PHY604 Lecture 22

November 21, 2023

Review: Mean value method

- Consider general integration problem: $I = \int_a^b f(x)dx$

- Average value of *f* in the range between *b* and *a* is:

$$\langle f \rangle \equiv \frac{1}{b-a}\int_a^b f(x)dx = \frac{I}{b-a}$$

- So, we can get the integral by finding the average of *f*:

$$I = (b-a)\langle f \rangle$$

- We can estimate the average by measuring *f*(*x*) at *N* points chosen at random between *a* and *b*
- Then:

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

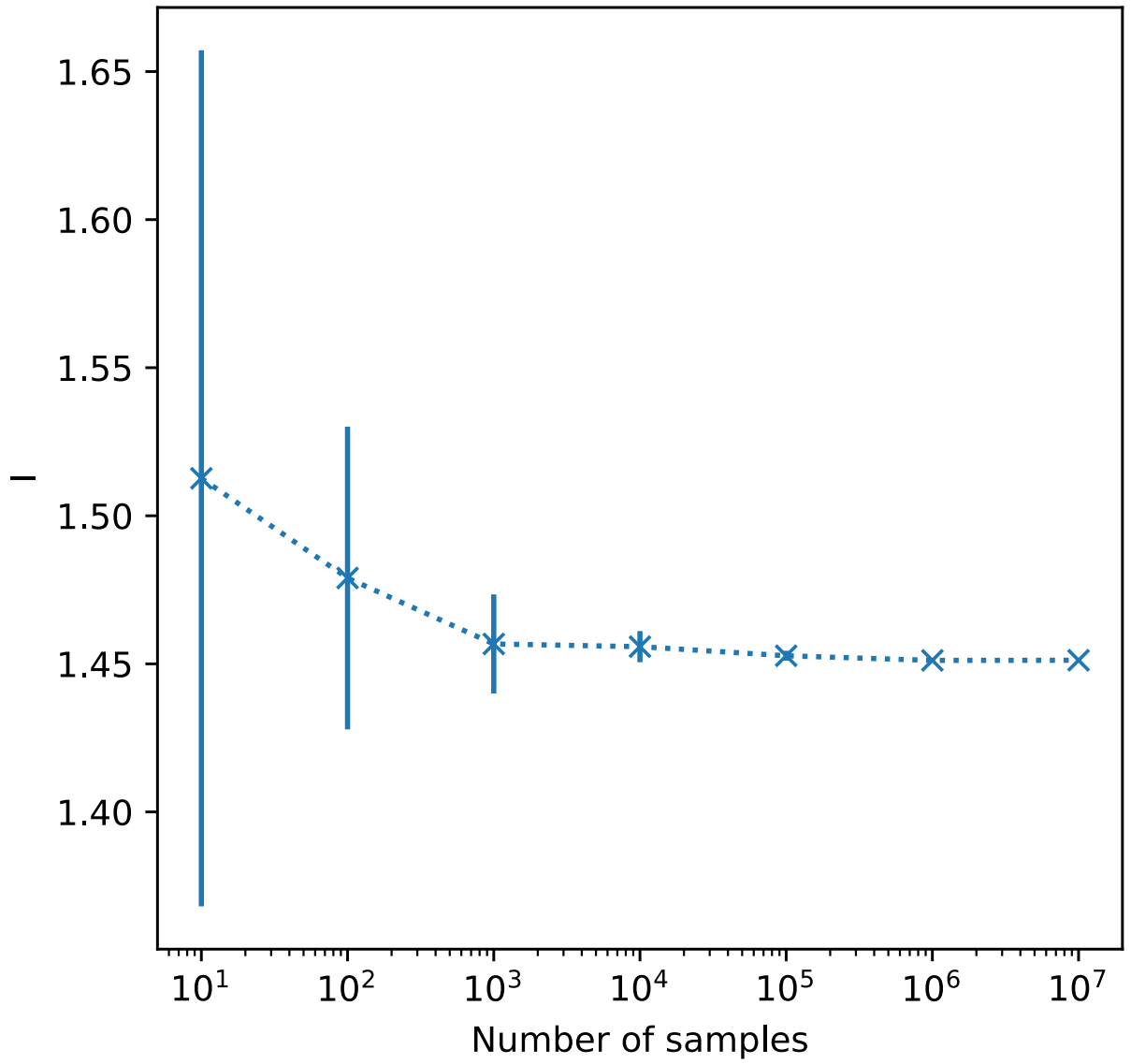
Review: Mean value method

- Equation:

$$I = \int_0^2 \sin^2 \left[\frac{1}{x(2-x)} \right] dx$$

- Errors:

$$I_{\text{error}} = (b-a) \frac{\sqrt{\text{var} f}}{\sqrt{N}}$$



Review: Importance sampling, 1D integral

- Putting everything together:

$$I \simeq \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{w(x_i)} \int_a^b w(x) dx$$

- Generalization of mean value method, which is where $w(x)=1$
- $w(x)$ can be any function that we choose
  - Can be chosen to remove pathologies in the integrand

- However, now we need to draw from a nonuniform distribution

Review: Monte Carlo simulation

- Any computer simulation that uses random numbers to simulate physical process
- We saw a few examples already: radioactive decay and Rutherford scattering
- Used in every branch of physics
  - Particularly important in statistical mechanics and many-body physics

Review: Monte Carlo simulation in stat mech

- Fundamental problem in statistical mechanics: Calculate expectation value of quantity of interest in thermal equilibrium

- Don't know the exact state of the system, only probability of occupying state *i* with energy E_i

$$P(E_i) = \frac{e^{-\beta E_i}}{Z}, \qquad Z = \sum_i e^{-\beta E_i}$$

- Then average value of observable *X*:

$$\langle X \rangle = \sum_i X_i P(E_i)$$

Review: Monte Carlo approach to expectation values

- We could choose *N* terms in the sum at random to add up:

$$\langle X \rangle \simeq \frac{\sum_{k=1}^{N} X_k P(E_k)}{\sum_{k=1}^{N} P(E_k)}$$

Needed to normalize the weighted average if not summing over *all* states

- This would not work well! Boltzmann probability is exponentially small for states $E_i \gg k_B T$

- Usually, most of the states are high energy, only a few contribute significantly

