Topics in Computational Chemistry, CHEM*7500/CHEM 750 Fall 2021

Week 3: The Monte Carlo Method, Programming with Python

Prof. Leanne D. Chen

September 29, 2021

Department of Chemistry

UNIVERSITY
of GUELPH

- Last week we introduced the *partition function*, which allows us to describe the statistical properties of a thermodynamic system.

$$Q(N, V, T) = \sum_j e^{-E_j(N,V)/k_B T} \tag{1}$$

- This function $Q(N, V, T)$ is the central function of the canonical ensemble, and allows us to calculate different macroscopic properties.

- For example, if we would like to calculate the average energy of an ensemble,

$$\overline{E}(N, V, T) = \frac{\sum_j E_j(N, V) e^{-E_j(N,V)/k_B T}}{\sum_j e^{-E_j(N,V)/k_B T}} = \frac{\sum_j E_j(N, V) e^{-E_j(N,V)/k_B T}}{Q} \qquad (2)$$

- Or, to obtain the average pressure, where pressure is equal to the negative of the partial derivative of the energy with respect to the volume,

$$\overline{p} = \sum_j p_j P_j = -\frac{\sum_j \left(\frac{\partial E_j}{\partial V}\right) e^{-E_j/k_B T}}{\sum_j e^{-E_j/k_B T}} \qquad (3)$$

UNIVERSITY$_{of}$GUELPH

Monte Carlo Simulations

Connection to Statistical Thermodynamics

- Now, we are going to write the energy explicitly as a Hamiltonian that depends on positions and momenta of all particles in a system:

$$Q = c \int e^{\left[-\mathcal{H}\left(\mathbf{r}^N, \mathbf{p}^N\right)/k_B T\right]} d\mathbf{p}^N d\mathbf{r}^N \tag{4}$$

- Here, $\mathbf{r}^N$ stands for the coordinates of all N particles, and $\mathbf{p}^N$ for the corresponding momenta.

- The function $\mathcal{H}\left(\mathbf{q}^N, \mathbf{p}^N\right)$ is the Hamiltonian of the system, which expresses the total energy of an isolated system as a function of the coordinates and momenta of all particles:

$$\mathcal{H} = \mathcal{K} + \mathcal{U} \tag{5}$$

where $\mathcal{K}$ is the kinetic energy of the system, and $\mathcal{U}$ is the potential energy.

Connection to Statistical Thermodynamics

- The classical analogue of Equation 2 is

$$\langle \mathcal{A} \rangle = \frac{\int \mathcal{A}\left(\mathbf{p}^N, \mathbf{r}^N\right) e^{-\beta \mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)} d\mathbf{p}^N d\mathbf{r}^N}{\int e^{-\beta \mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)} d\mathbf{p}^N d\mathbf{r}^N} \tag{6}$$

where $\beta = 1/k_B T$.

- In this equation, the observable $\mathcal{A}$ has been expressed as a function of coordinates and momenta.

- For the kinetic energy $\mathcal{K}$, which is a quadratic function of the momenta, this integral can be calculated analytically.

$$\int x^2 e^{cx} dx = e^{cx} \left(\frac{x^2}{c} - \frac{2x}{c^2} + \frac{2}{c^3} \right) \tag{7}$$

