

Principles of Monte Carlo Simulations

PhD Trial Lecture

Sigbjørn Løland Bore

Hylleraas Centre for Quantum Molecular Sciences,
Department of Chemistry, University of Oslo, Norway

Disputation, Oslo, 23 April 2020

Outline:

Two parts:

- ▶ Part 1: Mathematical principles
- ▶ Part 2: Physical principles

Outline:

Two parts:

- ▶ Part 1: Mathematical principles
- ▶ Part 2: Physical principles

Principles \neq Technicalities:

- ▶ Fundamental concepts
- ▶ Simple examples
- ▶ Master student level

Mathematical Principles

The objective of Monte Carlo (MC) simulations

What are Monte Carlo methods:

Broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

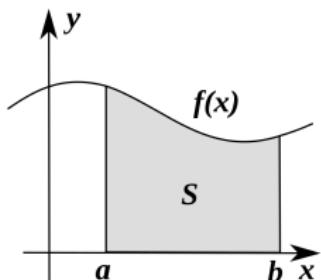
The objective of Monte Carlo (MC) simulations

What are Monte Carlo methods:

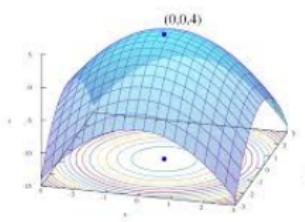
Broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

Three main uses:

Integration



Optimization



Generating
random draws

$$P(\Gamma) = ?$$

\downarrow

$$\Gamma_i$$

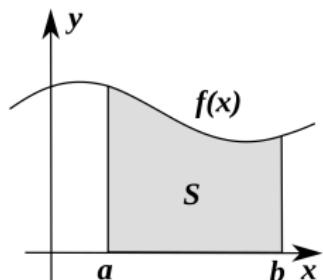
The objective of Monte Carlo (MC) simulations

What are Monte Carlo methods:

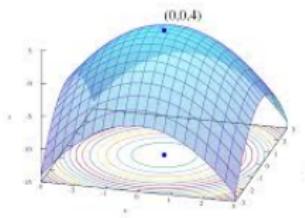
Broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

Three main uses:

Integration



Optimization



Generating
random draws

$$P(\Gamma) = ?$$

\downarrow

$$\Gamma_i$$

Integration

$$I = \int_a^b dx f(x),$$

Integration

$$I = \int_a^b dx f(x),$$

Analytic approach:

$$I = F(b) - F(a)$$

Integration

$$I = \int_a^b dx f(x),$$

Analytic approach:

$$I = F(b) - F(a)$$

Numerical quadrature methods:

$$I \simeq \frac{b-a}{2N} \sum_{i=1}^N (f(x_{i+1}) + f(x_i))$$

Integration

$$I = \int_a^b dx f(x),$$

Analytic approach:

$$I = F(b) - F(a)$$

Numerical quadrature methods:

$$I \simeq \frac{b-a}{2N} \sum_{i=1}^N (f(x_{i+1}) + f(x_i)) \simeq (b-a) \underbrace{\frac{1}{N} \sum_i^{N-1} f(x_{i+0.5})}_{\langle f \rangle}$$

Integration

$$I = \int_a^b dx f(x),$$

Analytic approach:

$$I = F(b) - F(a)$$

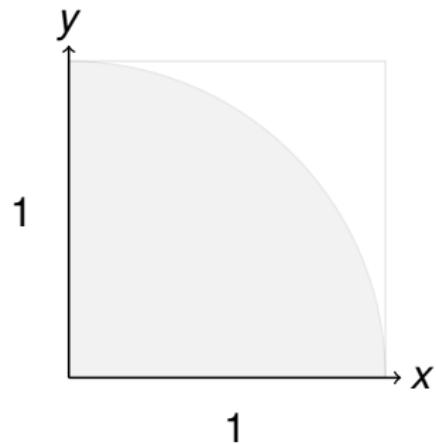
Numerical quadrature methods:

$$I \simeq \frac{b-a}{2N} \sum_{i=1}^N (f(x_{i+1}) + f(x_i)) \simeq (b-a) \underbrace{\frac{1}{N} \sum_i^{N} f(x_{i+0.5})}_{\langle f \rangle}$$

Monte Carlo approach:

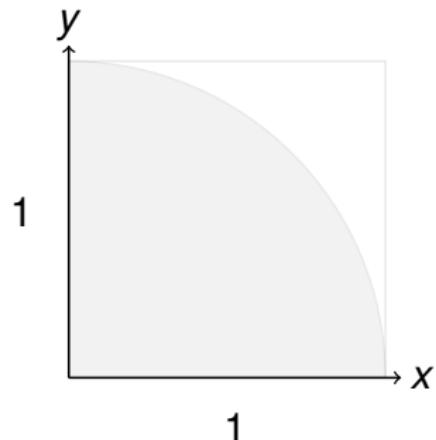
$$I = (b-a) \langle f \rangle, \quad I = V \langle f \rangle \text{ for multiple dimensions}$$

Integration example: The area of a circle (1)



$$A(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 < 1 \\ 0, & \text{if } x^2 + y^2 \geq 1 \end{cases}$$

Integration example: The area of a circle (1)

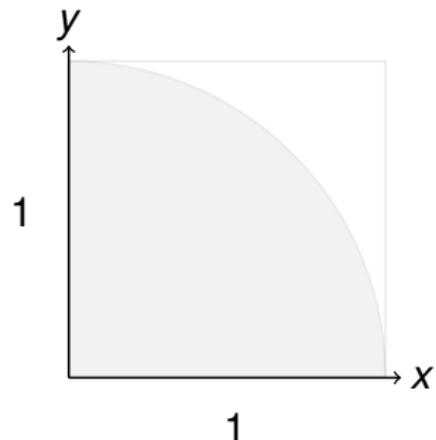


$$A(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 < 1 \\ 0, & \text{if } x^2 + y^2 \geq 1 \end{cases}$$

The area of a quarter of a circle $\times 4$:

$$I = 4 \int_0^1 dx \int_0^1 dy A(x, y)$$

Integration example: The area of a circle (1)

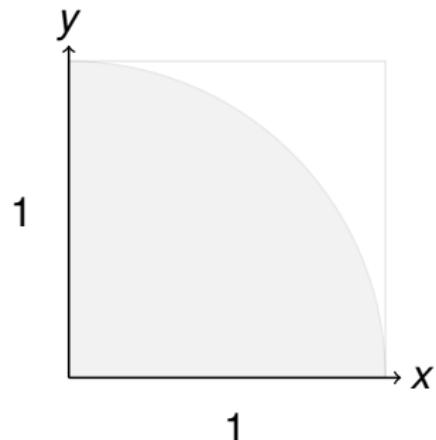


$$A(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 < 1 \\ 0, & \text{if } x^2 + y^2 \geq 1 \end{cases}$$

The area of a quarter of a circle $\times 4$:

$$I = 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy A(x, y) = 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy 1$$

Integration example: The area of a circle (1)

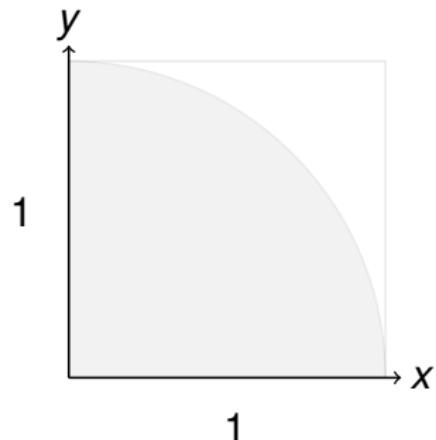


$$A(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 < 1 \\ 0, & \text{if } x^2 + y^2 \geq 1 \end{cases}$$

The area of a quarter of a circle $\times 4$:

$$\begin{aligned} I &= 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy A(x, y) = 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy 1 \\ &= 4 \int_0^1 dx \sqrt{1-x^2} = 2 \arcsin(1) \end{aligned}$$

Integration example: The area of a circle (1)



$$A(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 < 1 \\ 0, & \text{if } x^2 + y^2 \geq 1 \end{cases}$$

The area of a quarter of a circle $\times 4$:

$$\begin{aligned} I &= 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy A(x, y) = 4 \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy 1 \\ &= 4 \int_0^1 dx \sqrt{1-x^2} = 2 \arcsin(1) = 3.141592653\cdots = \underline{\underline{\pi}} \end{aligned}$$

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

We know: $V = L_x \times L_y = 1 \times 1 = 1$

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq \frac{4}{N} \sum_{i=1}^N A(x_i, y_i),$$

where x_i, y_i are drawn from a uniform random distribution.

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq 4 \frac{\text{points inside}}{\text{points inside} + \text{outside}},$$

where x_i, y_i are drawn from a uniform random distribution.

Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

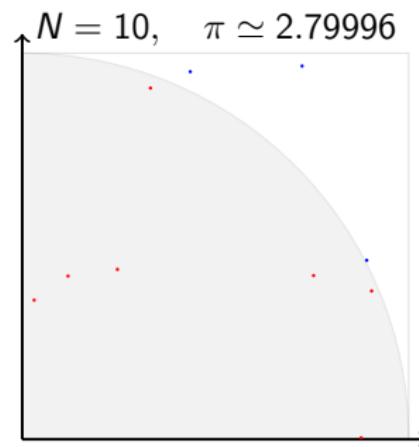
We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq 4 \frac{\text{points inside}}{\text{points inside} + \text{outside}},$$

where x_i, y_i are drawn from a uniform random distribution.



Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

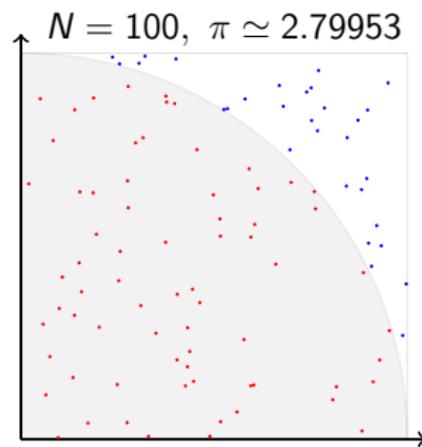
We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq 4 \frac{\text{points inside}}{\text{points inside} + \text{outside}},$$

where x_i, y_i are drawn from a uniform random distribution.



Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

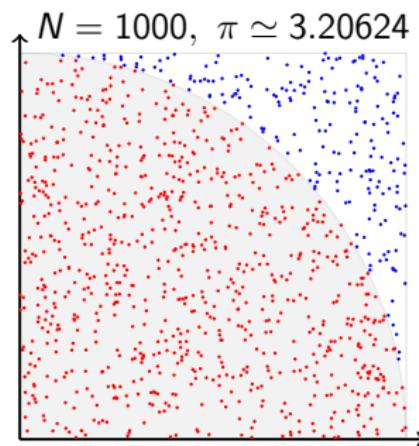
We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq 4 \frac{\text{points inside}}{\text{points inside} + \text{outside}},$$

where x_i, y_i are drawn from a uniform random distribution.



Integration example: The area of a circle (2)

MC estimate:

$$\pi = 4 \int_0^1 dx \int_0^1 dy A(x, y) \simeq 4 \langle A \rangle V$$

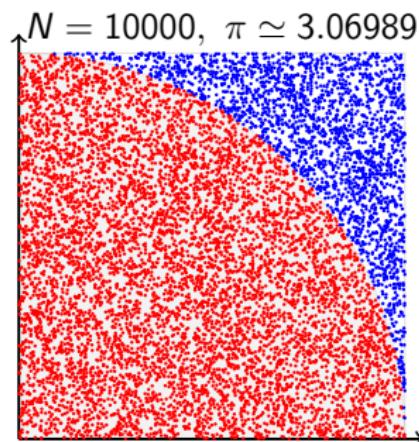
We know: $V = L_x \times L_y = 1 \times 1 = 1$

We don't know: $\langle A \rangle$

Random sampling of A :

$$\pi \simeq 4 \frac{\text{points inside}}{\text{points inside + outside}},$$

where x_i, y_i are drawn from a uniform random distribution.



Slow convergence, why should we use Monte Carlo?

Variance of the observable:

$$\text{Var}(A) \equiv \sigma_A^2 = \frac{1}{N-1} \sum_i (A(\Gamma_i) - \langle A \rangle)^2$$

Variance of the estimate:

$$\text{Var}(\langle A \rangle) = \text{Var}\left(\frac{1}{N} \sum_i A(\Gamma_i)\right) = \frac{\sigma_A^2}{N}$$

Standard deviation $\sim 1/\sqrt{N}$

Slow convergence, why should we use Monte Carlo?

Variance of the observable:

$$\text{Var}(A) \equiv \sigma_A^2 = \frac{1}{N-1} \sum_i (A(\Gamma_i) - \langle A \rangle)^2$$

Variance of the estimate:

$$\text{Var}(\langle A \rangle) = \text{Var}\left(\frac{1}{N} \sum_i A(\Gamma_i)\right) = \frac{\sigma_A^2}{N}$$

Standard deviation $\sim 1/\sqrt{N}$, but independent of the dimension!

Slow convergence, why should we use Monte Carlo?

Variance of the observable:

$$\text{Var}(A) \equiv \sigma_A^2 = \frac{1}{N-1} \sum_i (A(\Gamma_i) - \langle A \rangle)^2$$

Variance of the estimate:

$$\text{Var}(\langle A \rangle) = \text{Var}\left(\frac{1}{N} \sum_i A(\Gamma_i)\right) = \frac{\sigma_A^2}{N}$$

Standard deviation $\sim 1/\sqrt{N}$, but independent of the dimension!

Quadrature methods with N points per dimension:

Number of points needed $\sim N^D$

Number grows exponentially with the dimension!

