PHYS 142/242

Lecture 13: Markov chain Monte Carlo

Javier Duarte — February 7, 2025

Canonical ensemble

In the canonical ensemble, the probability p_n that the system is in a state of energy E_n is

$$p_n = \frac{1}{Z} e^{-E_n/(k_\mathrm{B}T)},$$

where $k_\mathrm{B} = 1.38 \times 10^{-23}\,\mathrm{J/K}$ is the Boltzmann constant and

$$Z = \sum_n e^{-E_n/(k_\mathrm{B}T)}$$

is the partition function.

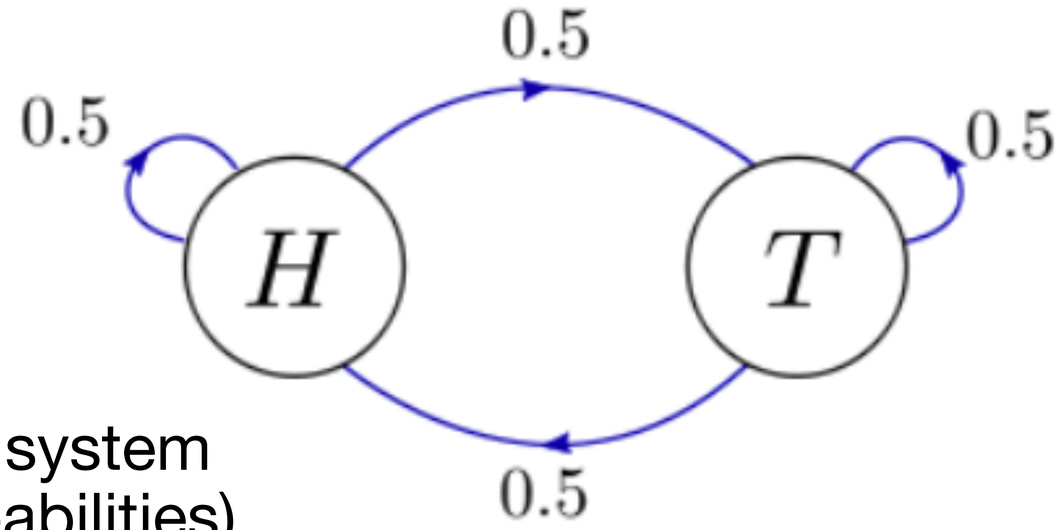
Convenient to define “inverse temperature” $\beta = 1/(k_\mathrm{B}T)$

Coin-flipping game

Suppose we have a coin which can be heads (H) or tails (T)

At each step, we flip the coin, producing a new state (H or T). This generates a sequence like "HTTHTHTTHH..."

Visualize the process with a **state diagram**



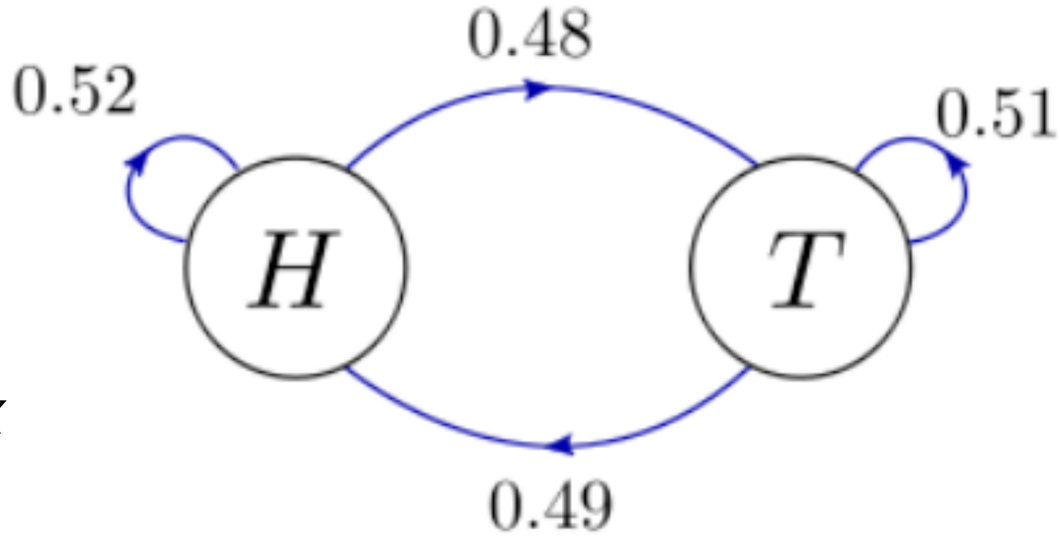
- Circles represent the possible states
- Arrows indicate the possible states that the system could transition to (with corresponding probabilities)
- Transition probabilities coming out of each state must sum up to one

