Abstract

One of the methods to find the global minimum of a potential energy surface of a molecular system is simulated annealing. The main idea of simulated annealing is to start your system at a high temperature and then slowly cool it down so that there is a chance for the atoms in the system to explore the different degrees of freedom and ultimately find the global minimum. Simulated annealing is traditionally used in classical Monte Carlo or in classical molecular dynamics. In molecular dynamics, one of the traditional methods was first implemented by Woodcock in 1971. In this method the velocities are scaled down after a given number of molecular dynamics steps, let the system explore the potential energy surface and scale down the velocities again until a minimum is found. In this work we propose to use a viscous friction term, similar to the one used in Langevin dynamics, to slowly bring down the temperature of the system in a natural way. We use drag terms that depend linearly or quadratically on the velocity of the particles. These drag terms will naturally bring the temperature of the system down and when the system reaches equilibrium they will vanish. Thus, imposing a natural criterion to stop the simulation. We tested the method in Lennard-Jones clusters of up to 20 atoms. We started the system in different initial conditions and used different values for the temperature and the drag coefficients and found the global minima of every one of the clusters. This method demonstrated to be conceptually very simple, but very robust, in finding the global minima.

1 Introduction

Geometry optimization of molecular systems in chemistry, physics and biology is a fundamental task that is needed to describe these systems and the properties that strongly depend on their structures. There is a wealth of methods devoted to this task ranging from gradient based steepest descent methods (Pulay, 1987) (Biring & Chaudhury, 2012), to statistically based simulated annealing (Kirkpatrick, Gelatt Jr., & Vecchi, 1983) (Xue, Luo, Shen, Li, & Zhao, 2010), to genetic algorithms (Kanters & Donald, 2014) (Johnston, 2003). The conjugated gradient based methods tend to be fast but there is no guarantee to reach the global minimum and one often finds one of the local minima. Simulated annealing is a method that is often used in the search for global minima and that is suitable for minimization of large scale problems. Simulated annealing is based on mimicking the process of cooling a metal, but its application goes way beyond geometry optimization. It can be applied to any minimization problem in which one can design a cost function. One of its successes is an effective
“solution” to the traveling salesman problem. Another popular method for optimization problems in general, is the use of genetic algorithms. In this class of methods one uses a class of initial guesses, the genes of the parents represented as binary strings, that generate new guesses (offsprings) that combine attributes (genes) from each parent and the principle of survival of the fittest is applied. There are several methods that are used as applied to geometry optimization. For clusters, in particular, the offsprings are generated by a crossover of the structure of each of the parents to create the offsprings. In (Johnston, 2003) the crossover process simply combines the parts of two parent clusters. The genetic algorithms based on this geometric crossover mechanism work well for very symmetric systems but are not very robust in finding more complicated structures. Other tools have been used in helping understand the topology of the potential energy surface as well as to help find the global minima. For instance, disconnectivity graphs (Czerminski & Elber, 1990) (Wales, 2010) are a 2-dimensional representation of the potential energy surface where the height vertical lines represent the energy of local minima. The minima are then connected at the highest point which represents the barrier that the system has to overcome to go from one minimum to another. For more effective minimizations a combination of multiple of the above techniques can be used.

In this work we propose to use molecular dynamics in a way similar to the simulated annealing or simulated quenching methods. We use an artificial viscous force that depends either linearly or quadratically with the velocity of the particles, in a way that is similar to the Langevin dynamics used to study Brownian motion of particles in fluids. For benchmarking purposes we apply it to the optimization of Lenard-Jones clusters as the global minima are readily available from the Wales group (Wales, et al., 1996). We also discuss how the use of a linear or quadratic dissipative force resembles the normal and exponential schedule of cooling in simulated annealing.

## 2 Methods

In this work we use classical molecular dynamics using the velocity Verlet algorithm (Verlet, 1967) (Verlet, 1968) to search for the global minima of Lenard-Jones clusters of up to 20 atoms. The use of molecular dynamics for geometry optimization is associated in general with simulated annealing or simulated quenching. The idea is to start the system at a given temperature and let it explore the potential energy surface. The system will spend more time near the deeper minima, hopefully the global minima. Then, similar to what is done in metallurgy the system is slowly cooled down by scaling the velocities of the particle. In Monte Carlo applications of simulated annealing there are typically two schedules for the temperature scaling, a logarithm one which is consistent with the Boltzmann distribution,

\[
T_k = T_0 \frac{\ln k}{\ln k_0},
\]

where, \(T\) is the temperature and \(k\) represents the step in the cooling schedule, and \(k_0\) is the starting step. This cooling schedule tends to be very long and therefore other cooling schedules have been proposed to speed up the optimization process. An exponential schedule has the advantage of being faster, but at the cost that there are no theoretical reasons to always delivers a true global minimum (Ingber, 1993). In this work we propose a quenching mechanism that has a physical interpretation by including a dissipative force in the dynamics, so that the total force on the system will be given by:

\[
\vec{F} = \sum_i - c_1 \vec{v}_i - c_2 \vec{v}_i \vec{v}_i - \nabla_i \mathcal{U}(r_1, r_2, ..., r_n),
\]

