

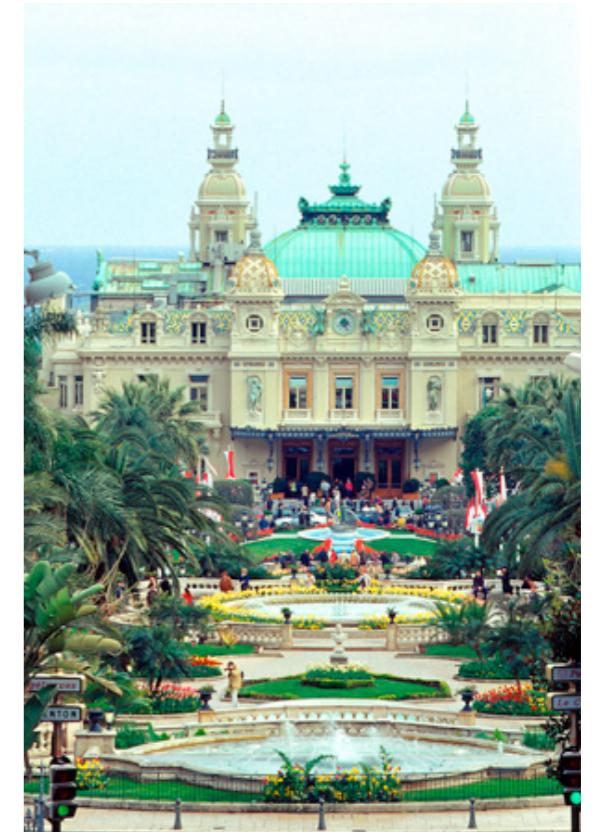
# Monte Carlo and molecular dynamics simulations

# Outline

- What is Monte Carlo simulation and what can it do
- Error sources and how to reduce them
- When is MC better than other methods?
- Molecular Dynamics
- Ensembles
- Project descriptions

# What is Monte Carlo simulation?

- MC is any algorithm based on random numbers
- Used for calculating multiple integrals, generate probability distributions, growth phenomena, etc
- Results are obtained as averages over random samples
- MC can be exact in the sense systematic-error-free
- Results are exact up to statistical errors
- Better than deterministic methods for large number of degrees of freedom

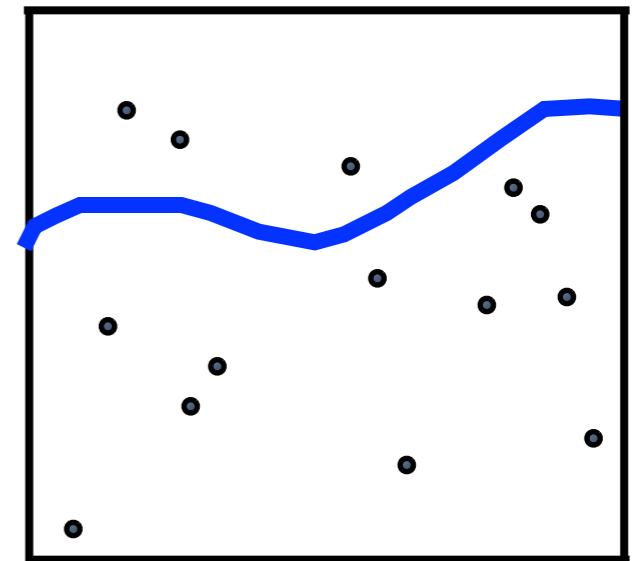


# Hit-and-miss method

- To calculate an integral as the area under the graph of a function:

1. Inscribe the graph in a box
2. Generate N random points in box
3. Each point hits or misses target area
4. Integral estimated from area ratio

$$\frac{\text{integral}}{\text{box area}} \sim \frac{\text{n.o. hits in target area}}{N}$$



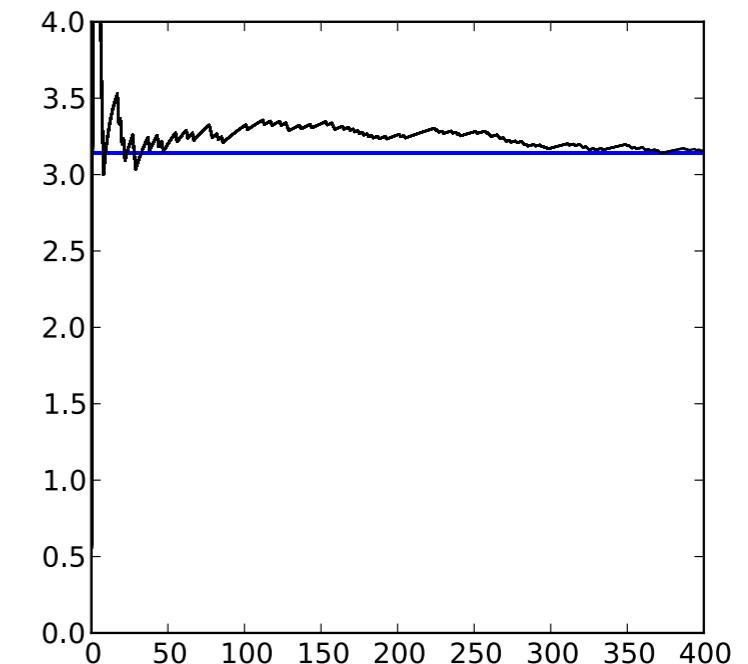
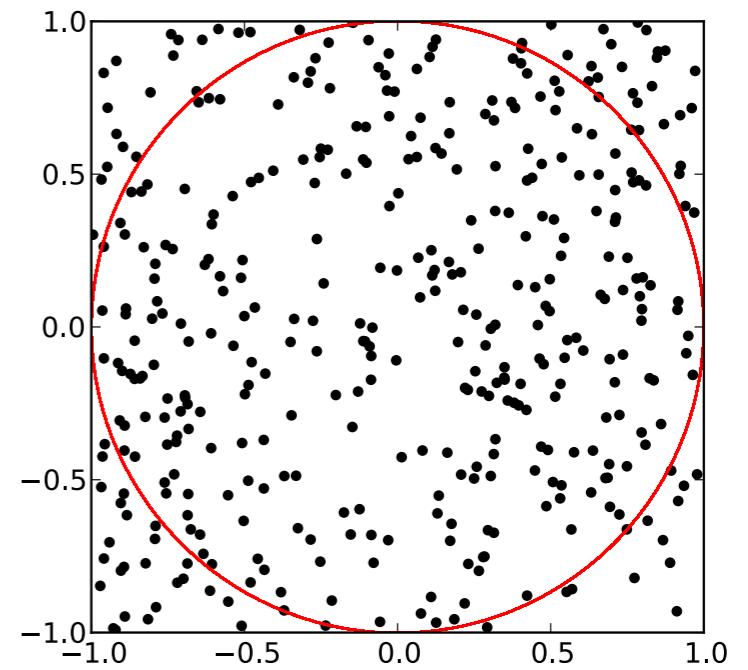
- To make this into a practical method we need to understand:
  5. how to generalize to calculate any integral
  6. how to calculate the statistical error
  7. what to do if the error is large