Connection to Statistical Thermodynamics

- However, for functions in the form $\mathcal{A}\left(\mathbf{r}^N\right)$, it is generally not possible to calculate these averages analytically.
- In other words, the potential energy function is usually very complex and has no closed-form integral.
- Thus, we must use numerical techniques to evaluate these integrals.
- How about numerical quadrature?
- Unfortunately, this becomes expensive very quickly, and is thus infeasible for anything except the smallest systems.
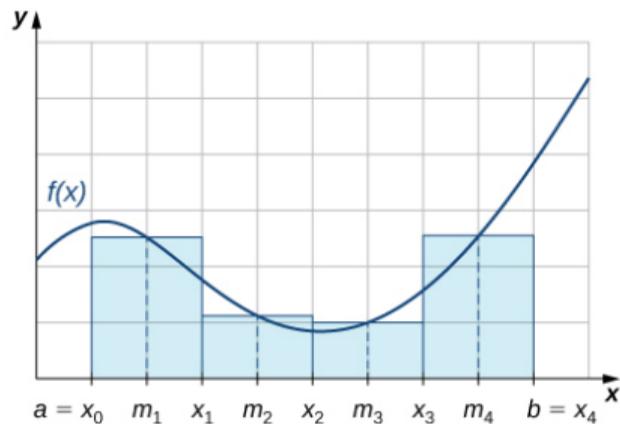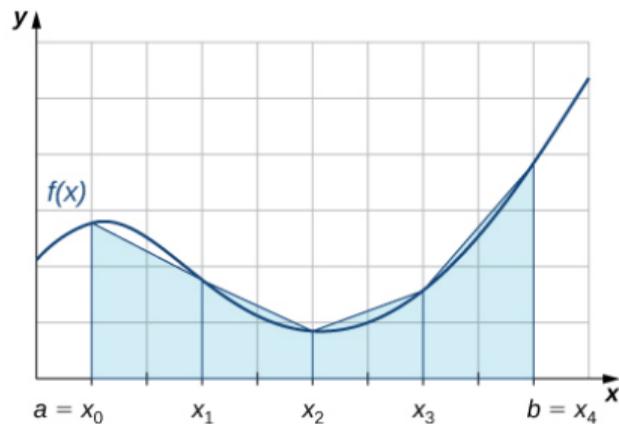
Source: LibreTexts

$$\int_a^b f(x)\mathrm{d}x \approx (b - a)f\left(\frac{a + b}{2}\right) \quad (8)$$

Source: LibreTexts

$$\int_a^b f(x)dx \approx (b-a)f\left(\frac{f(a)+f(b)}{2}\right) \quad (9)$$

Numerical Quadrature

- Suppose we would like to carry out the numerical quadrature by using a simple method such as the trapezoidal rule for a system containing N molecules.

- This means that we have to integrate in $3N$ dimensions, with m intervals on each axis, leading to m^{3N} evaluations total.

- Even for a very small system of 100 particles, with only 10 intervals along each axis, we would need to evaluate the integrand at 10^{300} points!

- For reference, the number of protons in the observable universe is approximately 10^{80}.

- Computations of such magnitude are simply not possible.

Numerical Quadrature

- Numerical quadrature works best on functions that are smooth over distances of the order of the mesh size.

- However, for most intermolecular potentials, the Boltzmann factor in Equation 6 is a rapidly varying function of the particle coordinates, thus a very small mesh spacing is required.

- More over, as an example, when evaluating the integrand for a dense liquid, for the overwhelming majority of points, the Boltzmann factor is vanishingly small.

