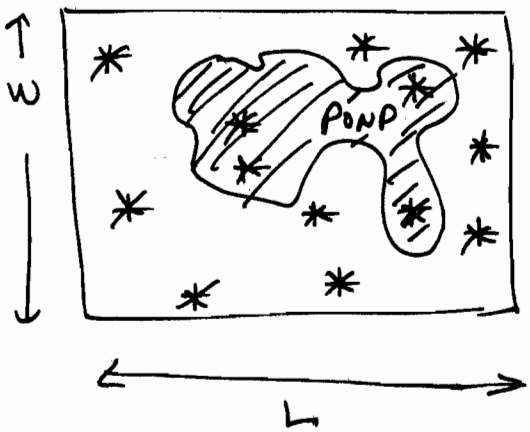
History

von Neumann, Ulam + Metropolis (1947) → Studies of diffusion of neutrons in fissionable matter. Coined the term "Monte Carlo".

Metropolis, Rosenbluth², Teller² (1953) → first MC simulation for equilibrium stat. mech. properties.

Monte Carlo Integration

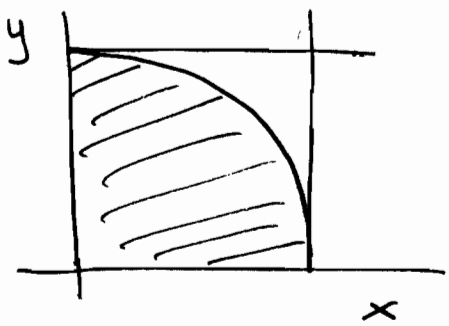


Consider a pond of arbitrary shape. Possible way: throw rocks at random, in area bounded by W×L rectangle

$$\text{Area of pond} = W \times L \times \frac{\text{hits in pond}}{\text{total hits}}$$

Finding an area is an <u>integration</u> problem, so that "Monte Carlo" is a way to perform multidimensional integrals.

E.g.



Find area of quarter circle (or value of π) by above method:

After 10^7 trials $\pi \simeq 3.1417$ (4 sig. figures)

The gain in accuracy is only 1 decimal point per 10x in the number of trials.

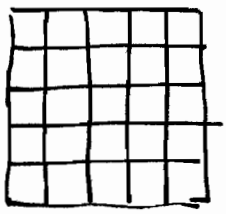
For a system of low dimensionality, analytical or direct numerical methods (e.g. Simpson's rule) work much better than Monte Carlo. E.g. for 10^4 function evaluations (intervals) Simpson's rule integration gives $\pi \simeq 3.141593$ (6 sig. figures)

Multidimensional Integrals

Canonical Partition function for $N \simeq 20$ particles

$$Q = \sum_{\substack{\text{pos}\\1}} \sum_{\substack{\text{pos}\\2}} \sum_{\substack{\text{pos}\\3}} \cdots \sum_{\substack{\text{pos}\\N}} \exp(-\beta \mathcal{H})$$

Even if each particle only has 25 possible possitions, one needs $\cancel{20^{25}} = \cancel{10^{32}}$ $25^{20} = 10^{28}$ function evaluations for a direct integration. Moreover, most of the configurations contribute very little to Q:



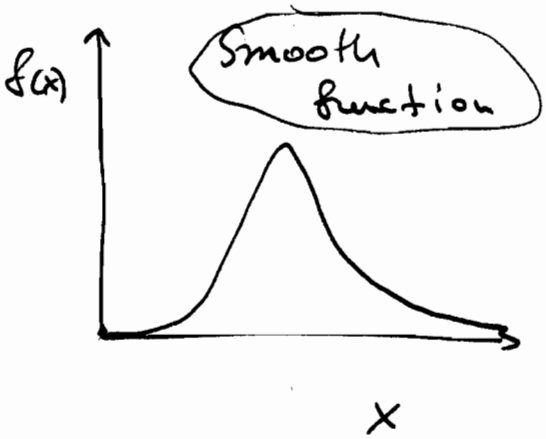
← 25 positions

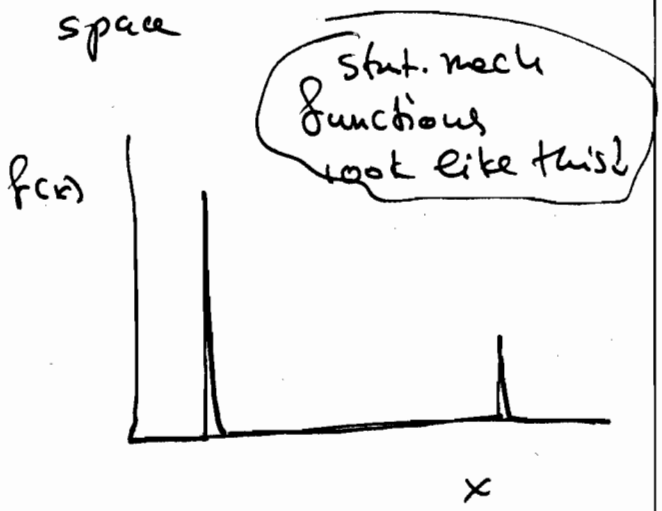
① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ... ⑳