# Example: calculate pi

- Inscribe a circle in a square
- Throw stones/darts at random
- Estimate pi from area ratio

$$\pi = \frac{4 \text{ circle area}}{\text{square area}} \sim 4 \frac{\text{hits in circle}}{\text{number of points}} + \text{statistical error}$$

- Requires throws distributed uniformly in the integration volume, which can be done easily in a box geometry but not in general



# Simple MC

- Calculate area as average of random samples

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(x) dx = \frac{1}{N} \sum_{i=1}^N f(x_i) \pm \frac{\sigma}{\sqrt{N}}$$

$$\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2$$

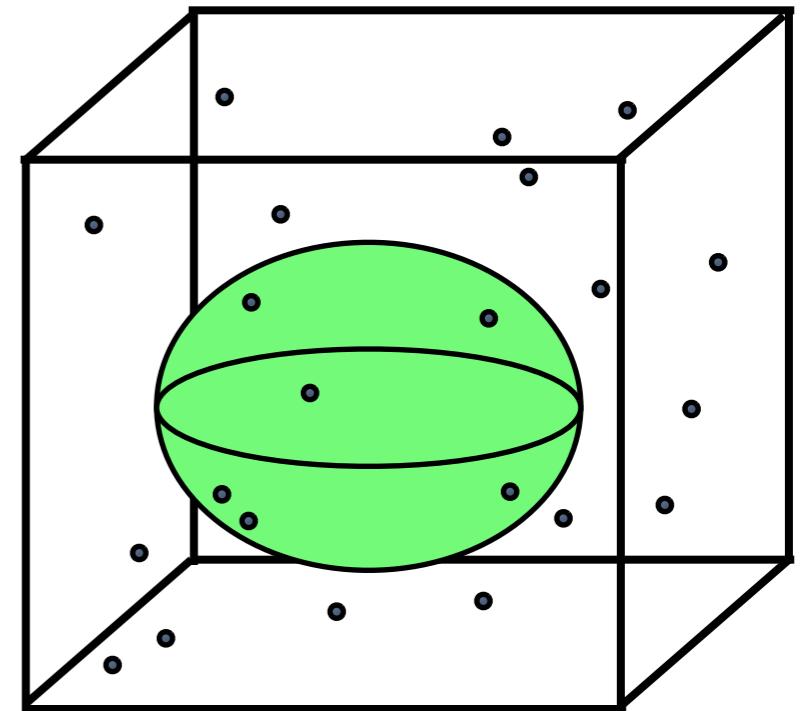
- Here  $x_i$  are independent uniform random numbers in  $[a,b]$
- Convergence is ok if the variance is small  
Otherwise convergence can be poor
- Not clear when the slow convergence is useful even with small variance!

# The hit-and-miss method can calculate any volume integral

$$\langle f \rangle = \frac{1}{V_{\text{box}}} \int f(\vec{r}) d^d r = \frac{1}{N} \sum_{i=1}^N f(\vec{r}_i) \pm \frac{\sigma}{\sqrt{N}}$$
$$\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2$$

set  $f=0$  outside  $V$

- Big variance main source of bad convergence
- Two sources of big variance:
  - I. Box volume too big
    - Try to make it similar to the integration volume
    - The difficulty is to generate random points uniformly distributed in the volume
  2. Function varies too rapidly (next topic)



# Standard error of the mean

$$\begin{aligned}\Delta^2 &= \left\langle \left[ \left( \frac{1}{N} \sum_{i=1}^N f_i \right) - \langle f \rangle \right]^2 \right\rangle = \\ &= \frac{1}{N^2} \sum_{i,j=1}^N \langle f_i f_j \rangle - \frac{2}{N} \sum_{i,j=1}^N \langle f_i \rangle \langle f_j \rangle + \langle f \rangle^2 = \\ &= \frac{1}{N^2} \sum_{i,j=1}^N \langle f_i \rangle \langle f_j \rangle + \frac{1}{N^2} \sum_{i=1}^N (\langle f^2 \rangle - \langle f \rangle^2) - 2\langle f \rangle^2 + \langle f \rangle^2 = \\ &= \frac{1}{N^2} \sum_{i=1}^N (\langle f^2 \rangle - \langle f \rangle^2) = \frac{\langle f^2 \rangle - \langle f \rangle^2}{N} = \frac{\sigma^2}{N}\end{aligned}$$

$$\Delta = \frac{\sigma}{\sqrt{N}}$$

$$\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2 \approx \frac{1}{N} \sum_{i=1}^N f_i^2 - \left( \frac{1}{N} \sum_{i=1}^N f_i \right)^2$$

- Note the crucial assumption of independent random samples:  $i \neq j \Rightarrow \langle f_i f_j \rangle = \langle f_i \rangle \langle f_j \rangle$
- For dependent random samples we instead have

$$\Delta = \frac{\sigma}{\sqrt{N_{\text{independent}}}}$$

# When is MC better than deterministic methods?

## Integration schemes

$$R_N = \sum_{i=1}^N f(x_i) \Delta x$$

$$M_N = \sum_{i=1}^N f\left(\frac{1}{2}[x_i + x_{i+1}]\right) \Delta x$$

$$\begin{aligned} T_N &= \sum_{i=1}^N \frac{1}{2} [f(x_i) + f(x_{i+1})] \Delta x = \\ &= \left[ \frac{1}{2} f(x_1) + \frac{1}{2} f(x_N) + \sum_{i=2}^{N-1} f(x_i) \right] \Delta x \end{aligned}$$

$$\begin{aligned} S_N &= \frac{1}{3} [f(x_1) + 4f(x_2) + 2f(x_3) + \dots \\ &\quad + 2f(x_{N-2}) + 4f(x_{N-1}) + f(x_N)] \Delta x \end{aligned}$$

$$= \left[ f(x_1) + f(x_N) + 4 \sum_{i=2,4,\dots}^{N-1} f(x_i) + 2 \sum_{i=3,5,\dots}^{N-2} f(x_i) \right] \Delta x$$