- Thus, we need a more efficient numerical technique to compute thermal averages in a simulation.
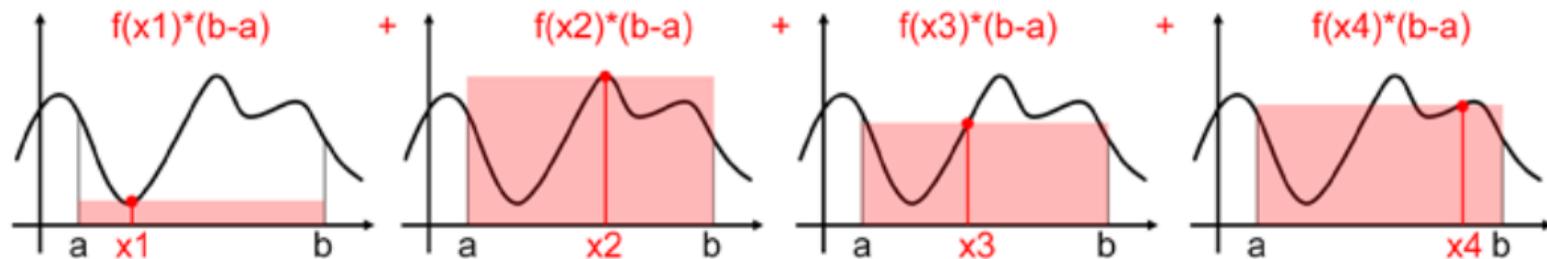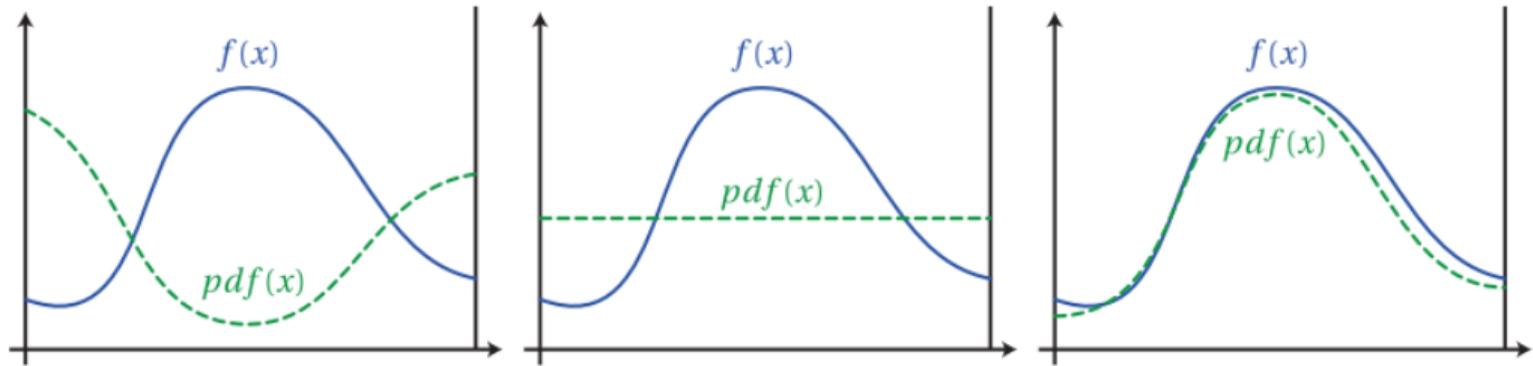
Source: Scratchapixel 2.0

Figure A.2: Comparison of three probability density functions. The PDF on the right provides variance reduction over the uniform PDF in the center. However, using the PDF on the left would significantly increase variance over simple uniform sampling.

Source: Wojciech Jarosz Dissertation

The Algorithm

- The prime purpose of the kind of Monte Carlo or Molecular Dynamics program that we shall be discussing is to compute equilibrium properties of classical many-body systems.
- From now on, we shall refer to such programs simply as MC or MD programs.
- It should be noted that there are many other applications of the Monte Carlo method (and, to a lesser extent, of the Molecular Dynamics method).

The Algorithm

1. Select a particle at random, and calculate its energy $U(\mathbf{r}^N)$.
2. Give the particle a random displacement, $\mathbf{r}' = \mathbf{r} + \Delta$, and calculate its new energy $U(\mathbf{r}'^N)$.
3. Accept the move from $\mathbf{r}^N$ to $\mathbf{r}'^N$ with probability

$$\text{acc}(\text{old} \to \text{new}) = \min\left(1, e^{-\beta\left[U(\mathbf{r}'^N) - U(\mathbf{r}^N)\right]}\right) \tag{10}$$

Boundary Conditions

- Monte Carlo and Molecular Dynamics simulations of atomic or molecular systems aim to provide information about the properties of a macroscopic sample.
- Yet, the number of degrees of freedom that can be conveniently handled in present-day computers ranges from a few hundred to a few million.
- Most simulations probe the structural and thermodynamic properties of a system of a few hundred to a few thousand particles.
- Clearly, this number is still far removed from the thermodynamic limit.

Boundary Conditions

- To be more precise, for such small systems it cannot be safely assumed that the choice of the boundary conditions has a negligible effect on the properties of the system.

- In fact, in a three-dimensional *N*-particle system with free boundaries, the fraction of all molecules that is at the surface is proportional to $N^{-1/3}$.

- For instance, in a simple cubic crystal of 1000 atoms, some 49% of all atoms are at the surface, and for 10^6 atoms this fraction has decreased to only 6%.