- **Need to use importance sampling!**

Today's lecture:
Monte Carlo simulation simulated annealing

- Monte Carlo simulation in Stat Mech:
  - The ideal quantum gas
  - The Ising model

- Simulated Annealing
  - Travelling salesman problem

Importance sampling for thermal average

- Choose nonuniform distribution to focus on this small set
- Define weighted average over states:

$$\langle g \rangle_w \simeq \frac{\sum_i w_i g_i}{\sum_i w_i}$$

- We choose: $g_i = X_i P(E_i)/w_i$
- So:

$$\left\langle \frac{X_i P(E_i)}{w_i} \right\rangle_w = \frac{\sum_i w_i X_i P(E_i)/w_i}{\sum_i w_i} = \frac{\sum_i X_i P(E_i)}{\sum_i w_i} = \frac{\langle X \rangle}{\sum_i w_i}$$

- Or:

$$\langle X \rangle = \left\langle \frac{X_i P(E_i)}{w_i} \right\rangle_w \sum_i w_i$$

Importance sampling for thermal average

$$\langle X \rangle = \left\langle \frac{X_i P(E_i)}{w_i} \right\rangle_w \sum_i w_i$$

- Evaluate by selecting *N* states randomly with nonuniform distribution:

$$\langle X \rangle \simeq \frac{1}{N} \sum_{k=1}^{N} \frac{X_k P(E_k)}{w_k} \sum_i w_i$$

Summed over all states

Summed over *N* samples

- Still need to choose w_i to bias us towards high-probability samples
  - Also, so that sum over all states *i* can be evaluated analytically

Weights for importance sampling

- Simple choice: $w_i = P(E_i)$
- Sums to 1 over all by definition
- Then we have:

$$\langle X \rangle \simeq \frac{1}{N} \sum_{k=1}^{N} X_k$$

- Thus, choose *N* states in proportion to their Boltzmann weights, and average *X* over them

Markov chain Monte Carlo

- Recall that: $P(E_i) = \frac{e^{-\beta E_i}}{Z}, \qquad Z = \sum_i e^{-\beta E_i}$

- Partition function requires a sum over all states that we are trying to avoid
- Can use a Markov chain to choose states with probability $P(E_i)$ without knowing the partition function:
  - Start with a state *i*
  - Generate a new state *j* by making a small change to *i*
  - Choice of new state is determined probabilistically by a set of transition probabilities T_{ij} that give probability for changing from state *i* to *j*
- If we chose T_{ij} correctly, probability of visiting any state on a step of the Markov chain is $P(E_i)$!

Transition probabilities in the MC

- We must end up in some state on every MC step, so:

$$\sum_j T_{ij} = 1$$

- Choose transition probabilities such that:

$$\frac{T_{ij}}{T_{ji}} = \frac{P(E_j)}{P(E_i)} = \frac{e^{-\beta E_j}/Z}{e^{-\beta E_i}/Z} = e^{-\beta(E_j - E_i)}$$

- I.e., choosing particular ratio of the probability to go from *i* to *j*, and *j* to *i*

- Partition function cancels out!

Transition probabilities in the MC

- If we have correct probability of being in a given state at one step, we will have the correct probability for all later steps
- To see this:
  - Suppose we find a set of T_{ij}'s that satisfy the previous conditions
  - Suppose the probability to be in state *i* on one particular step is $P(E_i)$
  - Then, probability to be in state *j* on the next step is:

$$\sum_i T_{ij} P(E_i) = \sum_i T_{ji} P(E_j) = P(E_j) \sum_i T_{ji} = P(E_j)$$

- Once we get a Boltzmann distribution over states, we will keep it
  - Boltzmann distribution is a fixed point of the Markov chain
- Can also prove that we will converge to Boltzmann distribution
  - See, e.g., Appendix D of Newman

Metropolis-Hastings accept/reject

- Still have not worked out what elements of T_{ij} are
  - Actually, many possible choices
- Most common choice: Metropolis-Hastings algorithm:
  - Choose the change between *i* and *j* from specified set of possible changes
    - Can be, e.g., chosen at random, uniformly
  - Accept or reject the new state with acceptance probability:

$$P_a = \begin{cases} 1 & \text{if } E_j \leq E_i \\ e^{-\beta(E_j - E_i)} & \text{if } E_j > E_i \end{cases}$$

  - I.e., definitely accept if energy is lowered (or equal); may still accept if energy is increased

Transition probabilities under Metropolis-Hastings

- Total probability to move from *i* to given *j* (if $E_j < E_i$)

$$T_{ij} = \frac{1}{M} e^{-\beta(E_j - E_i)}$$

Probability we accept

Probability we choose *j*

Transition probabilities under Metropolis-Hastings

- If $E_j > E_i$:

$$T_{ij} = \frac{1}{M} e^{-\beta(E_j - E_i)}, \quad T_{ji} = \frac{1}{M} \quad \Longrightarrow \quad \frac{T_{ij}}{T_{ji}} = e^{-\beta(E_j - E_i)}$$

- If $E_j \leq E_i$:

$$T_{ij} = \frac{1}{M}, \quad T_{ji} = \frac{1}{M} e^{-\beta(E_i - E_j)} \quad \Longrightarrow \quad \frac{T_{ij}}{T_{ji}} = e^{-\beta(E_j - E_i)}$$

- Thus, both consistent with :

$$\frac{T_{ij}}{T_{ji}} = \frac{P(E_j)}{P(E_i)} = \frac{e^{-\beta E_j}/Z}{e^{-\beta E_i}/Z} = e^{-\beta(E_j - E_i)}$$

Some comments about the Metropolis algorithm

- Note that many steps will not change the system
  - Still need to include in the sum
- The number of possible moves *M*, must be the same when going from *i* to *j* as *j* to *i*
- Moves must be chosen to get you to every state
  - Move set for which all states are accessible is called ergodic
- Will generally take some (unknown) time to equilibrate to Boltzmann distribution

Steps of Markov chain Monte Carlo:

- 1. Choose random starting state
- 2. Choose a move uniformly at random from set of moves
- 3. Calculate the acceptance probability
- 4. Accept or reject the move
- 5. Measure *X* in current state, add to sum
- 6. Go back to step 2

Example: Ideal gas

- Consider the quantum states of a particle or atom of mass *m* in cubic box of length *L*
- Energy of one particle given by:

Quantum numbers from 1 to infinity.

$$E(n_x, n_y, n_z) = \frac{\pi^2 \hbar^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2)$$