- In deterministic methods the convergence goes as  $N^{-k}$   
 $N=n.o.$  function evaluations
- Rectangular, midpoint methods:  $k=1$
- Trapez, Simpson methods:  $k=2$
- d-dimensional integral: error  $\sim N^{-k/d}$   
 $N^{1/d}$  points used in each dimension
- MC converges faster when  
 $N^{-k/d} > N^{-1/2}$  or  $d > 2k$
- MC wins for  $d > 4$
- In practice MC is the only good method for integrals with a large number of degrees of freedom
- Example: statistical mechanics

# Derivation of integration errors

- Integration in one mesh interval by Taylor's formula:

$$f(x_i + x) = f(x_i) + f'(x_i)x + \frac{1}{2} f''(x_i)x^2 + \frac{1}{3!} f'''(x_i)x^3 + \dots = \sum c_m x^m$$
$$\Rightarrow \int_{x_i}^{x_{i+1}} f dx = \sum_{m=0} d_m \Delta x^{m+1} \quad \text{where } \Delta x = x_{i+1} - x_i \text{ and } c_m, d_m \text{ are different constants}$$

- The rectangular and midpoint methods integrate a constant function exactly
- The error for one subinterval is therefore  $\sim \Delta x^2 \sim I/N^2$
- The error summed up over N intervals is  $\sim N/N^2 = I/N$
- The Trapezoid method integrates linear functions exactly
- The error for one subinterval is therefore  $\sim \Delta x^3 \sim I/N^3$
- The error summed up over N intervals is  $\sim N/N^3 = I/N^2$
- Simpson's rule can be obtained from Trapezoid by  $S_N = (4T_{2N} - T_N)/3$
- The error is therefore  $\sim I/N^2$  as for Trapezoid

# Metropolis algorithm

- The Metropolis algorithm is used to calculate averages of the form

$$\langle f \rangle = \frac{\int f(x)P(x)dx}{\int P(x)dx}$$

- Construct a sequence of random points that converges towards the target distribution
- The sequence of points is constructed as a Markov process:  
a random walk defined by constructing a transition probability  $T_{ij}$  from one value  $x_i$  to another value  $x_j$
- The distribution of points approaches  $P(x)$  after many steps
- A common choice is

$$T_{ij} = \min \left\{ 1, \frac{P(x_j)}{P(x_i)} \right\}$$

Metropolis algorithm:

- Choose a trial point  $x_j = x_i + d_i$  where  $d_i$  is a uniform random number in  $[-\delta, \delta]$
- If  $w = P(x_j)/P(x_i) > r$  where  $r$  is a random number in  $[0, 1]$ , then accept the change and set  $X_{i+1} = X_j$
- Otherwise reject the change and set  $X_{i+1} = X_i$
- Repeat from step 1 until enough points have been generated
- Skip  $N_0$  initial points to approach the target distribution  $P(x)$  and form averages as

$$\langle f \rangle \approx \frac{1}{N} \sum_{i=N_0}^{N_0+N} f(x_i)$$

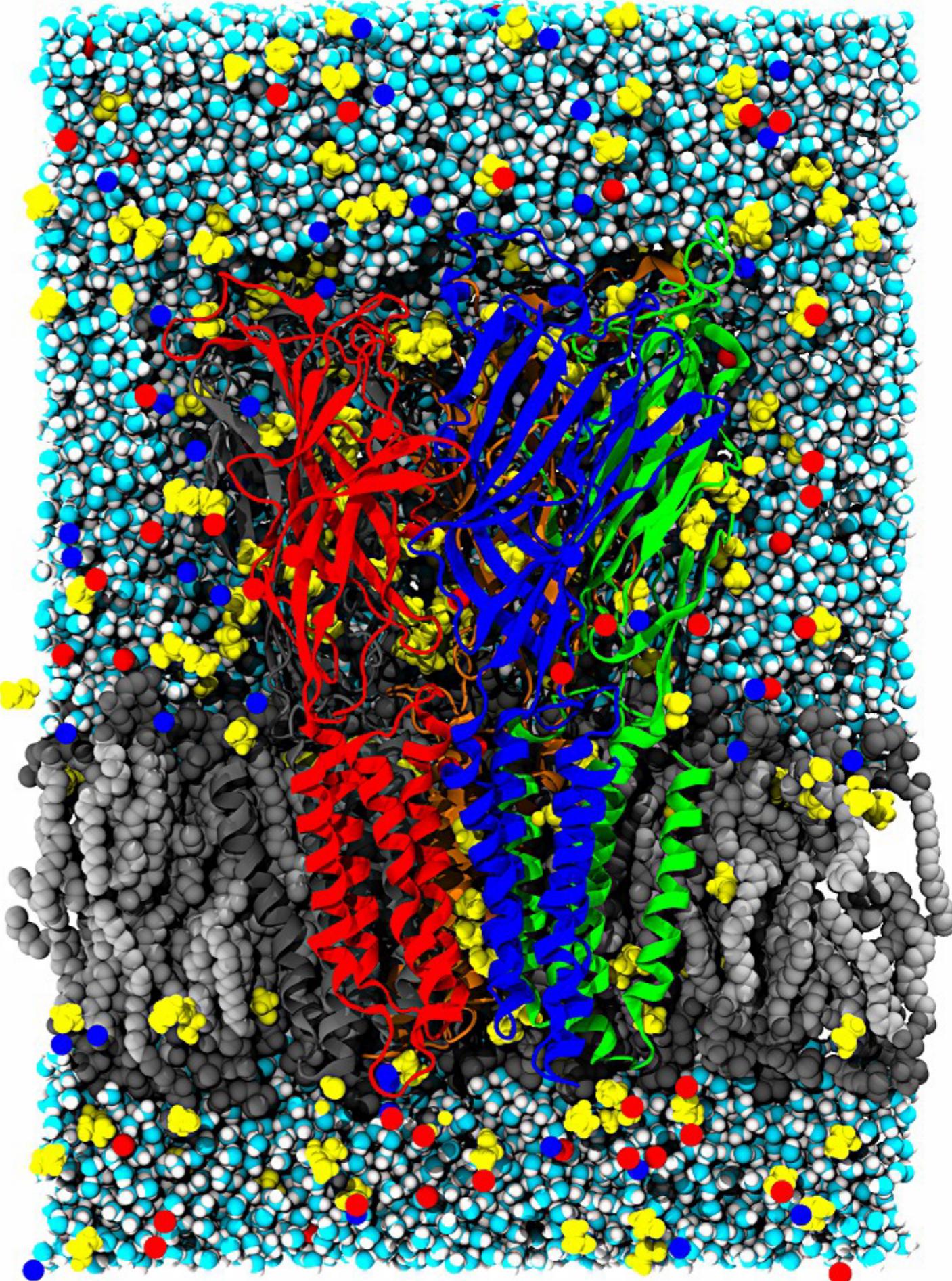
- A complication is that successive points are correlated which means that the statistical error is  $> \sigma / \sqrt{N}$