Boundary Conditions

- In order to simulate bulk phases, it is essential to choose boundary conditions that mimic the presence of an infinite bulk surrounding our *N*-particle model system.
- This is usually achieved by employing periodic boundary conditions.
- The volume containing the *N* particles is treated as the primitive cell of an infinite periodic lattice of identical cells.
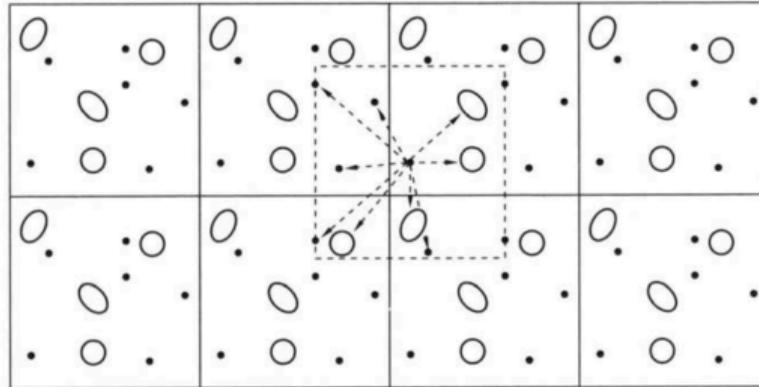
- A commonly used intermolecular pair potentials is the **Lennard-Jones potential**, which is typically used to describe non-bonded interactions in a simulation.

$$U^{lj}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \tag{11}$$

- In three dimensions, the average potential energy of any given atom *i* is given by

$$U_i = \frac{4\pi}{2} \int_0^\infty r^2 \rho(r) U(r) \mathrm{d}r \tag{12}$$

Simple Truncation

- The simplest method to truncate potentials is to ignore all interactions beyond r_c.
- The potential that is simulated is

$$U^{\text{trunc}}(r) = \begin{cases} U^{\text{lj}}(r), & r \leq r_c \\ 0, & r > r_c \end{cases} \tag{13}$$

- However, this may result in an appreciable error in our estimate of the potential energy of the true Lennard-Jones potential.
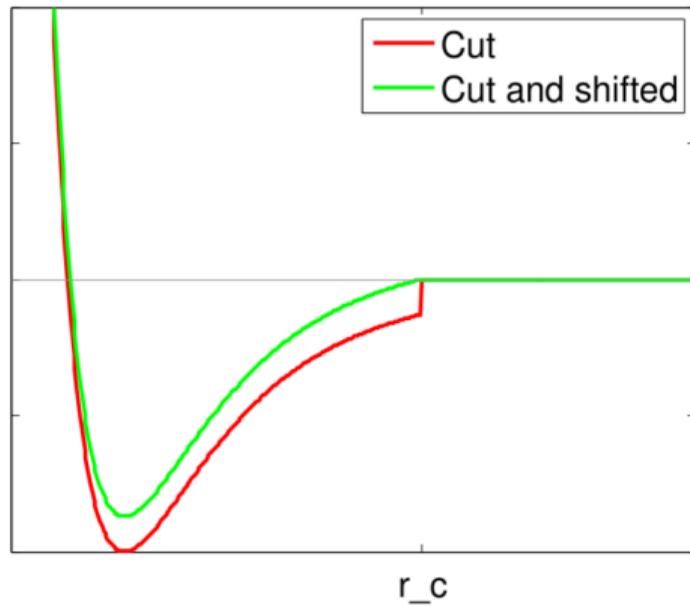- Also, the potential at the cutoff radius is discontinuous.

UNIVERSITY$_{of}$GUELPH

- In molecular dynamics simulations, it is common to use another procedure.
- The potential is truncated and shifted, such that the potential vanishes at the cutoff radius:

$$U^{\text{tr-sh}}(r) = \begin{cases} U^{\text{lj}}(r) - U^{\text{lj}}(r_c), & r \leq r_c \\ 0, & r > r_c \end{cases} \qquad (14)$$

- In this case, there are no discontinuities in the potential.

Source: Helsinki Physics

Python

We will now start the practical component of today's lecture.

Thank you!

Please feel free to unmute yourself or raise your hand to ask a question.

UNIVERSITY *of* GUELPH