Biased coin

Suppose the coin is biased to be more likely to land on the side it landed on before

State diagram looks like this. The probability of flipping a subsequent H or T is

$$p'_H = P(H|H)p_H + P(H|T)p_T$$
$$p'_T = P(T|H)p_H + P(T|T)p_T$$

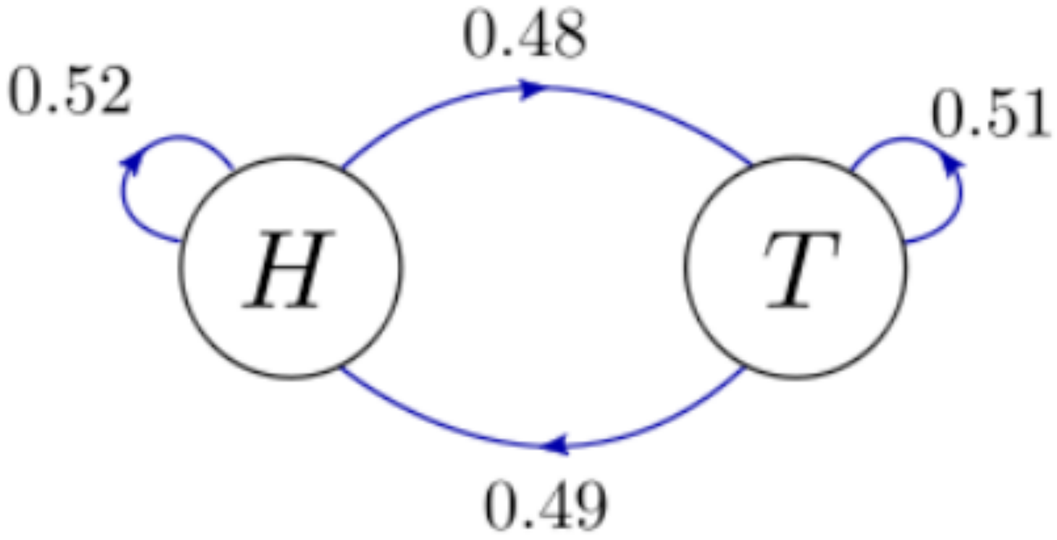


which we can write using the *transition matrix*

$$\begin{bmatrix} p'_H \\ p'_T \end{bmatrix} = \begin{bmatrix} P(H|H) & P(H|T) \\ P(T|H) & P(T|T) \end{bmatrix} \begin{bmatrix} p_H \\ p_T \end{bmatrix} = \begin{bmatrix} 0.52 & 0.49 \\ 0.48 & 0.51 \end{bmatrix} \begin{bmatrix} p_H \\ p_T \end{bmatrix}$$

Stationary state

After a large number of steps, the states probabilities might converge to a "stationary" distribution (π_H, π_T), such that they no longer change significantly on subsequent steps.



The stationary probabilities must satisfy

$$\begin{bmatrix} \pi_H \\ \pi_T \end{bmatrix} = \begin{bmatrix} P(H|H) & P(H|T) \\ P(T|H) & P(T|T) \end{bmatrix} \begin{bmatrix} \pi_H \\ \pi_T \end{bmatrix}$$

Note, this is an eigenvalue equation, with eigenvalue = 1!