# Correctness of Metropolis MC

- Requirements:
  - existence of stationary distribution
  - uniqueness of stationary distribution
- Simpler, but stricter requirement: detailed balance:
$$P(x'|x)P(x) = P(x',x) = P(x|x')P(x')$$
- Rewrite:
$$\frac{P(x'|x)}{P(x|x')} = \frac{P(x')}{P(x)}$$
- Split process into proposal distribution  $g$  and acceptance ratio  $A$ :
$$P(x'|x) = g(x'|x)A(x',x)$$
$$\frac{A(x',x)}{A(x,x')} = \frac{P(x')}{P(x)} \frac{g(x|x')}{g(x'|x)}$$
- Insert:
- Choose acceptance criterion
$$A(x',x) = \min \left( 1, \frac{P(x')}{P(x)} \frac{g(x|x')}{g(x'|x)} \right)$$

# Molecular Dynamics

Simulation of biomolecules can aid understanding of diseases and help design new drugs



# Newton's equation of motion

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = \mathbf{f}_i(\mathbf{x}) = -\nabla_i V(\mathbf{x})$$

- Can be used for simulating any type of (classical) particle dynamics, ranging from:
  - dynamics of atoms and electrons
  - to celestial dynamics

# Molecular interactions

Lennard-Jones potential models

Van der Waals interactions

Between atom i and j:

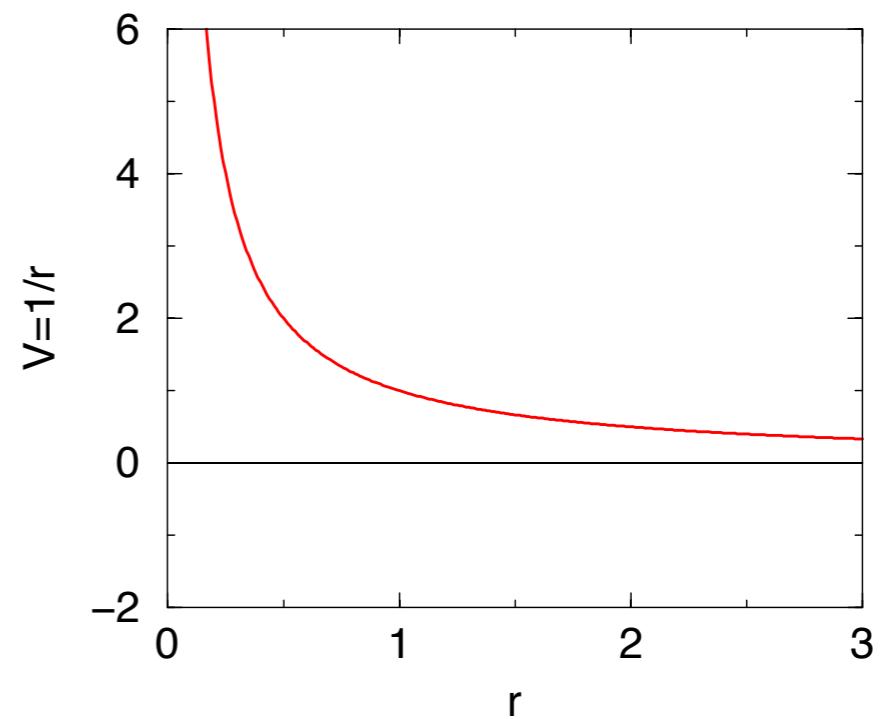
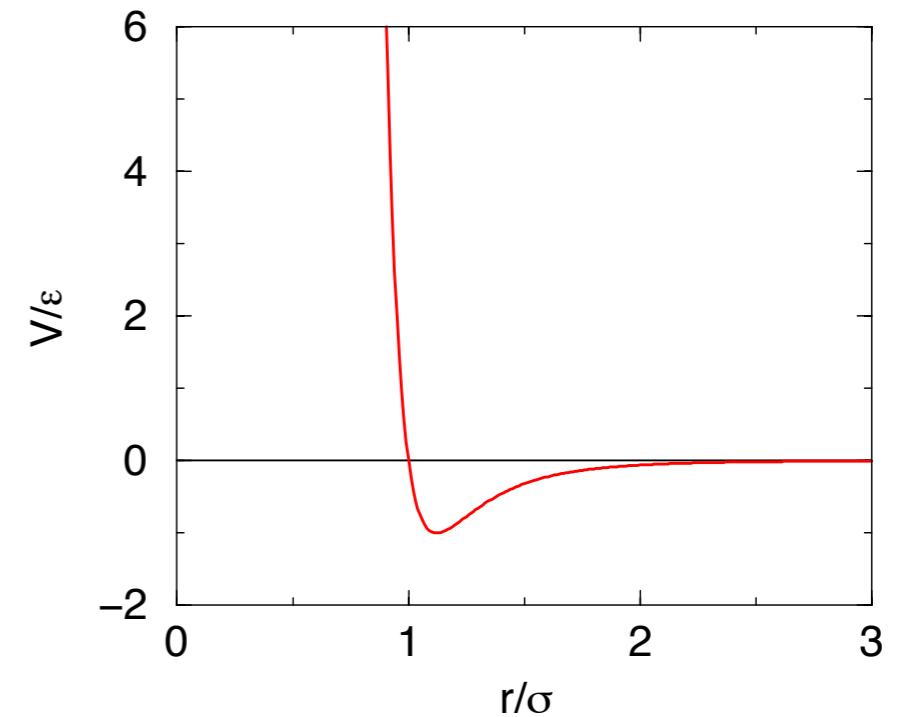
$$V_{ij}(r_{ij}) = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