Throw 20 particles at random

Assume no two particles can occupy same position in space. What fraction of configurations generated this way are allowable?

Things are better for less dense systems, but most of the systems of interest in stat. mechanics have this problem → function to be integrated has non-zero values only in small regions of phase space





<u>Importance Sampling</u> (also called Metropolis MC)

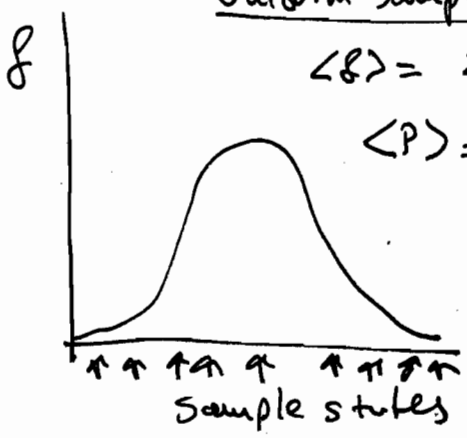
Overcomes sampling problem by generating states with a desired probability distribution, rather than uniformly.

This does not work for getting $\int f(x)dx$, but does work for getting properties that depend on $f(x)$:

<u>Uniform Sampling</u>

$$\langle f \rangle = \Sigma f_i / n$$

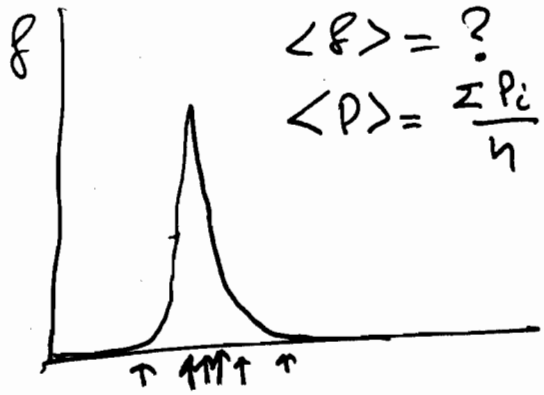
$$\langle P \rangle = \frac{\Sigma f_i P_i}{\Sigma f_i}$$



<u>Importance Sampling</u>

$$\langle f \rangle = ?$$

$$\langle P \rangle = \frac{\Sigma P_i}{n}$$



How do we generate states with the correct probability?

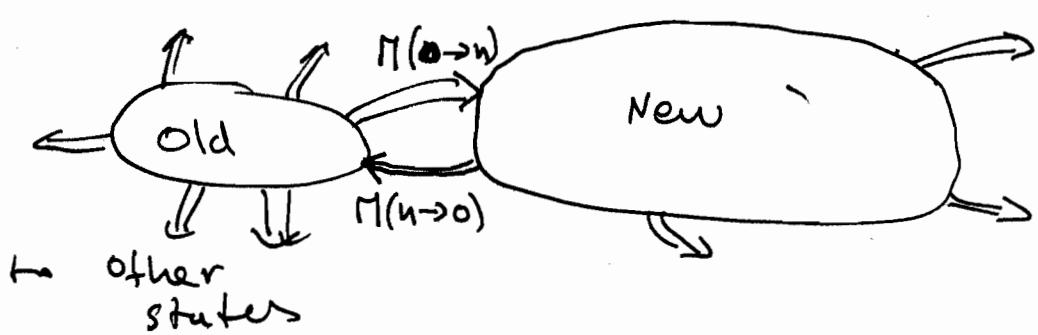
Consider the system of 20 particles in 25 positions - the number of possible states (for distinguishable particles) is $25^{20} \simeq 10^{32}$ states

Now consider two configurations Old (o) and New (n)

In a thought experiment, we can perform a MC simulation with a number of samples <u>much greater</u> than the number of states (samples $\gg 10^{32}$). In this simulation, we generate ~~random~~ uniform configurations. In the large sample, there are $N(o)$ and $N(n)$ realizations for the "old" and "New" configurations, with

$$\frac{N(o)}{N(n)} = \frac{\exp(-\beta u(o))}{\exp(-\beta u(n))} = \exp(-\beta \Delta u)$$ in the NVT ensemble!

