Basics of Monte Carlo simulations

Dr. Jinglei HU

Outline

- **History of Monte Carlo method**
- **Simple sampling Monte Carlo**
- **Importance sampling Monte Carlo**
  - Markov process
  - Detailed balance
  - Metropolis algorithm
- **Case study 2: 3D Lennard-Jones fluid**

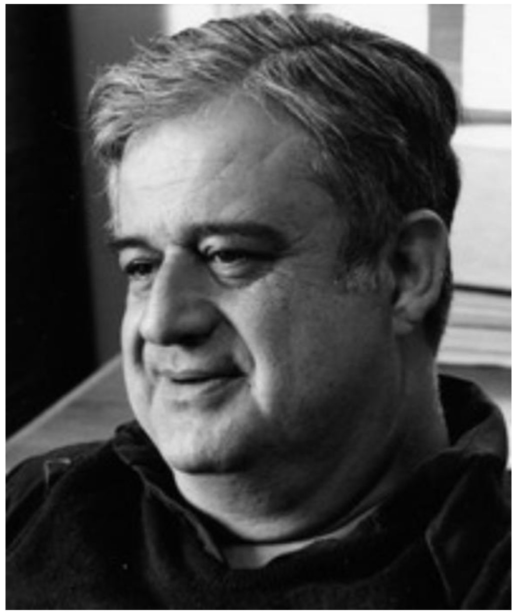
History of Monte Carlo methods

- Stanislaw Ulam (re)invented the method in ~1947 during his work on Manhattan project at the Los Alamos National Laboratory.
- John von Neumann programmed the ENIAC computer to perform MC calculations.
- Nicholas Metropolis coined the name Monte Carlo, and later together with Marshall Rosenbluth reported adopting Metropolis algorithm to calculate the equation of state of 2D hard spheres.


Stanislaw Ulam (1909-1984)

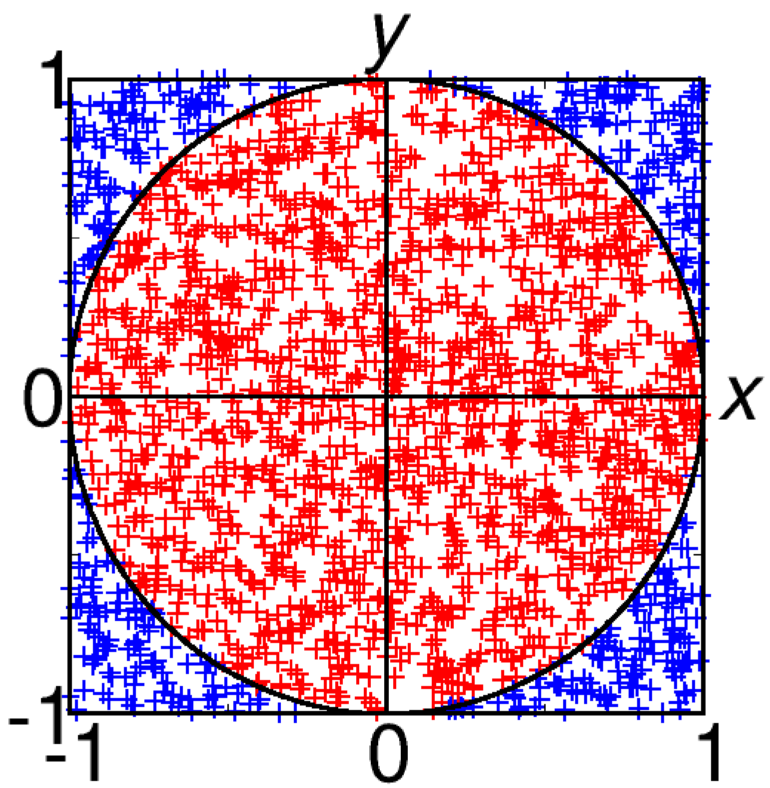

John von Neumann (1903-1957)


Nicholas Metropolis (1915-1999)

