Random processes and Monte Carlo Simulation

- Importance Sampling and Statistical mechanics.
- Markov Chain Monte Carlo.
- Simulated Annealing

- One of the problems in statistical mechanics is to calculate the average (or expectation) value of a quantity of interest in a physical system in thermal equilibrium at temperature T.

Importance Sampling and Statistical Mechanics

- One of the problems in statistical mechanics is to calculate the average (or expectation) value of a quantity of interest in a physical system in thermal equilibrium at temperature T.
- One in general doesn't know the exact state of the physical system.

Importance Sampling and Statistical Mechanics

- One of the problems in statistical mechanics is to calculate the average (or expectation) value of a quantity of interest in a physical system in thermal equilibrium at temperature T.
- One in general doesn't know the exact state of the physical system.
- Instead we know that at temperature T, a system will pass through a succession of states such that the probability of occupying a state i with energy E_i is:

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

Importance Sampling and Statistical Mechanics

- One of the problems in statistical mechanics is to calculate the average (or expectation) value of a quantity of interest in a physical system in thermal equilibrium at temperature T.
- One in general doesn't know the exact state of the physical system.
- Instead we know that at temperature T, a system will pass through a succession of states such that the probability of occupying a state i with energy E_i is:

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

- The average value of X:

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

- In rare few cases, this can be done analytically..

Importance Sampling and Statistical Mechanics

- In rare few cases, this can be done analytically..
- In most cases, this has to be done numerically – however a brute force summing will be out of question (in most cases) – example if one has 10^{23} gas molecules, and even if each had 2 quantum states then the total number of states is $2^{10^{23}}$!!!!

- In rare few cases, this can be done analytically..
- In most cases, this has to be done numerically – however a brute force summing will be out of question (in most cases) – example if one has 10^{23} gas molecules, and even if each had 2 quantum states then the total number of states is $2^{10^{23}}$!!!!
- We can take the approach that we will calculate the sum via Monte Carlo (random sampling). In that case, we will chose N random states, $k = 1 \ldots N$ and calculate:

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

The denominator is needed to normalize the weighted average correctly..

- Unfortunately, even this will not work correctly!

- Unfortunately, even this will not work correctly!
- In most cases, with $E_i \gg k_B T$, the Boltzmann probability will be exponentially small...

- Unfortunately, even this will not work correctly!
- In most cases, with $E_i \gg k_B T$, the Boltzmann probability will be exponentially small...
- That is most of the states we choose will have have very little contribution to the original average.

- Unfortunately, even this will not work correctly!
- In most cases, with $E_i \gg k_B T$, the Boltzmann probability will be exponentially small...
- That is most of the states we choose will have have very little contribution to the original average.
- Typically, only a small fraction of states will actually contribute significantly to the value of the sum – and choosing states at random, we are unlikely to pick these important states and hence get a poor estimate of the true value.

Importance Sampling and Statistical Mechanics

- Unfortunately, even this will not work correctly!
- In most cases, with $E_i \gg k_B T$, the Boltzmann probability will be exponentially small...
- That is most of the states we choose will have have very little contribution to the original average.
- Typically, only a small fraction of states will actually contribute significantly to the value of the sum – and choosing states at random, we are unlikely to pick these important states and hence get a poor estimate of the true value.
- But this is ideally suited for importance sampling!

Importance Sampling and Statistical Mechanics

- Importance sampling calculates the correct value of a sum using a set of samples drawn non-uniformly.

Importance Sampling and Statistical Mechanics

- Importance sampling calculates the correct value of a sum using a set of samples drawn non-uniformly.
- For any quantity g_i that depends on states i we can define a weighted average over states:

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

where w_i is any sets of weights we choose.

Importance Sampling and Statistical Mechanics

- Importance sampling calculates the correct value of a sum using a set of samples drawn non-uniformly.
- For any quantity g_i that depends on states i we can define a weighted average over states:

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

where w_i is any sets of weights we choose.

- Making the particular choice $g_i = X_i P(E_i)/w_i$:

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

- Importance sampling calculates the correct value of a sum using a set of samples drawn non-uniformly.
- For any quantity g_i that depends on states i we can define a weighted average over states:

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

  where w_i is any sets of weights we choose.
- Making the particular choice $g_i = X_i P(E_i)/w_i$:

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

- Then:

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

- We can evaluate this expression approximately by selecting a set of N sample states randomly but non-uniformly such that the probability of choosing a state i is:

$$p_i = \frac{w_i}{\sum_j w_j}$$

- We can evaluate this expression approximately by selecting a set of N sample states randomly but non-uniformly such that the probability of choosing a state i is:

$$p_i = \frac{w_i}{\sum_j w_j}$$

- In this case, we get:

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

Note that the first sum is only over the states that we sample, but the second sum is over all states i and has to be calculated analytically!

- Our goal is to choose w_i so that most of the samples are in the region where $P(E_i)$ is big and very few in the region where it is small.

