# **Computational Physics, TFYA90**

Modern Techniques in Materials Science Part I. Molecular dynamics (MD) and Monte Carlo (MC)

Lecture 3, November 9, 2020

Davide G. Sangiovanni

Office G408, Physics Building davide.sangiovanni@liu.se

Theoretical Physics, IFM, Linköping University

#### Monte Carlo is NOT another form of dynamics.

#### Monte Carlo is a SAMPLING method.

The two most important aspects to be decided in Monte Carlo approaches are:

1. WHICH POPULATION TO SAMPLE FROM. One needs to impose some constraints on the population of states sampled.

2. WITH WHAT PROBABILITY TO SAMPLE. Biased or unbiased sampling can make a huge difference in efficiency.

#### Monte Carlo usage in Science



Classical Monte Carlo (CMC) – used to obtain samples from a probability distribution to determine, for example, energy minimum structures (e.g., Ising spin model).

Simulation Monte Carlo (SMC) – algorithms used to evolve configurations based on various acceptance rules (e.g., Kinetic Monte Carlo).

Quantum Monte Carlo (QMC) – random walks can be used to determine quantummechanical energies.

Path-Integral Monte Carlo (PMC) – thermodynamics properties can be evaluated from quantum statistical mechanical integrals.

#### Molecular Dynamics or Monte Carlo?

In Molecular Dynamics, properties are evaluated by tracking them over time. For a given microscopic state, macroscopic properties are calculated as time averages.

These time averages, however, include only the states which occur during the time scale of the MD simulation. Several important issues arise:

1. The length of the MD run is finite (10s or 100s of  $\mu$ s with current supercomputers). There are processes/excitations which occur over longer times which would not be included in the MD time averages.

2. One needs to calculate properties averages, but there might not be interest in simulating, or knowledge of, the actual system dynamics (e.g., Ising spin model). A considerably less demanding technique (CPU-wise) can be used to do the job.

In both cases, MC statistical sampling could be the best choice.



Macroscopic conditions (constant V, T, N) translate as boundary conditions, or constraints, in the microscopic universe.

Microscopic systems are then defined by the fixed thermodynamic variables in the macroscopic world (NVE), (NVT), (NPT) etc.

There are two types of thermodynamic variables:

Extensive variables – scale with size of system (V, N).

Intensive variables – don't scale with size (T, P,  $\mu$ )

Intensive variables are the conjugates of extensive variables.

The constraint used to sample the microscopic system is set by the fixed extensive thermodynamic variable.

The sampling probability depends on the relevant Hamiltonian.

The Hamiltonian in microscopic space corresponds to the free energy function in macroscopic space.

The conjugate, extensive and intensive variables, always "work" in pairs.

First law of thermodynamics, in the energy formulation, yields for the work terms, using the conjugate pairs:

 $dU = TdS + (-PdV) + \mu dN + \dots$ 

S is extensive, T is intensive;TdS (heat flow term)V is extensive, P is intensive;PdV (mechanical work done term)N is extensive, μ is intensive;μdN (chemical work term)

One needs to always specify at least one variable for each pair of conjugate variables:

 $\begin{array}{c} \text{constant S or constant T} \\ \text{constant V or constant P} \\ \text{constant N or constant } \mu \end{array}$ 

This is how the constraints for the microscopic systems are defined.

This is also how the so-called valid thermodynamic ensembles are constructed.

In MC, thermodynamic quantities are averages over relevant set (population) of microscopic states (ensembles).

(NVE) – microcanonical ensemble (NVT) – canonical ensemble  $(\mu VT)$  – grand-canonical ensemble (NPT) – isothermal-isobaric ensemble

**Ensemble is the collection of all possible microscopic states** the system can be in, for a given macroscopic condition.

This defines the population of states, including relevant constraints, which must be sampled in MC simulations.

As ensembles are determined by the extensive variables kept constant, the simplest one to construct is the (NVE) microcanonical ensemble.

It is ideally suited for Newtonian mechanics in a system closed in a box. If the box is closed, N cannot change, the volume is again fixed, and in the case of Newtonian dynamics, the energy is fixed as well.

This is the reason why the (NVE) ensemble is the most natural ensemble for MD simulations. However, this is not the case in MC, where particle momenta are not involved.

In the NVE ensemble all states  $S(\mathbf{q}_i, \mathbf{p}_i)$  of energy "E" have equal probability to be visited. We say that there are "n" of such states. All other states "S<sub>k</sub>" (with energies  $E_k \neq E$ ) have zero probability to be visited. Thus, the probability that a specific state  $S(\mathbf{q}_i, \mathbf{p}_i)$  is occupied at a given instant is 1/n.