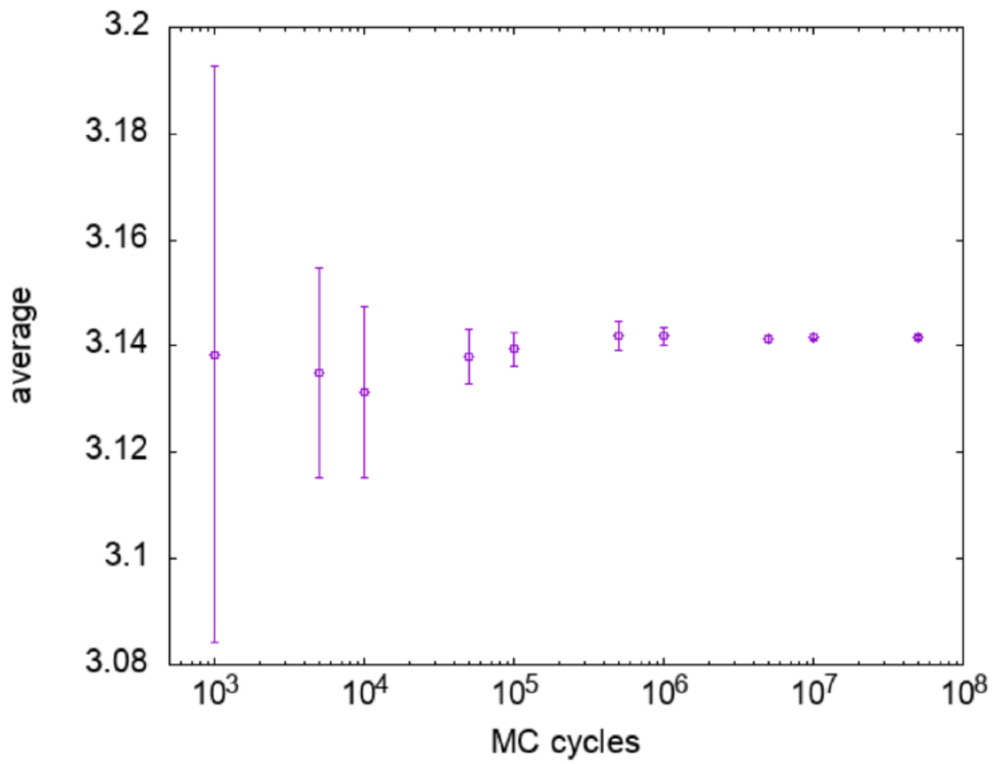

Marshall Rosenbluth (1927-2003)

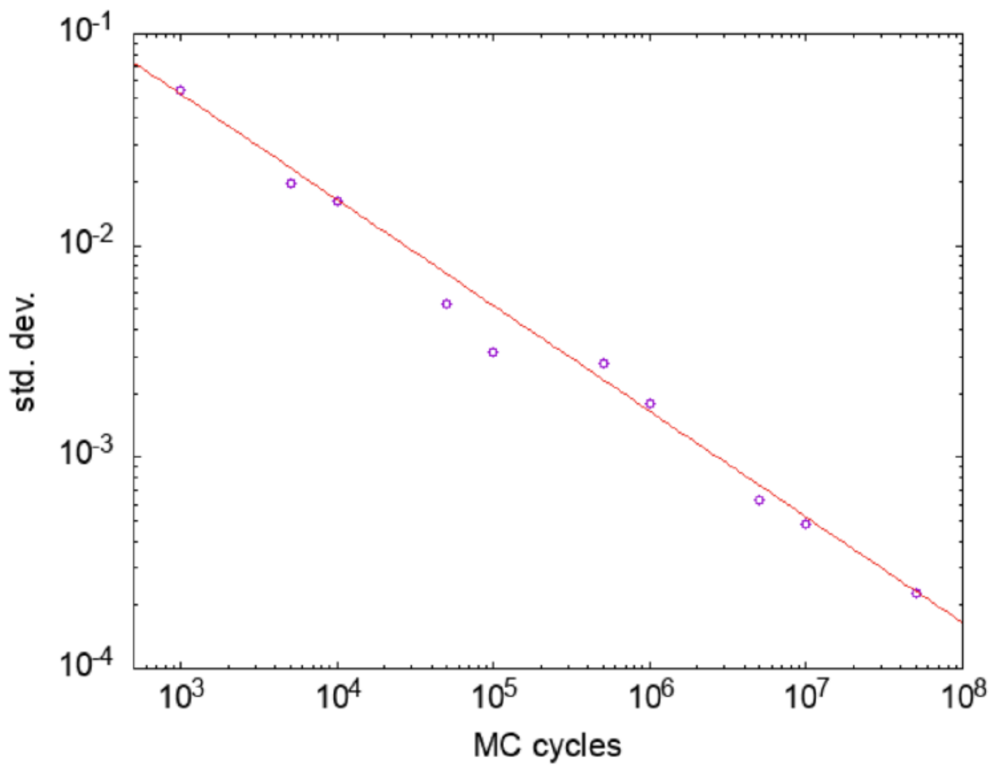
N. Metropolis & S. Ulam, *J. Am. Stat. Assoc.* 44(247): 335-341 (1949).
N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller & E. Teller, *J. Chem. Phys.* 21(6): 1087-1092 (1953).
N. Metropolis, *Los Alamos Science* Special Issue, 15: 125-130 (1987).
M. N. Rosenbluth, *AIP Confer. Proc.* 690(1): 22-30 (2003).