Importance Sampling and Statistical Mechanics

- Our goal is to choose w_i so that most of the samples are in the region where $P(E_i)$ is big and very few in the region where it is small.
- We can choose $w_i = P(E_i)$, in which case the above condition is satisfied as well as $\sum_i w_i = 1$.

- Our goal is to choose w_i so that most of the samples are in the region where $P(E_i)$ is big and very few in the the region where it is small.
- We can choose $w_i = P(E_i)$, in which case the above condition is satisfied as well as $\sum_i w_i = 1$.
- In other words, we just choose N states in proportion to their Boltzmann probabilities and take the average of X over them..

- Our goal is to choose w_i so that most of the samples are in the region where $P(E_i)$ is big and very few in the region where it is small.
- We can choose $w_i = P(E_i)$, in which case the above condition is satisfied as well as $\sum_i w_i = 1$.
- In other words, we just choose N states in proportion to their Boltzmann probabilities and take the average of X over them..
- Unfortunately, we are not done yet – The catch is that it is not easy to pick states with probability $P(E_i)$. This is because to calculate $P(E_i)$, we need to know the partition function, Z.

Markov Chain

- Remarkably, it turns out that we can choose states with probability $P(E_i)$, without knowing the partition function, using a device called Markov Chain.

- Remarkably, it turns out that we can choose states with probability $P(E_i)$, without knowing the partition function, using a device called Markov Chain.
- In this, we want states for the sum $\langle X \rangle = \frac{1}{N}\sum_{k=1}^{N} X_k$, which we will get by generating a Markov chain.

Markov Chain

- Remarkably, it turns out that we can choose states with probability $P(E_i)$, without knowing the partition function, using a device called Markov Chain.
- In this, we want states for the sum $\langle X \rangle = \frac{1}{N} \sum_{k=1}^{N} X_k$, which we will get by generating a Markov chain.
- Consider a single step in the process – Suppose that the previous state for the step before this one, was state i.

Markov Chain

- Remarkably, it turns out that we can choose states with probability $P(E_i)$, without knowing the partition function, using a device called Markov Chain.
- In this, we want states for the sum $\langle X \rangle = \frac{1}{N}\sum_{k=1}^{N} X_k$, which we will get by generating a Markov chain.
- Consider a single step in the process – Suppose that the previous state for the step before this one, was state i.
- For the new state, instead of choosing randomly, we will make some change (usually small) to the state i so as to create a new state.

- Remarkably, it turns out that we can choose states with probability $P(E_i)$, without knowing the partition function, using a device called Markov Chain.
- In this, we want states for the sum $\langle X \rangle = \frac{1}{N} \sum_{k=1}^{N} X_k$, which we will get by generating a Markov chain.
- Consider a single step in the process – Suppose that the previous state for the step before this one, was state i.
- For the new state, instead of choosing randomly, we will make some change (usually small) to the state i so as to create a new state.
- The choice of the new state is determined probabilistically by a set of transition probablilities T_{ij} that give the probability of changing from state i to j.

- If we choose T_{ij} correctly, we can arrange that the probability of visiting any particular state on any step of the Markov chain to be precisely the Boltzmann probability, $P(E_i)$.

Markov Chain

- If we choose T_{ij} correctly, we can arrange that the probability of visiting any particular state on any step of the Markov chain to be precisely the Boltzmann probability, $P(E_i)$.
- When we take many steps and generate the entire chain, the complete set of states that we move through is a correct sample of the Boltzmann distribution and we can average any quantity we like over these states.

- The trick lies in choosing T_{ij}:

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

and also

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

In other words, we are choosing a particular value for the ratio of the probability to go from i to j and the probability to go back from j to i.

Markov Chain

- The trick lies in choosing T_{ij}:

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

  and also

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

  In other words, we are choosing a particular value for the ratio of the probability to go from i to j and the probability to go back from j to i.
- The Boltzmann distribution is a fixed point of the Markov chain.

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.
- The most common choice is the choice that leads to the Metropolis algorithm.

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.
- The most common choice is the choice that leads to the Metropolis algorithm.
- Note that we are allowed to visit the same state more than once in the Markov chain..

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.
- The most common choice is the choice that leads to the Metropolis algorithm.
- Note that we are allowed to visit the same state more than once in the Markov chain..
- Suppose we generate a new state j after making some change in state i. We chose the particular change we make uniformly at random from a specified set of possible changes – called the move set.

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.
- The most common choice is the choice that leads to the Metropolis algorithm.
- Note that we are allowed to visit the same state more than once in the Markov chain..
- Suppose we generate a new state j after making some change in state i. We chose the particular change we make uniformly at random from a specified set of possible changes – called the move set.
- We will accept or reject the new state with probability P_a:

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

Metropolis algorithm

- While this is all good – we still need to figure out what T_{ij} should be.
- The most common choice is the choice that leads to the Metropolis algorithm.
- Note that we are allowed to visit the same state more than once in the Markov chain..
- Suppose we generate a new state j after making some change in state i. We chose the particular change we make uniformly at random from a specified set of possible changes – called the move set.
- We will accept or reject the new state with probability P_a:

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

- This scheme will satisfy all the criterion of the transition probability: $\frac{T_{ij}}{T_{ji}} = e^{-\beta(E_i - E_j)}$

Metropolis Algorithm

1. Choose a random starting state.

Metropolis Algorithm

1. Choose a random starting state.
2. Choose a move uniformly at random from an allowed set of moves.

Metropolis Algorithm

1. Choose a random starting state.
2. Choose a move uniformly at random from an allowed set of moves.
3. Calculate the probability, P_a to accept or reject the move.