Markov process and chain

A ***Markov process*** is a set of probabilistic rules that tell us how to choose a new state of the system, based only on the system's current state. If the system is currently in state n then the probability of choosing state m on the next step is given by the “transition probability” $P(m|n)$.

By repeatedly applying the Markov process, we move the system through a random sequence of states, $\{n^{(0)}, n^{(1)}, n^{(2)}, \ldots\}$ where $n^{(k)}$ denotes the state on step k.

This random sequence is called a ***Markov chain***.

Because the system must transition to some state,

$$\sum_m P(m|n) = 1 \quad \text{for all } n \in \{0,1,\ldots\}\,.$$

Markov process and chain

The probability to be found in state m at step $k+1$ depends only on step k

$$p_m^{(k+1)} = \sum_n P(m\,|\,n)\, p_n^{(k)}$$

or in terms of the transition matrix,

$$\begin{bmatrix} p_0^{(k+1)} \\ p_1^{(k+1)} \\ \vdots \end{bmatrix} = \begin{bmatrix} P(0\,|\,0) & P(0\,|\,1) & \cdots \\ P(1\,|\,0) & P(1\,|\,1) & \cdots \\ \vdots & \vdots & \end{bmatrix} \begin{bmatrix} p_0^{(k)} \\ p_1^{(k)} \\ \vdots \end{bmatrix},$$

The vector of stationary probabilities $[\pi_0, \pi_1, \pi_2, \ldots]$ is an eigenvector of the transition matrix, with eigenvalue 1.

We will use Markov processes to model thermodynamic systems, such that a stationary distribution represents the distribution of thermodynamic microstates in thermal equilibrium

Detailed balance

One way to figure out π_n is through the property of detailed balance.

Detailed balance requires that the rate at which the $n \to n+1$ transition occurs is the same as the rate at which the opposite transition occurs $n+1 \to n$.

If this holds for every pair of states, then the probability distribution is necessarily stationary.

$$P(n+1\,|\,n)\,\pi_n = P(n\,|\,n+1)\,\pi_{n+1} \quad \text{for all} \quad n \in \{0,\ldots,N\}\,.$$

Note: not all Markov chains obey detailed balance!

Example: 2D Ising model

Set of lattice sites Λ, usually a square grid

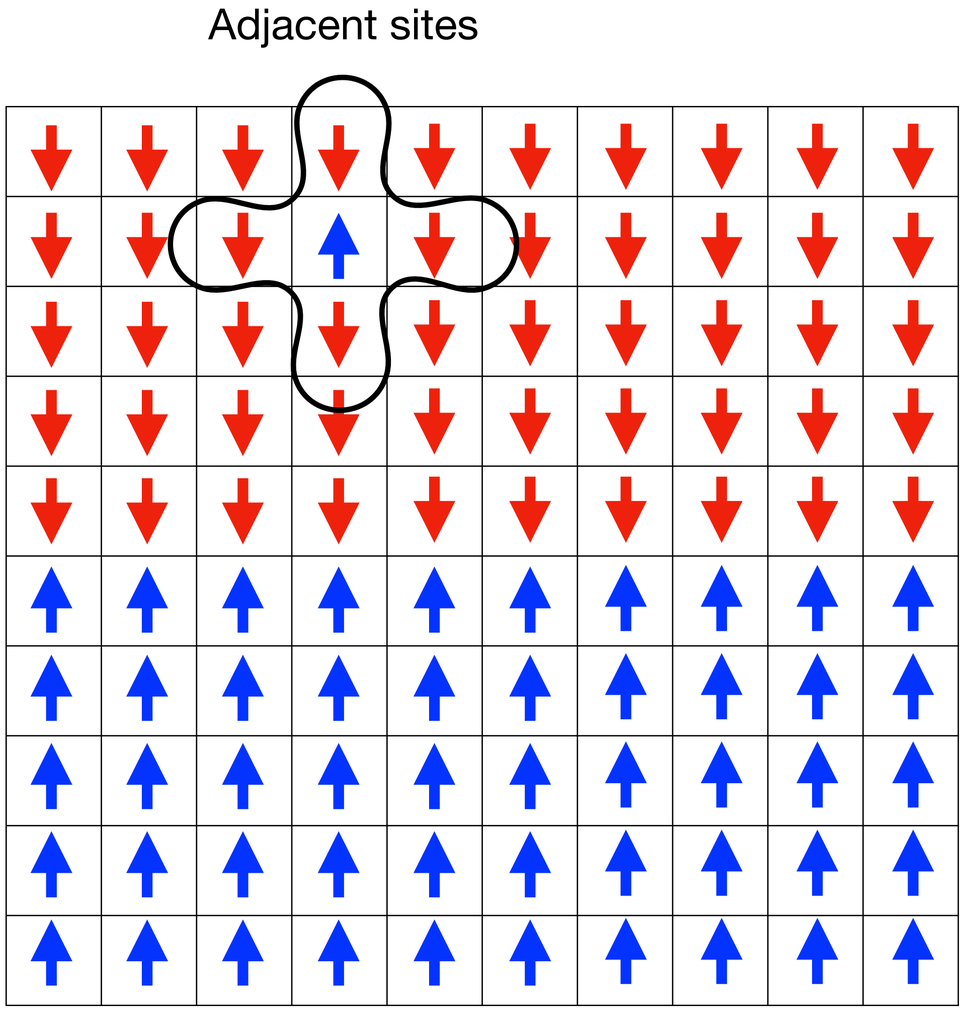
At each site, there is a discrete variable $\sigma_k \in \{-1, +1\}$, representing the site's spin