Coulomb potential models

electrostatics interactions

Between atom i and j:

$$V_{ij}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$



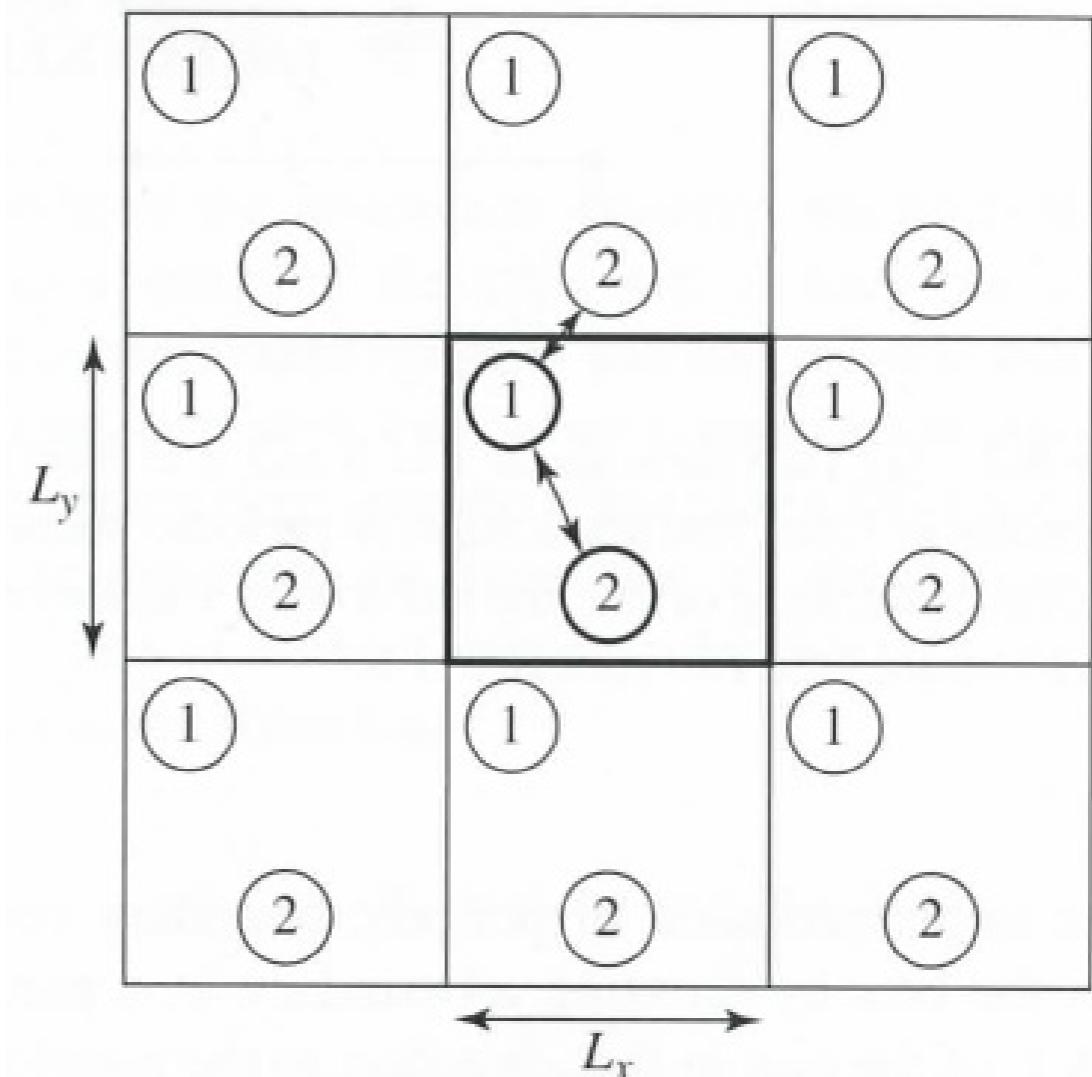
# Periodic boundary conditions

We can only simulate finite, small systems.

We want to avoid boundaries around the system, since boundary effects would dominate in most cases.

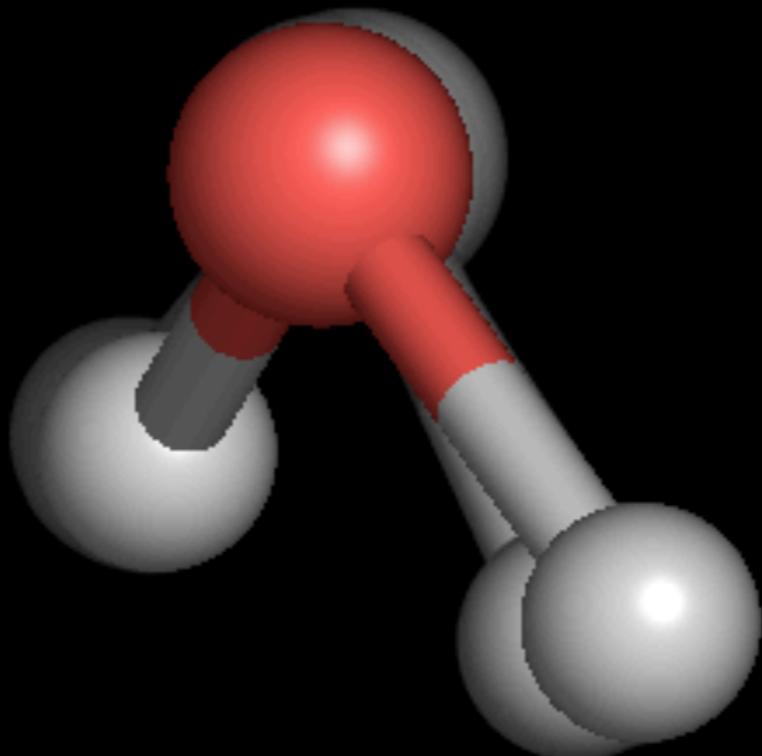
A solution:  
periodic boundary conditions