But we can do better: Importance sampling

Rewrite:

$$I = \int_V d\Gamma A(\Gamma) = \int \frac{A(\Gamma)}{P(\Gamma)} P(\Gamma) d\Gamma$$

But we can do better: Importance sampling

Rewrite:

$$I = \int_V d\Gamma A(\Gamma) = \int \frac{A(\Gamma)}{P(\Gamma)} P(\Gamma) d\Gamma$$

Random draw of Γ_i from P :

$$I \simeq \frac{1}{N} \sum_{i=1}^N \frac{A(\Gamma_i)}{P(\Gamma_i)}, \quad \Gamma_i \sim P(\Gamma_i)$$

But we can do better: Importance sampling

Rewrite:

$$I = \int_V d\Gamma A(\Gamma) = \int \frac{A(\Gamma)}{P(\Gamma)} P(\Gamma) d\Gamma$$

Random draw of Γ_i from P :

$$I \simeq \frac{1}{N} \sum_{i=1}^N \frac{A(\Gamma_i)}{P(\Gamma_i)}, \quad \Gamma_i \sim P(\Gamma_i)$$

Variance of integral:

$$\begin{aligned} \text{Var}(I) &= \frac{1}{N^2} \sum_i \left(\frac{A}{P} - \left\langle \frac{A}{P} \right\rangle \right)^2 \\ &= \frac{1}{N} \left(\left\langle \frac{A^2}{P^2} \right\rangle - \left\langle \frac{A}{P} \right\rangle^2 \right) \end{aligned}$$

But we can do better: Importance sampling

Rewrite:

$$I = \int_V d\Gamma A(\Gamma) = \int \frac{A(\Gamma)}{P(\Gamma)} P(\Gamma) d\Gamma$$

Random draw of Γ_i from P :

$$I \simeq \frac{1}{N} \sum_{i=1}^N \frac{A(\Gamma_i)}{P(\Gamma_i)}, \quad \Gamma_i \sim P(\Gamma_i)$$

Variance of integral:

$$\begin{aligned} \text{Var}(I) &= \frac{1}{N^2} \sum_i \left(\frac{A}{P} - \left\langle \frac{A}{P} \right\rangle \right)^2 \\ &= \frac{1}{N} \left(\left\langle \frac{A^2}{P^2} \right\rangle - \left\langle \frac{A}{P} \right\rangle^2 \right) \end{aligned}$$

Zero variance: $P(\Gamma_i) = A(\Gamma_i)/\langle A \rangle$, but we don't have $\langle A \rangle$!

Importance sampling: example

$$I = \int_{-\infty}^{\infty} dx e^{-x^2/\sqrt{2\pi}} = 1,$$

Importance sampling: example

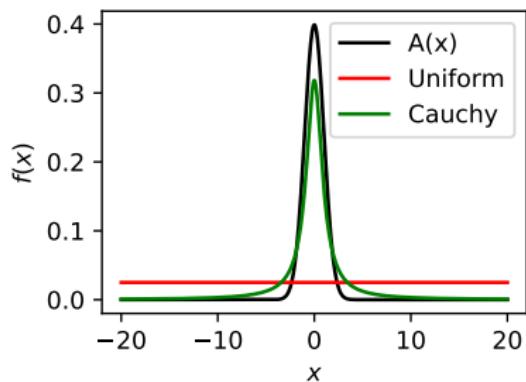
$$I = \int_{-\infty}^{\infty} dx e^{-x^2/\sqrt{2\pi}} = 1, \quad I_{MC} = \frac{1}{N} \sum_i^N \frac{e^{-x_i^2/\sqrt{2\pi}}}{P(x_i)}, \quad x_i \sim P$$

Importance sampling: example

$$I = \int_{-\infty}^{\infty} dx e^{-x^2/\sqrt{2\pi}} = 1, \quad I_{MC} = \frac{1}{N} \sum_i^N \frac{e^{-x_i^2/\sqrt{2\pi}}}{P(x_i)}, \quad x_i \sim P$$

$$P_{\text{Uni.}}(x) = \frac{1}{40}, \quad P_{\text{Cau.}}(x) = \frac{1}{\pi(1+x^2)}$$

Draw 1000 random numbers

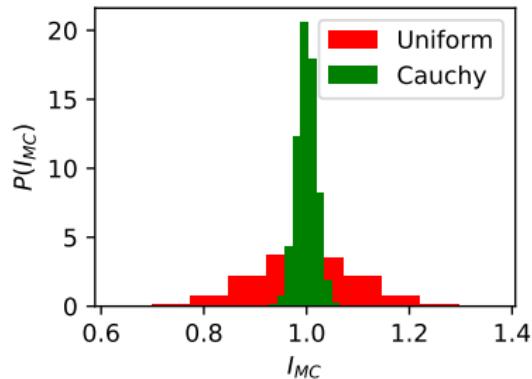
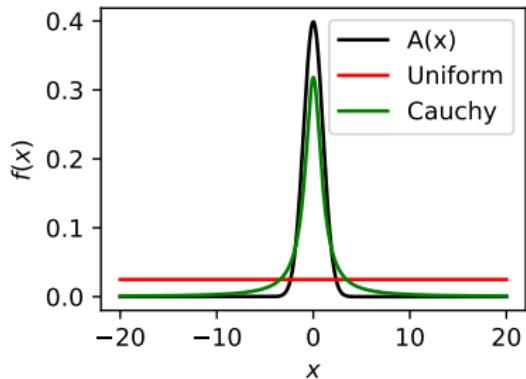


Importance sampling: example

$$I = \int_{-\infty}^{\infty} dx e^{-x^2/\sqrt{2\pi}} = 1, \quad I_{MC} = \frac{1}{N} \sum_i^N \frac{e^{-x_i^2/\sqrt{2\pi}}}{P(x_i)}, \quad x_i \sim P$$

$$P_{\text{Uni.}}(x) = \frac{1}{40}, \quad P_{\text{Cau.}}(x) = \frac{1}{\pi(1+x^2)}$$

Draw 1000 random numbers



Markov chains: The idea

Importance sampling \implies improved convergence

Markov chains: The idea

Importance sampling \implies improved convergence

But how do we generate Γ_i for any $P(\Gamma_i)$?

Markov chains: The idea

Importance sampling \implies improved convergence

But how do we generate Γ_i for any $P(\Gamma_i)$?

\implies Markov chains

Markov chains: The idea

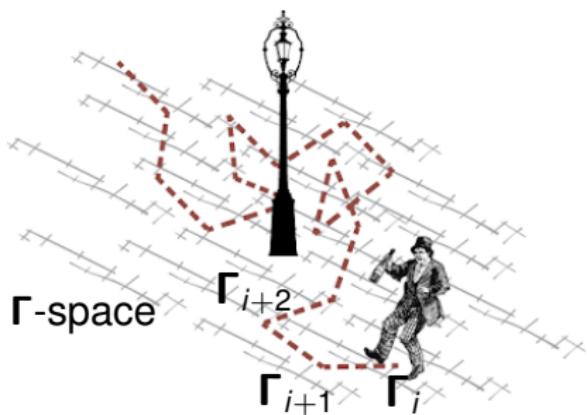
Importance sampling \implies improved convergence

But how do we generate Γ_i for any $P(\Gamma_i)$?

\implies Markov chains

What is a Markov chain? It is a random walk where:

$$P(\Gamma_{i+1} | \Gamma_i \dots \Gamma_1) = P(\Gamma_{i+1} | \Gamma_i)$$



Markov chains: The idea

Importance sampling \implies improved convergence

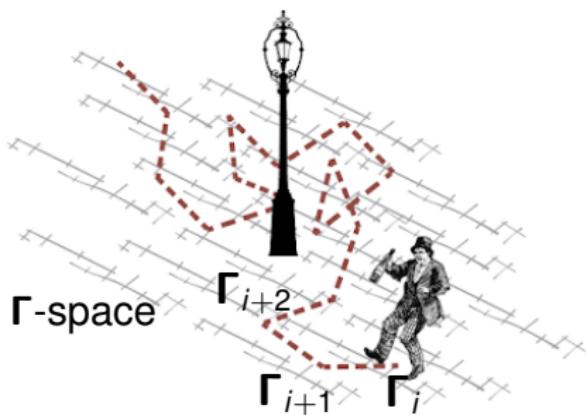
But how do we generate Γ_i for any $P(\Gamma_i)$?

\implies Markov chains

What is a Markov chain? It is a random walk where:

$$P(\Gamma_{i+1} | \Gamma_i \dots \Gamma_1) = P(\Gamma_{i+1} | \Gamma_i)$$

\implies The **future** is determined by the **present!** \equiv No memory



Markov chains: The idea

Importance sampling \implies improved convergence

But how do we generate Γ_i for any $P(\Gamma_i)$?

\implies Markov chains

What is a Markov chain? It is a random walk where:

$$P(\Gamma_{i+1} | \Gamma_i \dots \Gamma_1) = P(\Gamma_{i+1} | \Gamma_i)$$

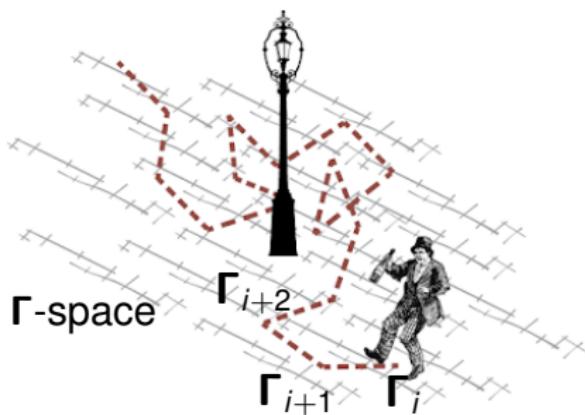
\implies The **future** is determined by the **present!** \equiv No memory

Goal:

Random walk in Γ -space

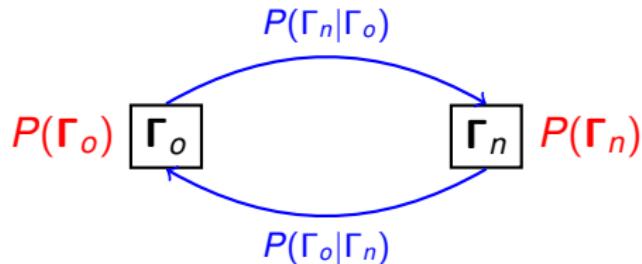
such that $\Gamma_i \sim P(\Gamma)$,

where P is decided by us!



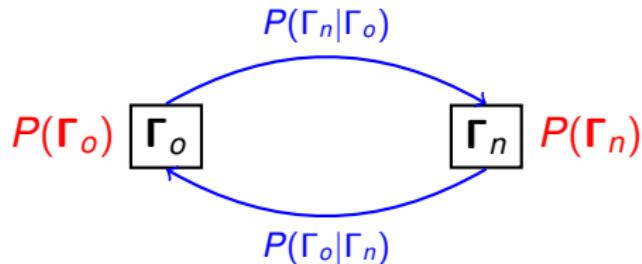
Markov chains: Equilibrium sampling

Flow of probability between states o and n :



Markov chains: Equilibrium sampling

Flow of probability between states o and n :

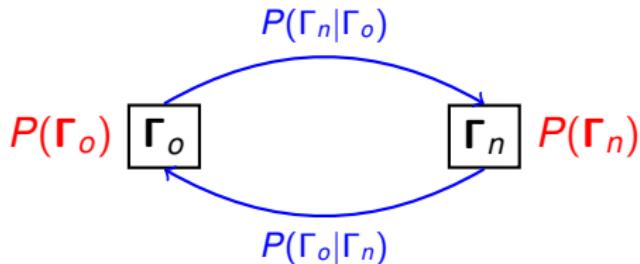


Detailed balance

$$P(\Gamma_n|\Gamma_o)P(\Gamma_o) = P(\Gamma_o|\Gamma_n)P(\Gamma_n)$$

Markov chains: Equilibrium sampling

Flow of probability between states o and n :



Detailed balance

$$P(\Gamma_n|\Gamma_o)P(\Gamma_o) = P(\Gamma_o|\Gamma_n)P(\Gamma_n)$$

⇒ No net probability flow ≡ Equilibrium sampling

MC Markov chains: The Metropolis algorithm

Decompose the transition probability between states o and n :

$$P(\Gamma_n | \Gamma_o) = \alpha(\Gamma_n, \Gamma_o) g(\Gamma_n | \Gamma_o),$$

(α : acceptance ratio, g : proposal distribution.)

MC Markov chains: The Metropolis algorithm

Decompose the transition probability between states o and n :

$$P(\Gamma_n | \Gamma_o) = \alpha(\Gamma_n, \Gamma_o) g(\Gamma_n | \Gamma_o),$$

(α : acceptance ratio, g : proposal distribution.)

Detailed balance $\implies \frac{\alpha(\Gamma_n, \Gamma_o)}{\alpha(\Gamma_o, \Gamma_n)} = \frac{P(\Gamma_n)g(\Gamma_o | \Gamma_n)}{P(\Gamma_o)g(\Gamma_n | \Gamma_o)}$

MC Markov chains: The Metropolis algorithm

Decompose the transition probability between states o and n :

$$P(\Gamma_n | \Gamma_o) = \alpha(\Gamma_n, \Gamma_o) g(\Gamma_n | \Gamma_o),$$

(α : acceptance ratio, g : proposal distribution.)