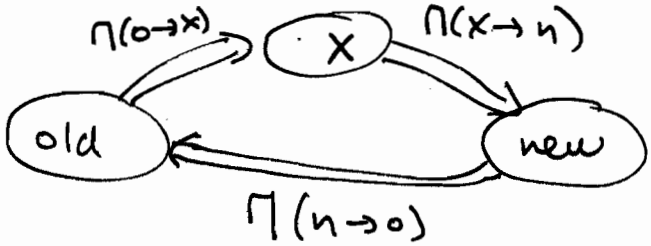
Now consider the case of non-uniform sampling. To ensure that configurations appear with the correct probability, one possible way is to require that in the evolution of configurations,



Detailed Balance Condition: $\boxed{N(o)\cdot \Pi(o\to n) = N(n)\cdot \Pi(n\to o)}$

In other words, we must have that the probabilities of going from $o \to n$ and $n \to o$ must be inversely proportional to the desired probability of finding each state.

Why is this only one possible way? → One could have this:



But, in practice, it is very hard to make sure that this type of simulation (that violates detailed balance) is correct

The probability of going from state o to state n has two parts:

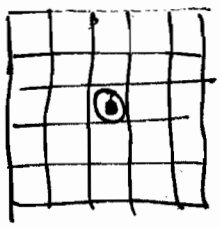
$$\Pi(o\to n) = \alpha(o\to n) \times \text{accept}(o\to n)$$

$\alpha(o\to n)$: probability of selecting state n from state o

accept$(o\to n)$: to be determined

$\alpha(o\to n)$ contains the rules of the simulation –

e.g.



{try to move particle to a new position to the left, right, top, or bottom of the old position}

The best choice for $\alpha(o \to n)$ is one that satisfies

<u>microscopic reversibility</u> $\alpha(o \to n) = \alpha(n \to o)$

so that the probability of trying state n from o is the same as the probability of trying o from n.

<u>E.g.</u> <u>attempted move</u> 1. select random particle, attempt to move it to the left or top position
↳ violates microscopic reversibility

2. select random particle, attempt to move to a randomly selected position → ok

3. select random particle, attempt to move it to a randomly selected <u>unoccupied</u> position → is this ok? [class problem]

Now we can derive the acceptance condition:

$$N(o) \cdot \Pi(o \to n) = N(n) \cdot \Pi(n \to o) \Rightarrow$$

$$N(o) \cdot \alpha(o \to n) \cdot \text{accept}(o \to n) = N(n) \cdot \alpha(n \to o) \cdot \text{accept}(n \to o)$$

$$\Rightarrow \boxed{\frac{acc(o \to n)}{acc(n \to o)} = \frac{N(n)}{N(o)} = \exp(-\beta[u(n) - u(o)])}$$

There are many possibilities for $acc(o \to n)$ that satisfy this equation – for example:

Metropolis Rule:

$$acc(o \to n) = 1 \quad \text{if } U(n) < U(o)$$

$$acc(o \to n) = \exp(-B[U(n) - U(o)]) \quad \text{if } U(n) \geq U(o)$$

Symmetrical (Barker) Rule:

$$acc(o \to n) = \frac{\exp(-BU(n))}{\exp(-BU(o)) + \exp(-BU(n))}$$

Example of appliation of these rules:

Consider a system with exactly 10 states, numbered

j:	1	2	3	4	5	6	7	8	9	10
$\exp(-BU(j))$	1	2	3	4	5	6	7	8	9	10

} probability is equal to index

What is the partition function? (Class Problem)

What is the equilibrium probability of state 1?

Now, let's make an importance sampling simulation.

Transition rule? $\alpha(o \to n) = 1/10$ for <u>any</u> pair of states o and n (including $o = n$)

$\Pi(o \to n)$ Metropolis

$o \downarrow$ \ $\to n$	1	2	3	...	10
1	1/10	1/10	1/10	...	1/10
2	1/20	3/20	1/10	...	1/10
3	1/30	2/30	...		1/10
⋮					
10	1/100	2/100	...		11/20

$\Pi(o \to n)$ Barker

$o \downarrow$ \ $\to n$	1	2	3	...	10
1	0.252	$\frac{2}{30}$	$\frac{3}{40}$	...	$\frac{10}{110}$
2	1/30	0.371	$\frac{3}{50}$	...	$\frac{10}{120}$
3	1/40				
⋮					
10	1/110		...		0.719

Barker's probability of remaining in the same state is higher than for Metropolis acceptance. In general, convergence from an arbitrary starting state to equilibrium is faster with Metropolis acceptance.

