Lecture 9 - The Monte Carlo method

- Stochastic vs deterministic processes
- Metropolis MC
- Applications of MC in Quantum Chemistry

Properties as Ensemble Averages

$$\langle A \rangle = \int \int A(q,p) P(q,p) \mathrm{d}q \mathrm{d}p$$

From statistical mechanics:

$$\begin{aligned} P(q,p) &= Q^{-1} e^{-E(q,p)/k_B T} \\ Q &= \int \int e^{-E(q,p)/k_B T} \mathrm{d}q \mathrm{d}p \end{aligned}$$

Q – system partition function.

Metropolis Monte Carlo

Metropolis at al (1953): Instead of choosing configuration randomly then weighting them with $exp(-E/k_B T)$, we choose configurations with a probability $exp(-E/k_B T)$ and weight them evenly.

$$\langle A \rangle = \frac{1}{X} \sum_{i=1}^{X} A(q_i)$$

Can be only used only for systems in equilibrium.

MC Applications in QC

- Many problems are difficult (example: properties of liquids).
- Property averaging.
- Reactions.
- Ground state geometries (usually with simulated annealing).

Reaction Studies

- common mathematical formulation used for biochemical system modeling: deterministic.
- chemical species as continuous valued concentrations .
- chemical interactions represented as ordinary differential equations.
- Biological reactions involve lower concentrations: Stochastic effects need to be accounted for.
- Metropolis sampling cannot be used (systems are not in equilibrium).
- Use dynamic Monte-Carlo.

Stochastic Reaction Modeling

- Stochastic behavior described by chemical master equation (Gillespie, 1992).
- CME: often an infinite set of ordinary differential equations; is difficult to solve analytically.
- Use Monte-Carlo.
  1. stochastic simulation algorithm (SSA) predicts the execution of each individual reaction event for the given system (Gillespie, 1976).
  2. First reaction method (FRM); Direct method (DM) (Gillespie, 1977).
  3. Next reaction method (NRM) (Gibson and Bruck (2000))
  4. Optimized direct method (ODM) (Cao et al. (2004)).
- Valuable review: Comp. Biol. Chem. 30, 39-49, 2006.

Simulated Annealing

- A procedure of locating a ground state geometry by gradual decrease of temperature.
- Allows big conformation changes in the beginning but smaller and smaller as you get closer to the minimum.

Genetic Algorithms

- Yet another way of sampling “difficult” potential surfaces.
- Two basic operations: mutations (Monte-Carlo type modifications) and “gene” mixing (hoping for best of breed).
- ☐ Conformation searching, protein folding, homology based protein structure determination, NMR peak assignment and laser pulse design.

Computational Aspects of MC

- Only energy needs to be evaluated for each configuration.
- New configurations generated by making small random differences in bond lengths and angles.
- Trivial to parallelize.
- While configuration can be generated sequentially, there is no notion of time.
- Result quality depends on the quality of the random numbe generator (not a problem nowadays).

Other Technical Details of MC

- Size of the perturbing step: large step gives low acceptance radio - few points in the distribution space sampled.
- Small step opposite: many changes accepted but part of the sampled phase space is small.
- Optimum step: fastest convergence of properties.
- Empirical: such changes that give every second step accepted on average.

Monte-Carlo Summary

- With proper setup, large parts of the phase space can be sampled.
- But some aspect are problematic: eg. chirality change.
- In general, correlation processes difficult to follow with MC (protein folding).