Where, the first two terms are analogous to the linear and quadratic drag of a viscous fluid and the third term is the force derived from the potential energy. It is clear that there is no energy conservation...
due to the dissipative nature of the viscous forces. The simplicity of the method is illustrated in Figure 1, where we show that the tendency of the dynamic represented by the forces in Eq. (2) is that the system will naturally fall in a minimum. The velocity of the particles will naturally decrease due to the viscous forces and bring the system to a state of zero forces that will be a minimum. In order to increase the chances to fall into a global minimum we employ the same strategies of simulated annealing or quenching. We bring the system to a high temperature and allow it to cool down. The cooling schedule is controlled by the value of the parameters $c_1$ and $c_2$. One of the big differences of this method with simulated annealing is that the cooling is taking place at each molecular dynamic step rather than every so often. The linear drag cooling schedule follows an exponential decay that can be derived by analytically solving the problem for a free particle with a given initial speed that enters a region of linear drag. If one considers only the quadratic drag, the cooling schedule will follow a $1/t^2$ dependence.

![Figure 1](image_url)

**Figure 1** – Schematics of the minimization process under the drag assisted simulated annealing method. The white circles represent the total energy of the system, while the black ones represent the potential energy. Each arrow represents a time step, or a series of steps. Upon convergence the total energy is equal to the potential energy of the systems. $\{R\}$ represents the set of the positions of all atoms in the system.

The use of a drag or viscous force resembles the work of Biswas and Hamman where they combined Langevin molecular dynamics with simulated annealing to optimize clusters of silicon (Biswas & Hamann, 1986). This is where the analogy ends, as the use of the viscous term in Langevin molecular dynamics is to achieve ergodicity by applying a rapidly fluctuating force and a viscous damping force. The careful balance of these two forces assures that the mean square velocity has the proper asymptotic value proportional to $k_B T$. In this work the role of the drag term is to slow the particles down and bring the system to a minimum of the potential energy of the system. The natural application of the method is to obtain the equilibrium geometry of molecules. However, like many minimization techniques, the method can be generalized to any optimization problem by
creating a fictitious dynamics where the forces are going to be a gradient of the cost function you are trying to optimize.

We implemented both the linear and quadratic forces and tested on Lenard-Jones clusters in MD units, \( m=k_\beta=\epsilon=\sigma=1 \). We studied the dependence with the values of \( c_1 \) and \( c_2 \), and determined the strategies to find the global minima. We started from different initial guesses to test the robustness of the method. In the next section we present our results and discussion.

3 Results

In Table 1 and in Figure 2 we present the energy and point group symmetry of Lenard-Jones clusters of up to 20 atoms. The global minima are in complete agreement with those from the Wales group database and references therein (Wales, et al., 1996). Analysis of Figure 2 shows that in many cases the cluster with \( n+1 \) atoms can be seen as the \( n \) atom cluster plus 1 atom. The sequences where this takes place are \( n = 2-5, n = 7-17, n = 18-20 \). This kind of trend has been observed in most metallic clusters. It is then natural when trying to find a global minimum to explore the addition of an atom in different sites of the \( n \) atom cluster.

| N  | Point Group | Energy (This work) | Energy (Exact) |
|----|-------------|--------------------|----------------|
| 2  | D\(_{4h}\)   | -1.000000          | -1.000000      |
| 3  | D\(_{3h}\)   | -3.000000          | -3.000000      |
| 4  | T\(_d\)      | -6.000000          | -6.000000      |
| 5  | D\(_{3h}\)   | -9.103852          | -9.103852      |
| 6  | O\(_h\)      | -12.712062         | -12.712062     |
| 7  | D\(_{5h}\)   | -16.505384         | -16.505384     |
| 8  | C\(_s\)      | -19.821489         | -19.821489     |
| 9  | C\(_{2v}\)   | -24.113360         | -24.113360     |
| 10 | C\(_{3v}\)   | -28.422532         | -28.422532     |
| 11 | C\(_{2v}\)   | -32.765970         | -32.765970     |
| 12 | C\(_{5v}\)   | -37.967600         | -37.967600     |
| 13 | I\(_h\)      | -44.326801         | -44.326801     |
| 14 | C\(_{3v}\)   | -47.845157         | -47.845157     |
| 15 | C\(_{2v}\)   | -52.322627         | -52.322627     |
| 16 | C\(_s\)      | -56.815742         | -56.815742     |
| 17 | C\(_2\)      | -61.317995         | -61.317995     |
| 18 | C\(_{5v}\)   | -66.530949         | -66.530949     |
| 19 | D\(_{5h}\)   | -72.659782         | -72.659782     |
| 20 | C\(_{2v}\)   | -77.177043         | -77.177043     |