Simple sampling: MC estimation of π

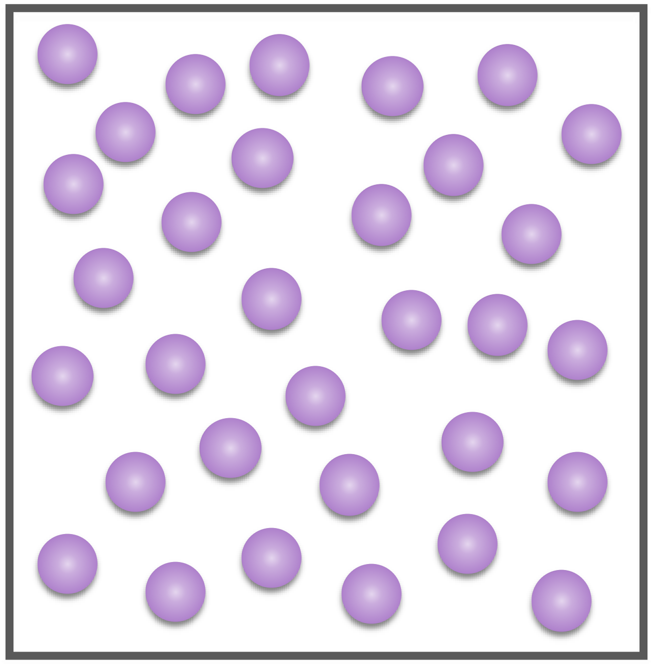


- Exact value: $\pi = 4A_{\mathrm{circle}}/A_{\mathrm{square}}$
- MC estimate: $\pi \approx 4{\color{red}N_{\mathrm{red}}}/({\color{red}N_{\mathrm{red}}} + {\color{blue}N_{\mathrm{blue}}})$
- **How many MC cycles do we need to get a good estimate?**





Simple sampling fails …



224 hard discs

- Dimension of the **configuration space**: 448
- Each dim. sampled by 100 random points
- Total number of random configurations: 100^{448} !
- Random config. may be statistically insignificant

Importance sampling

- Idea: to only count the important configurations, i.e. microstates with considerable statistical weight
- Recall the basics of statistical mechanics

observable: $\langle \mathcal{O} \rangle = \dfrac{\int \mathrm{d}\mathbf{r}^N \mathcal{O}(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_\mathrm{B}T}}{\int \mathrm{d}\mathbf{r}^N e^{-U(\mathbf{r}^N)/k_\mathrm{B}T}}$

config. part of partition function: $Z = \int \mathrm{d}\mathbf{r}^N e^{-U(\mathbf{r}^N)/k_\mathrm{B}T}$

probability density function: $p(\mathbf{r}^N) = e^{-U(\mathbf{r}^N)/k_\mathrm{B}T}/Z$

$$\langle \mathcal{O} \rangle = \int \mathcal{O}(\mathbf{r}^N) \underline{p(\mathbf{r}^N) \mathrm{d}\mathbf{r}^N}$$