Detailed balance $\implies \frac{\alpha(\Gamma_n, \Gamma_o)}{\alpha(\Gamma_o, \Gamma_n)} = \frac{P(\Gamma_n)g(\Gamma_o | \Gamma_n)}{P(\Gamma_o)g(\Gamma_n | \Gamma_o)}$

The Metropolis Algorithm:

$$\alpha(\Gamma_n, \Gamma_o) = \text{MIN} \left(1, \frac{P(\Gamma_n)g(\Gamma_o | \Gamma_n)}{P(\Gamma_o)g(\Gamma_n | \Gamma_o)} \right)$$

MC Markov chains: The Metropolis algorithm

Decompose the transition probability between states o and n :

$$P(\Gamma_n | \Gamma_o) = \alpha(\Gamma_n, \Gamma_o) g(\Gamma_n | \Gamma_o),$$

(α : acceptance ratio, g : proposal distribution.)

$$\text{Detailed balance} \implies \frac{\alpha(\Gamma_n, \Gamma_o)}{\alpha(\Gamma_o, \Gamma_n)} = \frac{P(\Gamma_n)g(\Gamma_o | \Gamma_n)}{P(\Gamma_o)g(\Gamma_n | \Gamma_o)}$$

The Metropolis Algorithm:

$$\alpha(\Gamma_n, \Gamma_o) = \text{MIN} \left(1, \frac{P(\Gamma_n)g(\Gamma_o | \Gamma_n)}{P(\Gamma_o)g(\Gamma_n | \Gamma_o)} \right)$$

Symmetric case $g(\Gamma_o | \Gamma_n) = g(\Gamma_n | \Gamma_o)$:

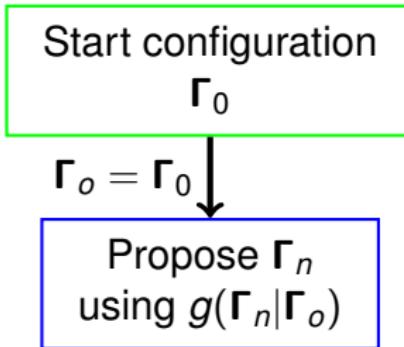
$$\alpha(\Gamma_n, \Gamma_o) = \text{MIN} \left(1, \frac{P(\Gamma_n)}{P(\Gamma_o)} \right)$$

Metropolis algorithm (2)

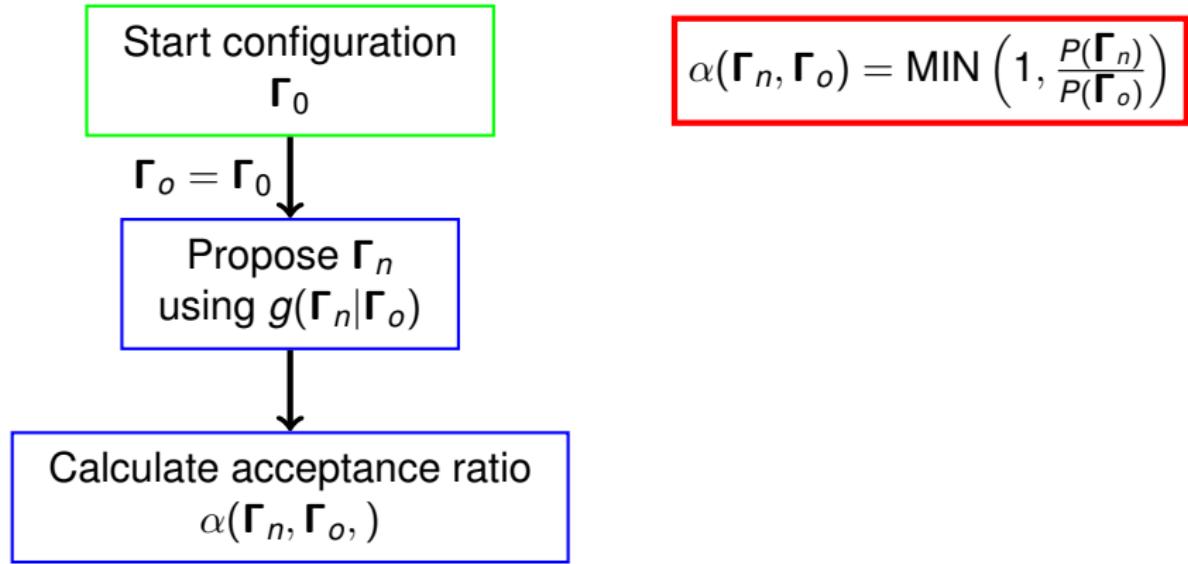
Start configuration

$$\Gamma_0$$

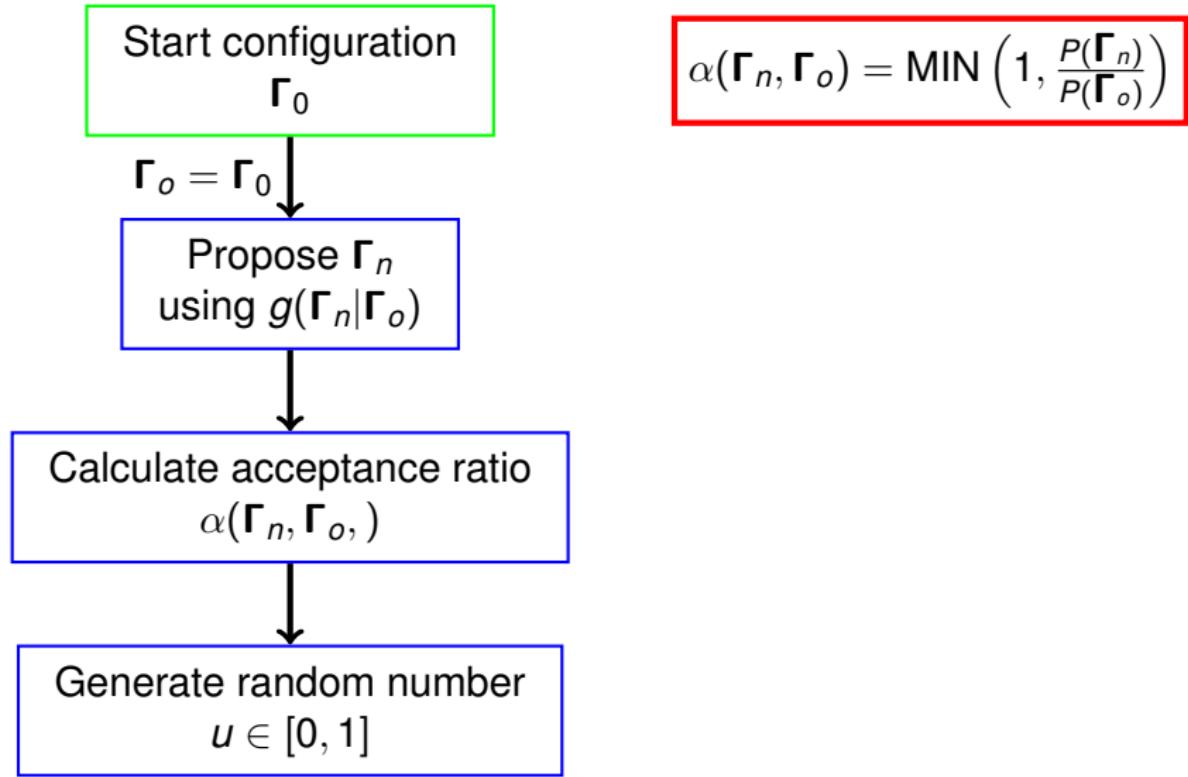
Metropolis algorithm (2)



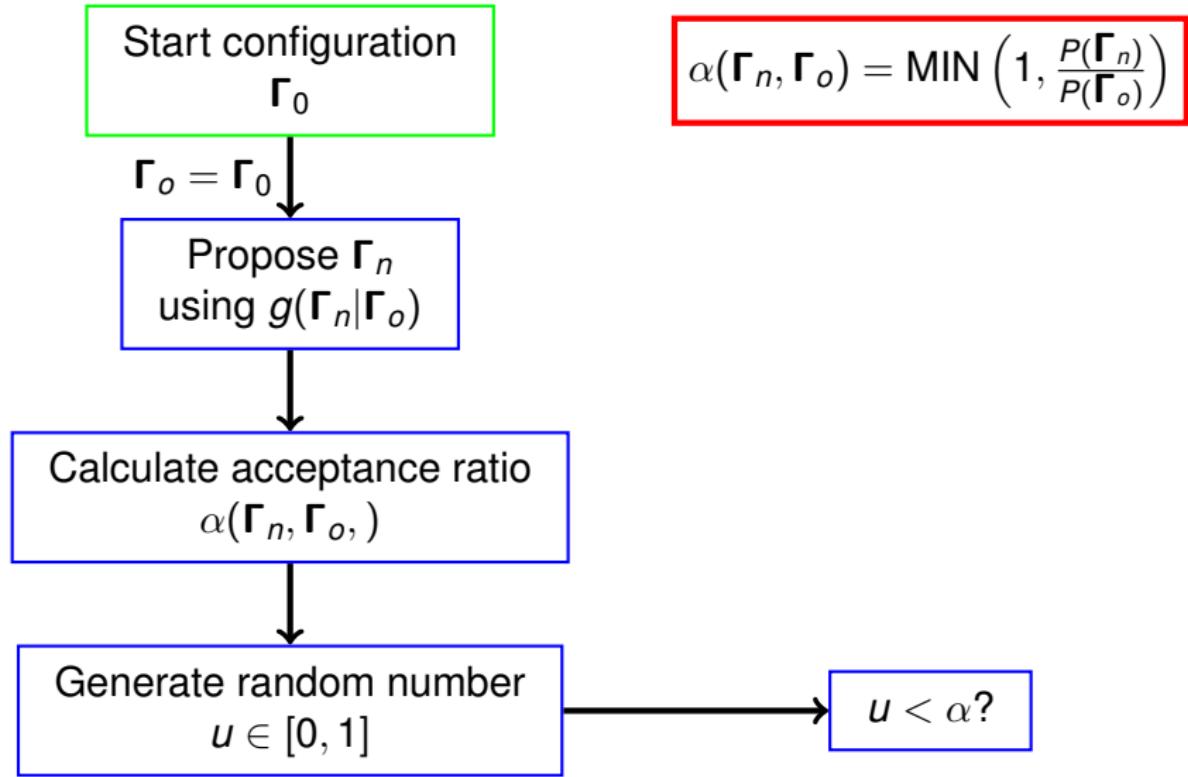
Metropolis algorithm (2)



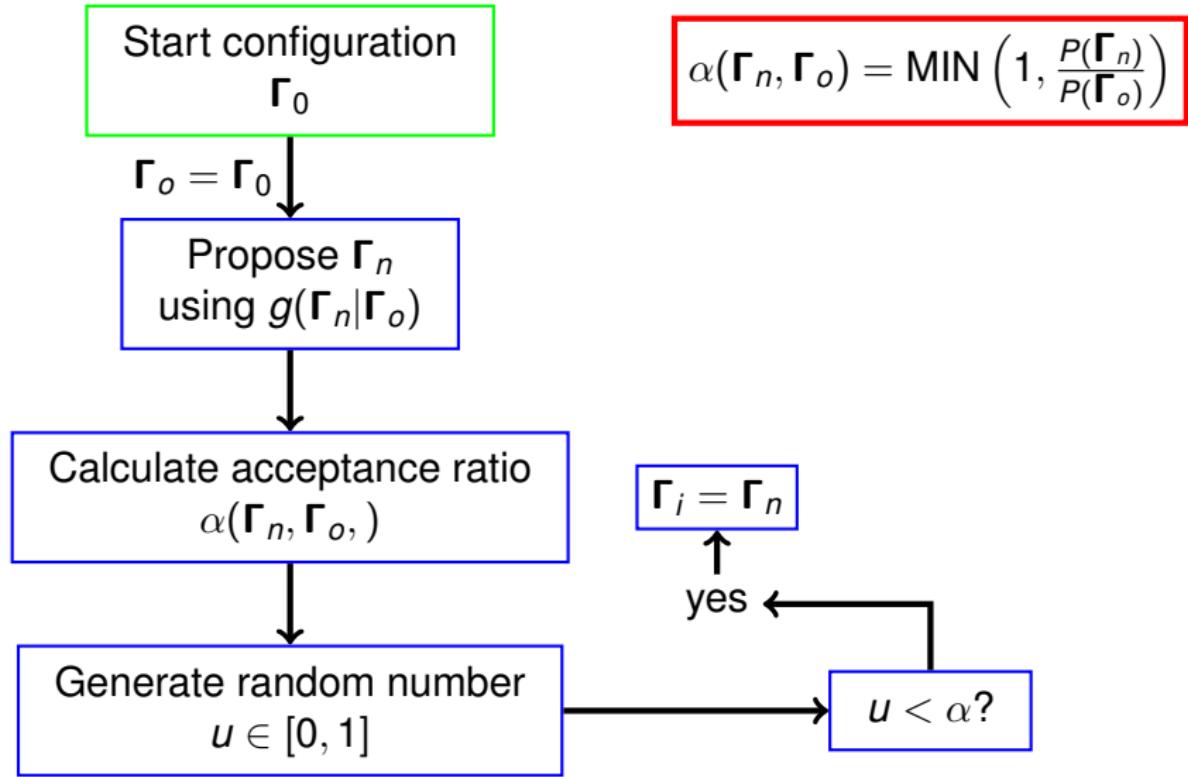
Metropolis algorithm (2)