Metropolis Algorithm

1. Choose a random starting state.
2. Choose a move uniformly at random from an allowed set of moves.
3. Calculate the probability, P_a to accept or reject the move.
4. With probability P_a, accept the move – ie system changes to the new state; OR reject the move – ie system stays in the current state.

Metropolis Algorithm

1. Choose a random starting state.
2. Choose a move uniformly at random from an allowed set of moves.
3. Calculate the probability, P_a to accept or reject the move.
4. With probability P_a, accept the move – ie system changes to the new state; OR reject the move – ie system stays in the current state.
5. Measure the value of the quantity of interest and add it to the running sum.

Metropolis Algorithm

1. Choose a random starting state.
2. Choose a move uniformly at random from an allowed set of moves.
3. Calculate the probability, P_a to accept or reject the move.
4. With probability P_a, accept the move – ie system changes to the new state; OR reject the move – ie system stays in the current state.
5. Measure the value of the quantity of interest and add it to the running sum.
6. Go to step 2.

Metropolis Algorithm

- The steps where you reject the move that dont change the state of the system do count as steps.

Metropolis Algorithm

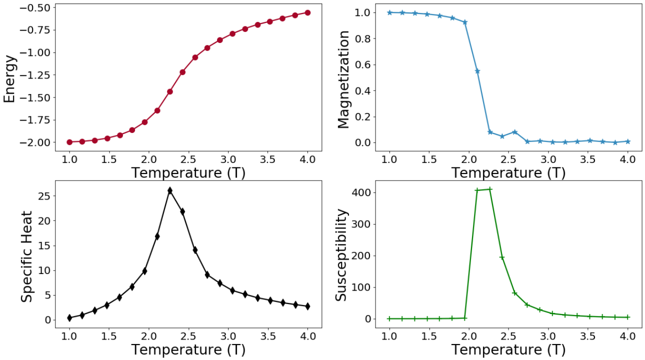
- The steps where you reject the move that dont change the state of the system do count as steps.
- The number of moves that you chose to take you from i to j should be the same as from j to i.

Metropolis Algorithm

- The steps where you reject the move that dont change the state of the system do count as steps.
- The number of moves that you chose to take you from i to j should be the same as from j to i.
- One has to choose a move set such that every possible state is reachable.

- The steps where you reject the move that dont change the state of the system do count as steps.
- The number of moves that you chose to take you from i to j should be the same as from j to i.
- One has to choose a move set such that every possible state is reachable.
- That the Markov chain will go to Boltzmann distribution is proved but, how long will it to take to reach equillibrium is not known.

$$E = -\sum_{\langle ij \rangle} s_i s_j$$



Simulated Annealing

- If we are interested in the ground state energy of a system (classical or quantum), we can use simulated annealing.

Simulated Annealing

- If we are interested in the ground state energy of a system (classical or quantum), we can use simulated annealing.
- The probability that the system is in a state i is:

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

Simulated Annealing

- If we are interested in the ground state energy of a system (classical or quantum), we can use simulated annealing.
- The probability that the system is in a state i is:

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

- Let us assume that the system has a single unique ground state and let us choose our energy scale such that $E_i = 0$ in the ground state and $E_i > 0$ for all others.

Simulated Annealing

- If we are interested in the ground state energy of a system (classical or quantum), we can use simulated annealing.
- The probability that the system is in a state i is:

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

- Let us assume that the system has a single unique ground state and let us choose our energy scale such that $E_i = 0$ in the ground state and $E_i > 0$ for all others.
- Then in the limit $T \to 0$, $\beta \to \infty$ and $e^{-\beta E_i} \to 0$ except in the ground state where $e^0 = 1$.

- If we are interested in the ground state energy of a system (classical or quantum), we can use simulated annealing.
- The probability that the system is in a state i is:

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

- Let us assume that the system has a single unique ground state and let us choose our energy scale such that $E_i = 0$ in the ground state and $E_i > 0$ for all others.
- Then in the limit $T \to 0$, $\beta \to \infty$ and $e^{-\beta E_i} \to 0$ except in the ground state where $e^0 = 1$.
- This in this limit:

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

Simulated Annealing

- One of the ways to find the ground state is then tp slowly lowering the temperature, and see what state the system lands in.

Simulated Annealing

- One of the ways to find the ground state is then tp slowly lowering the temperature, and see what state the system lands in.
- As one lowers the temperature, the system should land in the ground-state!