Example of flow of probabilities with the system with 10 states – The starting state is state 3, which we represent as [0 0 1 0 0 0 0 0 0 0]

Metropolis

Trial										
0	0	0	1	0	0	0	0	0	0	0
1	0.033	0.067	0.200	0.100	0.100	0.100	0.100	0.100	0.100	0.100
2	0.024	0.049	0.083	0.089	0.102	0.114	0.124	0.132	0.139	0.145
3	0.02	0.04	0.062	0.079	0.097	0.113	0.128	0.142	0.154	0.165
4	0.019	0.038	0.057	0.075	0.093	0.111	0.128	0.145	0.16	0.174
5	0.018	0.037	0.055	0.074	0.092	0.11	0.128	0.145	0.162	0.178
6	0.018	0.037	0.055	0.073	0.091	0.109	0.128	0.145	0.163	0.18
7	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.146	0.163	0.181
8	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.164	0.182
OO	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.164	0.182

Barker

0	0	0	1	0	0	0	0	0	0	0
1	0.025	0.040	0.454	0.057	0.063	0.067	0.070	0.073	0.075	0.077
2	0.025	0.046	0.225	0.075	0.086	0.095	0.103	0.109	0.115	0.12
3	0.023	0.044	0.128	0.079	0.093	0.106	0.118	0.128	0.137	0.145
4	0.021	0.041	0.087	0.078	0.095	0.11	0.124	0.137	0.149	0.16
5	0.019	0.039	0.069	0.077	0.094	0.111	0.127	0.141	0.155	0.168
6	0.019	0.038	0.061	0.075	0.093	0.111	0.128	0.144	0.159	0.173
7	0.019	0.037	0.057	0.074	0.093	0.11	0.128	0.145	0.161	0.176
8	0.018	0.037	0.056	0.074	0.092	0.11	0.128	0.145	0.162	0.178
9	0.018	0.037	0.055	0.073	0.092	0.11	0.128	0.145	0.163	0.18
10	0.018	0.036	0.055	0.073	0.091	0.11	0.128	0.145	0.163	0.18
11	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.163	0.181
12	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.146	0.163	0.181
13	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.164	0.181
14	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.164	0.182
OO	0.018	0.036	0.055	0.073	0.091	0.109	0.127	0.145	0.164	0.182

In general, mc moves need to be

(a) symmetric, $\alpha(o \to n) = \alpha(n \to o)$

(b) ergodic: should allow sampling of all states

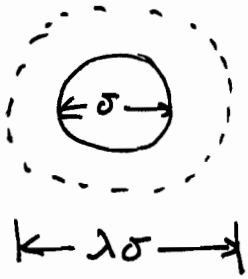
(c) efficient: good balance of cost vs. acceptance ratio

Example of MC for continuous-space systems

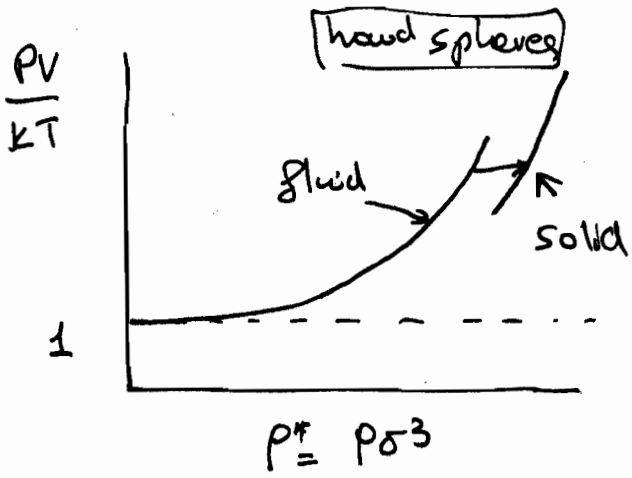
Lattice models are fast and allow studies of large systems, but in most cases of engineering interest continuous-space models are more appropriate.