- Strategy: to generate *M* microstates according to the Boltzmann distribution $p(\mathbf{r}^N)$

$$\boxed{\langle \mathcal{O} \rangle \approx \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}_i}$$

Markov process for generating chain of states

- *Stochastic* process that has *no memory*, i.e. the probability for getting to a new state depends only on the present state, but NOT on the history how the present state is reached.
- Process is fully defined by the **transition-probability** matrix

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1M} \\ P_{21} & P_{22} & \dots & P_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ P_{M1} & P_{M2} & \dots & P_{MM} \end{bmatrix}$$

with the requirements:

- each transition probability from *i* to *j* : $\mathbf{0 \leq P_{ij} \leq 1}$
- sum of each row is 1: $\sum_j \mathbf{P_{ij} = 1}$
- prob. of staying in the present state could be non-zero: $\mathbf{P_{ii} \neq 0}$

How to construct the Markov transition-probability matrix?

Stationary distribution of Markov process

- For a system in equilibrium, the transitions into and out of any state must be balanced:

$$\sum_j p_j P_{ji} = \sum_j p_i P_{ij} = p_i$$

  p is the **stationary distribution** that does not change under further application of the transition-probability matrix ***P***. *Once the equilibrium is reached, the distribution will not change any more.*

- An arbitrary initial distribution will converge to the stationary (equilibrium) distribution ***p*** if one repeats the Markov process, given that ***P*** is **irreducible** and **aperiodic**.
  - irreducible: any state is reachable from any other state in finite steps, i.e. $[P^m]_{ij} > 0$ for any *ij* pair – condition of ergodicity 各态历经
  - aperiodic: if there exists a state *i* such that $\boldsymbol{P_{ii} > 0}$, ***P*** is aperiodic. E.g., $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has a stationary distribution $p = (0.5 \;\; 0.5)$, but $p^* \cdot P^m$ alternates between p^* and $\mathrm{I}\text{-}p^*$.

Condition of detailed balance

- Detailed balance ***p_j P_{ji} = p_i P_{ij}*** is a sufficient but unnecessary condition for
$$\sum_j p_j P_{ji} = \sum_j p_i P_{ij} = p_i$$
and widely used to construct the transition-probability matrix ***P***.

- Under the **conditions of ergodicity and detailed balance**, one uses the Markov process to generate a chain of states. If the probability of remaining in a state is non-zero, i.e. ***P_{ii}* ≠ 0**, one will end up with states that obey the equilibrium distribution ***p***.

Example of a three-state system

- Equilibrium distribution: $p = (p_1\, p_2\, p_3) = (0.25\ 0.5\ 0.25)$

- For a transition-probability matrix

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.5 & 0.5 & 0 \end{pmatrix}$$

  that satisfies the detailed-balance condition $\boldsymbol{p_j P_{ji} = p_i P_{ij}}$,

$$p \cdot P = (0.25\ 0.5\ 0.25) = p$$

- ***p*** is the **stationary** or **limiting distribution** of the transition-probability matrix.

Example of a three-state system

- Start with an initial distribution $p' = (0 \; 0.05 \; 0.95)$

- Application of the transition-probability matrix

$$p' \cdot P = (0 \; 0.05 \; 0.95) \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.5 & 0.5 & 0 \end{pmatrix} = (0.4875 \; 0.5 \; 0.0125)$$

$$p' \cdot P^{10} = (0.249536 \; 0.5 \; 0.250464)$$

$$p' \cdot P^{20} = (0.25 \; 0.5 \; 0.25)$$

leads to the stationary distribution.