#### Regarding other statistical ensembles like NVT, NPT, and $\mu VT$

#### How to sample?

The probability of states in any ensemble is proportional to  $e^{-\beta H}$ , where H is the Hamiltonian and  $\beta = 1/k_BT$ .

 $p \sim exp(-\beta H)$ 

This probability has to be normalized by the partition function Z, which is the sum of probabilities over all states v in the ensemble:

$$Z = \sum_{\nu} \exp(-\beta H_{\nu}) \qquad \qquad p_{\nu} = \frac{\exp(-\beta H_{\nu})}{\sum_{\nu} \exp(-\beta H_{\nu})} = \frac{\exp(-\beta H_{\nu})}{Z}$$

 $p_v$ , called the probability distribution function (PDF), yields in this manner the correct probability to sample essentially in any ensemble, provided one knows H.

The microscopic H should include everything that fluctuates in the system. Its correct form can be obtained by taking a Legendre transformation of the entropy of the system (H is essentially a Legendre transform), obtained from 1<sup>st</sup> law:

$$dS = \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN + \dots$$

Note the conjugate pairs in the entropy formulation: (1/T, U), (-P/T, V),  $(\mu/T, N)$ .

The Hamiltonian, which corresponds to the relevant free energy function in macroscopic space, can be obtained for each microscopic ensemble.

**Canonical Ensemble (NVT)** 
$$dE = TdS + (-PtV) + \mu dN + ...$$
  
First law becomes:  $dS = \frac{1}{T}dE$  or  $d\left(S - \frac{1}{T}E\right) = 0$ 

the relevant free energy is F = E - TS, which is the Helmholtz free energy, and the Legendre transform of entropy yields: -F/T = S - E/T.

The Hamiltonian will thus contain only the -E/T term, and the PDF for the canonical (NVT) ensemble takes the following form:

$$p_{\nu}^{\text{NVT}} = \frac{\exp(-\beta E_{\nu})}{\sum_{\nu} \exp(-\beta E_{\nu})} \qquad \beta = 1/(k_{\text{B}}T)$$

 $dE = TdS + (-PdV) + \mu V + \dots$ Isothermal-Isobaric Ensemble (NPT) First law becomes:  $dS = \frac{1}{T}dE + \frac{P}{T}dV$  or  $d\left(S - \frac{1}{T}E - \frac{P}{T}V\right) = 0$ The free energy in this case is: F = E - TS + PVand the Legendre transform of entropy takes the form: -F/T = S - E/T - PV/T. From this, only the -(E + PV)/T erm is taken in the Hamiltonian and the PDF for the isothermal-isobaric (NPT) ensemble becomes:  $p_{\nu}^{\text{NPT}} = \frac{\exp[-\beta(E_{\nu} + PV_{\nu})]}{\sum \exp[-\beta(E_{\nu} + PV_{\nu})]}$  $\beta = 1/(k_B T)$ 

 $dE = TdS + (-PdV) + \mu dN + \dots$ **Grand-canonical Ensemble (µVT)** First law becomes:  $dS = \frac{1}{T}dE - \frac{\mu}{T}dN$  or  $d\left(S - \frac{1}{T}E + \frac{\mu}{T}N\right) = 0$ The free energy for fixed  $(\mu, V, T)$  becomes:  $F = E - TS - \mu N$ and the Legendre transform of entropy for this ensemble is:  $-F/T = S - E/T + \mu N/T.$ Again, from this one takes only the  $-(E - \mu N)/T$  erm in the Hamiltonian and the PDF for the isothermal-isobaric  $(\mu VT)$  ensemble becomes:  $p_{\nu}^{\mu VT} = \frac{\exp[-\beta(E_{\nu} - \mu N)]}{\sum \exp[-\beta(E_{\nu} - \mu N)]}$ 

# **Monte Carlo Integration**

Originally, Monte Carlo was used as an integration method.



Ъ An illustration of Monte Carlo integration. In this example, the domain *D* is the inner circle and the domain E is the square. Because the square's area (4) can be easily calculated, the area of the circle ( $\pi$ \*1.0<sup>2</sup>) can be estimated by the ratio (0.8) of the points inside the circle (40) to the total number of points (50), yielding an approximation for the circle's area of 4\*0.8 = 3.2 ≈ π.

The procedure was easily extended to functions of two variables and multiple integrals. It is called **SIMPLE SAMPLING**.