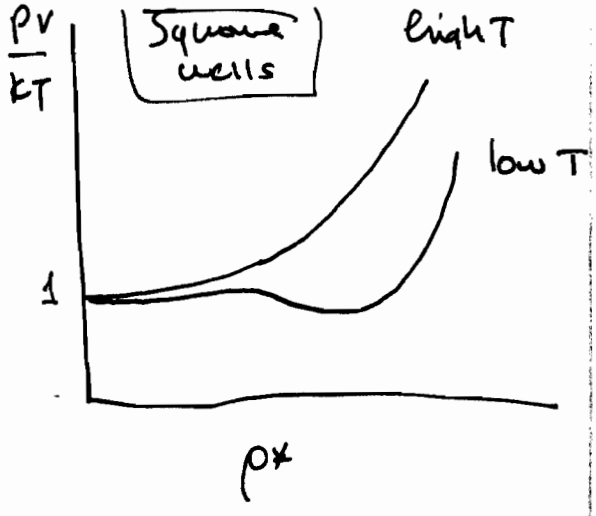
Example of simple continuous-space models:

hard spheres $\quad \mathcal{U}(r) = \begin{cases} \infty & r \le \sigma \\ \phi & r > \sigma \end{cases}$



square well system $\quad \mathcal{U}(r) = \begin{cases} \infty & r \le \sigma \\ -\epsilon & \sigma < r \le \lambda\sigma \\ \phi & r > \lambda\sigma \end{cases}$

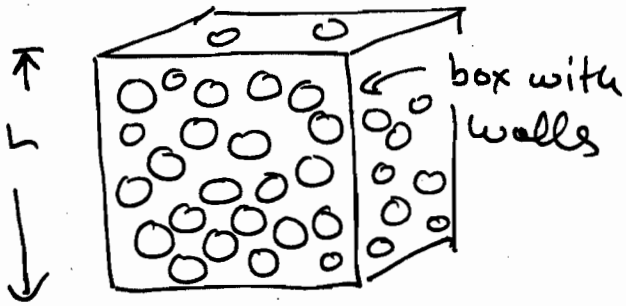




Hard spheres and square-well systems have been used to model proteins, colloids, polymers ...

How do we set up an NVT MC simulation of such a system?

A. Boundary conditions



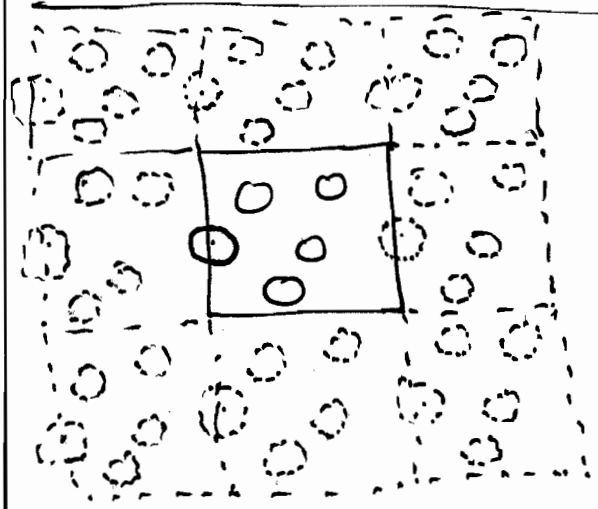
In 3 dimensions –

for $L = 10\sigma$, fraction of particles within 2σ from wall $= 1 - \frac{6^3}{10^3} = 80\%$

$L = 20\sigma \rightarrow$ fraction $= 50\%$

But – for $L = 20\sigma$ at $\rho^* \simeq 0.5$, need to keep track of 4,000 particles (too many)

A.1 Periodic boundary conditions



Duplicate central "box" in all directions.

Particles leaving from one side "come back" on the opposite end.

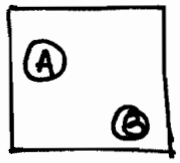
Periodic boundary conditions replace a finite system with an infinite periodic system – in most cases this has small (but measurable) effects on the properties relative to the properties of an infinite system ("bulk" or macroscopic)

A.2 Minimum Image Convention

Once we have an infinite system, we should start worrying about calculating interactions – do we need to take an infinite number of particles into account?

A very useful rule is the following:

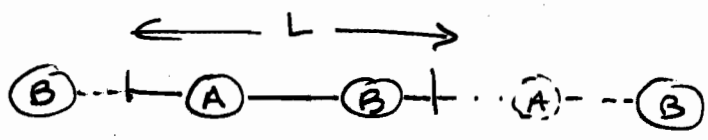
<u>If</u> interactions between two particles are zero beyond a distance <u>less than</u> $L/2$, then we only need to take into account interactions between "minimum image" pairs



Given particles A and B in central box, there is <u>at most</u> one image of B that is within $L/2$ distance from A.