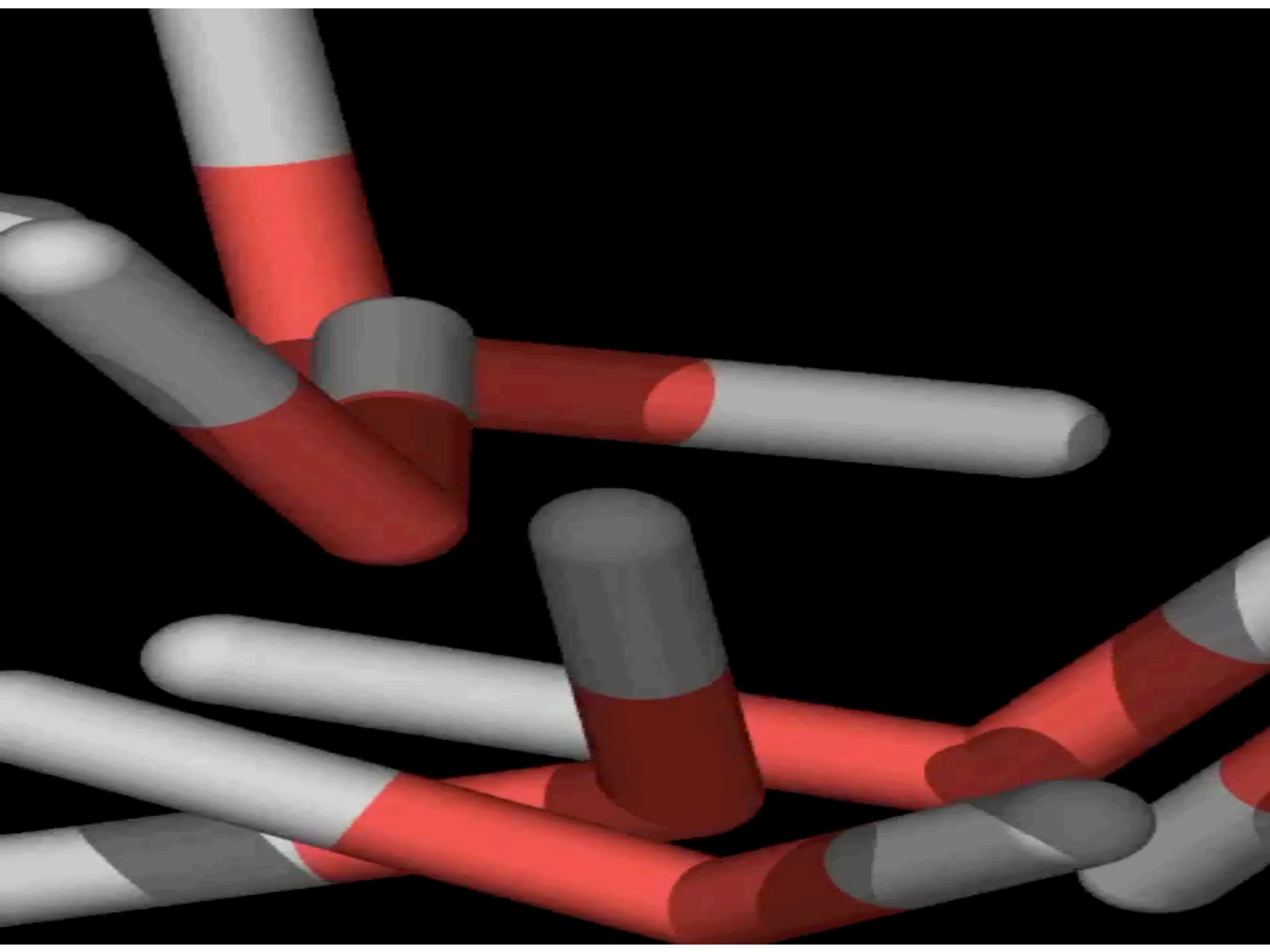
in the example here, particle 1 interacts with particle 2 over the periodic boundary (shorter distance than within the same cell)



# Taking some time steps

8 fs





# MD is chaotic

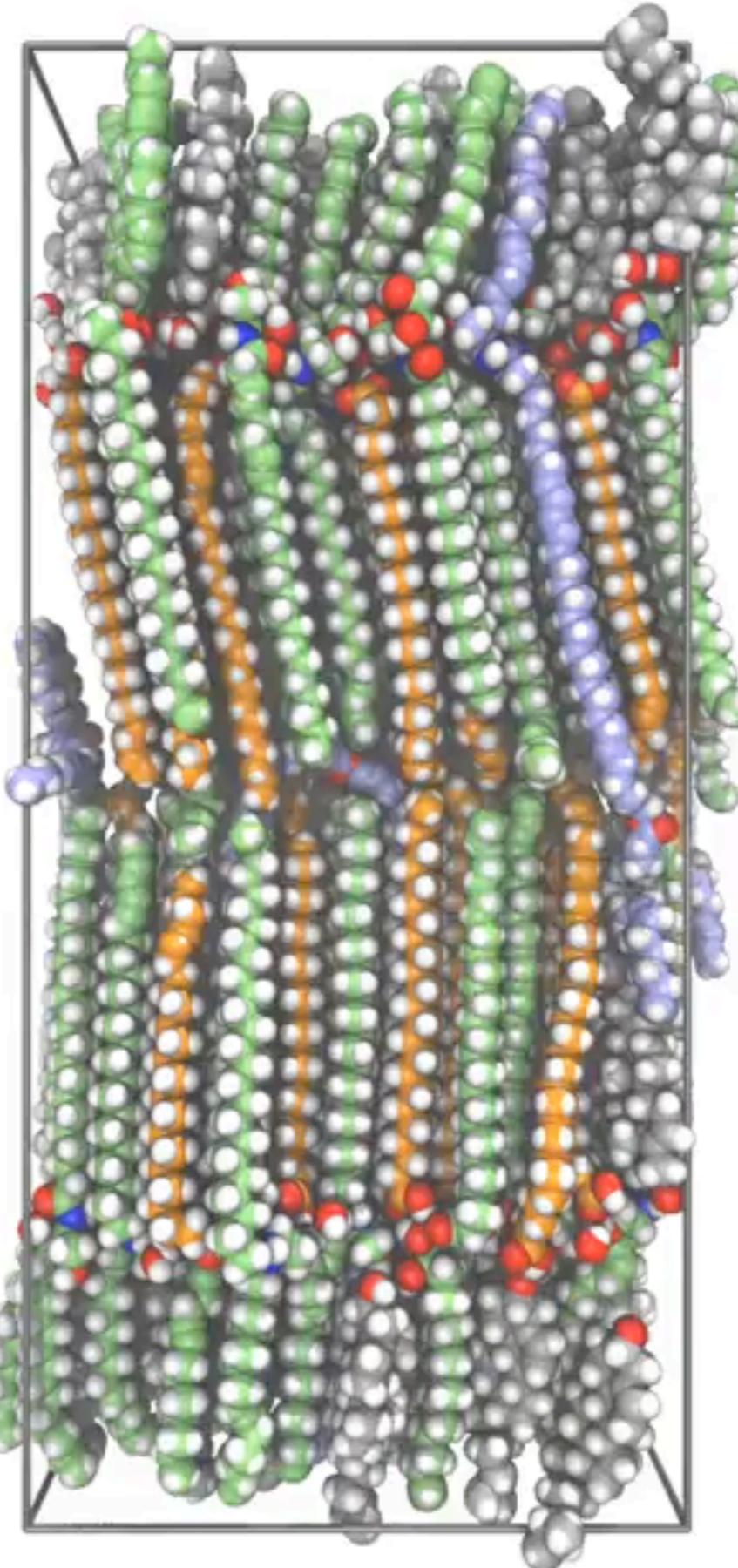
- Since MD involves non-linear equations for many particles, you can expect chaotic behavior
- However, we are almost never interested in following the trajectory if one (or more) particle over a long time.
- Rather we are interested in more global equilibrium and/or dynamic properties.

