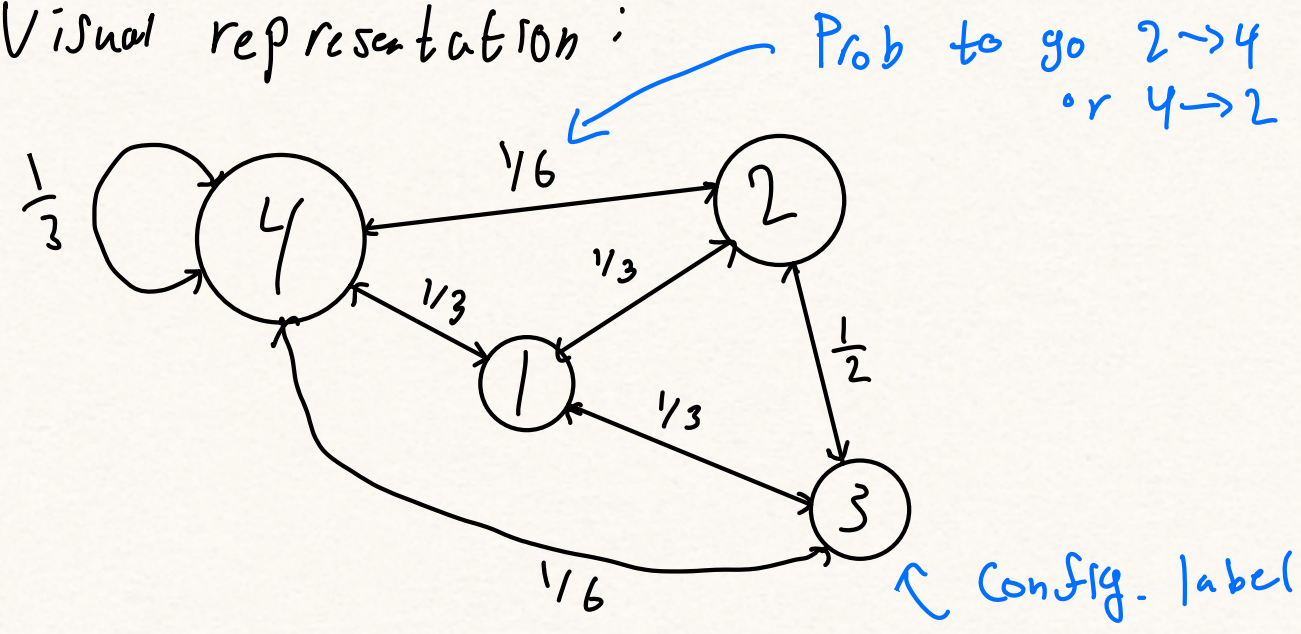
Classical Monte Carlo

Computational Statistical Physics Lucas Böttcher Hans Herrmann

- **Definition:** A **Markov chain** is a stochastic process describing a sequence of possible events in which the probability of each event depends only on the current state

Visual representation:

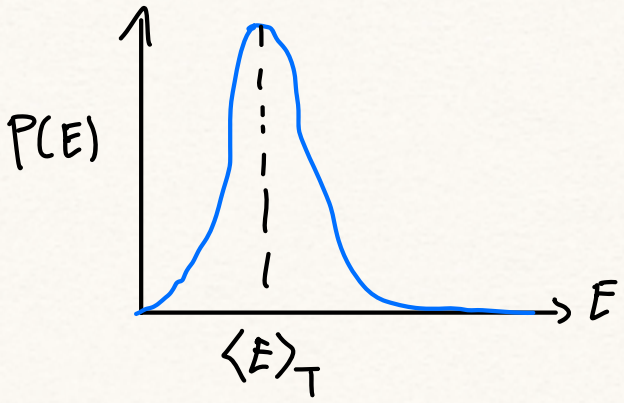


In general: $P(1 \to 2) \neq P(2 \to 1)$

Why Markov chain sampling?

Because it allows us to more efficiently sample the important subsets of phase space!

Assume distribution has the form:



Clearly "LHS" is more important to sample.

Uniform sampling thinks RHS equally important to LHS => very inefficient.