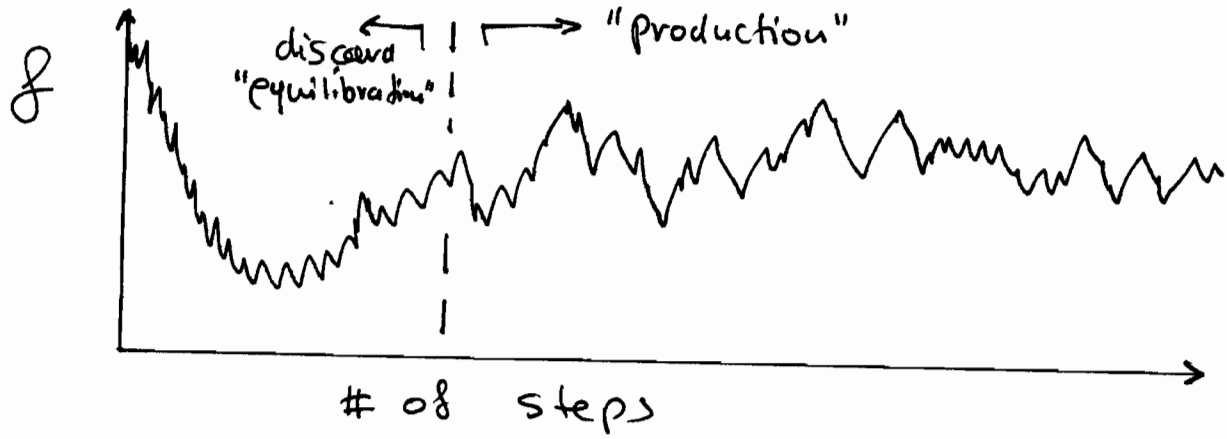
How do we find this?

In 1 dimension:



Analysis of Simulation Results

At the end of a simulation, one has a list of observed values as a function of MC (or MD) step – e.g.



How do we obtain estimates for $\langle f \rangle$ <u>and</u> (equally important) an uncertainty estimate?

For $\langle f \rangle$ itself, one needs to obtain an average of values obtained after an initial "equilibration" period is discarded, to eliminate artifacts of initial configuration,

$$\langle f \rangle = \frac{\sum f_v}{\#\text{ of "production" steps}}$$

The "simulation uncertainty" is <u>not</u> the standard deviation of the $\{f_v\}$ values – Recall fluctuation expressions,

$$\langle (\delta \underline{u})^2 \rangle = k_B T^2 \underline{C}_v$$

so the standard deviations (fluctuations) are thermodynamic (extensive) properties

To obtain an estimate for the statistical uncertainties, we need to take into account the correlations among successive samples

Elementary Statistics

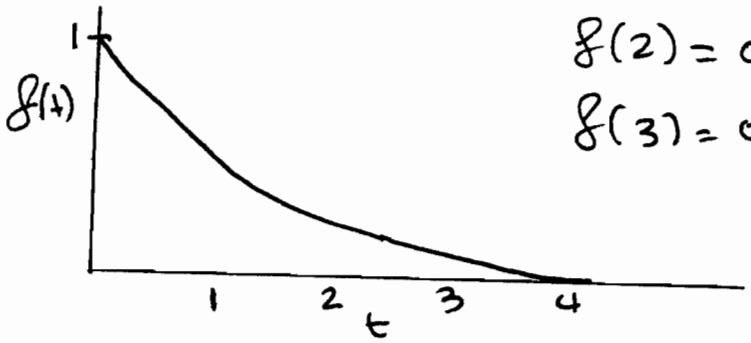
Consider a normally distributed random variable with mean value $\bar{x}_\infty$ and std. deviation σ_∞

Draw N independent samples of x: $\left\{ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{matrix} \right\}$ as $N \to \infty$

$$\bar{x}_N \to \bar{x}_\infty$$
$$\sigma_N \to \sigma_\infty$$

The normalized difference between sample mean $\bar{x}_N$ and true mean $\bar{x}_\infty$, $t = \dfrac{\bar{x}_N - \bar{x}_\infty}{s}$

where $s = \sigma_N / \sqrt{N}$ follows Student's t distribution w/ ∞ degrees of freedom:



$f(2) = 0.048$
$f(3) = 0.003$ $\Big\} \to$ multiply sample variance by 2 for 95% confidence, 3 for 99.7%

→ In practice: Split production period of run into N "independent" blocks (typically 10–100).