# Studying equilibrium properties with MD or MC

- Molecular dynamics simulations are often used to study equilibrium properties:
  - Structures of liquids, crystals, polymer melts, proteins
  - Structures of interfaces, e.g. water liquid-vapor interfaces
  - Distributions of, for instance, ions in liquids
  - Structural fluctuations of proteins and polymers (this is not dynamics!)
  - Phase diagrams (Monte Carlo is often better suited)
  - ...
- For all this the “dynamics” part of MD is only used for sampling

# Skin

- Your skin forms a barrier for molecules exiting or entering your body
- But it would be useful to administer drugs through the skin
- How do molecules move through the skin?

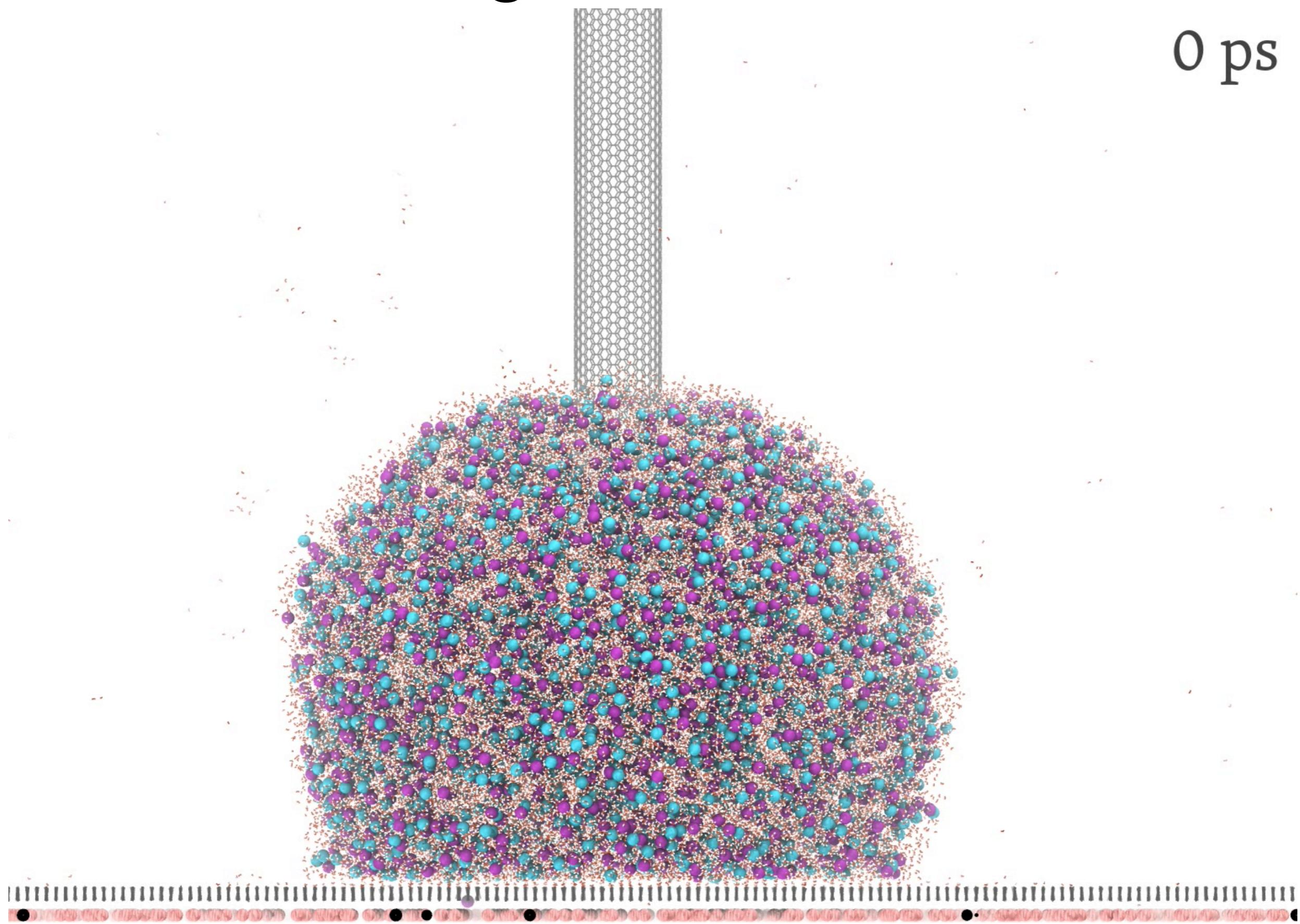


# Studying dynamic properties with MD

- Some examples of Molecular dynamics used for dynamics:
  - Diffusion
  - Viscosity
  - Dynamics of phase transitions
  - Figuring out how proteins work in our bodies
  - Drugs binding to proteins (more kinetics than dynamics)
  - Liquids spreading on surfaces
  - ...

# Electrowetting of water + Na<sup>+</sup> + Cl<sup>-</sup>

0 ps



# Ensembles in statistical mechanics

- Microcanonical ensemble: system isolated from environment, energy  $E=\text{constant}$ , particle number  $N=\text{constant}$
- Canonical ensemble: divide a microcanonical ensemble into two parts, called the "system" and the "heat bath or reservoir", that can exchange energy but not particles.  $E\neq\text{constant}$ , instead  $T=\text{constant}$
- Grand canonical ensemble: the system and reservoir can exchange energy and particles.