# **Simple Sampling in MC (Simple MC)**

The aim in MC simulations is to calculate average thermodynamic properties,  $\langle A(r^N) \rangle$ , which can be done by evaluating multidimensional integrals over the 3N degrees of freedom in an N particle system:

 $\langle A(r^N) \rangle = \int A(r^N)p(r^N)dr^N$ 

where  $p(r^N)$  is the appropriate PDF in the respective ensemble. Here, one can concentrate on the (NVT) ensemble for two reasons:

1. ALL OTHER ENSEMBLES follow the same rational/approach as in (NVT).

#### 2. The NVT ensemble is the natural choice for MC simulations.

In MD, Newton's EOM lead naturally to energy conservation, hence the NVE selection.

In MC, until recently, it was not possible to perform calculations in the NVE ensemble due to the absence of kinetic energy. Temperature, however, can be easily kept constant in the PDF, and the NVT-MC is simplest to implement.

# Simple Sampling in MC (Simple MC)

As shown, the PDF in the NVT ensemble takes the form:

$$p(\vec{r}^{N}) = \frac{\exp\left[-\beta E(\vec{r}^{N})\right]}{\int \exp\left[-\beta E(\vec{r}^{N})\right] d\vec{r}^{N}}$$

These integrals cannot be evaluated analytically or numerically. Typical schemes for 3N-dimensional integrals require m<sup>3N</sup> function evaluations, where m is the number of points required to evaluate the integral in each dimension.

In simple MC, a large number of trial configurations  $r^N$  are generated and the integrals are replaced by summations over a finite number of configurations:

$$\left\langle \mathbf{A}(\vec{\mathbf{r}}^{N}) \right\rangle = \frac{\sum_{i=1}^{N_{\text{trial}}} \mathbf{A}_{i}(\vec{\mathbf{r}}^{N}) \exp\left[-\beta \mathbf{E}_{i}(\vec{\mathbf{r}}^{N})\right]}{\sum_{i=1}^{N_{\text{trial}}} \exp\left[-\beta \mathbf{E}_{i}(\vec{\mathbf{r}}^{N})\right]}$$

## **Simple Sampling in MC (Simple MC)**

With simple sampling, each trial configuration  $\vec{r}^N$  corresponds to a randomly chosen state (point) v. If one randomly picks M states, they need to be weighted with the correct probability  $p_v$ :

$$\langle A \rangle = \sum_{\nu=1}^{M} p_{\nu} A_{\nu}$$
  $p_{\nu} = \frac{\exp(-\beta E_{\nu})}{\sum_{\nu=1}^{M} \exp(-\beta E_{\nu})} = \frac{\exp(-\beta E_{\nu})}{Z_{M \neq NVT}}$ 

Simple sampling does not work. The reason is states are picked essentially in proportion to their degeneracy. The higher the energy, the more states at that energy. One samples a great number of states but not the relevant ones.

This random, unbiased sampling of states yields too many configurations with low weight, or very small Boltzmann factor, which make very little contribution to the average which needs to be calculated.

One needs to bias the sampling method: **IMPORTANCE SAMPLING.** 

# **Importance Sampling**

In importance sampling points are chosen according to the anticipated importance of the value to the function (contribution it makes) and weighted by the inverse of the probability of choice.

The difference is that in importance sampling one no longer uses a simple average of all points sampled. In importance sampling the sampling is biased by the use of a weighted average.

## **Importance Sampling**

In MC, importance sampling translates into biasing the sampling towards the important, relevant, low energy states.

In other words, one picks states with a probability proportional to  $exp(-\beta E)$ , instead of randomly picking them and weighing them later by a probability.



Relevant thermodynamic ensembles fluctuate around states with low energy, so importance sampling is used in MC to sample mostly these states.

#### **Markov Chains**

Used to construct probability weighted samples.

A Markov chain is a sequence of trials in which the outcome of successive trials depends only on the immediately preceding trial.

The procedure ensures that one "walks" through the phase space and "visit" each state with proper probability.

In Markov chains, a new state is accepted only if it is more "favorable" than the existing state. For simulations of ensembles, this usually means that the new trial state is lower in energy.

In MC simulations Markov chains are required to accurately determine the properties of the system in the finite time available for simulation. The role of Markov chains is to sample those states which make the most significant contributions to the calculated thermodynamic averages.

# **Metropolis Sampling**

Generates Markov chains which construct the probability weighted sample to explore the thermodynamical behavior around the energy minimum.

Metropolis sampling biases the generation of configurations towards those which make the most significant contribution to the integral/average of interest.

It generates states with a probability of  $exp[-\beta E(r^N)]$  and counts each of them equally.

This is in contrast to the simple MC integration/sampling, which generates states with equal probability and assigns them a weight of  $exp[-\beta E(r^N)]$ .

