An Overview on Simulated Annealing Vs Adaptive Mutation Simulated Annealing

Srijeeta Talukder*

ABSTRACT

Stochastic optimization algorithms are the global search engine that can find the global minima unequivocally. Most of such algorithms are based on some physical processes or natural phenomena. Simulated Annealing (SA) is one of the very popular stochastic search techniques, which has been applied in different domains of science for problem-solving. For chemists, the most common use of SA or any other stochastic optimizers is to find the geometry of molecules to which the optimizer needs to explore the very rugged multi-minima surface. Searching such surfaces to find global minima (which indicates the most stable geometry of the molecule) is not a trivial job. Generally, SA has a tendency to premature convergence, however, if one incorporates an adaptive mutation technique to get the step length for each search steps the optimizer would work in a much better way. This kind of SA is named as Adaptive Mutation Simulated Annealing (AMSA). The present paper portray an overview on SA and AMSA and a comparison on their workability is also be presented.

I. Introduction

Optimization is something that we are doing continuously knowingly or unknowingly to place ourselves in a best possible situation. Apart from the philosophical aspect, as a chemist we always need optimization algorithm to find molecular structure as we know that the stable molecular structures are none other than a minimum in potential energy surface (PES). The dimension of a potential energy surface is quite large even for a small molecule as we know that for a molecular system of N atom the dimensionality of PES is 3N-6 (for non-linear system). These multi-dimensional PES is constists of many number of minima. The stable molecular structure has most resemblance with the global minimum or deepest minimum sturucture on the PES. Thus to get the deepest minimum structure is not a trivial job. There are many optimization techniques available and a useful way to catgorise them is to divide them as deterministic approach or stochastic search processes. Deterministic approaches [1] are initial point dependents and generally fail to get the global minimum for a multi-dimentional multi-minima surface, whereas in stochastic search processes algorithm has some in-build property by which it can surmount the potential barrier to find global solution.

The stochastic optimization algorithms are generally algorithmic replica of a natural processes. There are plenty numbers of such optimizers. Some popular stochstic optimizers are Monte Carlo based algorithm [2], Particle swam opotimization [3] etc.

In the presernt study I want to portray an overview Simulated Annealing [4,5], a Monte Carlo algorithm based stochastic optimizer and its comparison with Adaptive Mutation Simulated Annealing [6].

Department of Chemistry,
Adamas University, Kolkata, 20016, India.
**Corresponding email:*
srijeeta1.talukder@adamasuniversity.ac.in

Stochastic ooptimization algorithms, so also SA or AMSA would not only be used in molecular structure determination [7,8,9], rather examples are there where different kinds of problems can be cast as optimization problem by defining onjective function properly, such as reaction path determination [10,11,12], control dynamical study [13,14], kinetic parameters detrrmination [15] etc.

II. Simulated Annealing (SA):-

The stochastic optimization algorithms are generally inspired by some natural selection or physical processes. SA mimics the thermodynamic annealing protocal. In thermodynamic annealing the system is slowly cool down to attain the most stable thermodynamic state, whereas in SA system is simulated well to obtain the global minima. An algorithmic temperature is defined, known as Annealing Temperature.

Like any other optimization algorithm SA also moves with respective to an objective function or cost function which is being minimised during the simulation. A move during simulation is accoiunted by exploration of optimization surface, which is generally very rugged.

$$x_{i+1} = x_i + \Delta r \dots\dots\dots\dots\dots\dots\dots\dots (1)$$

Here, r is the random number and Δ is the maximum amount of change that is allowed. During simulation the move will be accepted if the cost function is got a decrease, however if it increases the move is not rejected in a straigt forward manner. A Metropolis test would be performed to decide whether the step is accepted or rejected.

$$P_M = \exp\left|-{}^{diff}/_{KT}\right| \dots\dots\dots\dots\dots\dots\dots (2)$$

$$diff = Cost_{i+1} - Cost_i \dots\dots\dots\dots\dots\dots\dots\dots (3)$$

$diff$ is the difference in cost in two successive moves, T_{at} is the algorithmic annealing temperature and K is a constant. P_M is the Metropolis probability, clearly from the eq. 2, if the cost for $i+1$th step decreases the value of P_M is greater than 1 i.e, straightway accepted. But if the same is increases, i.e $diff$ is positive then value of P_M is inbetween 0 to 1. Depending upon T_{at} the value of P_M is guided. If $T_{at}$is high P_M is close to 1 i.e moves with higher cost have high probability to be accepted.