In order to assess the ability of the method to reach the global minimum we started the systems from different configurations and we did not use the common practice of starting the guess of the cluster with \( n+1 \) atoms by adding an atom to different positions of the cluster with \( n \) atoms. To illustrate this fact we present the starting and final configurations of a few of the clusters in Figures 3-6. In Figure 3 we present the initial guesses for the 4 atom cluster. We started from two different planar structures and in both cases they converged to the global minimum in \( 2.5 \times 10^6 \) molecular dynamics steps. In this case we started at a temperature in MD units of 0.04 and a linear drag coefficient of 0.6. The MD simulation ran continuously until the system reached equilibrium. In
traditional simulated annealing at each temperature the system would run for a few thousand steps and after it reaches equilibrium at that temperature the velocities would be scaled down to a corresponding lower temperature. The process would be repeated until the temperature is below a certain threshold.

**Figure 2** - Global minima of Lennard-Jones clusters of up to 20 atoms.

In Figure 4, we present two different initial guesses for the 7 atom cluster. Once again we started from planar configurations. The initial temperatures were between 0.04 and 0.06 MD units and the linear drag coefficients were between 0.6 and 0.9. Both cases converged in fewer than $2 \times 10^6$ MD steps. In Figure 5, we show three different guesses for the 9 atom cluster that converged directly to the global minimum. Once again we started with planar geometries. The initial temperatures were also in the range of 0.04 to 0.06 and the linear drag coefficients were between 0.6 and 0.9. This was the last case in which we used planar structures as initial guesses. For cluster with more than 9 atoms the planar configurations were too energetically unfavorable and the cluster would dissociate into smaller fragments. In Figure 6 we present one of the initial configurations of the 20 atom cluster that
converged directly into the global minimum. The initial configuration was 3 dimensional but was completely unrelated to the global minimum. Once again the method converged into the global minimum. In this case we used both a linear and a quadratic drag term. The quadratic term will dominate in the high temperature cases, while the linear term is responsible for the final convergence of the method. We used values of 0.1 and 0.01 for \( c_1 \) and \( c_2 \), respectively. This drag assisted implementation of simulated annealing has components of the traditional simulated annealing and the steepest descent methods. In the initial steps, at high temperatures, the method will allow for a full exploration of the potential energy surface, as the temperatures drops, due to the viscous terms added in Eq. (2), the method will converge to the nearest minimum as represented in Figure 1.

![Figure 3](image1.png)

**Figure 3** - Different initial guesses a) and b) and the final converged c) structure of the 4 atom cluster.

![Figure 4](image2.png)

**Figure 4** - Different initial guesses a) and b) and the final converged c) structure of the 8 atom cluster.

![Figure 5](image3.png)

**Figure 5** - Different initial guesses a), b) and c) and the final converged d) structure of the 9 atom cluster.
As indicated in the last paragraph, if the initial temperature is not high enough, or the drag term is too big, the method might converge to a local minimum. In Figure 7 we present one of the few cases we encountered in this study where we found a local minimum before the method converged to the global one. Figure 7a) shows a different planar structure used as initial guess for the 9 atom cluster. This led to a $C_2v$ local minimum that is a bicapped pentagonal bipyramid. This configuration has an energy of -23.1708 which is higher than the value of -24.113360 of Table 1. The initial temperature was 0.08 and we used a linear drag coefficient of 0.6. To reach the global minimum we used the local minimum as an initial guess and then used a higher initial temperature. In this case we used a temperature of 1 and kept the same drag coefficient. The initial MD run used $1.5 \times 10^6$ MD steps and the additional run used $1.5 \times 10^6$ additional steps to converge to the global minimum.