A spin configuration $\sigma = \{\sigma_k\}_{k \in \Lambda}$

For any two adjacent sites $\langle ij \rangle$ (no double counting) there is a spin interaction J_{ij} and at every site there may be an external field h_j

$$E(\sigma) = -\sum_{\langle ij \rangle} J_{ij}\sigma_i\sigma_j - \mu \sum_j h_j\sigma_j$$

Simplification: $J = J_{ij} > 0$ and $h_j = 0$



Energy

Given the simplified Hamiltonian

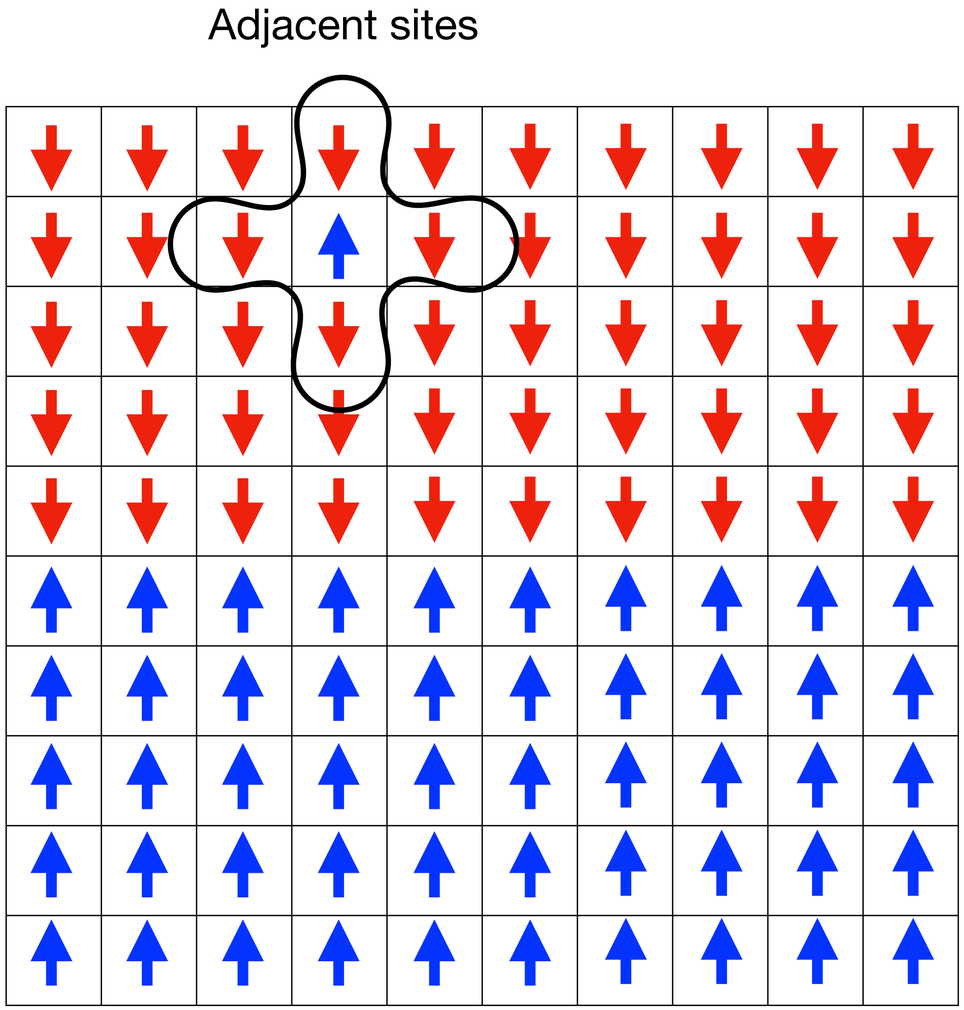
$$E(\sigma) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