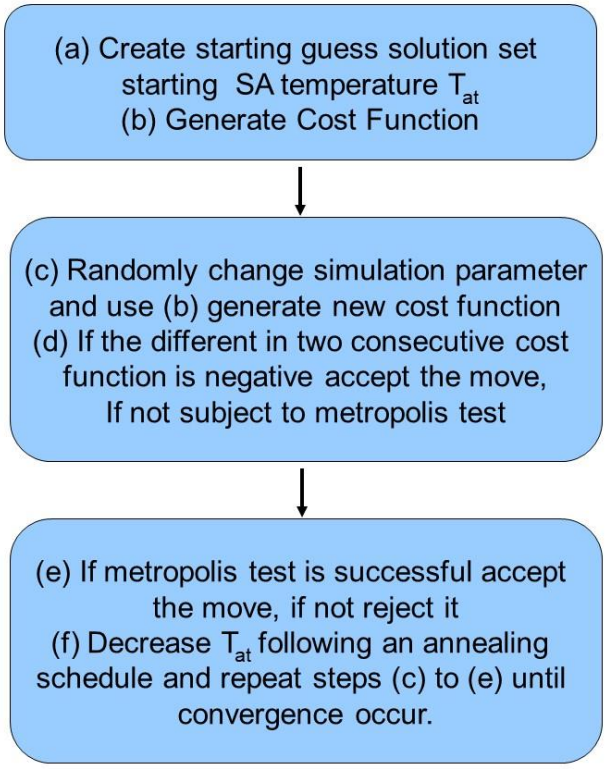

Fig 1: Flowchart for SA

III. Adaptive Mutation Simulated Annealing (AMSA):-

The philosophy of AMSA is also based on the physical process of thermodynamics Annealing. However this search technique has a smart in built mechanism to mutate the step size optimization. Basically AMSA uses Metropolis sampling in a same way like SA, however it uses the count of Metropolis of acceptance in order to mutate the step length which eventually control the Metropolis Probability (P_M) and so that the search direction.

During search process initially (or when the temperatrure is high) the step size should be large so that most of the surface would be explored and eventually when the algorithm finds a direction of convergence the step size should be decreased to get fine tunning of the variables. In SA this change in step length has to be done manually, however in AMSA, an algotithm has been inplemented to mutate the same by its own.

The Metropolis accepatnce is the key for the adaptive mutation of the step length [6]. If the P_M is very high, that means the search process may trap in a basin which may cause premature convergence, so the step length should be higher to reduce the P_M. One can use the following equation to increase the step length

$$\Delta = \Delta \times (1.0 + r) \dots\dots\dots\dots (4)$$

Where, r is therandum number between 0 to 1. Then according to the equation (4) the upper limit of the increment in step length (Δ) is 100%. Similarly if P_M is very low one must decrease the Δ in order to reduce unnecessary exploration. Very low P_M means that the optimizer heads nowhere. This would not be a good situation either just like very high P_M. Then the Δ may be reduced by the following manner

$$\Delta = \frac{\Delta}{(1.0 + r)} \dots\dots\dots\dots (5)$$

So the step length got halved if the r is at its maximum value.

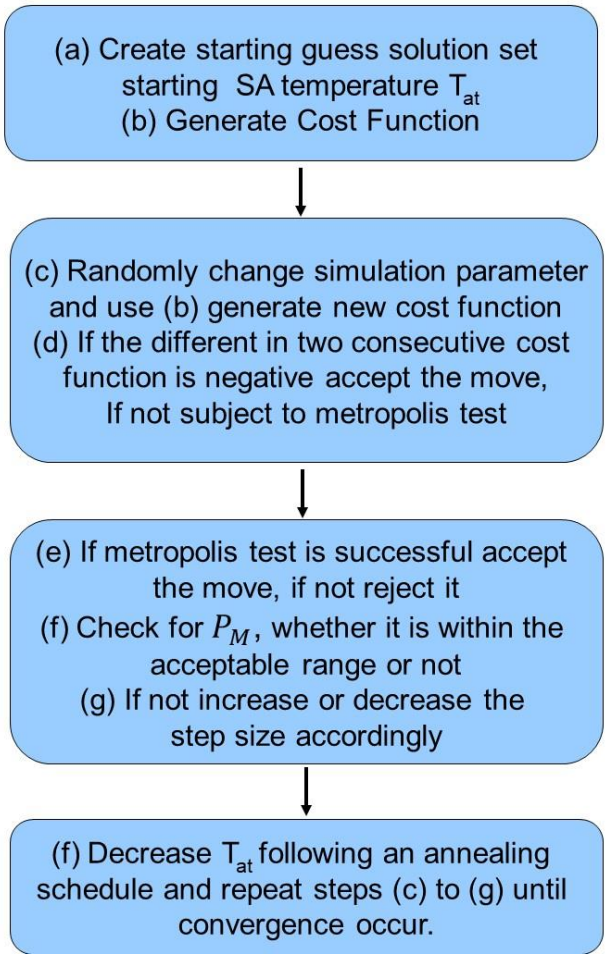

Fig. 2 Flowchart for AMSA

By equation (4) and (5) the Δ is adjusted and for a definite temperature in a way one may got the correct Δ. With decrease in temperature the Δ would also get decresed adaptively. The scheme reduces the human effeort and also human error as the algorithm would find the correctstep size depending upon the situation during optimization. The acceptable range of P_M may be varied but a preferable one would be in between 15% to 35%.

IV. Comparison

A simulation has been performed to make a comparison between the SA and AMSA putting all other optimization parameters same. Silver cluster of size 10 has been taken as model system [17]. In the following figures the optimization profiles for SA and AMSA are presented.