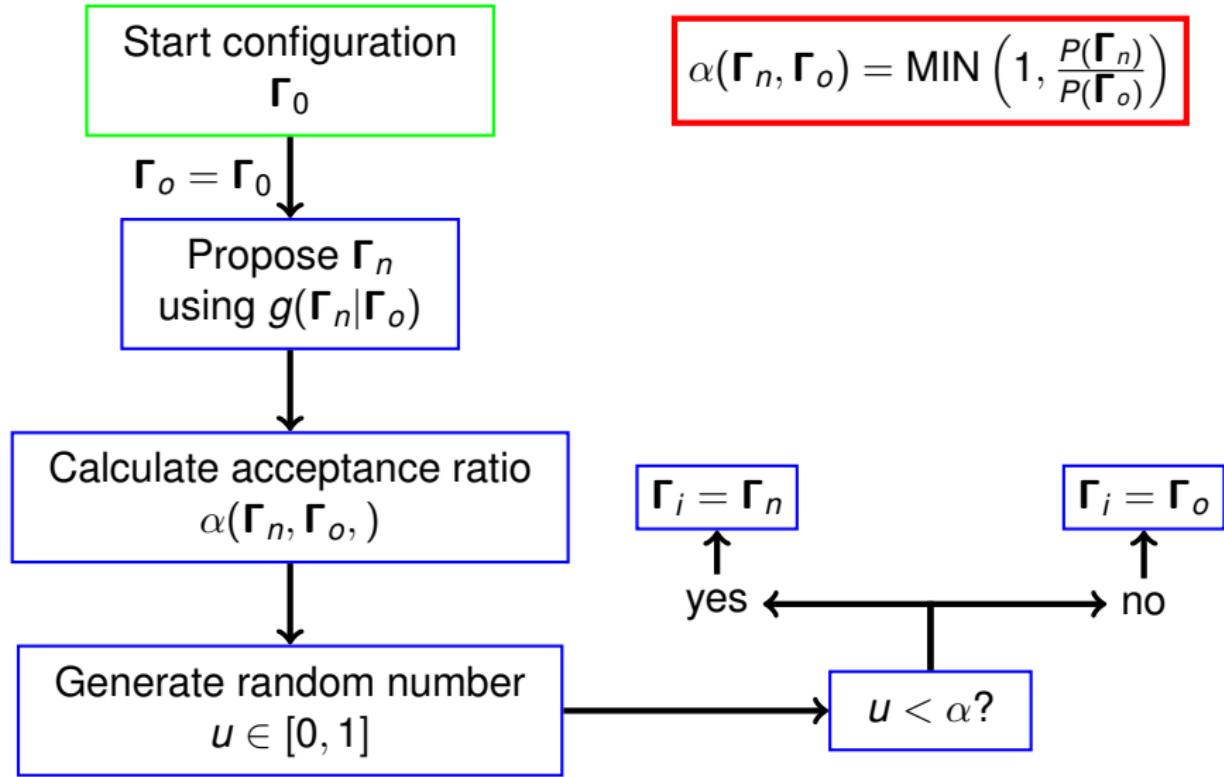
Metropolis algorithm (2)



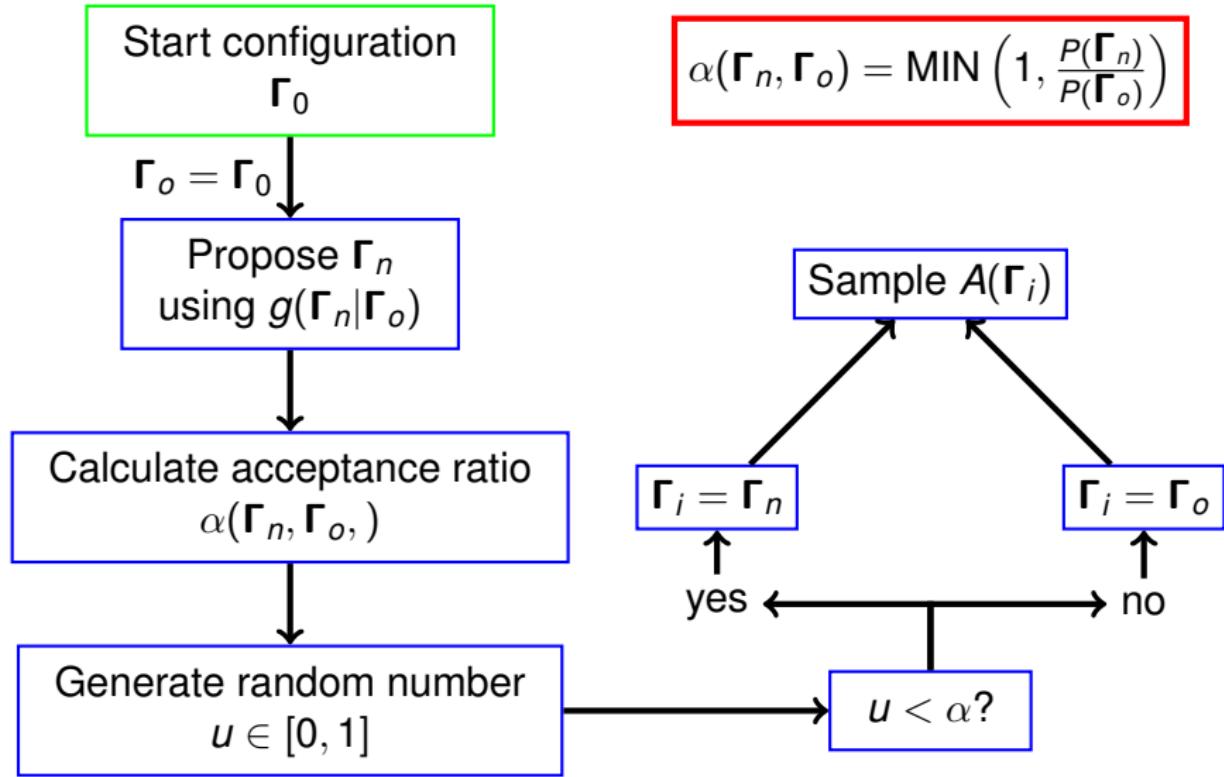
Metropolis algorithm (2)



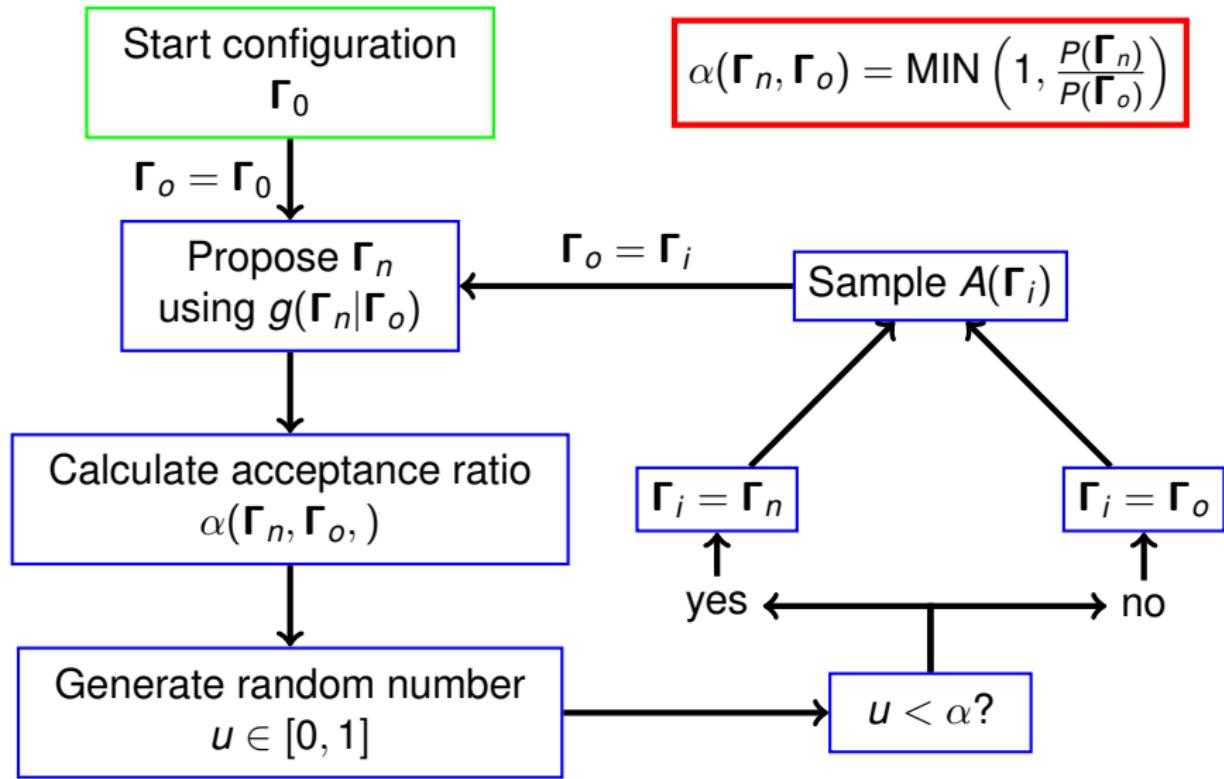
Metropolis algorithm (2)



Metropolis algorithm (2)



Metropolis algorithm (2)



$$\alpha(\Gamma_n, \Gamma_o) = \text{MIN} \left(1, \frac{P(\Gamma_n)}{P(\Gamma_o)} \right)$$

Simple example integral:

$$I = \int_{-\infty}^{\infty} dx \ x^2 e^{-x^2/\sqrt{2\pi}} = 1$$

Importance sampling using:

$$A(x) = x^2, \quad P(x) = e^{-x^2/\sqrt{2\pi}}$$

Simple example integral:

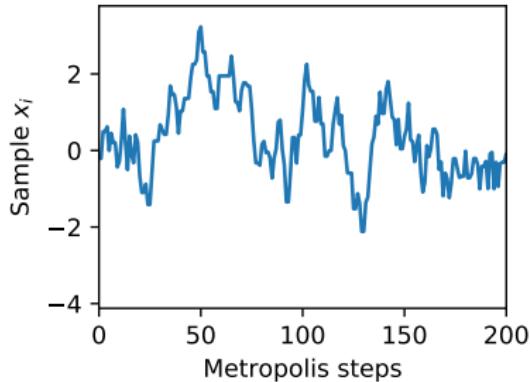
```
1  def P(x):
2      return norm.pdf(x)
3  # Initialize start
4  x_o=0
5  avg_A=0.0
6  P_o=P(x_o)
7  # MC loop
8  for i in range(N_MC_STEPS):
9      # Propose a move
10     x_n = x_o + (np.random.random()-0.5)*2
11     # Calculate acceptance ratio
12     P_n = P(x_n)
13     acc_ratio = np.min((1.,P_n/P_o))
14     if(acc_ratio>np.random.random()):
15         # Accept
16         x_o=x_n
17         P_o=P_n
18     # Gather data
19     x_i=x_o
20     avg_A += x_i**2/N_MC_STEPS
```

Simple example integral:

$$I = \int_{-\infty}^{\infty} dx \ x^2 e^{-x^2/\sqrt{2\pi}} = 1$$

Importance sampling using:

$$A(x) = x^2, \quad P(x) = e^{-x^2/\sqrt{2\pi}}$$

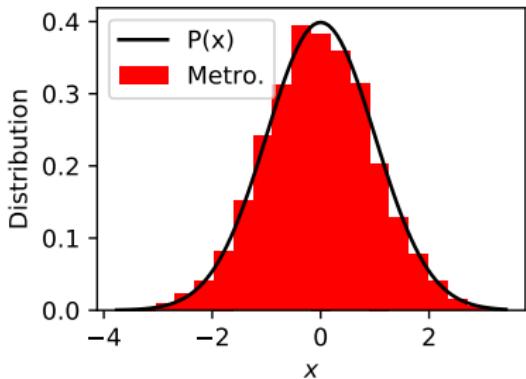
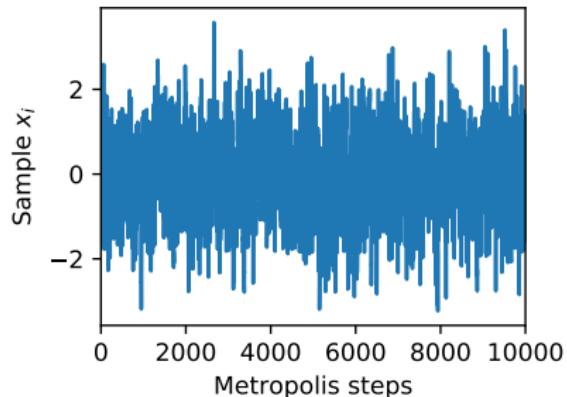


Simple example integral:

$$I = \int_{-\infty}^{\infty} dx \ x^2 e^{-x^2/\sqrt{2\pi}} = 1 \simeq 1.039$$

Importance sampling using:

$$A(x) = x^2, \quad P(x) = e^{-x^2/\sqrt{2\pi}}$$



Physical Principles

N particle configuration:

$$\Gamma = \mathbf{r}^N, \dot{\mathbf{r}}^N$$

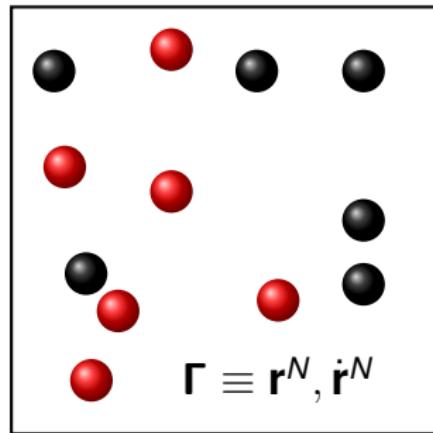
System with energy:

$$E = K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N)$$

Properties

$$\langle A(\mathbf{r}^N, \dot{\mathbf{r}}^N) \rangle_{NVT} = ?$$

NVT Ensemble



(N : Number of particles. V : Volume. T : Temperature.
 $\mathbf{r}^N, \dot{\mathbf{r}}^N$: Particle positions and velocities)

The Boltzmann distribution

$$P(E_i) = \frac{e^{-E_i/k_b T}}{\sum_j e^{-E_j/k_b T}}$$



Ludwig Boltzmann

The Boltzmann distribution

$$P(E_i) = \frac{e^{-E_i/k_b T}}{\sum_j e^{-E_j/k_b T}}$$

$k_b T \rightarrow 0$:

- ▶ Few configurations visited
(energy dominated)



Ludwig Boltzmann

The Boltzmann distribution

$$P(E_i) = \frac{e^{-E_i/k_b T}}{\sum_j e^{-E_j/k_b T}}$$

$k_b T \rightarrow 0$:

- ▶ Few configurations visited
(energy dominated)

$k_b T \gg 0$:

- ▶ Many configurations
visited
(entropy dominated)



Ludwig Boltzmann

The Boltzmann distribution

$$P(E_i) = \frac{e^{-E_i/k_b T}}{\sum_j e^{-E_j/k_b T}}$$

$k_b T \rightarrow 0$:

- ▶ Few configurations visited
(energy dominated)

$k_b T \gg 0$:

- ▶ Many configurations
visited
(entropy dominated)



Ludwig Boltzmann

For a particle system:

$$P(\mathbf{r}^N, \dot{\mathbf{r}}^N) = \frac{e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}}{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}},$$

Formulation as statistical physics problem

A system of N particles at temperature T in a volume V :

$$\langle A \rangle = \frac{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N A(\mathbf{r}^N, \dot{\mathbf{r}}^N) e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}}{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}},$$

Formulation as statistical physics problem

A system of N particles at temperature T in a volume V :

$$\langle A \rangle = \frac{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N A(\mathbf{r}^N, \dot{\mathbf{r}}^N) e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}}{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}},$$

Important simplification: Configurational mean

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int d\dot{\mathbf{r}}^N e^{-K(\dot{\mathbf{r}}^N)/k_b T} \int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{\int d\dot{\mathbf{r}}^N e^{-K(\dot{\mathbf{r}}^N)/k_b T} \int_V d\mathbf{r}^N e^{-U(\mathbf{r}^N)/k_b T}},$$

Formulation as statistical physics problem

A system of N particles at temperature T in a volume V :

$$\langle A \rangle = \frac{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N A(\mathbf{r}^N, \dot{\mathbf{r}}^N) e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}}{\int_V d\mathbf{r}^N d\dot{\mathbf{r}}^N e^{-(K(\dot{\mathbf{r}}^N) + U(\mathbf{r}^N))/k_b T}},$$

Important simplification: Configurational mean

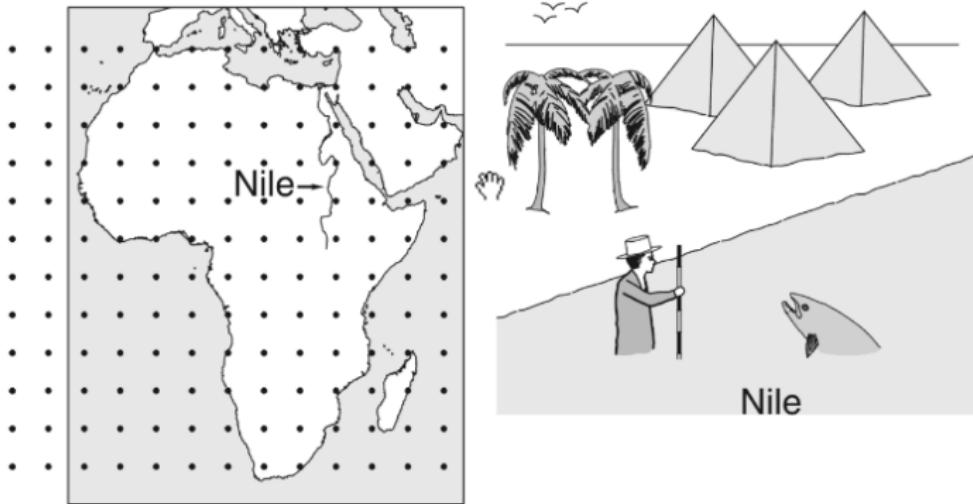
$$\langle A(\mathbf{r}^N) \rangle = \frac{\int d\dot{\mathbf{r}}^N e^{-K(\dot{\mathbf{r}}^N)/k_b T} \int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{\int d\dot{\mathbf{r}}^N e^{-K(\dot{\mathbf{r}}^N)/k_b T} \int_V d\mathbf{r}^N e^{-U(\mathbf{r}^N)/k_b T}},$$

Recast our problem as:

$$\Gamma \equiv \mathbf{r}^N, \quad P(\mathbf{r}^N) = \frac{e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)}, \quad Q(N, V, T) \equiv \int_V d\mathbf{r}^N e^{-U(\mathbf{r}^N)/k_b T}$$

Motivation for Monte Carlo simulations

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)},$$



Importance sampling by metropolis algorithm

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)},$$

Challenges for this problem:

1. No random configuration generator for arbitrary $P(\mathbf{r}^N)$
2. $Q(N, V, T)$ is unknown

Importance sampling by metropolis algorithm

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)},$$

Challenges for this problem:

1. No random configuration generator for arbitrary $P(\mathbf{r}^N)$
2. $Q(N, V, T)$ is unknown

Solution: **The Metropolis algorithm** (symmetric)

1. Works for any P

Importance sampling by metropolis algorithm

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)},$$

Challenges for this problem:

1. No random configuration generator for arbitrary $P(\mathbf{r}^N)$
2. $Q(N, V, T)$ is unknown

Solution: **The Metropolis algorithm** (symmetric)

1. Works for any P
2. Algorithm does not require $Q(N, V, T)$

Importance sampling by metropolis algorithm

$$\langle A(\mathbf{r}^N) \rangle = \frac{\int_V d\mathbf{r}^N A(\mathbf{r}^N) e^{-U(\mathbf{r}^N)/k_b T}}{Q(N, V, T)},$$

Challenges for this problem:

1. No random configuration generator for arbitrary $P(\mathbf{r}^N)$
2. $Q(N, V, T)$ is unknown

Solution: **The Metropolis algorithm** (symmetric)

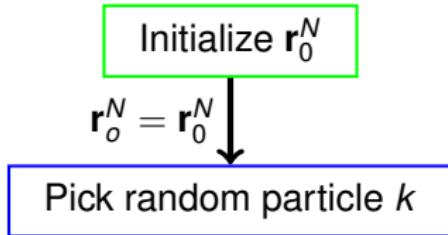
1. Works for any P
2. Algorithm does not require $Q(N, V, T)$

$$\begin{aligned}\alpha(\mathbf{r}_n^N, \mathbf{r}_o^N) &= \text{MIN}(1, P(\mathbf{r}_n^N)/P(\mathbf{r}_o^N)) \\ &= \text{MIN}(1, \frac{e^{-U(\mathbf{r}_n^N)/k_b T}/Q(N, V, T)}{e^{-U(\mathbf{r}_o^N)/k_b T}/Q(N, V, T)}) \\ &= \text{MIN}(1, e^{-(U(\mathbf{r}_n^N) - U(\mathbf{r}_o^N))/k_b T}), \text{ or } \text{MIN}(1, e^{-\Delta U/k_b T})\end{aligned}$$

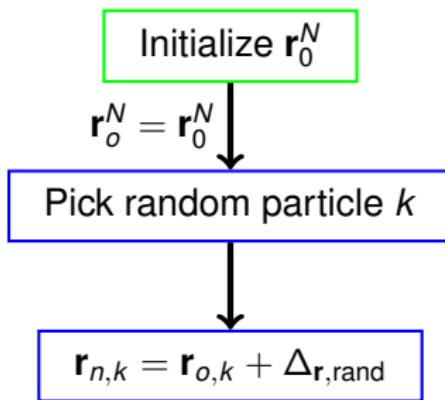
Metropolis algorithm for particle systems

Initialize \mathbf{r}_0^N

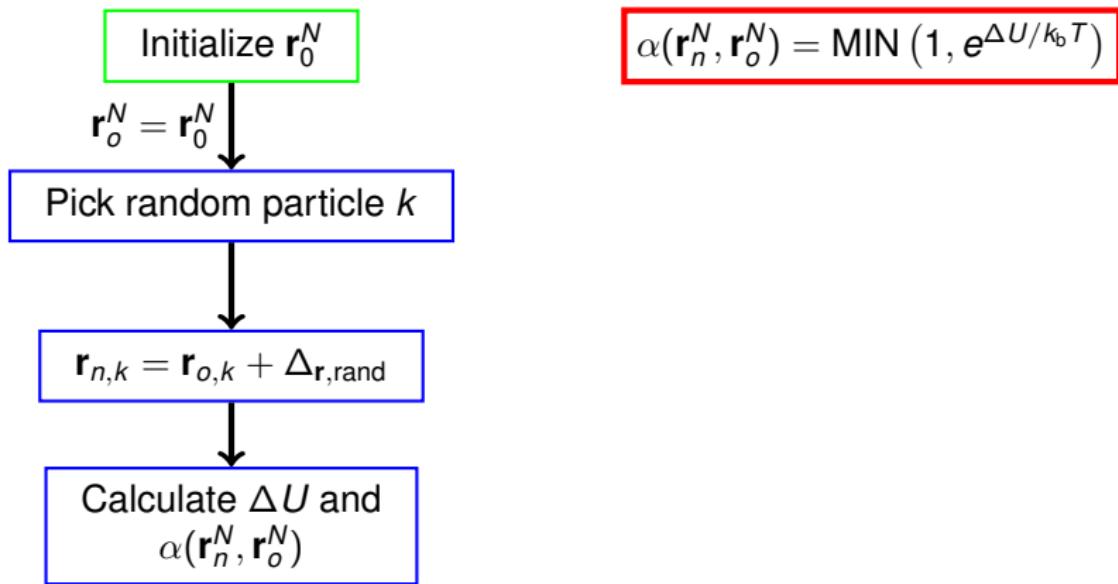
Metropolis algorithm for particle systems



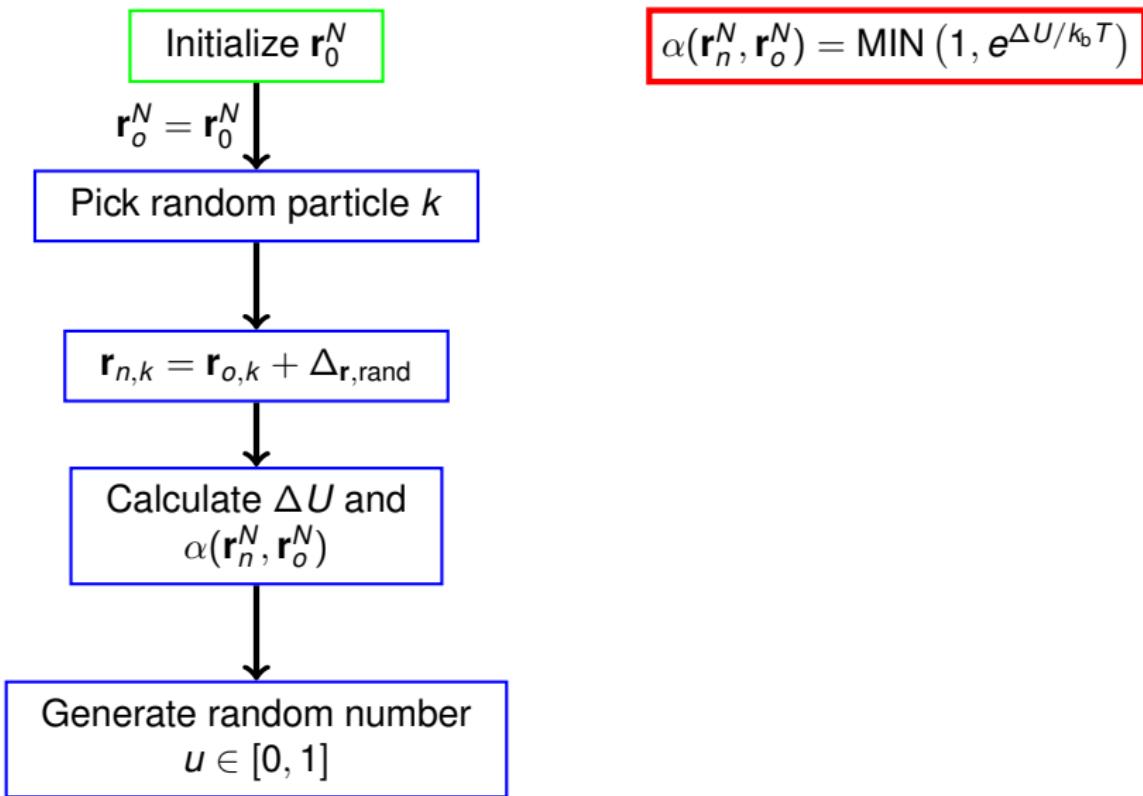
Metropolis algorithm for particle systems



