveling salesperson" problem efficiently, i.e., faster than the conventional SA (CSA). Here we have performed a comparison between CSA and GSA in order to determine whether or not GSA is a convenient optimization procedure to predict the ground-state geometry of molecular systems. Specifically, we chose silicon clusters (Si_n , $n=4-7$ and 10) as test cases. GSA performed extremely well in these test cases. Next we used GSA to perform geometry optimization of Si_{20} . We

were able to find an original geometry that, according to our tight-binding model, is lower in energy than those previously described in the literature.

Prediction of ground-state geometries of large molecules is hindered by time-consuming procedures, namely, the quantum-mechanical calculation of total energies and the large number of local minima. Currently, simulated annealing is one of the most promising techniques to tackle this class of problems because transitions out of local minima are allowed. To make use of SA, one must provide the following elements: (i) a criterion to describe possible system configurations, specifically, for molecular geometry optimization, simply a set of atomic coordinates that can be either Cartesian or internal; (ii) a generator of changes in configuration, in our case, random changes in atomic coordinates; (iii) an objective function, whose optimization is the goal of the procedure (here the total energy E has to be minimized); and (iv) a fictitious temperature T as well as an annealing schedule that controls how T is reduced during optimization. Here CSA was implemented using a stepwise-exponential decrease of temperature. For a fixed temperature T , the atomic coordinates were changed according to a Gaussian distribution whose standard deviation is α . Whenever five consecutive steps were rejected, the value of α was halved. This proceeds until α reaches $\alpha_0/16$; then the temperature is readjusted to a lower value according to an exponential cooling schedule and α returns to α_0 .

According to CSA, a new configuration was promptly accepted if it produced a decrease in total energy. Otherwise, Boltzmann¹⁰ statistics is applied, i.e., a higher-energy configuration could be accepted with probability

$$P \propto \exp(\Delta E/KT).$$

The pseudo-Boltzmann factor ($K = -1/\ln 0.5$) was adjusted to produced 50% acceptance whenever $\Delta E = T$. This way we associate some physical meaning with the fictitious temperature T . The CSA's performance is strongly dependent on parameters such as the initial and final fictitious temperatures, cooling scheme, and step length.¹¹

On the other hand, GSA takes advantage of Tsallis's statistics.¹² Thus the rearranged configuration is accepted with probability

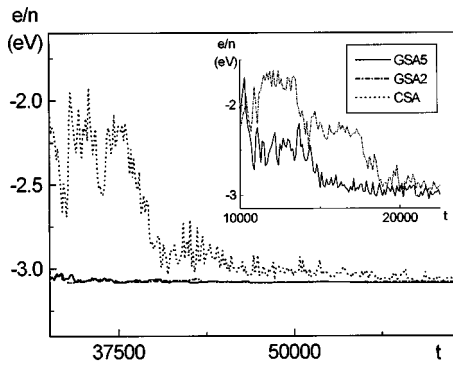


FIG. 1. Time evolution of the geometry optimization of Si_6 using the CSA (dotted line), GSA2 (dashed line), and GSA5 (solid line). The cohesive energy per atom as a function of time (total-energy calculation is shown). The inset shows the early stages of the simulation to stress the differences between the GSA2 and GSA5.

$$P = [1 - (1 - q)\Delta E/kT]^{1/(1-q)}.$$

One can prove that in the limit for $q \rightarrow 1$, Boltzmann statistics is recovered. Penna⁸ has pointed out that values of $q < 1$, accelerate the optimization procedure, but large absolute values of q lead to large roundoff errors. Therefore, to obtain a compromise between speed and accuracy, we chose to test $q = -2$ (GSA2) and $q = -5$ (GSA5). In order to maintain consistency, the same procedure was applied to reduce temperature and to generate new configurations for both CSA and GSA.