- Ideal gas: no interactions between particles
  - Energy is sum of individual particles:

$$E = \sum_{i=1}^{N} E(n_x^{(i)}, n_y^{(i)}, n_z^{(i)})$$

Move set for ideal gas

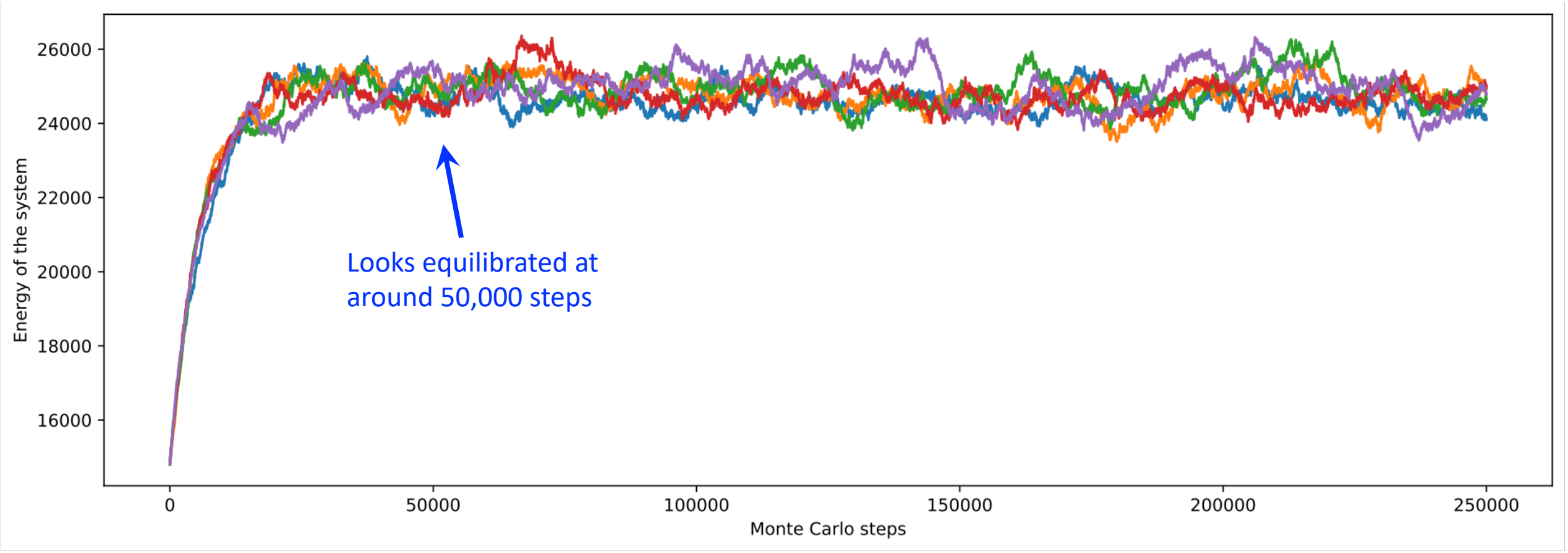
- Choose set of all moves of a single atom to one of the six "neighboring" states where n_x, n_y, or n_z differ by +/- 1
- Each Monte Carlo step, choose a random particle, chose a quantum number, change it by +/- 1
- Change in total energy just the change for single particle since there are no interactions
  - E.g., increase or decrease n_x of atom *i* by one:

$$\Delta E = \frac{\pi^2\hbar^2}{2mL^2}[(n_x \pm 1)^2 + n_y^2 + n_z^2] - \frac{\pi^2\hbar^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2)$$

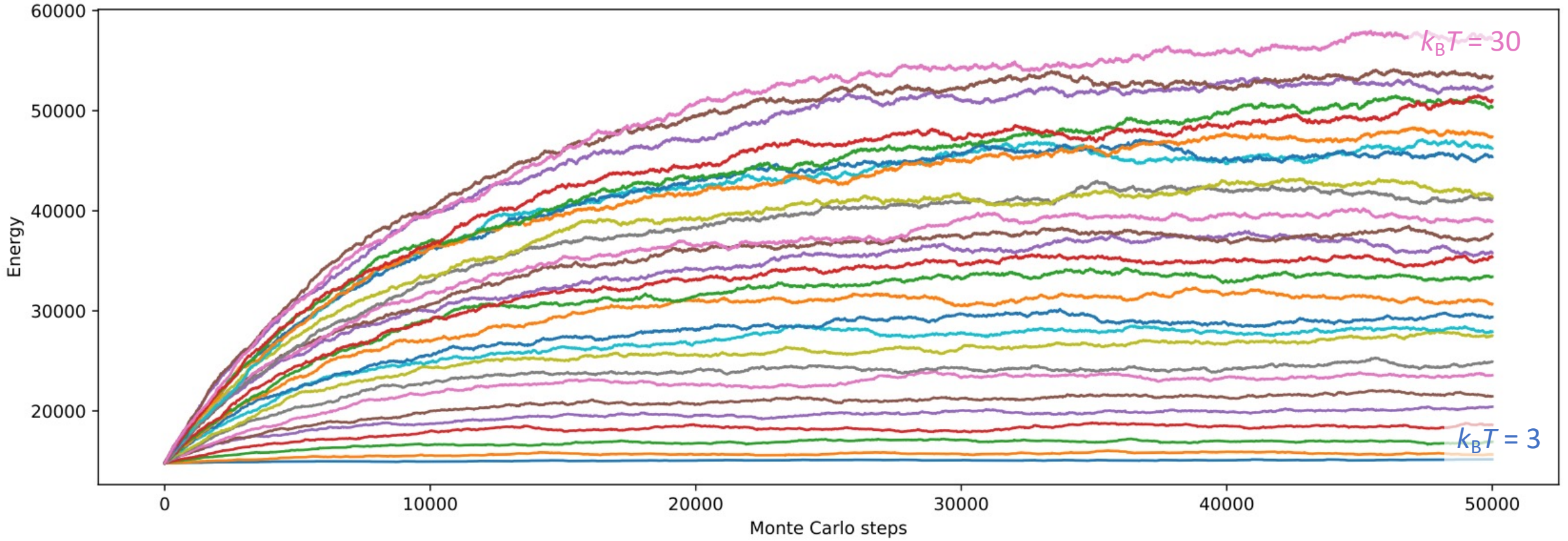
$$= \frac{\pi^2\hbar^2}{2mL^2}[(n_x \pm 1)^2 - n_x^2] = \frac{\pi^2\hbar^2}{2mL^2}(\pm 2n_x + 1)$$