Metropolis algorithm for particle systems

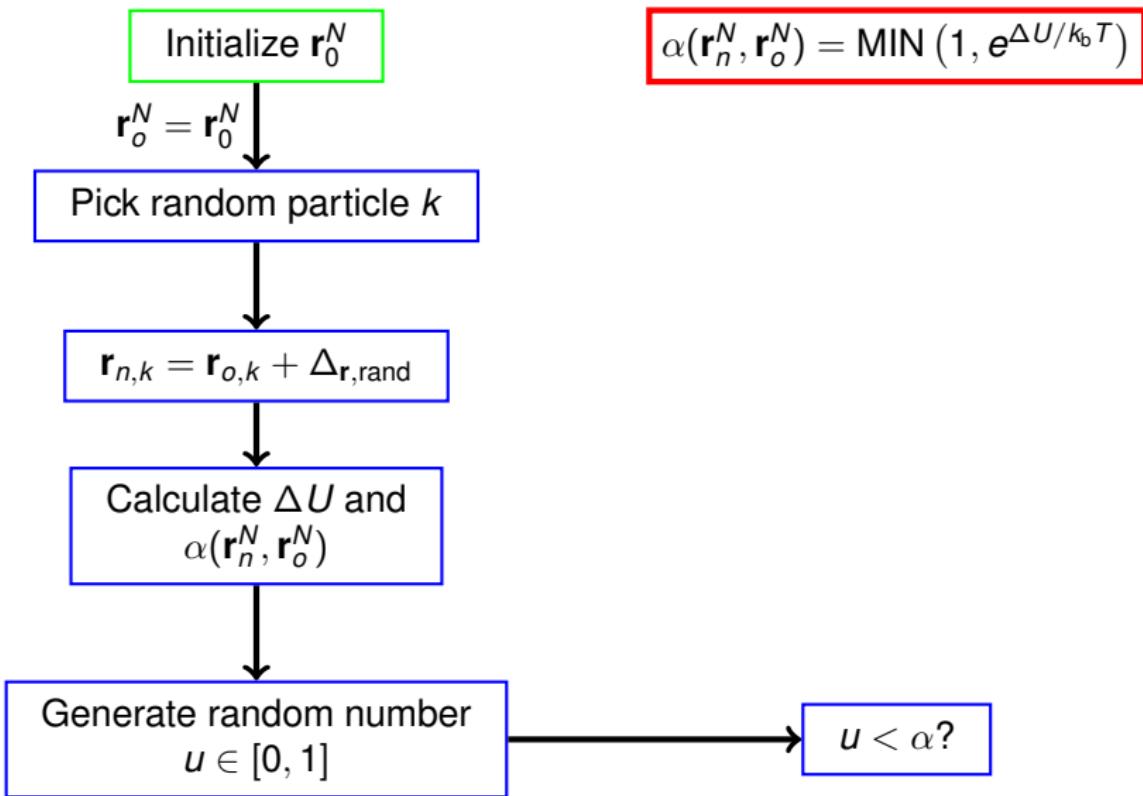


Metropolis algorithm for particle systems

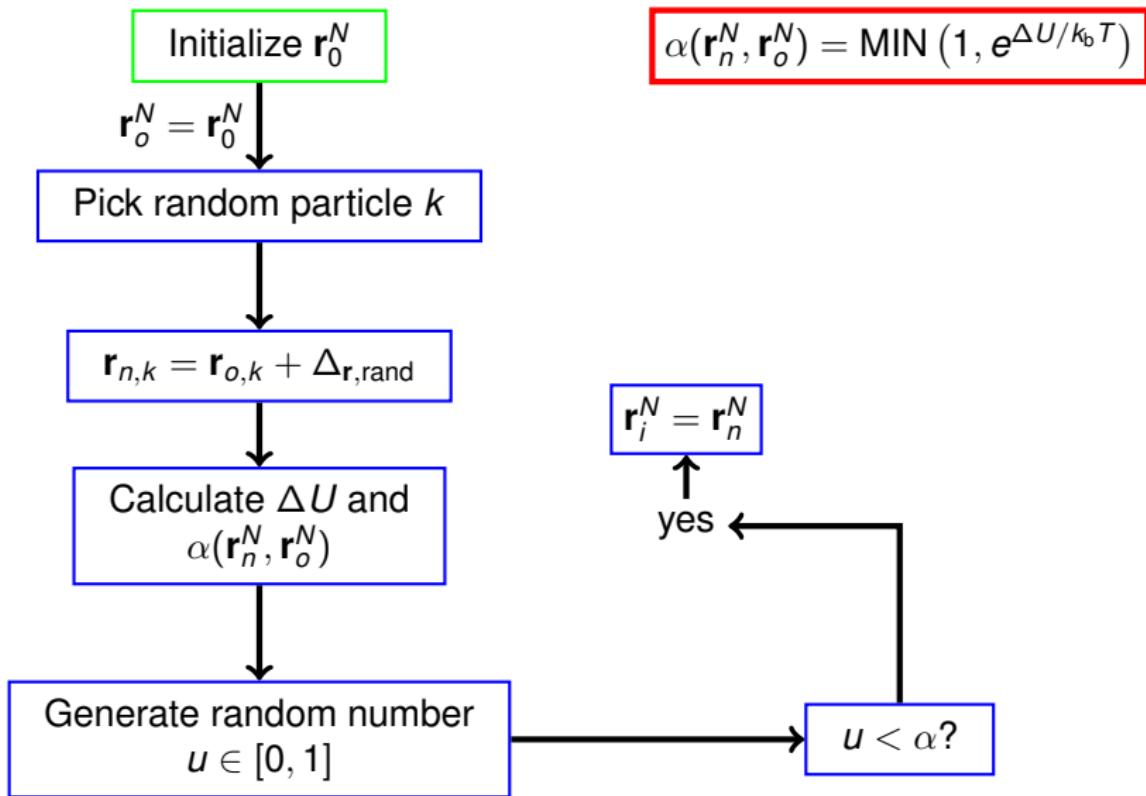


$$\alpha(\mathbf{r}_n^N, \mathbf{r}_o^N) = \text{MIN}(1, e^{\Delta U / k_b T})$$

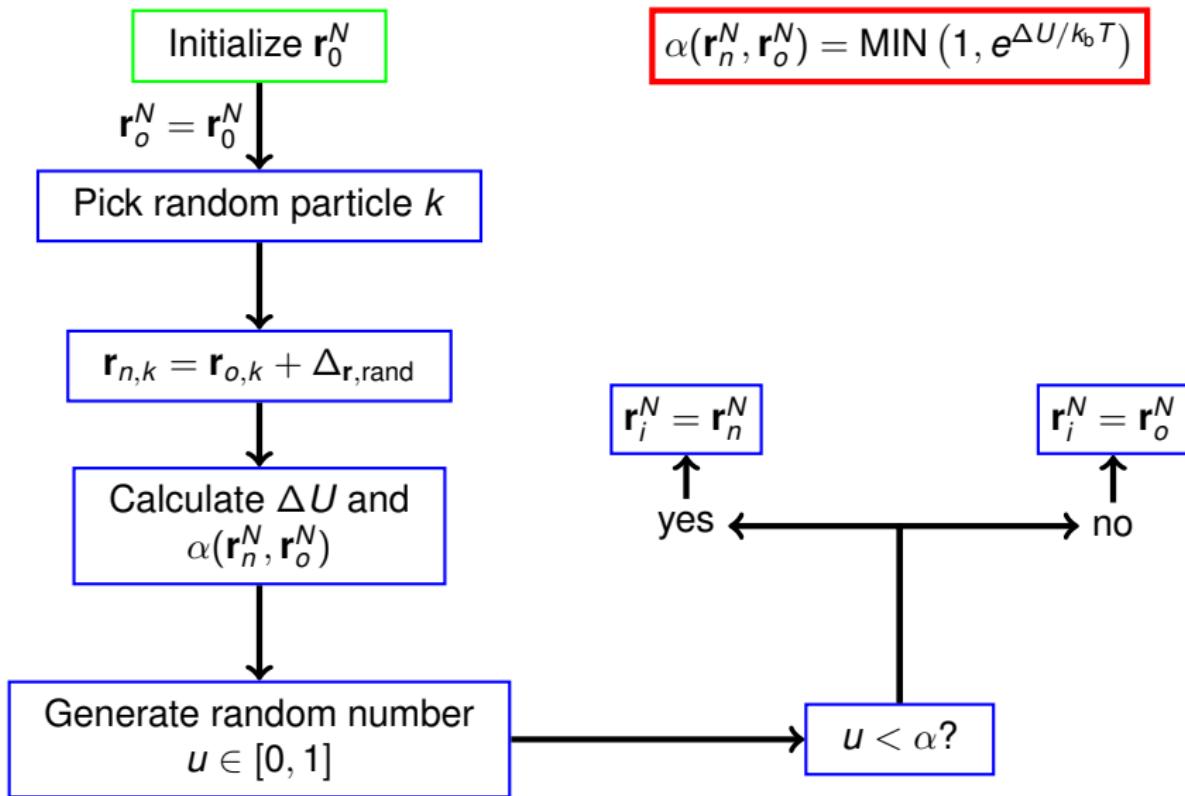
Metropolis algorithm for particle systems



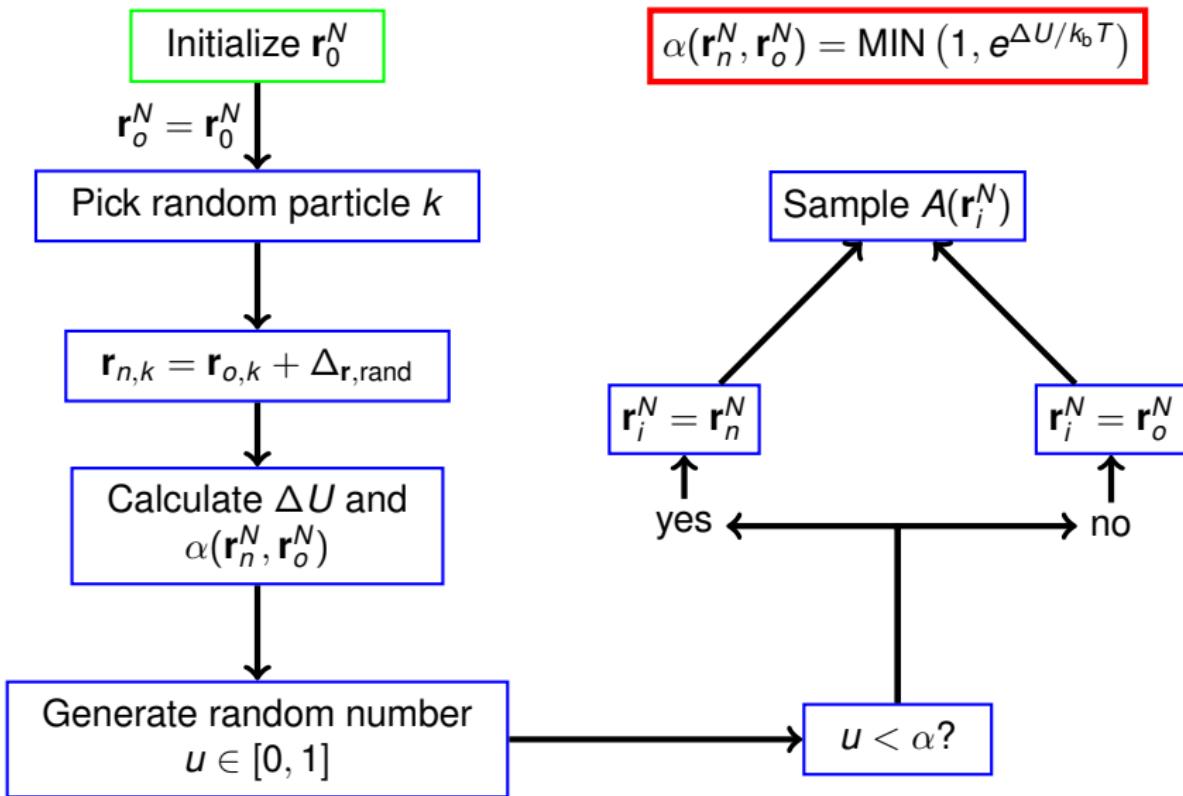
Metropolis algorithm for particle systems



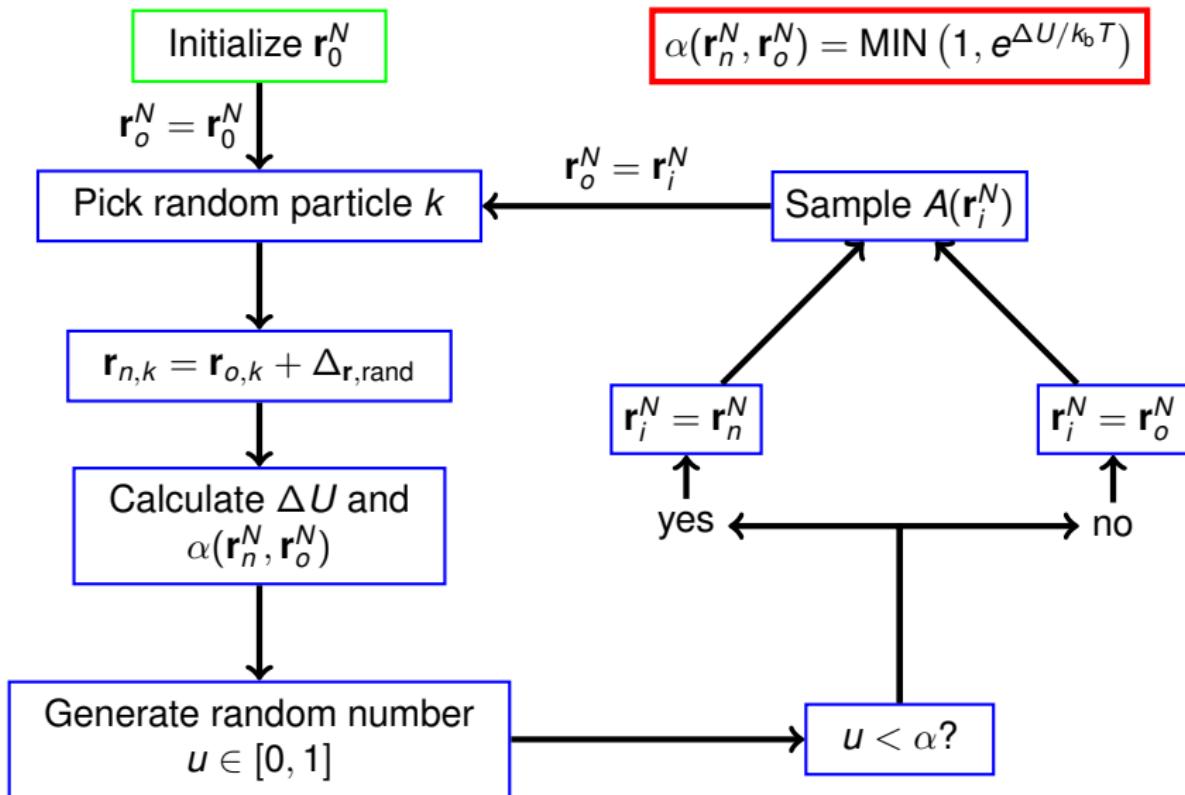
Metropolis algorithm for particle systems