To implement the Markov chain sampling we start in a configuration X and propose a new configuration with probability $T(X \to Y)$

Clearly we demand:

1. $\sum_Y T(X \to Y) = 1$ (Normalization)

2. $T(X \to Y) > 0 \quad \forall \; Y$ (ergodicity)

3. $T(X \to Y) = T(Y \to X)$ (reversibility)

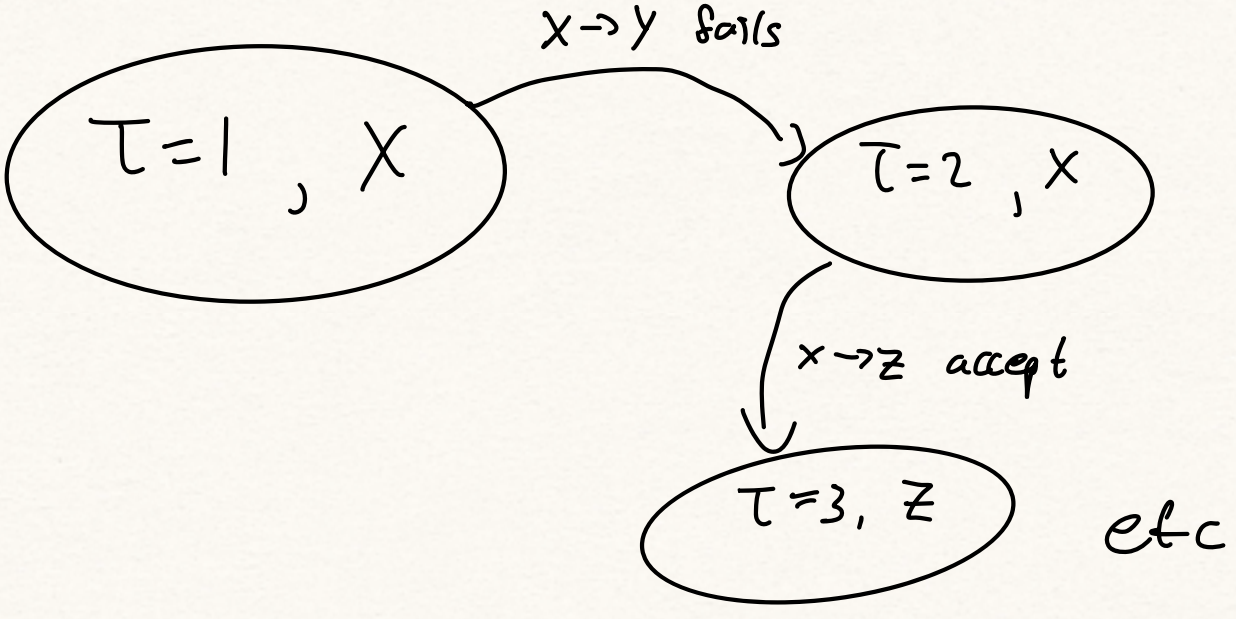
Reminder: We need ergodicity to replace time average with ensample average

To implement importance sampling we can not accept every proposed state. So also introduce an acceptance probability $A(X \to Y)$. The probability that a a state Y is proposed and accapted is then

$$w(x \to y) = T(x \to y) A(x \to y)$$

Note: $A(x \to y)$ is the important player as T can be uniform or just implemented with for loops

Also introduce a virtual time τ that keeps track of each step in the Markov chain.



! Ergodicity can be stated as "should be possible to reach any state Y from X within a finite # τ"

Let us introduce $P(X,\tau)$ = "Prob. to be in Config. X at time τ"

Also introduce some demands on W

1. $W(X \to Y) > 0 \quad \forall X, Y$ (ergodicity)
2. $\sum_Y W(X \to Y) = 1$ (unity)
3. $\sum_Y P(Y,\tau-1)\, W(Y \to X) = P(X,\tau)$ (homogeneity)

It's called homogeneity because $W(X \to Y)$ does not depend on time τ. One can view W as a time-step matrix: 3. $\leftrightarrow$ $P(\tau) = P(\tau-1)\, W$ (left-vector)

With this we can determine the dynamics of $P(X,\tau)$:

$$\frac{dP(X,\tau)}{d\tau} \equiv P(X,\tau) - P(X,\tau-1)$$

$$= \sum_Y P(Y,\tau-1)\, W(Y \to X) \text{ by (3)}$$

$$= P(X,\tau-1)\cdot 1 = \{(2)\} = P(X,\tau-1) \sum_Y W(X \to Y)$$

$$= \sum_Y P(X,\tau-1)\, W(X \to Y)$$

$\Rightarrow$ Master equation!

$$\boxed{\frac{dP(X,\tau)}{d\tau} = \sum_Y P(Y)\, W(Y \to X) - \sum_Y P(X)\, W(X \to Y)}$$

Ergodicity is demanded from Stat. Mech but is equivalent to the matrix W_{xy} being irreducible and aperiodic (Aperiodicity is guaranteed by $W(x \to x) \neq 0$ since the acceptance can fail)

By the Fundamental thm of Markov Chains there is a unique stationary distribution P_{st}.