- Note: Reject moves that try to make $n < 1$

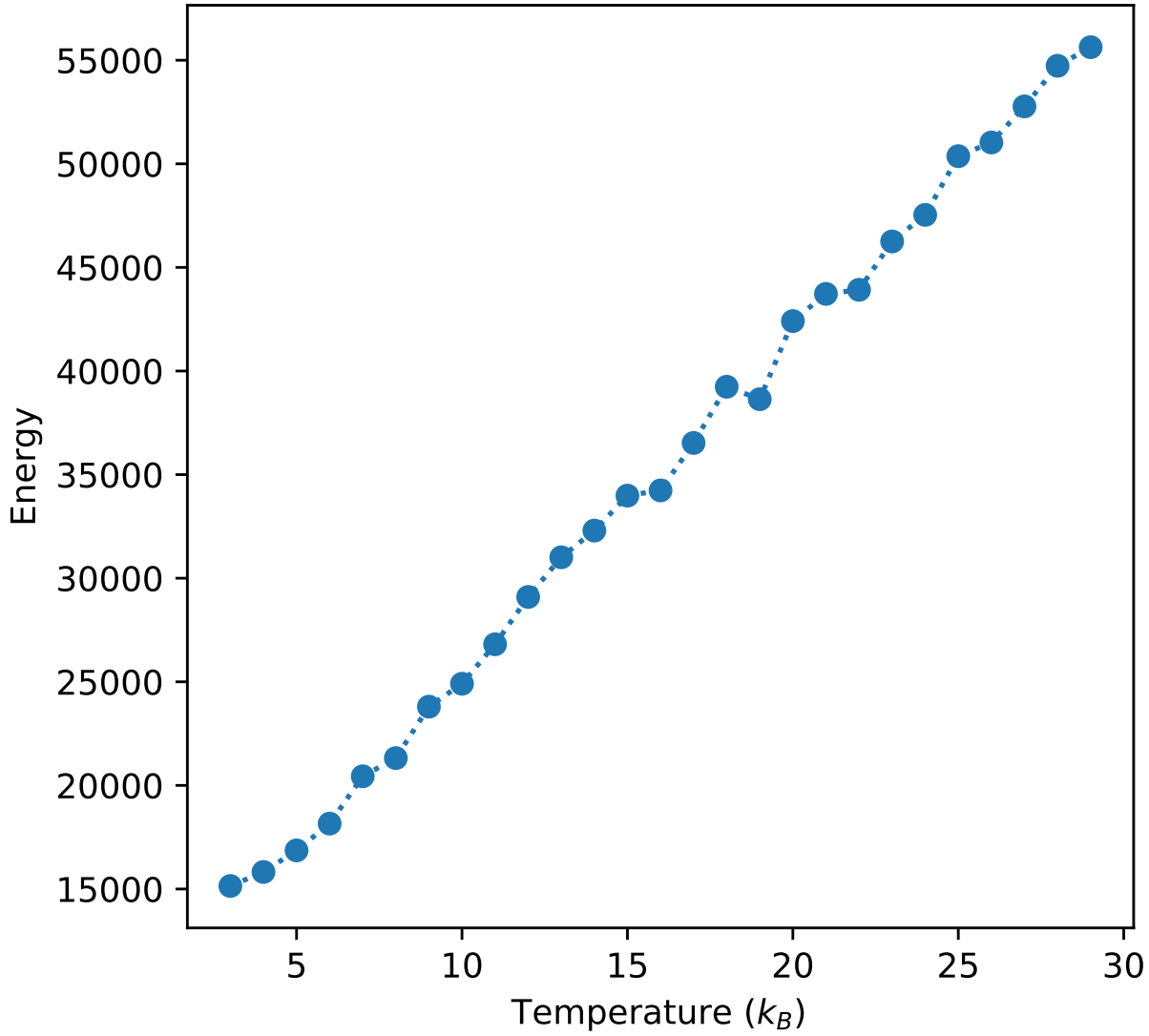
Monte Carlo simulation of ideal gas



Monte Carlo simulation of ideal gas: Dependence on T



Monte Carlo simulation of ideal gas: E vs. T



Example: The Ising model

- The Ising model is a classic model in statistical physics for describing magnetic systems

- Describe a magnetic material as classical spins on a lattice
  - Spins can only point up or down
  - Energy is given by:

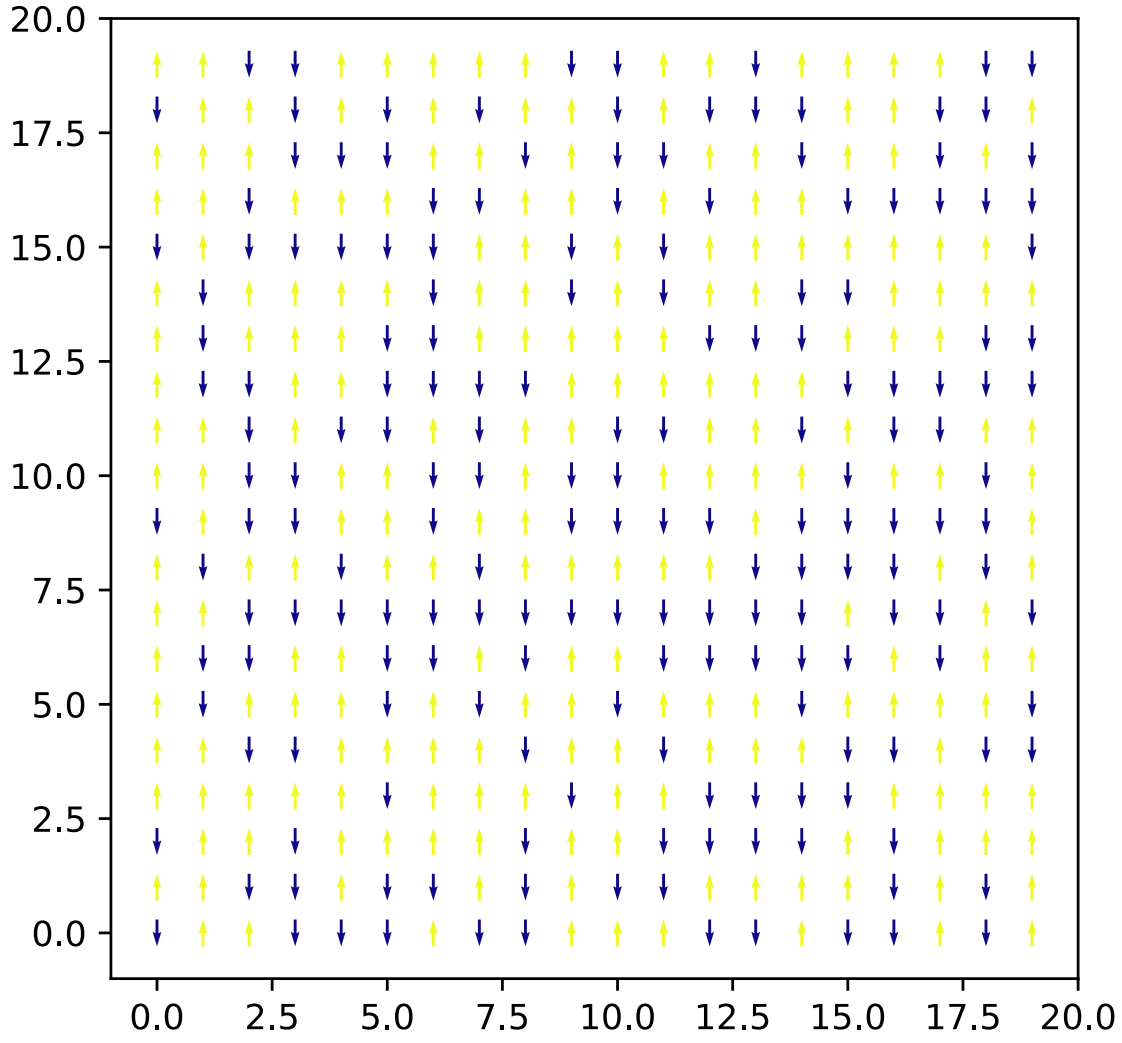
$$E = -J \sum_{\langle i,j \rangle} s_i s_j$$

  - Where $\langle i, j \rangle$ indicate neighboring spins
  - *J* is the interaction strength
    - If J > 0 aligned spins are preferred
    - If J < 0 antialigned spins are preferred

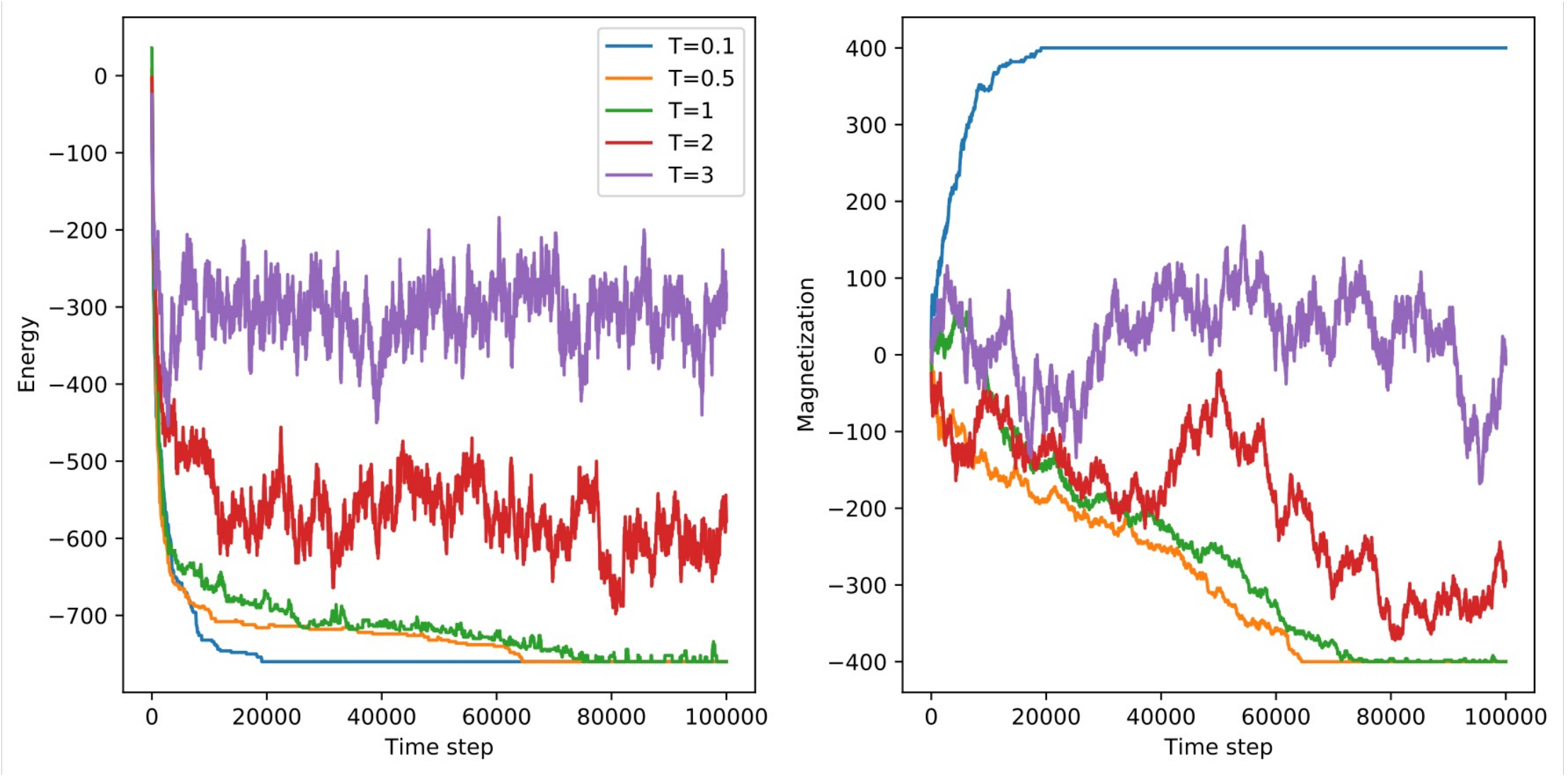
Monte Carlo simulation: Ising model on square lattice