In Figure 8, we show another case that converged to a local minimum. In Figure 8a) we have a 3 dimensional initial guess of the 13 atom cluster. The simulations started with an initial temperature of 5 and a drag coefficient of 0.12, and after $4.5 \times 10^6$ steps it converged to the local minimum in Figure 8b). This local minimum, with total energy of -41.3944, is geometrically close to the icosahedral global minimum for the 13 atom cluster. However, instead of capping the 5 fold ring of the 12 atom cluster the extra atom bonded to one of its 3 atom faces, and this resulted in an unfavorable local minimum that is 2.9324 higher than the global minimum. In this case we found that in addition to raising the temperature to 1 we also needed to increase the linear drag coefficient to 0.24. This seems to be the general strategy to find the global minimum. In the current study we had the benefit of having the benchmarks from the Wales group to check our results. In cases where the global minima are unknown we should always check the minima found by rerunning the method using the found minima as starting points and a higher initial energy to explore the potential energy surface. We should also use the general strategy of adding an atom to the $n$ atom cluster as an initial guess of the $n+1$ cluster or removing an atom from the $n$ atom cluster to find other minima of the $n-1$ case.
4 Conclusions

We presented a modification of the traditional simulated annealing to be used in conjunction with classical molecular dynamics for the optimization of the ground state geometry of molecules. The method uses linear and quadratic drag terms to smoothly reduce the temperature of the system until it reaches equilibrium. Unlike the traditional simulated annealing, where the system is kept at a constant temperature for a fixed number of steps and then reduced every so often until the system reaches equilibrium, the schedule of temperature change is regulated by the drag coefficients. The linear drag cooling schedule follows an exponential and the quadratic drag follows a cooling schedule with a $1/t^2$ dependence.

The method was applied successfully to find the global minima of Lenard-Jones clusters with up to 20 atoms. We used several different initial guesses and in most cases we found the global minimum directly. The typical starting temperature in MD units was 0.04 and the linear drag term of about 0.6. The quadratic term, when employed was two orders of magnitude smaller. The program converged in about $2.5 \times 10^6$ MD steps. In a few cases we found local minima. In these cases we used the local minimum as the starting point for a new MD simulation and typically increased the initial temperature and the drag coefficients so that the system can explore the potential energy surface and quickly converge to the global minimum. The second run converged typically within $2 \times 10^6$ MD steps.

As illustrated in Figure 1, the system combines the strategies of traditional simulated annealing implementations where the system is initially allowed to explore the potential energy surface in order to find the global minimum and those of steepest descent methods (at low temperatures) where the system will converge to the closest minimum. The method is robust, as we used very unfavorable initial conditions and in most cases found the global minimum in one MD run. Although the method is physically related to geometry optimization it can also be employed for general optimization problems where the positions of the atoms will be replaced by the optimization parameters and the total energy will be replaced by the cost function one wants to optimize.

5 Acknowledgements

We would like to thank for the financial support of Northeastern Illinois University Student Center for Science Engagement.

References

Biring, S., & Chaudhury, P. (2012). A comparative study on full diagonalization of Hessian matrix and Gradient-only technique to trace out reaction path in doped noble gas clusters using stochastic optimization. *Chemical Physics*, 400, 198-206.

Biswas, R., & Hamann, D. R. (1986). Simulated Annealing of Silicon Atom Clusters in Langevin Molecular Dynamics. *Phys. Rev.*, 34(2), 895.

Czernicki, R., & Elber, R. (1990). Reaction path study of conformational. *J. Chem. Phys.*, 5580-5601.

Ingber, L. (1993). Simulated Annealing: Practice vs. Theory. *J. Math. Compu. Modelling*, 18(11), 29-57.

Johnston, R. L. (2003). Evolving better nanoparticles: Genetic Algorithms for optimising cluster geometries. *Dalton Trans.*, 4193-4207.
Kanters, R., & Donald, K. (2014). Searching for Unique Low Energy Minima of Structures Using a Novel Implementation of a Genetic Algorithm. *Journal of Chemical Theory and Computation, 10*, 5729-5737.

Kirkpatrick, S., Gelatt Jr., C., & Veči, M. P. (1983). Optimization by Simulated Annealing. *Science*, 671-680.

Pulay, P. (1987). Derivative Methods in Quantum Chemistry. *Adv. Chem. Phys*, 241-286.

Verlet, L. (1967). Computer 'Experiments' on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, 159, 98-103.

Verlet, L. (1968). Computer 'Experiments' on Classical Fluids. I. Equilibrium Correlation Functions. *Phys. Rev.*, 165, 201-214.

Wales, D. J. (2010). Energy landscapes: some new horizons. *Current Opinion in Stuct. Biol.*, 3-10.

Wales, D. J., Doye, J. P., Dullweber, A., Hodges, M. P., Calvo, F., Hernandez-Rojas, J., et al. (1996). *Lennard-Jones Clusters*. Retrieved October 28, 2014, from The Cambridge Cluster Database: http://www-wales.ch.cam.ac.uk/CCD.html

Xue, T., Luo, J., Shen, S., Li, F., & Zhao, J. (2010). Lowest-Energy structures of Cationic P$_{2m+1}^+$ (m=1-12) clusters from first-principles simulated annealing. *Chemical Physics Letters, 485*, 26-30.