Theor Chem Acc (2000) 103:225–227
DOI 10.1007/s002149900053

Theoretical
Chemistry Accounts
© Springer-Verlag 2000

Perspective

Perspective on "Equation of state calculations by fast computing machines"

Metropolis N, Rosenbluth AE, Rosenbluth MN, Teller AH, Teller E (1953) J Chem Phys 21: 1087–1092

William L. Jorgensen

Department of Chemistry, Yale University, New Haven, CT 06520–8107, USA

Received: 6 April 1999 / Accepted: 14 April 1999 / Published online: 14 July 1999

Abstract. An overview is given of the paper by Metropolis et al. that has formed the basis for Monte Carlo statistical mechanics simulations of atomic and molecular systems.

Key words: Monte Carlo – Statistical mechanics – Metropolis sampling

1 Background

For many-particle systems in classical statistical mechanics, the key numerical problem is the solution of the configurational integral that appears in the averages for a property Q (Eqs. 1, 2).

$$\langle Q \rangle = Q_k + \int Q(\vec{X}) P(\vec{X})\, \mathrm{d}\vec{X} \tag{1}$$

$$P(\vec{X}) = \exp[-\beta E(\vec{X})] \Big/ \int \exp[-\beta E(\vec{X})]\, \mathrm{d}\vec{X}. \tag{2}$$

These equations are for the canonical NVT ensemble where $P(\vec{X})$ is the Boltzmann factor, $E(\vec{X})$ is the total potential energy, $\beta = 1/kT$, and the integrals are taken over all possible geometrical configurations, $\vec{X}$, of the system. Q_k represents the contribution from the kinetic energy, which is taken as separable from the configurational contribution. For N particles, $\vec{X}$ has about $3N$ dimensions or coordinates. In addition, at liquid or solid densities most arbitrary choices of $\vec{X}$ would have overlapping particles with low probability, $P(\vec{X})$, and, therefore, they would contribute little to the average. Consequently, a brute-force Monte Carlo solution of the configurational integral by random selection of configurations, $\vec{X}$, becomes impractical for more than a few particles with typical potential-energy functions.

2 Key contribution

In their classic paper, Metropolis et al. [1] recognized that a practical solution for the configurational integrals could be obtained by a modified Monte Carlo procedure where "instead of choosing configurations randomly, then weighting them with $\exp(-E/kT)$, we choose configurations with a probability $\exp(-E/kT)$ and weight them evenly." With this procedure and converting the integral to a sum over discrete configurations, Eqs. (1 and 2) are simplified to Eq. (3), where L is the number of configurations considered and $\vec{X}_{\mathrm{M}}$ indicates a Metropolis-sampled configuration.

$$\langle Q \rangle = Q_k + 1/L \sum_{i}^{L} Q_i(\vec{X}_{\mathrm{M}}) \tag{3}$$

The Metropolis algorithm involves generation of a chain of configurations. When the system is at configuration i, an attempted move to another configuration j is considered. The attempted move can be accepted and j is the next configuration in the chain, or the move is rejected, i repeats in the chain and another attempted move is tried. The decision to move from i to j is determined from a probability $p = \pi_j/\pi_i$, where for the NVT ensemble the Boltzmann factor (Eq. 4)

$$\pi_j = \exp(-\beta E_j) \tag{4}$$

is appropriate. Then, if $p \geq 1$, which for the NVT ensemble means $E_j \leq E_i$ (the potential energy went down), the move is accepted. If $p < 1$, p is compared to a random number x between 0 and 1, and if $p \geq x$, the move is still accepted, otherwise j is rejected and i repeats. In summary, configuration j is accepted with a probability $\min[1, \pi_j/\pi_i]$. The chain of configurations generated in this way provides the Boltzmann-weighted configurations that are needed in Eq. (3). It is apparent that the Metropolis sampling, while allowing the energy to go up, focuses on low-energy configurations, which contribute the most to the Boltzmann averages. Nevertheless, care is required to guarantee convergence of the calculations and different computed properties converge at different rates [2, 3].

One could imagine that the Metropolis sampling algorithm could have been conceived by inspection from considering a simple one-particle system with two states, r and s. If we let $\varepsilon_r - \varepsilon_s = kT \ln 3$, for example, then the Boltzmann factor $\pi_r/\pi_s = 1/3$ and the algorithm needs to sample state s 3 times on average for every time it samples state r. This could be achieved by a walk where when in state r, the next step is back to state s, and when in state s, state s needs to repeat an average of 2 times before a transition to r. For example, the following walk would work.

State r : *X* *X* *X* *X* *X* *X* *X* .....
State s : *XXX* *XXX* *XXX* *XXX* *XXX* *XXX* *XXX*

However, to make the process more general and random, the decision on whether to repeat s should be based on a random number that is compared to π_r/π_s. A transition from a lower-energy state, s, to a higher-energy one, r, needs to occur for π_r/π_s of the attempted moves.

3 Refinements, sampling and some details

Another issue is the size, N, of a sample that would be adequate to represent a liquid. To this end, Metropolis et al. [1] also introduced "periodic boundary conditions", which lead to simulating a liquid by explicitly considering a relatively small number of molecules, about 100–1000. Though they only treated hard disks in two dimensions in their paper, "extension to three dimensions is obvious." In three dimensions, the molecules are typically in a cubic or orthorhombic cell, which is surrounded by 26 images of itself. If on moving a molecule to create a new configuration it passes through a wall, then an image of it reenters the central cell through the opposite face. Thus, one just needs to keep track of the contents of the central cell.