Specifically, we chose small silicon clusters (Si_n , $n = 4-7$ and 10) to test GSA. To make the test particularly stringent, we have selected a starting geometry very different from the (known) ground-state configuration. Thus, as it is well known that silicon clusters form rather compact structures, in-line geometries were chosen as starting points for

the optimization procedures. The binding energy was calculated by the tight-binding¹³ model because it combines a quantum-mechanical treatment of the electronic structure with a very efficient computational procedure. Notice that our optimizations are limited to the realm of this model. The most time-consuming step in the procedure is the total-energy calculation; thus we use one such calculation as the time unit. Figure 1 shows the time evolution of the simulation for Si_6 . It is clear that both GSA2 ($q = -2$) and GSA5 ($q = -5$) find the ground-state configuration faster than CSA. Figure 1 also shows a distinct oscillatory pattern for each simulation. More frequent and stronger oscillations mean higher rates of acceptance of rearrangements that increase the energy. Clearly, CSA accepts this kind of move more frequently than GSA2 or GSA5. This is relevant because it defines the change of the algorithm to avoid being trapped in a local minimum, but it slows down the calculation. The inset permits a comparison between GSA2 and GSA5. Particularly in the early stages of the optimization procedure, GSA2 approaches the desired minimum more slowly because it clearly oscillates more often than GSA5. This pattern is typical for all the tests performed. As the number of minima depends dramatically on the number of atoms, it is necessary to test the performance of GSA for other clusters.

Table I compares the time required to obtain a certain energy level (50%, 75%, or 90% of the ground-state energy) with CSA, GSA2, and GSA5 for Si_n , $n = 4-7, 10$ clusters. Due to the random character of simulated annealing, we only considered that an energy level was reached if the mean energy remained below that level for at least 300 time steps.

For all tests GSA2 and GSA5 outperformed CSA, i.e., they find the global minimum faster. As GSA2 has a higher acceptance rate of rearrangements that increase the energy than GSA5, it converges slower than GSA5. All entries in Table I represent the average of a set of ten simulations, each

TABLE I. Results obtained during the geometry optimizations of Si_n , $n = 4-7$ and 10.

Cluster	Method	Number of time steps required to reach a percentage of ground-state energy		
		50%	75%	90%
Si_4	CSA	1900	21 000	28 000
	GSA2	1300	7300	13 100
	GSA5	500	2400	6800
Si_5	CSA	17 000	24 500	32 000
	GSA2	2600	13 000	18 200
	GSA5	1100	11 200	18 900
Si_6	CSA	13 200	33 300	39 700
	GSA2	2200	15 300	18 700
	GSA5	800	10 600	15 200
Si_7	CSA	20 500	34 200	36 400
	GSA2	4500	19 200	21 900
	GSA5	4200	14 400	21 000
Si_{10}	CSA	7900	36 600	48 300
	GSA2	4100	23 400	30 200
	GSA5	200	16 200	22 000

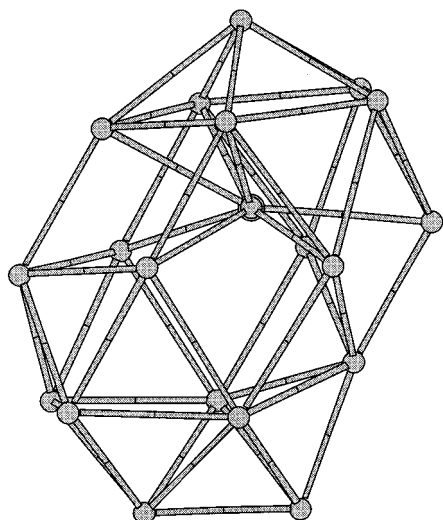


FIG. 2. Alternative structure of the ground-state geometry of a 20-atom silicon cluster.

one using a different random sequence. The three methods showed a very high (>80%) efficiency to obtain the ground-state energy (geometry) for the test systems. This indicates that the performance enhancement obtained with GSA is not followed by a noticeable decrease in efficiency.