The energy of the configuration here is
$E = -152J$

Assume periodic boundary conditions

Hint: number of adjacent site pairs (no double counting) is

$$N_{\text{adjacent pairs}} = 2|\Lambda|$$



Probability

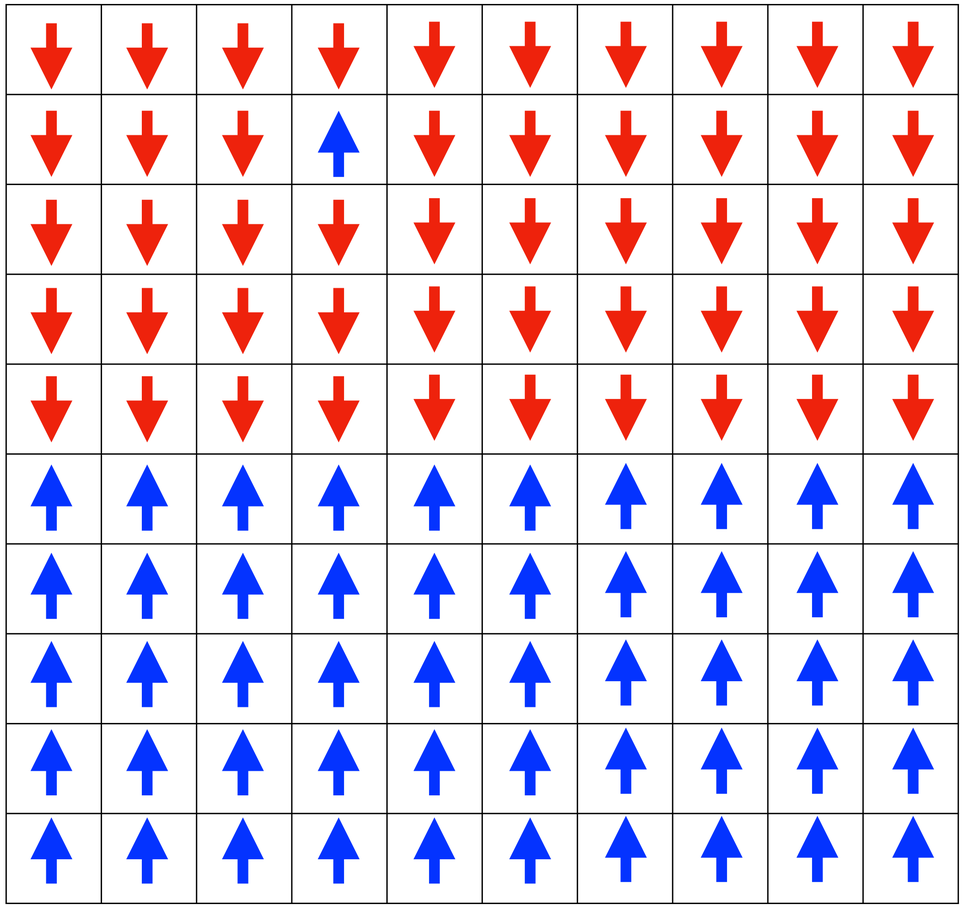
The probability of any given configuration follows the Boltzmann distribution