With rigid molecules, a move normally consists of picking one molecule at random, translating it randomly in all three Cartesian directions and randomly rotating it about one randomly chosen Cartesian axis. For flexible molecules, both the external and internal degrees of freedom need to be sampled. The external ones are handled in the same way as for rigid molecules. The internal ones are usually treated by representing the molecule by a Z matrix (r, v, φ internal coordinates). Upon an attempted move, the variable entries in the Z matrix are changed randomly within specified ranges and the molecule is rebuilt [4]. The use of internal coordinates also facilitates the enforcement of constraints, i.e., specification of bond lengths, bond angles and dihedral angles that are fixed, and avoids the introduction of additional terms in Eq. (4) with space-fixed coordinate systems [3]. It is clear that the acceptance rate is affected by the choices of ranges for the molecular translations and rotations; reasonable convergence behavior is normally obtained by adjusting the ranges such that 20–50% of the attempted moves are accepted.

An advantage of Monte Carlo simulations is that modified sampling procedures can potentially be devised to enhance convergence. One example is the preferential sampling of solvent molecules near a solute by attempting to move them more frequently than more distant solvent molecules [5]. The only caveat is to be sure that the intrinsic probabilities, i.e., in the absence of the Boltzmann factor, of states i and j are properly reflected. If this is not the case, then π_j in Eq. (4) needs to include an appropriate Jacobian term representing the phase-space volume for state j. For example, with an atomic liquid, if one wanted to sample in spherical rather than Cartesian coordinates, the required Jacobian is $r_j^2 \sin\varphi_j$, and the acceptance probability would be $\min[1, (r_j^2 \sin\varphi_j / r_i^2 \sin\varphi_i)\exp(-\beta(E_j - E_i))]$. Another attractive point is that it is straightforward to execute Monte Carlo simulations in the NPT ensemble by including volume terms in Eq. (4) to yield Eq. (5) and by allowing the volume of the central cell to vary [2, 3].

$$\pi_j = \exp(-\beta H_j)V_j^N = \exp[-\beta(E_j + PV_j - Nk_\mathrm{B}T \ln V_j)] \tag{5}$$

The principal potential pitfall is that there is no ultimate guarantee that convergence of a property or of the simulation as a whole has been reached. Pathological cases can be constructed. For example, one could set up a simulation for a liquid secondary amide at 25°C with all of the monomers initially in the cis conformation. Since the rotational barrier is about 20 kcal/mol for conversion of cis to trans, under normal Monte Carlo conditions a conversion would be an exceedingly rare event. The system could be equilibrated and the energy seemingly converged, but the simulated liquid is metastable since with an energy difference of about 2.5 kcal/mol, 98.5% of the monomers should be trans. However, in general, for typical organic liquids with standard force fields, convergence of most key properties such as heat of vaporization, density, and radial distribution functions is not problematic [6]. In fact, Monte Carlo simulations have been shown to be more efficient than molecular dynamics for liquid hexane [7]. Such Monte Carlo calculations are trivial to set up with modern software [4] and can be executed for one state point in a few hours on Pentium-II-based personal computers [8]. For dilute solutions of a single flexible solute in a solvent, convergence of the conformational properties of the solute can understandably be more taxing as statistics are only being accumulated on one molecule rather than on N.

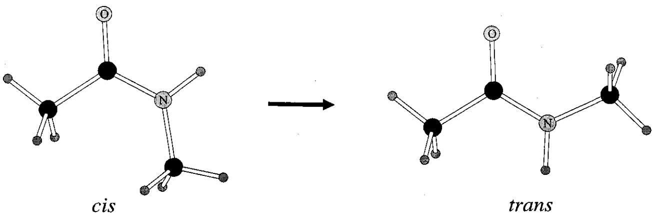

Figure 1.

Overall, Metropolis et al. introduced the sampling method and periodic boundary conditions that remain at the heart of Monte Carlo statistical mechanics simulations of fluids. This was one of the major contributions to theoretical chemistry of the twentieth century.

References

1. Metropolis N, Rosenbluth AE, Rosenbluth MN, Teller AH, Teller E (1953) J Chem Phys 21: 1087–1092
2. (a) Barker JA, Henderson D (1976) Rev Mod Phys 48: 587–671; (b) Jorgensen WL (1983) J Phys Chem 87: 5304–5314; (c) Kalos MH, Whitlock PA (1986) Monte Carlo methods. Wiley, New York; (d) Levesque D, Weiss JJ (1995) Top Appl Phys 71: 121–204; (e) Jorgensen WL (1998) In: Schleyer PvR (ed) Encyclopedia of computational chemistry, vol 3. Wiley, New York, pp 1754–1763
3. Allen MP, Tildesley DJ (1987) Computer simulations of liquids. Clarendon, Oxford
4. Jorgensen WL (1999) BOSS, version 4.1. Yale University
5. Owicki JC, Scheraga HA (1977) Chem Phys Lett 47: 600–602
6. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118: 11225–11236
7. Jorgensen WL, Tirado-Rives J (1996) J Phys Chem 100: 14508–14513
8. Tirado-Rives J, Jorgensen WL (1996) J Comput Chem 17: 1385–1386