Statistical uncertainties $\propto \dfrac{1}{\sqrt{\#\text{ of steps}}}$ $\Big\}$ ×10 in CPU time → ×3 in accuracy

NVT-ensemble Monte Carlo

Metropolis rule : $\mathcal{P}_{acc} = \min\{1, \exp(-\beta \Delta \underline{u})\}$

Since $\mathcal{P}_v \propto \exp(-\beta \underline{u}_v)$

Trial moves: (a) microscopically reversible
(b) ergodic
(c) efficient

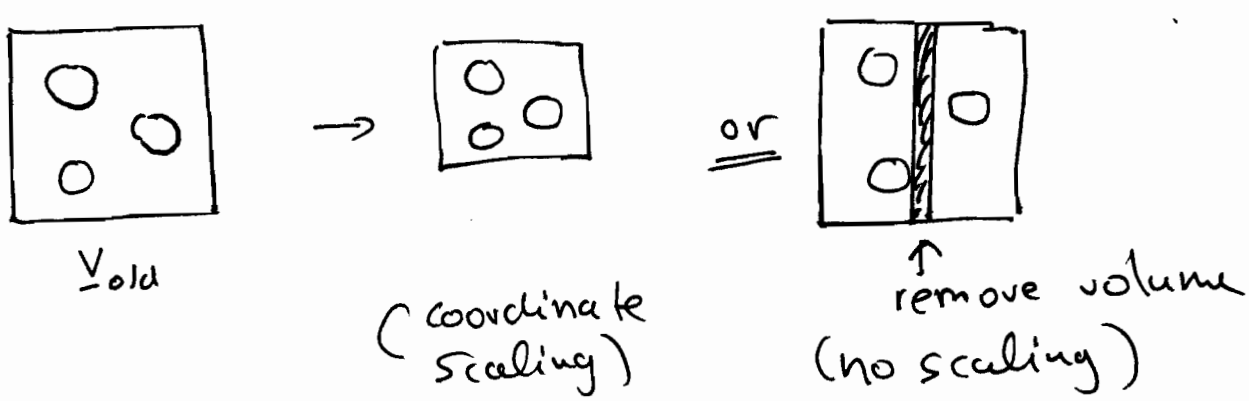
* Particles are picked one at a time at random * Attempt to displace by δr uniformly distributed between 0 and δr_{max}

δr_{max} controls acceptance ratio ; best $\sim 30\%$

NPT-ensemble Monte Carlo

Recall that $\mathcal{P}_v \propto \exp(-\beta \underline{u}_v - \beta P \underline{V}_v)$

In addition to particle displacements, need volume fluctuations



Second approach would also work for lattice models.

For <u>coordinate scaling</u>, must have

$$\mathcal{P}_{old} \propto \underline{V}^N \exp(-\beta \underline{U}_{old} - \beta P \underline{V}_{old})$$

"entropic" term to take into account different # of possible positions in different volumes

$$\mathcal{P}_{new} \propto (\underline{V} + \Delta \underline{V})^N \exp(-\beta \underline{U}_{new} - \beta P \underline{V}_{new})$$

$$\frac{\mathcal{P}_{new}}{\mathcal{P}_{old}} = \left(\frac{\underline{V} + \Delta \underline{V}}{\underline{V}}\right)^N \exp(-\beta \Delta \underline{U} - \beta P \Delta \underline{V}) \quad \text{(A)} \qquad \mathcal{P}_{acc} = \min\left(1, \frac{\mathcal{P}_{new}}{\mathcal{P}_{old}}\right)$$

(metropolis)

For <u>volume element creation / annihalation</u>, the particles do not move, so no $\underline{V}^N$ term is present:

$$\frac{\mathcal{P}_{new}}{\mathcal{P}_{old}} = \exp(-\beta \Delta \underline{U} - \beta P \Delta \underline{V}) \quad \text{(B)}, \quad \mathcal{P}_{acc} = \min\left(1, \frac{\mathcal{P}_{new}}{\mathcal{P}_{old}}\right)$$

Let's determine the equation of state followed by a system of non-interacting particles that undergo a "virtual simulation":

case (A) $\rightarrow$ must have (for <u>all</u> $\Delta \underline{V}$)

$$\left(\frac{\underline{V} + \Delta \underline{V}}{\underline{V}}\right)^N \exp(-\beta \cancel{\Delta \underline{U}} - \beta P_{eq} \Delta \underline{V}) = 1$$

no interactions

$$\boxed{\mathcal{P}^{+}_{acc} = \mathcal{P}^{-}_{acc} \text{ for } P = P_{eq}}$$

$$\Rightarrow \left(1+\frac{\Delta \underline{V}}{\underline{V}}\right)^N = \exp(\beta P_{eq} \Delta \underline{V}) \Rightarrow$$