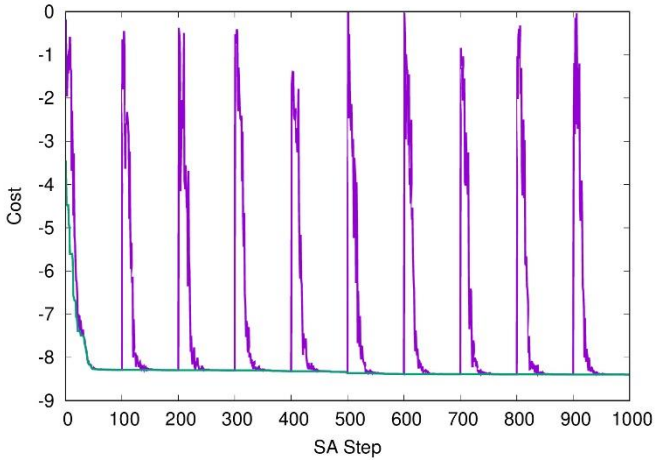

Fig 3: Optimization profile for SA

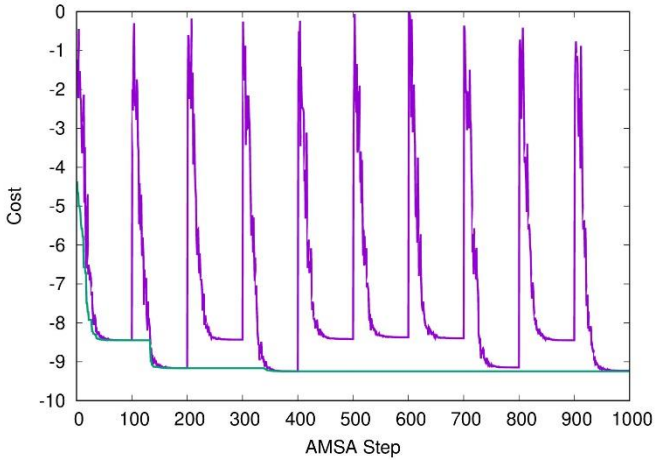

Fig 4: Optimization profile for AMSA

Both Fig.3 and Fig.4 The green lines denote the best cost (solution) and the violet lines represent the latest

cost. The latest cost periodically moves with the annealing schedule. The zigzag nature in violet lines account the stochasticity of the algorithms. Both the simulation run for 1000 optimization step with same annealing schedule. The initial step size for both the cases are same, however in AMSA the step size changes adaptively. It is clear from the figures that in SA the convergence occurs at higher cost than that of in AMSA. Rather in AMSA the global solution is got.

V. Conclusion

The efficiency of AMSA is definitely much better than SA atleast in the mentioned case. However there is lots of scope to test the efficiency of AMSA. I am using AMSA in different system which become eventually very usefull and practically the problem of premature convergence occurs in SA can be overcomed by using AMSA.

References

[1] T. Haukaas Deterministic Optimization Methods, www.inrisk.ubc.ca.
[2] M. J. Field, J. Chem. Phys. 103 3621 (1995).
[3] F. Weigend and A. Ahlrichs, Phil. Trans. R. Soc A 368 1245 (2010).
[4] K. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi Science 220, 671 (1983).
[5] K. S. Kirkpatrick, J. Stat. Phys. 34, 975 (1984).
[6] S. K. Biring, R. Sharma and P.Chaudhury, J. Math. Chem. 52, 368 (2014).
[7] W.J. Pullan, J. Comput. Chem. 18, 1096 (1997).
[8] C. Gatti, P. Fantucci and G. Pacchioni, Theor. Chim. Acta 72, 433 (1987).
[9] H. Handschuh, C. Cha, P.S. Bechthold, G. Ganteför and W. Eberhardt, J. Chem. Phys. 102, 6406 (1995).
[10] P. Chaudhury, S.P. Bhattacharyya, J. Mol. Str. (Theochem) 429 (1998) 175
[11] P. Chaudhury, S.P. Bhattacharyya, Int. J. Quant. Chem. 76 (2000) 61.
[12] P. Chaudhury, S.P. Bhattacharyya, W. Quapp, Chem. Phys. 253 (2000) 295.
[13] P. Naskar, S. Talukder, S. Ghosh, P. Chaudhury, Int. J. Quantum Chem. 2019, 119, e25927.
[14] S. Ghosh, S. Talukder, S. Sen, P. Chaudhury, Mol. Phys. 2015, 113, 3826.
[15] Srijeeta Talukder, Shrabani Sen, Ralf Metzler, Suman K. Banik and Pinaki Chaudhury, J. Chem. Sci. 125 1619 (2013).
[16] P. Naskar, R. Roy, S. Talukder and P. Chaudhury, Mol. Phys. 116 2172 (2018).
[17] X. Shao, X.Liu, and W. Cai, J. Chem. Theo. Comp. 1 762 (2005).