Metropolis algorithm for particle systems



Metropolis algorithm for particle systems

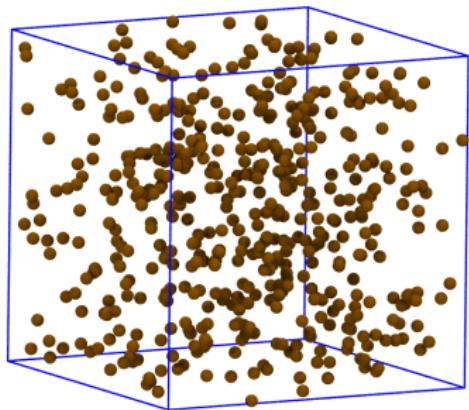
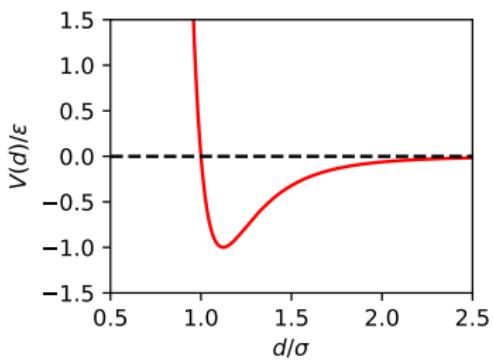


The Lennard-Jones model

$$U(\mathbf{r}^N) = \sum_{i < j} 4\epsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right)$$

Pauli excl. Attraction

Molecular system



NVT

Goal: Estimate statistical properties of the system

MC algorithm

```
1  for step in range(mc_steps):
2      # Pick a particle
3      k = np.random.randint(0,N)
4      r_o    = r[k]
5      # Move that particle
6      r_n = r_o + (np.random.random(size=3)-0.5)
7
8      # Pairs to that particle
9      r_pairs = np.delete(r,k, axis=0)
10
11     # Difference in energy
12     dU = LJ(r_n,r_pairs) - LJ(r_o,r_pairs)
13
14     # Metropolis acceptance ratio
15     acc_ratio = min(1, np.exp(-kbt_inv*dU))
16
17     #Accept?
18     if(np.random.random() < acc_ratio):
19         r[k] = r_n
20         E = E + dU
21     # Store particle positions
```

Three phases of Argon

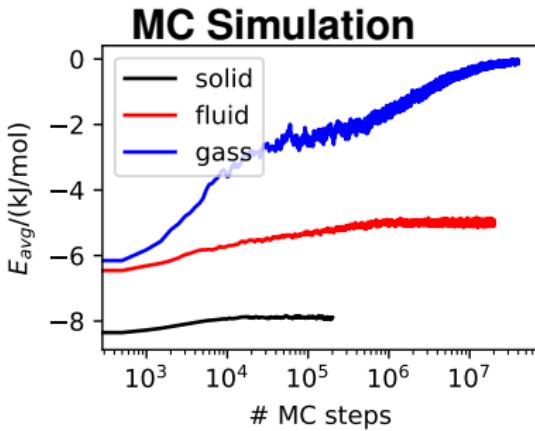
Table: Model for Argon

ϵ/J	1.65×10^{-21}
σ/m	3.4×10^{-10}

Three phases of Argon

Table: Model for Argon

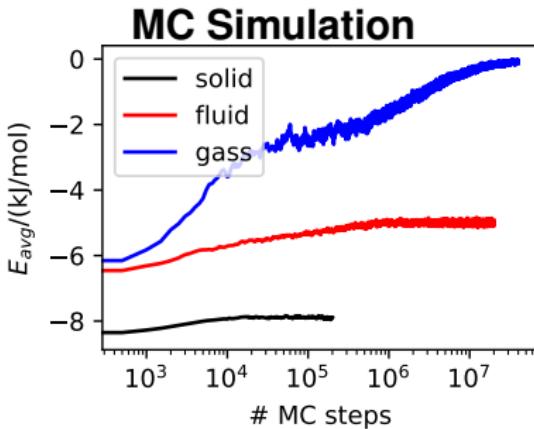
ϵ/J	1.65×10^{-21}
σ/m	3.4×10^{-10}



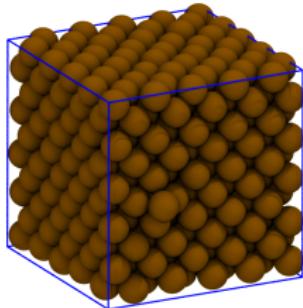
Three phases of Argon

Table: Model for Argon

ϵ/J	1.65×10^{-21}
σ/m	3.4×10^{-10}



Solid phase

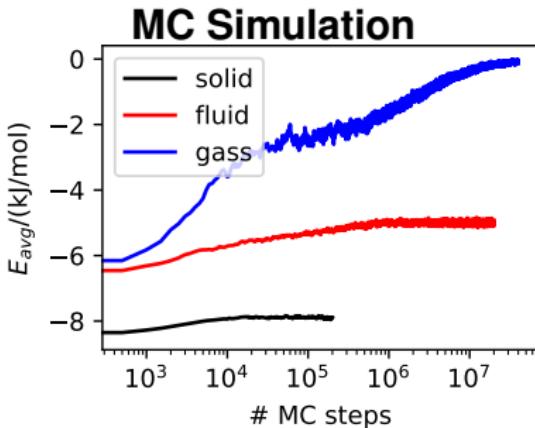


$$T = 50 \text{ K}, \rho = 1.8 \text{ g cm}^{-3}$$

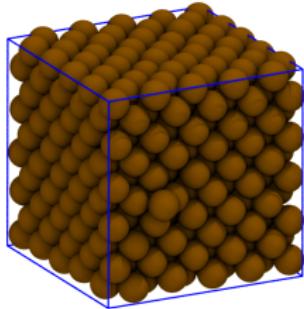
Three phases of Argon

Table: Model for Argon

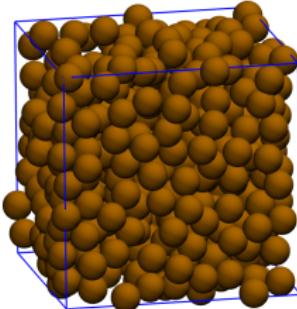
ϵ/J	1.65×10^{-21}
σ/m	3.4×10^{-10}



Solid phase



Fluid phase



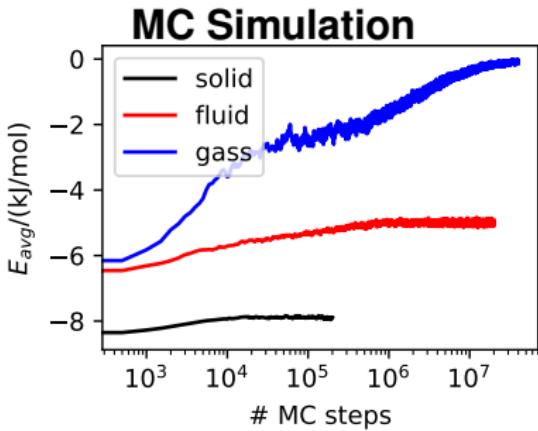
$$T = 50 \text{ K}, \rho = 1.8 \text{ g cm}^{-3}$$

$$T = 120 \text{ K}, \rho = 1.7 \text{ g cm}^{-3}$$

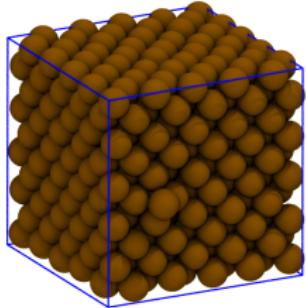
Three phases of Argon

Table: Model for Argon

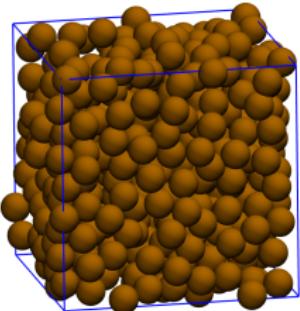
ϵ/J	1.65×10^{-21}
σ/m	3.4×10^{-10}



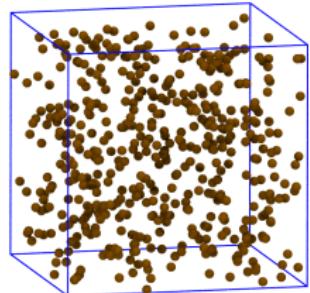
Solid phase



Fluid phase



Gas phase



$$T = 50 \text{ K}, \rho = 1.8 \text{ g cm}^{-3}$$

$$T = 120 \text{ K}, \rho = 1.7 \text{ g cm}^{-3}$$

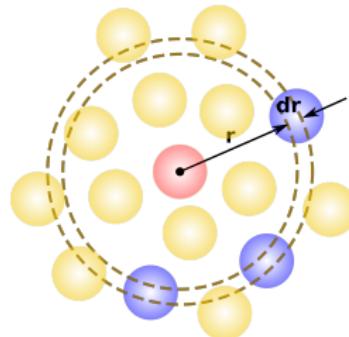
$$T = 600 \text{ K}, \rho = 0.033 \text{ g cm}^{-3}$$

Quantification of phase behavior

Radial distribution function:

$$A(\Gamma) = g(r) = \frac{1}{\rho} \langle \delta(r_{ij} - r) \rangle$$

Can be verified through scattering experiments.

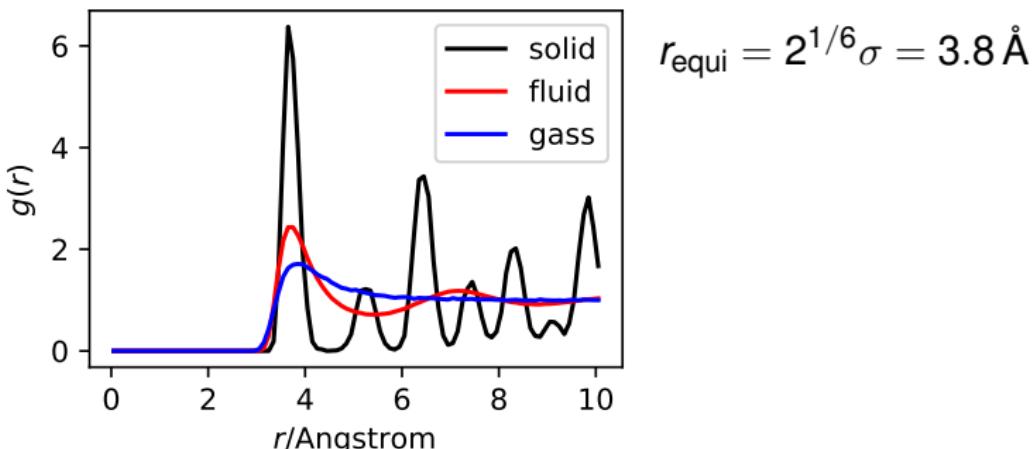
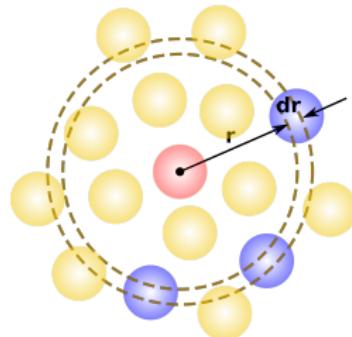


Quantification of phase behavior

Radial distribution function:

$$A(\Gamma) = g(r) = \frac{1}{\rho} \langle \delta(r_{ij} - r) \rangle$$

Can be verified through scattering experiments.



$$r_{\text{equi}} = 2^{1/6} \sigma = 3.8 \text{ \AA}$$

Other ensembles: *NPT*

Most experiments are done under constant pressure and temperature

Other ensembles: *NPT*

Most experiments are done under constant pressure and temperature

New probability with dependence on volume:

$$P(\mathbf{s}^N, V) = \frac{V^N e^{-PV/k_b T} e^{-U(\mathbf{s}^N, V)/k_b T}}{\int dV V^N e^{-PV/k_b T} \int d\mathbf{s}^N e^{-U(\mathbf{s}^N, V)/k_b T}}$$

(\mathbf{s}^N : Scaled coordinates.)

Other ensembles: *NPT*

Most experiments are done under constant pressure and temperature

New probability with dependence on volume:

$$P(\mathbf{s}^N, V) = \frac{V^N e^{-PV/k_b T} e^{-U(\mathbf{s}^N, V)/k_b T}}{\int dV V^N e^{-PV/k_b T} \int d\mathbf{s}^N e^{-U(\mathbf{s}^N, V)/k_b T}}$$

(\mathbf{s}^N : Scaled coordinates.)

Same algorithm as *NVT*, but with an extra MC move:

$$V_{i+1} = V_i + \Delta V$$

Other ensembles: *NPT*

Most experiments are done under constant pressure and temperature

New probability with dependence on volume:

$$P(\mathbf{s}^N, V) = \frac{V^N e^{-PV/k_b T} e^{-U(\mathbf{s}^N, V)/k_b T}}{\int dV V^N e^{-PV/k_b T} \int d\mathbf{s}^N e^{-U(\mathbf{s}^N, V)/k_b T}}$$

(\mathbf{s}^N : Scaled coordinates.)