- 20 x 20 square lattice of spins
- Initialized randomly
- Move set: Flip a random spin
  - If it lowers the energy, accept
  - If it raises the energy, accept with Boltzmann probability
- Can also monitor magnetization

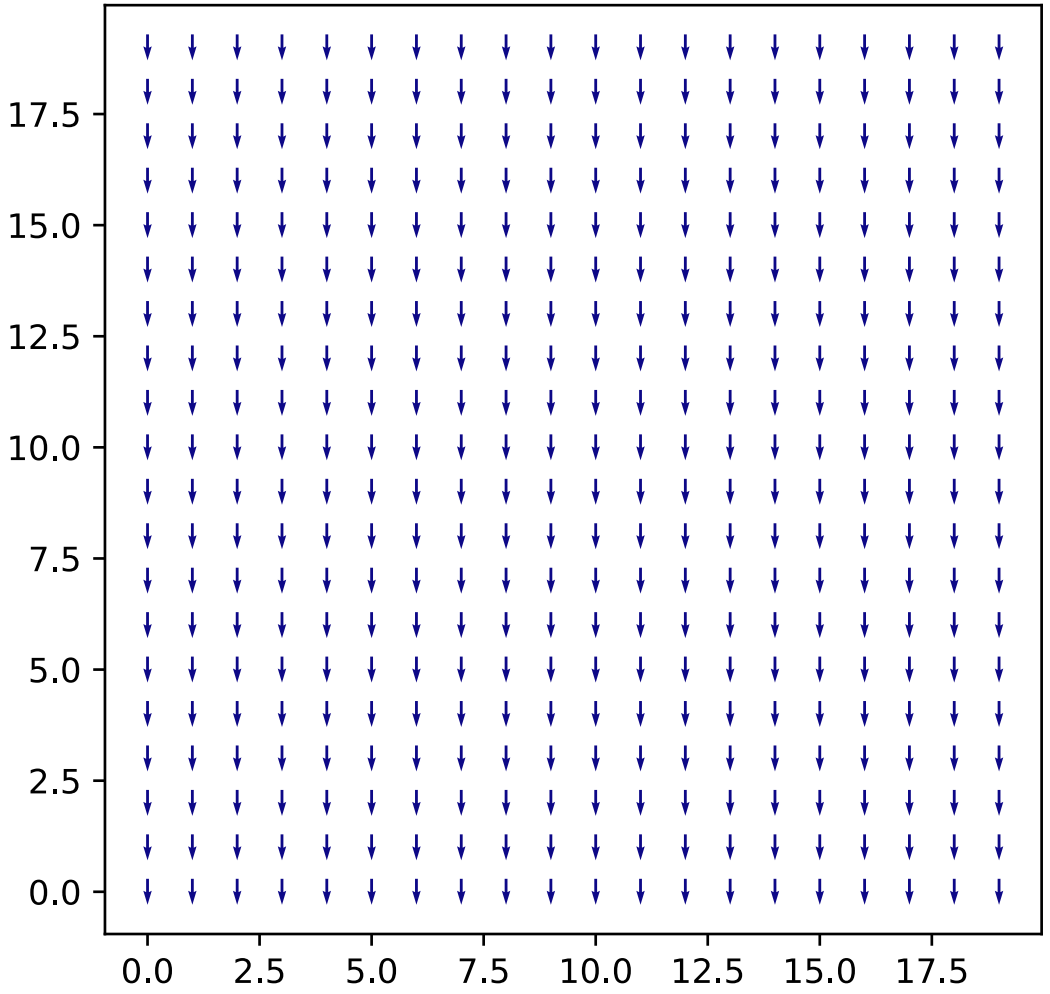
$$M = \sum_i s_i$$



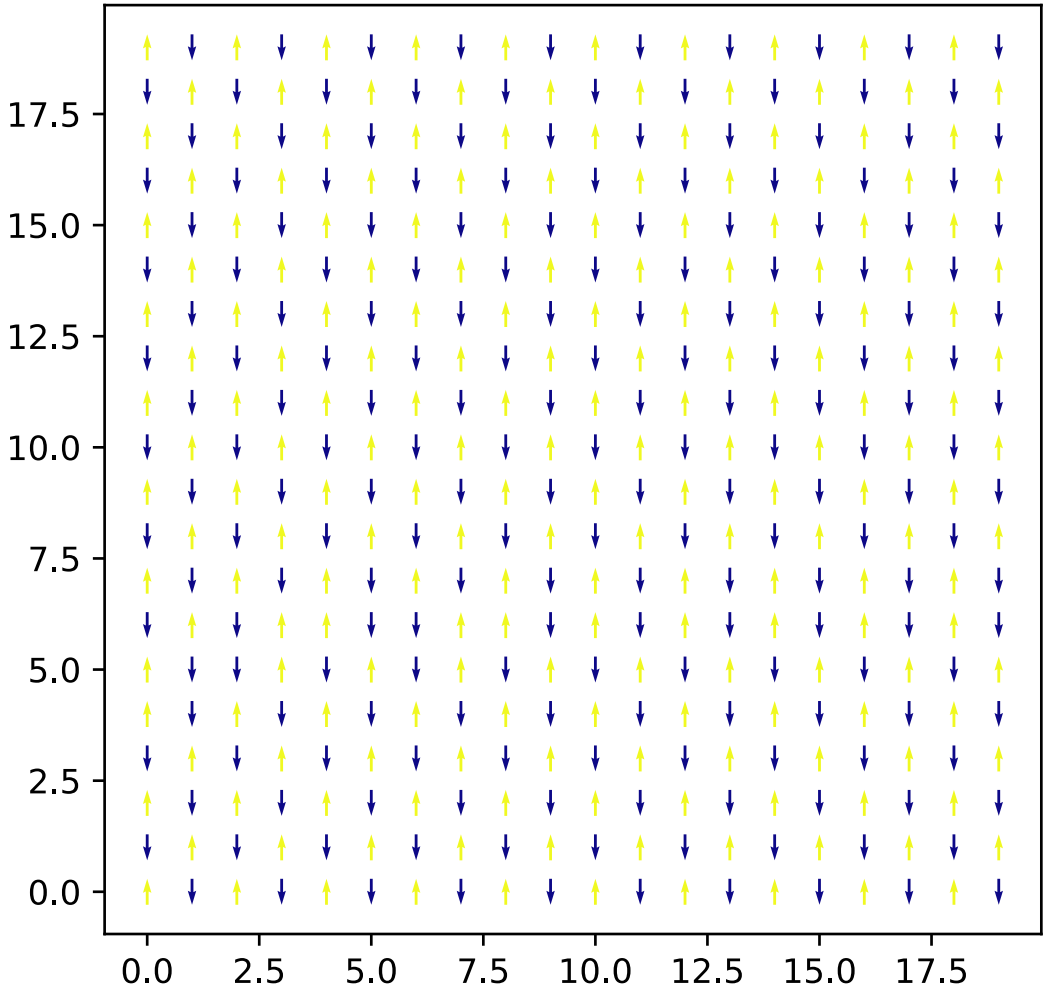
Ising model on square lattice versus T



Ferromagnetic for J > 0,



Antiferromagnetic for J < 0



Today's lecture:
Monte Carlo simulation simulated annealing

- Example of Monte Carlo simulation:
  - The ideal quantum gas
  - The Ising model

- Simulated Annealing
  - Travelling salesman problem

Simulated annealing

- Monte Carlo methods can also be used for numerical optimization
- Optimization methods discussed so far only give local minima
- Global optimization problems are very challenging
- Simulated annealing borrows ideas from statistical physics to tackle this problem

Statistical mechanics for optimization

- Recall the Boltzmann probability:

$$P(E_i) = \frac{e^{-\beta E_i}}{Z}, \qquad Z = \sum_i e^{-\beta E_i}$$

- Assume we have single, unique ground state
- Choose energy scale so that the ground state configuration is 0 energy
- If we cool the system to $T = 0$, then the probability distribution is:

$$P(E_i) = \begin{cases} 1 & \text{if } E_i = 0 \\ 0 & \text{if } E_i > 0 \end{cases}$$

- By cooling the system, we can find the ground state

Statistical mechanics for optimization

- We can use the same strategy (cooling the system) for finding the minimum of a function