Metropolis algorithm

- Transition probability = **probability for selecting a transition** X **probability for accepting the transition**:

$$p_i \underline{P_{ij}} = p_j \underline{P_{ji}}$$

$$p_i \underline{g_{ij} A_{ij}} = p_j \underline{g_{ji} A_{ji}}$$

Metropolis solution for the acceptance ratio:

$$\boxed{A_{ij} = \min\left\{1, \frac{p_j g_{ji}}{p_i g_{ij}}\right\}, A_{ji} = \min\left\{1, \frac{p_i g_{ij}}{p_j g_{ji}}\right\}}$$

which reduces **for symmetric selection probability** $g_{ij} = g_{ji}$ to

$$\boxed{A_{ij} = \min\{1, p_j/p_i\}, A_{ji} = \min\{1, p_i/p_j\}}$$

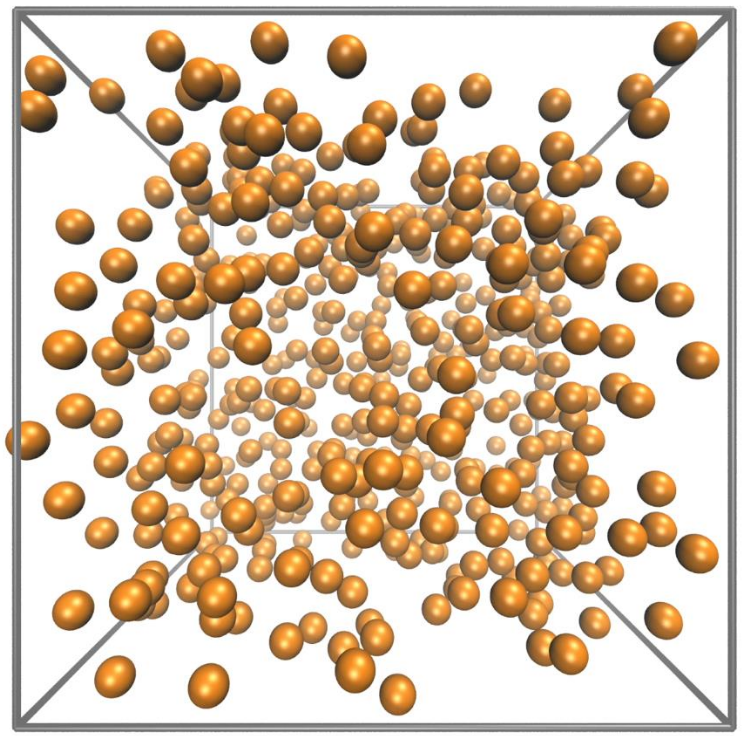
- Other but less common solutions exist, e.g., Glauber algorithm

$$A_{ij} = \frac{p_j g_{ji}}{p_j g_{ji} + p_i g_{ij}} = \frac{1}{1 + \frac{p_i g_{ij}}{p_j g_{ji}}}$$

R. J. Glauber, *J. Math. Phys.* 4(2): 294-307 (1963).

Case study: MC simulations of 3D LJ fluid

- Forward move to generate a new state:
  - choose a particle x, prob. = $1/N$
  - translate x around old position, prob. $\propto 1/(2d)^3$
  - accept the move with prob. = $\min\{1, e^{-\beta(U_n-U_o)}\}$