Same algorithm as *NVT*, but with an extra MC move:

$$V_{i+1} = V_i + \Delta V$$

Acceptance ratio:

$$\alpha = \text{MIN} \left(1, e^{-\beta [U(\mathbf{s}^N, V_{i+1}) - U(\mathbf{s}^N, V_i) + P(V_{i+1} - V_i) - N\beta^{-1} \log(\frac{V_{i+1}}{V_i})]} \right)$$

NPT example: Verification of equation of state

One-to-one relationship between thermodynamic parameters

$$P = f(N, V, T)$$

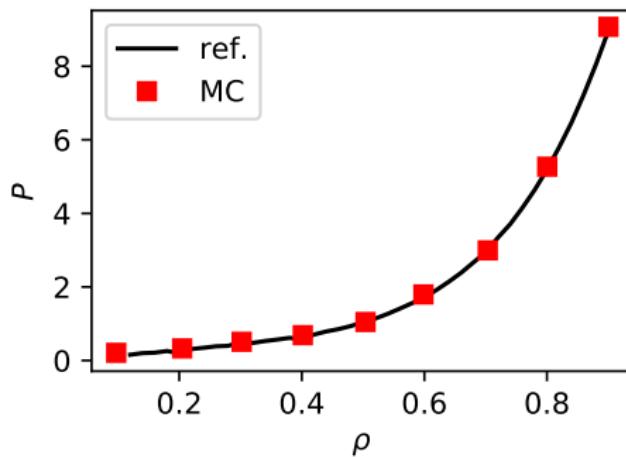
D. Frenkel and B. Smit, Elsevier 1 (2001)

J. Johnson et al., Mol. Phys. 73, 591 (1993)

NPT example: Verification of equation of state

One-to-one relationship between thermodynamic parameters

$$P = f(N, V, T)$$

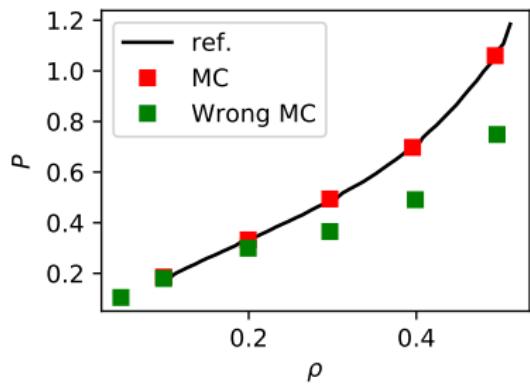


(Reported in Lennard-Jones units)

Violation of sampling principles

Improper sample weighing:

- ▶ Using only accepted configurations



(Reported in Lennard-Jones units)

D. Frenkel and B. Smit, Elsevier 1 (2001)
J. Johnson et al., Mol. Phys. 73, 591 (1993)

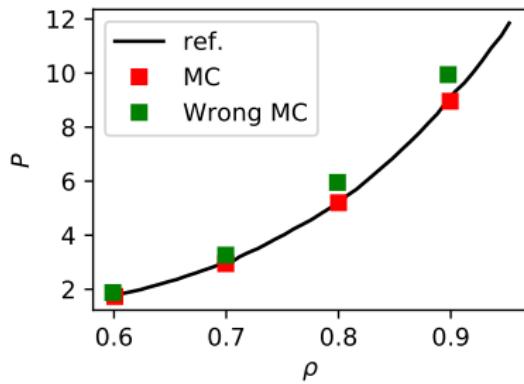
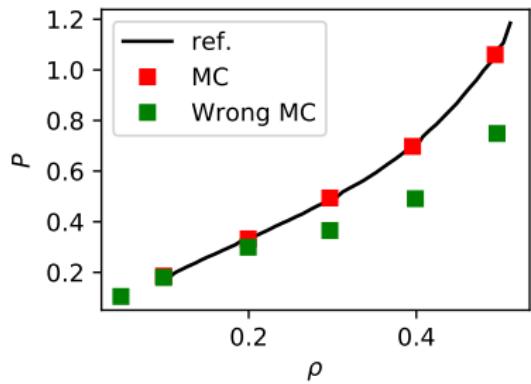
Violation of sampling principles

Improper sample weighing:

- ▶ Using only accepted configurations

Breaking detailed balance:

- ▶ $\Delta x_i = \text{randf}(0, 1)$

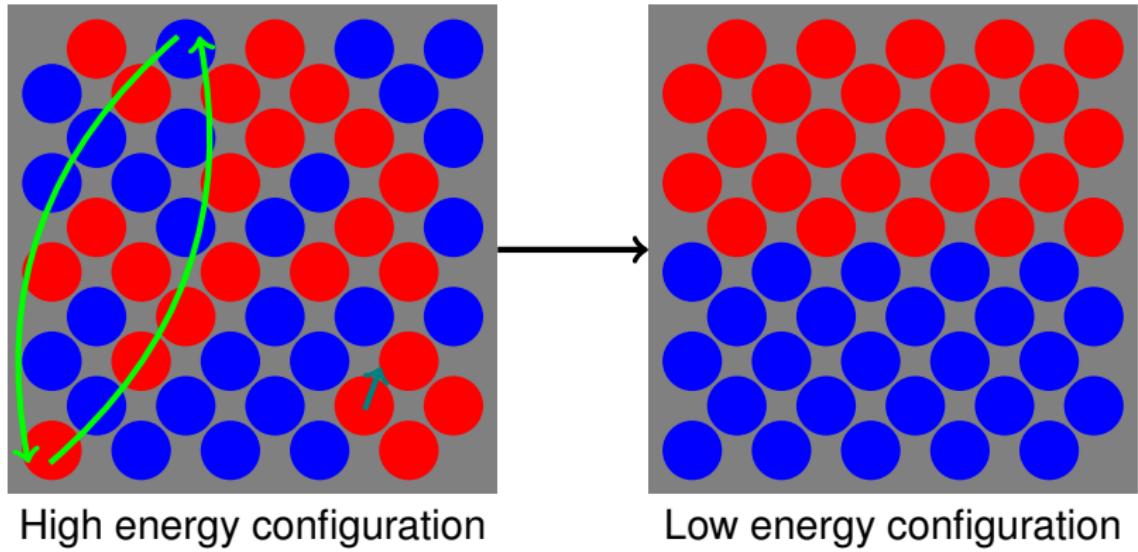


(Reported in Lennard-Jones units)

D. Frenkel and B. Smit, Elsevier 1 (2001)

J. Johnson et al., Mol. Phys. 73, 591 (1993)

The benefits of unphysical MC moves



Slow diffusion in solid and glassy phases

- ▶ Random spatial move: **Slow sampling**
- ▶ Random identity exchange: **Accelerated sampling**

Can MC describe dynamics?



Diffusion process

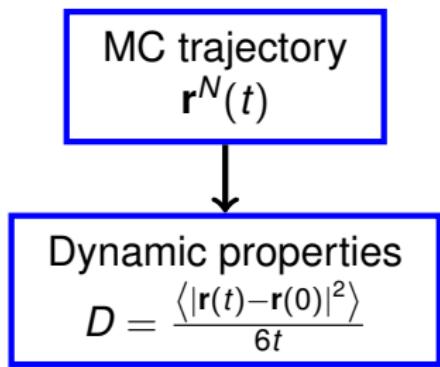
Can MC describe dynamics?

MC trajectory
 $\mathbf{r}^N(t)$



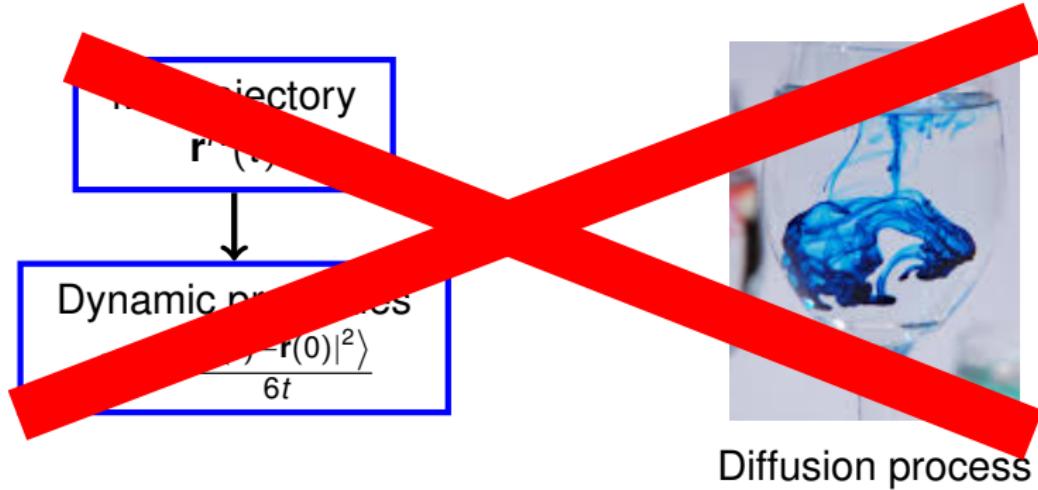
Diffusion process

Can MC describe dynamics?

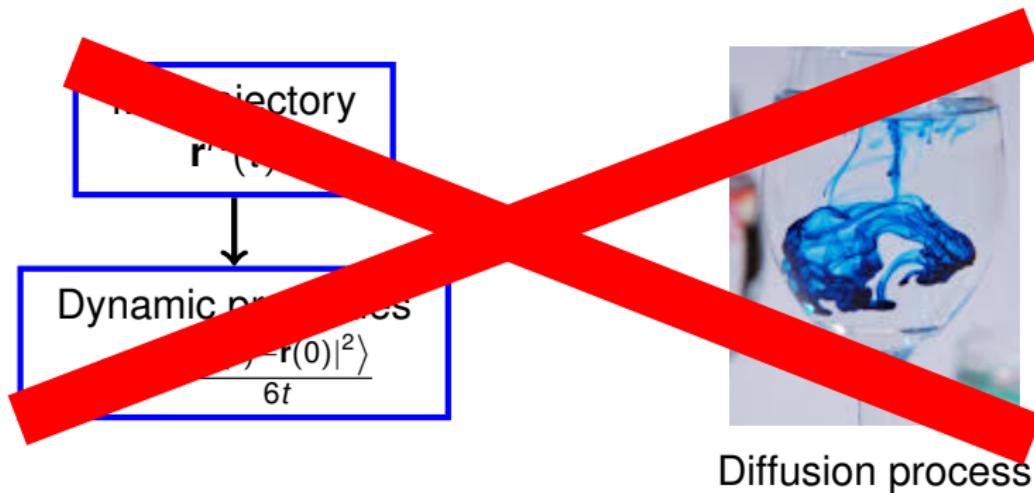


Diffusion process

Can MC describe dynamics?



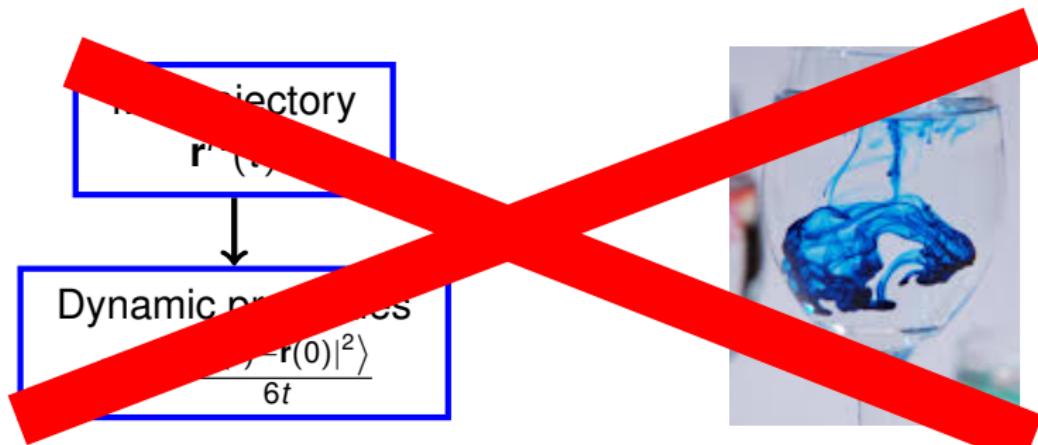
Can MC describe dynamics?



Reasons:

- ▶ No time in MC
- ▶ Unphysical steps
- ▶ Often non-Boltzmann

Can MC describe dynamics?



Diffusion process

Reasons:

- ▶ No time in MC
- ▶ Unphysical steps
- ▶ Often non-Boltzmann

But there are ways:

- ▶ Kinetic and dynamic MC
 - ▶ Master equation approach
 - ▶ Based on known rates

General considerations on the MC approach

Strengths

- ▶ No force-computation
- ▶ Unphysical MC moves
- ▶ Easy to implement

General considerations on the MC approach

Strengths

- ▶ No force-computation
- ▶ Unphysical MC moves
- ▶ Easy to implement

Weaknesses

- ▶ Particle overlap
- ▶ No dynamics
- ▶ Long-range interactions

General considerations on the MC approach

Strengths

- ▶ No force-computation
- ▶ Unphysical MC moves
- ▶ Easy to implement

Weaknesses

- ▶ Particle overlap
- ▶ No dynamics
- ▶ Long-range interactions

When is MC a good choice?

- ▶ Equilibrium properties
- ▶ Gas phase
- ▶ Benefit from unphysical moves
- ▶ Local interactions

General considerations on the MC approach

Strengths

- ▶ No force-computation
- ▶ Unphysical MC moves
- ▶ Easy to implement

Weaknesses

- ▶ Particle overlap
- ▶ No dynamics
- ▶ Long-range interactions

When is MC a good choice?

- ▶ Equilibrium properties
- ▶ Gas phase
- ▶ Benefit from unphysical moves
- ▶ Local interactions
- ▶ Hard particles