$$p(\sigma) = \frac{1}{Z} \exp\left(-\frac{E(\sigma)}{k_\mathrm{B}T}\right)$$

where $Z = \sum_\sigma \exp\left(-\frac{E(\sigma)}{k_\mathrm{B}T}\right)$

is the partition function.

With 10×10 sites $2^{100} \approx 10^{30}$ possible states

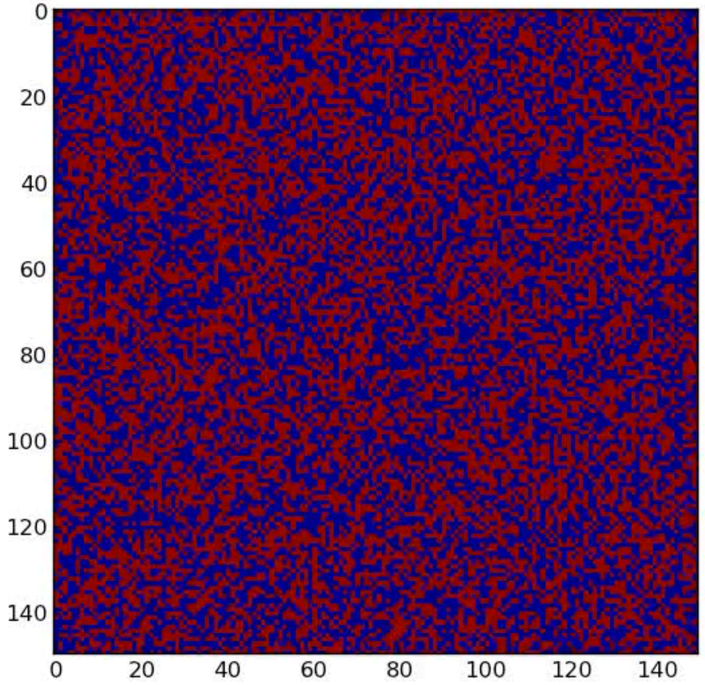


Ising model phase transition

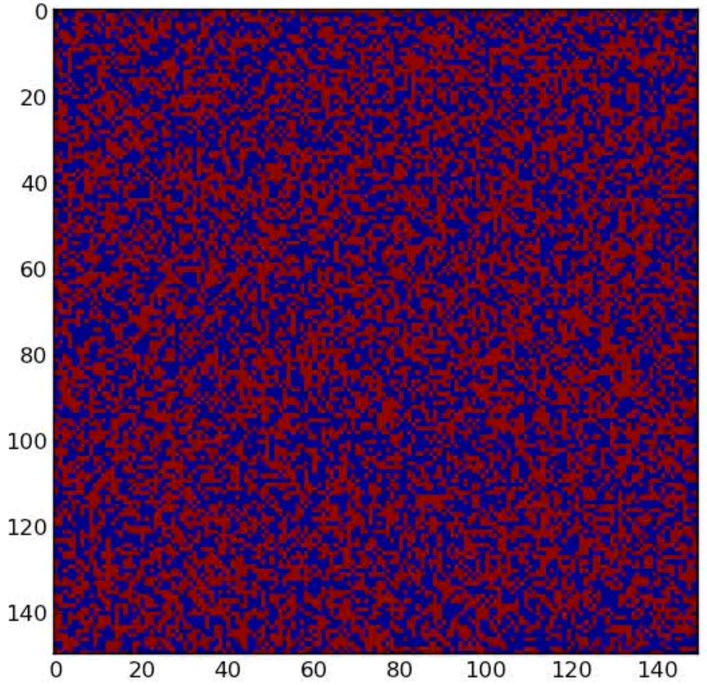
With 150×150 lattice, $2^{22500} \approx 10^{6773}$ possible states!