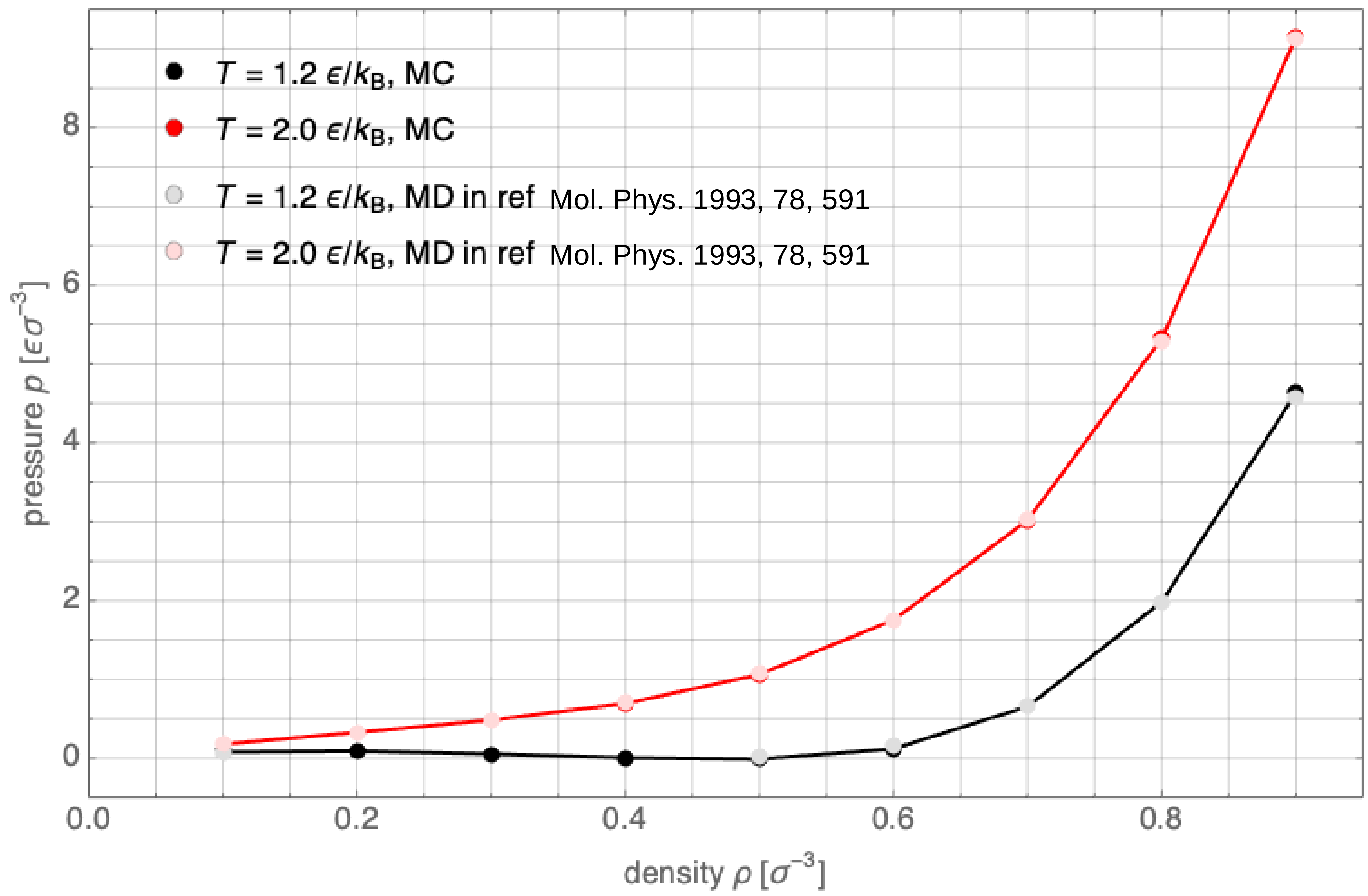
- Reverse move following the same protocol:
  - prob. of choosing particle x = $1/N$
  - prob. of translating the particle back $\propto 1/(2d)^3$
  - prob. of accepting the move = $\min\{1, e^{-\beta(U_o-U_n)}\}$

- Check condition of ergodicity

- Check detailed-balance condition

$$\frac{e^{-\beta U_\mathrm{o}}\mathrm{d}\mathbf{r}_\mathrm{o}}{Z}\cdot\underline{\frac{1}{N}\cdot\frac{1}{(2d)^3}}\cdot\underline{\min\{1, e^{-\beta(U_\mathrm{n}-U_\mathrm{o})}\}} \overset{?}{=} \frac{e^{-\beta U_\mathrm{n}}\mathrm{d}\mathbf{r}_\mathrm{n}}{Z}\cdot\underline{\frac{1}{N}\cdot\frac{1}{(2d)^3}}\cdot\underline{\min\{1, e^{-\beta(U_\mathrm{o}-U_\mathrm{n})}\}}$$

Case study: MC simulations of 3D LJ fluid



Remarks on MC simulations

- A MC move often involves a *local update* of the current configuration, in contrast to the global update in molecular dynamics simulations.

- Be careful with *rotational moves* in MC simulations.

- There are MC schemes that violate detailed-balance condition, but can still correctly sample the equilibrium distribution, e.g., sequential updating in Ising model.

- MC simulations can be used for the measurement of dynamical properties in a system if its dynamics is stochastic in nature and the MC moves are not unphysical.

V. I. Manousiouthakis & M. W. Deem, *J. Chem. Phys.* 111(6): 2753-2756 (1999).