Taylor expand

$$\cancel{1} + \frac{N \cancel{\Delta \underline{V}}}{\underline{V}} = \cancel{1} + \beta P_{eq} \cancel{\Delta \underline{V}} \Rightarrow \boxed{P_{eq} = \frac{N k_B T}{\underline{V}}}$$

Case (B): $\mathcal{P}^{+}_{acc} = \mathcal{P}^{-}_{acc} \Rightarrow$

$$\exp(-\beta \cancel{\Delta \underline{U}} - \beta P_{eq} \Delta \underline{V}) = \underbrace{\left(1 - \frac{\Delta \underline{V}}{\underline{V}}\right)^N}_{\text{probability of space being removed not containing any particles}}$$

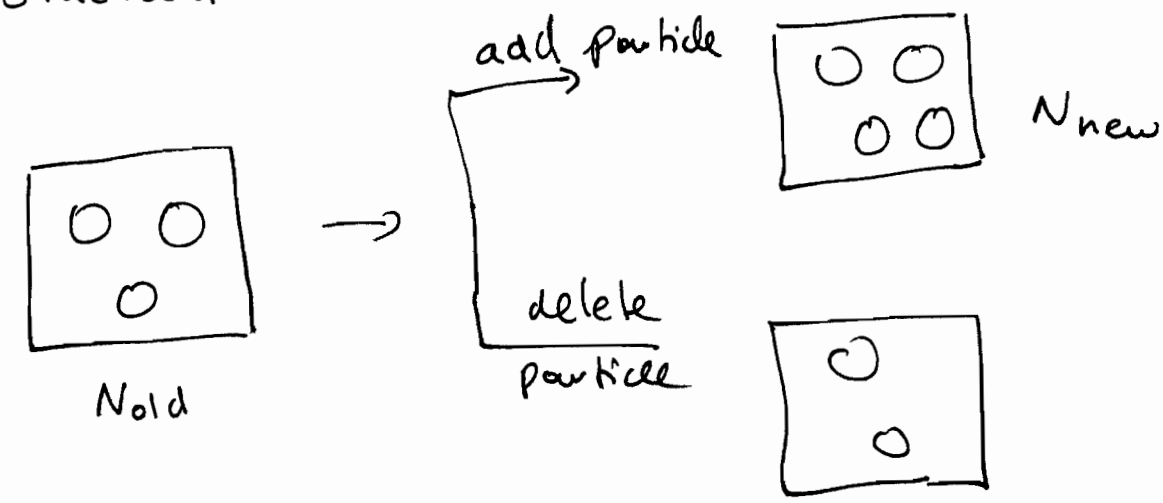
no inter.

$$\Rightarrow \cancel{1} - \beta P_{eq} \cancel{\Delta \underline{V}} = \cancel{1} - \frac{N \cancel{\Delta \underline{V}}}{\underline{V}} \Rightarrow \boxed{P_{eq} = \frac{N k_B T}{\underline{V}}}$$

Grand-Canonical ($\mu \underline{V} T$) Monte Carlo

$$\mathcal{P}_v \propto \exp(-\beta \underline{U}_v + \beta \mu N_v)$$

Particles are created/destroyed to sample fluctuations in N



Since particles have access to complete volume, must take into account "entropic" term:

$$\mathcal{P}_v \propto \frac{V^N}{N!} \exp(-\beta \underline{U}_v + \beta\mu N_v)$$

$$\frac{P_{new}}{P_{old}} = \begin{cases} \text{add} & \dfrac{V}{N+1} \exp(-\beta \Delta \underline{U} + \beta\mu) \\ \text{delete} & \dfrac{N}{V} \exp(-\beta \Delta \underline{U} - \beta\mu) \end{cases}$$

$$\mathcal{P}_{acc} = \min\left\{1, \frac{P_{new}}{P_{old}}\right\}$$

(as usual)

Acceptance of addition/removal step cannot be controlled!

Let's obtain the chemical potential-density relationship for a system of non-interacting particles:

$$\mathcal{P}^{+}_{acc} = \mathcal{P}^{-}_{acc} \Rightarrow \quad 1 = \frac{N_{eq}}{V} \exp(-\beta \cancel{\Delta \underline{U}} - \beta\mu) \Rightarrow$$

no interactions

$$\Rightarrow \quad \exp(\beta\mu) = \rho_{eq} \Rightarrow \quad \boxed{\beta\mu = \ln \rho_{eq}}$$