# Microcanonical ensemble

- Microcanonical ensemble: system isolated from environment  
energy  $E=\text{constant}$ , particle number  $N=\text{constant}$
- Basic postulate of statistical mechanics (Gibbs): all accessible states are equally probable

$$P_s = \begin{cases} 1/\Omega & \text{if state } s \text{ is accessible} \\ 0 & \text{else} \end{cases}, \quad \Omega = \text{number of accessible states}$$

- Ensemble average of  $A$ :  $\langle A \rangle = \sum_{s=1}^{\Omega} A_s P_s$   
(ergodic hypothesis: ensemble average = time average)

# Canonical ensemble

- Canonical ensemble: divide a microcanonical ensemble into two parts, called the "system" and the "heat bath or reservoir", that can exchange energy, but not particles.
- The temperature  $T$ , equilibrium Boltzmann distribution  $P$ , partition function  $Z$ , free energy  $F$ , entropy  $S$ , and ensemble average  $\langle A \rangle$  are given by:

$$P_s = \frac{1}{Z} e^{-\beta E_s}, \quad \beta = 1/kT, \quad Z = \sum_s e^{-\beta E_s}, \quad F = -kT \ln Z = E - TS, \quad S = k \ln \Omega(E), \quad \langle A \rangle = \sum_s A_s P_s$$

# Changing ensemble in MD

- By default molecular dynamics conserves the total energy and thus you get a microcanonical ensemble
- How to obtain a canonical ensemble?  
Add a thermostat
- How to obtain a constant-NpT ensemble?  
Add a thermostat and a barostat
- Note: thermostat and barostat need carefully designed algorithms and implementations to obtain the correct ensemble!

# Thermostats in MD

- Andersen thermostat: Every n-th step (n is a parameter):
  - choose a particle at random or loop over all atoms
  - re-assign random velocities from a Boltzmann distribution, take  $v_{x/y/z}$  from Gaussian distribution with  $\sigma^2 = k_B T$ ,  $\mu = 0$  m
  - Advantage: very simple to implement Disadvantage: strong disruption of dynamics
- Bussi thermostat:
  - Scale the velocities and add noise to the velocities
- Nosé-Hoover thermostat:
  - Add the kinetic energy as a degree of freedom in the Hamiltonian Advantage: smaller disruption of dynamics
  - Disadvantage: can be tricky to implement and control

# Metropolis method: system in contact with a heat bath

- A heat bath is considered to be so much bigger than the system so that it can give or take energy to the system without changing its temperature
- Generate trial moves, accept a move if  $e^{-\Delta E/kT} > [\text{random number in } (0,1)]$
- Non-equilibrium states appear initially in the simulation and show up as a transient in the measured energy and other quantities
- Such transient states must not be included when measuring ensemble average, in order to avoid bias in the estimates of the equilibrium properties
- After the initial transient has decayed, equilibrium is reached and the distribution of energy and other quantities becomes stationary
- In equilibrium the fluctuations of the system are distributed according to the Boltzmann distribution

# Energy fluctuations

- In a thermal system the potential energy fluctuations as the system samples different states/configurations
- The energy fluctuations are measured by the heat capacity, and the relative energy fluctuations tend to zero in the limit of a large number of particles:

$$C_v = \frac{\delta E}{\delta T} = \frac{(\Delta E)^2}{k_B T^2} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} \sim N \Rightarrow \frac{\Delta E}{E} \sim \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \rightarrow 0$$

where the subscript  $v$  on  $C_v$  indicates constant volume, there is also  $C_p$  which is at constant pressure

- Definition of heat capacity: the amount of heat to be supplied to a given mass of a material to produce a unit change in its temperature

# Computing the pressure

- Using the mechanical definition of the pressure as momentum flux, the, instantaneous, pressure is:

$$P = \frac{2}{V}(E_{\text{kin}} - \Xi)$$

where the virial of the force is defined as:

$$\Xi = -\frac{1}{2} \sum_{i < j} F_{ij} r_{ij}$$

(this is the derivative of the potential with respect to the volume)

# Examples of interacting models

- The simplest non-ideal gas model is the hard rod/disk/sphere model in 1/2/3 dimensions respectively
- Consider particles with radius  $a$  without long range interactions but with an infinite hard-core repulsion if they come closer than  $2a$
- At low density the system is a liquid, and at high density a packed solid is obtained
- Computer simulations suggest that there is a solid-liquid transition at some intermediate density for hard disks and spheres
- A common model for a finite-range interacting gas is given by a Lennard-Jones potential:  
$$V(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]$$
- This model can describe solid, liquid and gas phases and the transitions between these phases are conveniently studied by MC

# References

- Courses at physics:
  - Computational Physics (SI2530) MD+MC (discontinued?)
  - Biophysics (SI2720), focus on proteins, contains MD (labs)
  - 2 listed + 2 non-listed Statistical Physics courses
  - PhD-level: Advanced Molecular Dynamics (SI3450)
- Good book:
  - “Understanding Molecular Simulation” Frenkel & Smit
- Bussi thermostat: <https://arxiv.org/abs/0803.4060>
- Fancy pictures of simulated bio-molecules (our competitor):  
<http://www.ks.uiuc.edu/Gallery/Science-New/>

## Projects: Monte Carlo

3.1) Use the Metropolis method to calculate:

$$\langle x \rangle = \frac{\int_0^\infty x e^{-x} dx}{\int_0^\infty e^{-x} dx}$$

Here  $P(x) = e^{-x}$ . Use probability  $P(x) = 0$  for  $x < 0$ . Try different values of the parameter delta in the range between 0.01 and 10. Study  $\sigma/\sqrt{N}$  and compare with the actual difference to the exact answer. (note:  $N_0$  is not critical when  $x_0 = 0$ )

# Projects: Molecular Dynamics

- 3.2 a) Implement the LJ potential and force (with  $\sigma = 1$  and  $\epsilon = 1$ ) in the template. A correct implementation should give good conservation of the total energy. Use a temperature of 1 and increase the time step until things go very wrong. Study the quality of the integration by monitoring the drift in the total energy for several different time steps just before things go wrong.
- 3.2 b) Run a LJ simulation with initial velocities at a temperature of 0.2. What happens with the kinetic and potential energy?
- 3.2 c) Implement an Andersen thermostat that thermalises all particles simultaneously at a fixed step interval and run simulations at  $T = 1$  and 0.2. What differences in collective behavior do you observe between 1 and 0.2 at long times? A physics question: can you explain what you see?
- 3.2 d) Use MD with the thermostat to calculate the average energy and heat capacity between  $T = 0.2$  and 1. Make sure the results are sufficiently converged (choose your own definition of sufficient). Can you explain the behavior of the heat capacity from the behavior of the system that you observe in the animations?

## Projects: Molecular Dynamics (continued)

3.2 e) Now we will look at the pressure. Can you explain the differences in pressure between the different temperatures? Also run with a 4 times as large unit cell by doubling L. What happens to the pressure at different T? How does the pressure compare with that of an ideal gas?