Phase transition occurs at a critical temperature: ordered phase at low T, disordered phase at high T. How can we compute this?

High T



Low T



Ising model phase transition

Often interested in quantities like the average magnetization, i.e. first we average over the lattice

$$\sigma_{\text{ave}} = \frac{1}{|\Lambda|} \sum_{k \in \Lambda} \sigma_k$$

then we take a thermodynamic average

$$\langle \sigma_{\text{ave}} \rangle = \sum_{\sigma} \sigma_{\text{ave}} p(\sigma)$$

Markov chain MC

Trying to calculate the partition function directly by summing over all states is prohibitively time-consuming.

Instead the MCMC method selectively samples the states.

To accomplish this, we design a Markov process whose stationary distribution is equal to the correct distribution of probabilities.

Once we have this Markov process, we can generate a long Markov chain and calculate moving averages of our desired quantities like $\langle \sigma_{\text{ave}} \rangle$.

If the Markov chain is long enough, this will converge to the true expectation value

Metropolis-Hastings algorithm

Design a Markov process to match a given stationary distribution π_n: ***Metropolis-Hastings algorithm*** based on detailed balance

$$P(n \mid m)\,\pi_m = P(m \mid n)\,\pi_n \quad \text{for all } m, n$$

1. On step k, the system is in state n. Randomly choose a candidate state m by making an unbiased random step through the space of possible states.

2. Compare the probabilities π_n and π_m

   If $\pi_m \geq \pi_n$, accept the candidate

   If $\pi_m < \pi_n$, accept the candidate with probability π_m / π_n. Otherwise, reject the candidate

3. If the candidate is accepted, the state in step $k+1$ is$m$. Otherwise, the state in step $k+1$ remains n

4. Repeat

Checking detailed balance

Let's verify that the stationary distribution of the Markov process satisfies detailed balance.

Consider two states a, b with $\pi_a \leq \pi_b$.

Starting from a, suppose we choose candidate step $a \to b$ with some probability q. Then, the transition probability for $a \to b$ is q times the acceptance probability 1.

Starting from b, we will choose candidate step $b \to a$ with the same probability q (assuming this is unbiased). Then, the transition probability for $b \to a$ is q times the acceptance probability π_a / π_b.

$$\left.\begin{aligned} P(b \,|\, a) &= q \\ P(a \,|\, b) &= q \frac{\pi_a}{\pi_b} \end{aligned}\right\} \quad \Rightarrow \quad P(a \,|\, b)\, \pi_b = P(b \,|\, a)\, \pi_a \quad \text{(detailed balance)}$$

Stepping through state space

One way of thinking about the Metropolis algorithm is that it takes a scheme for performing an unbiased random walk through the space of possible states (represented by our candidate choices), and converts it into a scheme for performing a biased random walk

The biased random walk corresponds to a Markov process with the stationary distribution we are interested in

Applying MCMC to the 2D Ising model

To apply the MCMC method, we design a Markov process using the Metropolis algorithm discussed above. In the context of the Ising model, the steps are as follows:

1. On step k, randomly choose one of the spins i and consider a candidate move which consists of flipping that spin $\sigma_i \rightarrow -\sigma_i$
2. Calculate the change in energy that would result from flipping spin i, relative to $k_{\mathrm{B}}T$, i.e. the quantity:

$$\frac{\Delta E}{k_{\mathrm{B}}T} = -\left[\frac{J}{k_{\mathrm{B}}T}\sum_{\langle ij\rangle}\sigma_j\right]\Delta\sigma_i,$$

where $\Delta\sigma_i$ is the change in σ_i due to the spin flip $\Delta\sigma_i = -2\sigma_i$

If $\Delta E \leq 0$, accept the spin flip

If $\Delta E > 0$, accept the spin flip with probability $\exp(-\Delta E/k_{\mathrm{B}}T)$. Otherwise, reject the flip.

3. Update the moving average of $\langle\sigma_{\mathrm{ave}}\rangle$.

4. Repeat.