# **Metropolis Monte Carlo**

- 1. Assign initial position to particles & calculate U.
- 2. *Move one particle randomly* & calculate new U' and  $\Delta U=U'-U$ .
- 3. If  $\Delta U < 0$  *accept* move.
- 4. If  $\Delta U > 0$  *accept* move if  $\xi < \exp[-\beta \Delta U]$ ;  $\xi \in (0,1)$  random number.
- 5. If *move rejected* take the old configuration as the new one
  - repeat 2 4 procedure for another arbitrarily chosen particle.
- 6. For each new configuration *evaluate* <<u>A</u>>.
- 7. Repeat the whole procedure a few million times for adequate statistic





#### 2D Ising spin model: example of Metropolis MC simulation



https://www.ibiblio.org/e-notes/Perc/ising.htm

https://ruihaoqiu.github.io/MC-Magnetic-Phase-Transition/



# Diffusion







- Pick an atom at random
- Pick a hop direction



- Pick an atom at random
- Pick a hop direction
- Calculate  $\exp(-\Delta E_b/k_BT)$



- Pick an atom at random
- Pick a hop direction
- Calculate  $\exp(-\Delta E_b/k_BT)$
- If  $(\exp(-\Delta E_b/k_BT)$ >random number) do the hop

# **Kinetic Monte Carlo**

Consider all hops simultaneously

















#### **Optimization techniques – Simulated Annealing (SA)**

General method to find the global minimum of a given function in a large space. Name analogue to the metallurgy technique of heating and controlled cooling, to increase crystal size, reduce defects.

SA is extensively used in conjunction with MD as well as MC and requires:

- 1. A description of possible system configurations.
- 2. A generator of random changes in the configurations.
- 3. An objective function (typically energy).
- 4. A control parameter (typically T) and an annealing schedule:
  - Slow cooling (annealing), one can find global ground state (real world).
  - Fast cooling (quenching), one can find local minimum (glass).

In MD, following steps are needed to implement SA:

- choose annealing schedule (cooling rate).
- solve Newton's EOM for each atom.
- control T via coupling to heat bath (Andersen, Nosé-Hoover thermostats).

Downhill energy moves with Steepest Descent, Conjugate Gradient methods.

Adaptive simulated annealing: Metropolis Monte Carlo coupled to fictious thermostat



Flowchart of Metropolis scheme coupled with a thermostat T.

# Additional notes

c<sub>i</sub> is the "degeneracy" of A<sub>i</sub> value

Generalization of arithmetic mean value <A>

$$=\frac{c\_1 \cdot A\_1 + c\_2 \cdot A\_2 + c\_3 \cdot A\_3}{c\_1 + c\_2 + c\_3} \longrightarrow \begin{array}{c} \sum\_{i} \(c\_i \cdot A\_i\) \\ =\underline{\sum\\_{i} c\\_i} \end{array}$$

 $c_1 = 7$  $c_2 = 4$  $c_3 = 5$ 

Let's represent all possible states of a system with

a square grid. Let's assume that all states  $s_i$  (square boxes) can be sampled with equal probability (e.g., each state has equal energy and the simulation ensemble is NVE).

The probability  $p_{Ai}$  that a state has a value "A<sub>i</sub>" of a property "A" is given by



The denominator corresponds to the "partition function"

Given that all states (boxes) can be sampled with equal probability, The expectation value (ensemble average) of A, when the system is in equilibrium corresponds to the arithmetic average:

$$\_{ens} = \sum\_i \(p\_{Ai} \cdot A\_i\)$$

This is the case for the NVE ensemble: the partition function is simply "N", the number of microstates of equal energy

All possible states of the system



# **1.1. Metropolis Sampling**

In statistical physics one can find the average of a property  $A(\{\underline{r}\})$  that is a function of the coordinates  $\{r\}$  of N particles, in a system that is in thermodynamic equilibrium, Probability that a state "r" i

$$\langle A \rangle = \frac{\int d^{3N} r A(\{\underline{r}\}) \exp[-U(\{\underline{r}\})/kT]}{\int d^{3N} r \exp[-U(\{\underline{r}\})/kT]}.$$

Probability that a state "r" is visited during NVT sampling given by Boltzmann statistics

Canonical "partition function" Z

The calculation involves averaging the dynamical variable of interest, A, which depends on the positions of all the particles in the system, over an appropriate thermodynamic ensemble. Often the canonical ensemble is chosen; one with a fixed number of particles, volume and temperature, N, V, and T. In this case the configurations are weighted by the Boltzmann factor  $\exp[-U(\{\underline{r}\})/kT]$ , where U is the potential energy of the system, and k the Boltzmann constant. Integration is over the positions of all particles (3N coordinates). The denominator in Eq. (1) is needed for normalization, and is an important quantity in its own right, because the Helmholtz free energy can be obtained from it (for a system with the independent variables, N, V, and T).