Next, we submitted GSA to a more demanding calculation. We used GSA to find the ground-state geometry (according to the tight-binding model) of Si_{20} . This is a particularly interesting problem because structural candidates previously have been studied through *ab initio* methods.^{14–16} As *ab initio* methods are CPU time consuming, these models were inspired in some kind of physical insight, specifically symmetry constrains,¹⁴ bulk structure,¹⁵ or reconstruction surface analogies.¹⁶ This restricts drastically the configuration hyperspace scanned. Unfortunately, it is well known that an unbiased procedure is fundamental in searching for the ground-state geometry. GSA coupled to tight-binding model for silicon may give this unbiased perspective.

The GSA algorithm was able to find a different minimum-energy structure for Si_{20} , shown in Fig. 2. It is formed by

TABLE II. Relative total energy per atom differences. The proposed models were relaxed under their respective symmetry constraints through the same tight-binding method. We set the energy of the present geometry model to zero to compare with those previously proposed.

Model	Differences in total energy per atom (eV)
1 ^a	0.68
2 ^b	0.46
3 ^c	0.25
This work	0

^aReference 16.

^bReference 15.

^cReference 14.

three puckered planes with five, six, and five silicon atoms, respectively, capped by two atoms on the bottom and one on the top. Furthermore, it contains an inner atom close to the six-atom plane.

This structure is an alternative to the ground-state geometry of a 20-atom silicon cluster. Using the tight-binding approach, we compared the present structure with those previously proposed (see Table II). The proposed model represents the lower-energy configuration for a 20-atom silicon cluster within the tight-binding model. Although this method is not *ab initio*, it has been able to describe the main features of silicon systems.

In conclusion, we have shown that the recently introduced GSA is faster than CSA in predicting the ground-state configuration of silicon clusters. Also, we have tested the capability of GSA to obtain an alternative for the ground-state geometry of Si_{20} . Our method proved to be efficient in scanning intelligently a wide configuration hyperspace, moving away from deep local minima to find a different structural model for a 20-atom silicon cluster. This model is, within the limits of our total-energy calculation method, lower in energy than previously proposed models.

M.R.L. acknowledge support from the Conselho Nacional de Pesquisa e Desenvolvimento.

*Also at Department of Physics, UNESP, Guaratinguetá, Brazil, 12500-000.

¹Y. Kostrowicki and H. A. Scheraga, *J. Phys. Chem.* **95**, 4114 (1991).

²E. Polak, *Computational Methods in Optimization* (Academic, New York, 1971).

³W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

⁴S. Kirkpatrick, D. C. Gellar, and M. P. Vecchi, *Science* **220**, 671 (1983).

⁵D. G. Bounds, *Nature (London)* **329**, 215 (1987).

⁶L. T. Wille and J. Venik, *J. Phys. A* **18**, L419 (1985).

⁷S. Geman and D. Geman *IEEE Trans. Pattern. Anal. Mach. Intell. PAMI-6*, 721 (1984).

⁸T. J. P. Penna, *Phys. Rev. E* **51**, R1 (1995).

⁹C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).

¹⁰N. Metropolis, A. Rosenbluth, M. Rosenbuth, M. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

¹¹F. Montonya and J. M. Dubois, *Europhys. Lett.* **22**, 79 (1993).

¹²D. A. Stariolo and C. Tsallis, in *Annual Review of Computational Physics II*, edited by D. Stauffer (World Scientific, Singapore, 1994).

¹³K. Laasonen and Nieminen, *J. Phys. C* **2**, 1509 (1990); M. Menon and K. R. Subbaswamy, *Phys. Rev. B* **47**, 12 754 (1993).

¹⁴J. C. Grossman and L. Mitos, *Phys. Rev. Lett.* **74**, 1323 (1995).

¹⁵E. Kaxiras and K. A. Jackson, *Z. Phys. D* **26**, 346 (1993).

¹⁶E. Kaxiras and K. A. Jackson, *Phys. Rev. Lett.* **71**, 727 (1993).