  - Take the value of the function to be the “energy”
  - Take the values of independent variables to define a state of the system

- But how can we avoid getting trapped in a local minima?
  - Energy of all nearby states are higher in energy, will not accept moves for low *T*

- Solution: “Anneal” by cooling slowly so system can find its way to the global minimum
  - Guaranteed to converge to global minimum if we cool slowly enough (often not possible)

Simulated annealing approach

- Choose $k_B T$ to be significantly greater than the typical energy change from a singe Monte Carlo move
  - Then:

$$\beta(E_j - E_i) \ll 1 \quad \Longrightarrow \quad P_a \simeq 1$$

  - Most moves accepted, state of the system rapidly randomized

- Make a cooling "schedule," e.g.:

$$T = T_0 e^{-t/\tau}$$

- Choice of τ require some trial and error, slower cooling is more likely to find ground state, but simulation takes longer

Example: Travelling salesman problem

- Find the shortest route that visits a given set of locations on a map
- One of the most famous optimization problems (NP hard)
- We will assume the salesman can travel between the *N* points on the map in straight lines (i.e., the world is flat)
  - *N* cities are chosen at random in a 2D square of unit length
- Want to minimize total distance travelled over the tour:

$$D = \sum_{i=0}^{N-1} |\mathbf{r}_{i+1} - \mathbf{r}_i|$$

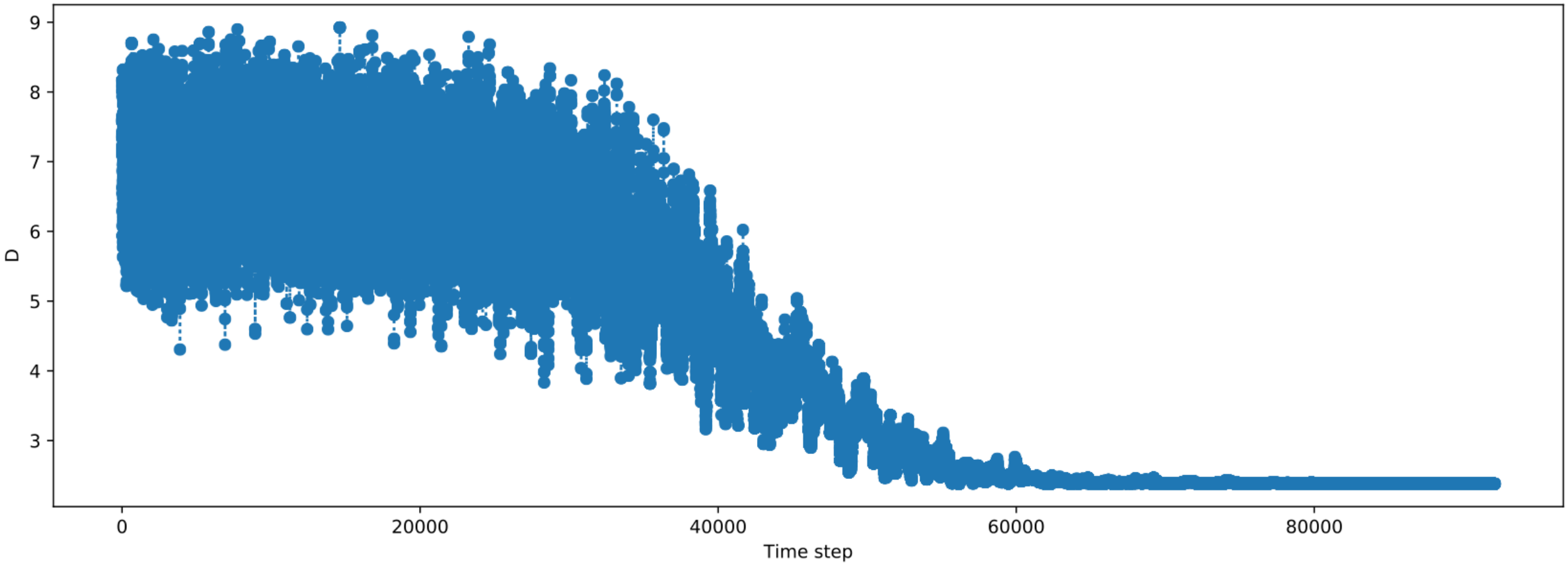
Markov chain Monte Carlo for traveling salesman