To make method be usefull for physics we assert that P_{st} is equal to the equilibrium distribution

$$P_{st} \overset{!}{=} P_{eq}$$

The continuity eq. then reads

$$0 = \sum_{y \neq x} P_{eq}(y) W(y \to x) - \sum_{y \neq x} P_{eq}(x) W(x \to y)$$

A sufficient condition is:

$$\boxed{P_{eq}(y) W(y - x) = P_{eq}(x) W(x - y)}$$

Detailed - balance equation

- Metropolis sampling

Plugging in $W(x \to y) = T(x \to y) A(x \to y)$ and using our condition $T(x \to y) = T(y \to x)$ and $T > 0$ detailed balance reduces to:

$$P_{eq}(y) A(y \to x) = P_{eq}(x) A(x \to y)$$

Any acceptance scheme that we could cook up would work, but one of the most famous one is the so called Metropolis (also called $M(RT)^2$)

$$A(x \to y) = \min\left\{1, \frac{P_{eq}(y)}{P_{eq}(x)}\right\}$$

If we have Boltzmann statistics (classical S.M)

$$\Rightarrow \quad A(x \to y) = \min\left\{1, e^{-\beta \Delta E}\right\}$$

So the configuration is always accepted if it lowers energy. This way we sample the most important states.

Example: Ising model

One MC-Step for Ising lattice:

1. Pick a site by random
2. Calculate the energy change if we flip the spin
3. Accept if $\Delta E \leq 0$, otherwise accept with probability $e^{-\beta \Delta E}$
4. Make any easy updates (such as energy)

Correlation time and relaxation time

Each time we flip a spin we have a new configuration
But clearly the two configurations are strongly correlated!
Need to sample only say k configurations and k large enough such that $Cor[X[\tau+k], X[\tau]] \leq$ threshold

Formally: To look at approach to equilibrium use non-linear correlation func.:

$$\Phi_{th}(\tau) = \frac{\langle \sigma(\tau) \rangle - \langle \sigma(\infty) \rangle}{\langle \sigma(\tau_0) \rangle - \langle \sigma(\infty) \rangle}$$

$$\tau_{th} = \int_0^\infty \Phi_{th}(\tau) d\tau$$

(usually E_{ne} decreases exponentially $\Rightarrow \Phi_{ne}(\tau) = e^{-\frac{\tau}{\tau_{ne}}}$)

Similarly for correlation time we first do sweeps until we reached equilibrium and then calculate

$$\overline{\Phi}_{eq} = \frac{\langle \sigma(\tau)\sigma(\tau_0)\rangle - \langle \sigma(\tau_0)\rangle^2}{\langle \sigma(\tau_0)^2\rangle - \langle \sigma(\tau_0)\rangle^2}$$

At $\tau = \tau_0$ clearly $\overline{\Phi}_{eq} = 1$

As $\tau \to \infty$ σ gets decorrelated and

$$\langle \sigma(\tau)\sigma(\tau_\infty)\rangle = \langle \sigma(\tau)\times\sigma(\tau_0)\rangle \simeq \langle \sigma(\tau_0)\rangle^2$$

↑ already in equilibrium

So $\langle \sigma(\tau)\rangle \simeq \langle \sigma(\tau_0)\rangle$

(Our notation is: $\sigma \equiv \frac{1}{N}\sum_i \sigma_i$)

Critical slowing down

The relaxation times diverge close to T_c

as $\tau(T) \sim |T - T_c|^{-z}$

For finite systems it can't diverge but reaches a maxima

$$\tau(T_c) \sim L^{z/\nu}$$ ν - critical exponent

So steps to sample uncorrelated samples:

1. Reach thermal equilibrium (discard $c \cdot \tau_{th}$ steps)
2. Sample every $c\tau_{eq}$ steps
3. If $T = T_c$ $\tau \to L^{z/\nu}$