$$D = \sum_{i=0}^{N-1} |\mathbf{r}_{i+1} - \mathbf{r}_i|$$

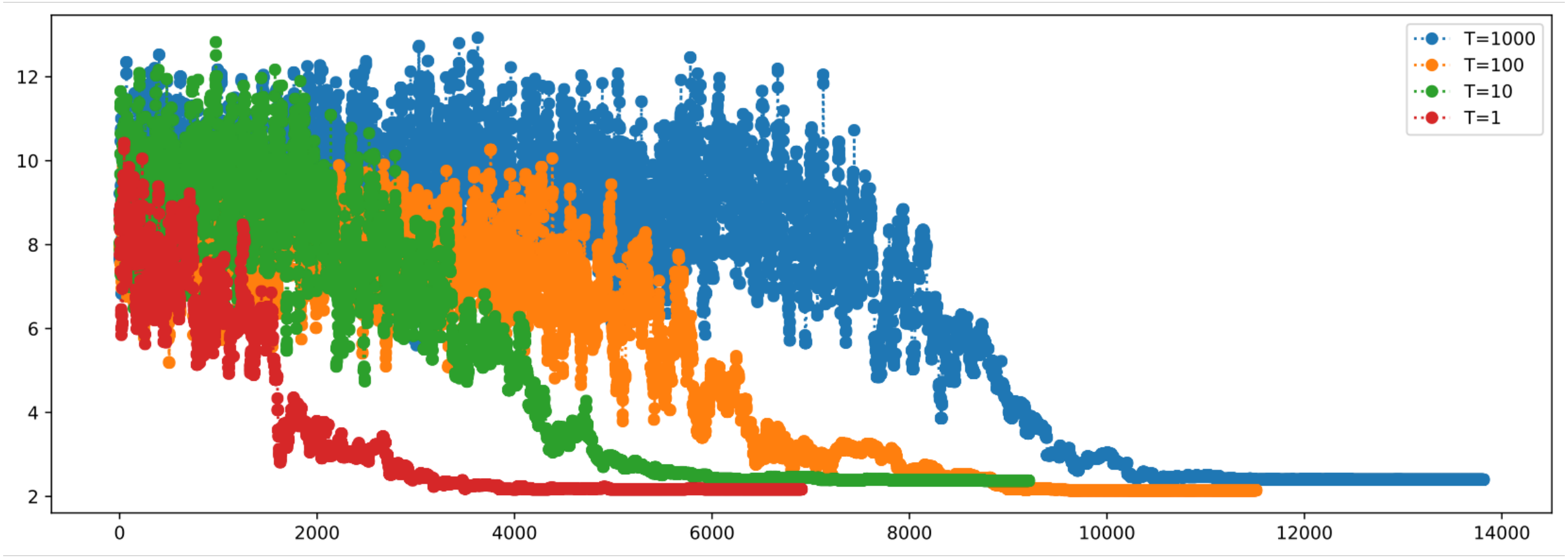
- Minimize *D* over set of all possible tours

- First set up an initial tour
- Then choose from set of moves: Swap pairs of cities
  - Accept if swap shortens the tour
  - If it lengthens the tour, accept with Boltzmann probability, energy replaced by distance D

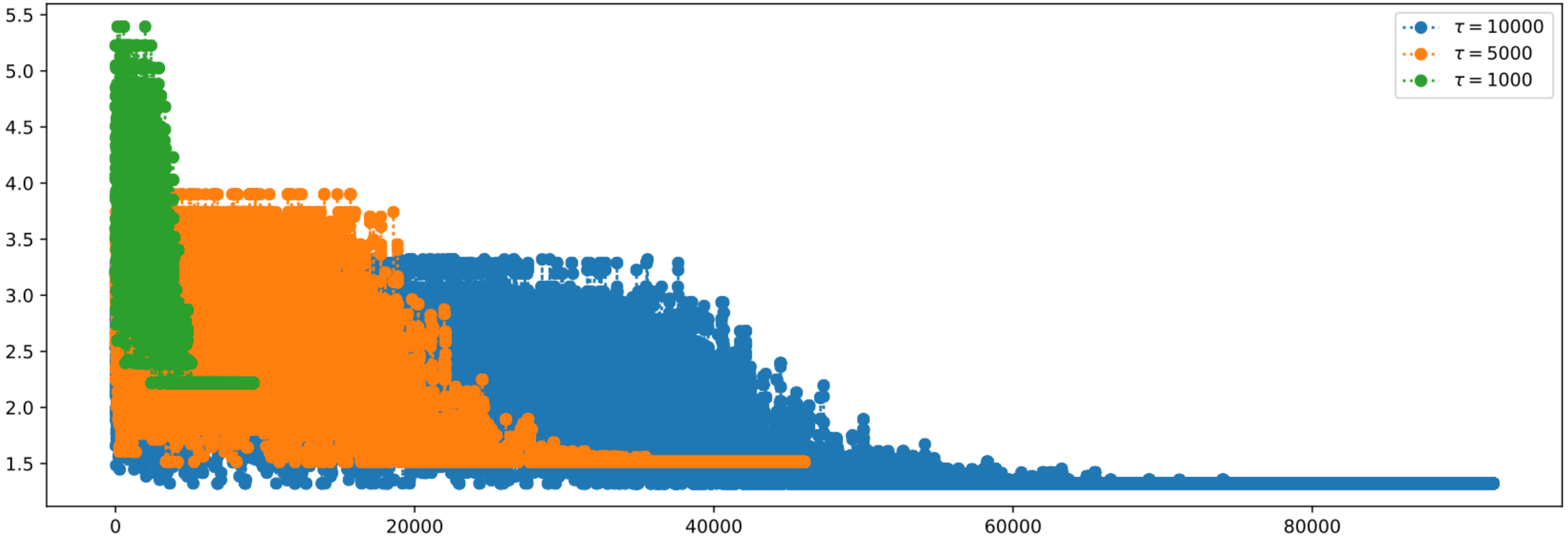
Simulated annealing for traveling salesman



Simulated annealing for traveling salesman: Different starting temperatures



Simulated annealing for traveling salesman: Different cooling rates temperatures



After class tasks and the rest of the semester

- Homework 5 is posted, due Nov. 28, 2023
- Final projects: Send topics if you have not already!
- Readings:
  - Newman Sec. 10.3